Lu et al. *Journal of Inequalities and Applications* (2016) 2016:62
DOI 10.1186/s13660-016-0994-3

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH

Open Access

New *a posteriori* error estimates for *hp* version of finite element methods of nonlinear parabolic optimal control problems

Zuliang Lu[1,2,3*], Hongyan Liu[4], Chunjuan Hou[5] and Longzhou Cao[1]

*Correspondence: zulianglux@126.com
[1]Key Laboratory of Signal and Information Processing, Chongqing Three Gorges University, Chongqing, 404000, P.R. China
[2]Key Laboratory for Nonlinear Science and System Structure, Chongqing Three Gorges University, Chongqing, 404000, P.R. China
[3]Research Center for Mathematics and Economics, Tianjin University of Finance and Economics, Tianjin, 300222, P.R. China
Full list of author information is available at the end of the article

Abstract

In this paper, we investigate residual-based *a posteriori* error estimates for the *hp* version of the finite element approximation of nonlinear parabolic optimal control problems. By using the *hp* finite element approximation for both the state and the co-state variables and the *hp* discontinuous Galerkin finite element approximation for the control variable, we derive *hp* residual-based *a posteriori* error estimates for both the state and the control approximation. Such estimates, which are apparently not available in the literature, can be used to construct a reliable *hp* adaptive finite element approximation for the nonlinear parabolic optimal control problems.

MSC: 49J20; 65N30

Keywords: residual-based *a posteriori* error estimates; nonlinear parabolic optimal control problems; *hp* version of finite element method

1 Introduction

Recently, optimal control problems have attracted substantial interest due to their applications in atmospheric pollution problems, exploration and extraction of oil and gas resources, and engineering. They must be solved with efficient numerical methods. The *hp* version of the finite element method is an important numerical method for the optimal control problems governed by partial differential equations. There have been extensive studies of the convergence of the finite element approximation for optimal control problems. Let us mention two early papers devoted to linear optimal control problems by Falk [1] and Geveci [2]. A systematic introduction of the finite element method for optimal control problems can be found in [3–5], but there are much less published results on this topic for *hp*-finite element methods for optimal control problems. The adaptive finite element method has been investigated extensively and became one of the most popular methods in the scientific computation and numerical modeling. In [6], the authors studied *a posteriori* error estimates for adaptive finite element discretizations of boundary control problems. *A posteriori* error estimates and adaptive finite element approximation for parameter estimation problems have been obtained in [7, 8]. The adaptive finite element approximation is among the most important means to boost the accuracy and efficiency of the finite element discretization. There are three main versions in adaptive finite element approximation, *i.e.*, the *p*-version, *h*-version, and *hp*-version. The *p*-version of finite

Springer

element methods uses a fixed mesh and improves the approximation of the solution by increasing the degrees of piecewise polynomials. The h-version is based on mesh refinement and piecewise polynomials of low and fixed degrees. In the hp-version adaptation, one has the option to split an element (h-refinement) or to increase its approximation order (p-refinement). Generally speaking, a local p-refinement is the more efficient method on regions where the solution is smooth, while a local h-refinement is the strategy suitable on elements where the solution is not smooth in [9]. There have been many theoretical studies as regards the hp finite element method in [9–16].

To the best of our knowledge, there are many h-versions of adaptive finite element methods for optimal control problems in [17, 18]. In fact, comparable literature for high order elements such as the hp-version of the finite element method for optimal control problems is rather limited. For the constrained optimal control problem governed by elliptic equations, the authors have derived *a posteriori* error estimates for the hp finite element approximation in [19]. The purpose of this work is to derive *hp a posteriori* error estimates for optimal control problems governed by nonlinear parabolic equations.

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 $\|v\|_{m,p}^p = \sum_{|\alpha|\le m}\|D^\alpha v\|_{L^p(\Omega)}^p$, 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 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 $s = \infty$. 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 [20].

The paper is organized as follows. In Section 2, we shall construct the hp finite element approximation for nonlinear parabolic optimal control problems. In Section 3, we derive *hp a posteriori* error estimates for the optimal control problems. In the last section, we briefly give conclusions and some possible future work.

2 The *hp* finite element of nonlinear parabolic optimal control

In this section, we study the hp finite element method and the backward Euler discretization approximation of convex optimal control problems governed by nonlinear parabolic equations. We shall take the state space $W = L^2(0,T;Y)$ with $Y = H_0^1(\Omega)$, the control space $X = L^2(0,T;U)$ with $U = L^2(\Omega_U)$ and $H = L^2(\Omega)$ to fix the idea. Let B be a linear continuous operator from X to $L^2(0,T;Y)$. We are interested in the following nonlinear parabolic optimal control problems:

$$\min_{u\in K}\left\{\frac{1}{2}\int_0^T \left(\|y-y_d\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega_U)}^2\right)dt\right\}, \tag{2.1}$$

$$y_t - \operatorname{div}(A\nabla y) + \phi(y) = f + Bu, \quad x \in \Omega, t \in (0,T], \tag{2.2}$$

$$y(x,t) = 0, \quad x \in \partial\Omega, t \in (0,T], \tag{2.3}$$

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

where Ω and Ω_U are bounded open sets in $\mathbb{R}^2$ with a Lipschitz boundary $\partial\Omega$ and $\partial\Omega_U$, K is a set defined by $K = \{v \in X : \int_0^T\int_{\Omega_U} v\,dx\,dt \ge 0\}$, and $f, y_d \in L^2(0,T;H)$, $y_0(x) \in V = H_0^1(\Omega)$, and $A(\cdot) = (a_{i,j}(\cdot))_{2\times 2} \in (C^\infty(\overline{\Omega}))^{2\times 2}$, such that there is a constant $c > 0$ satisfying

$\xi^t A\xi \geq c\|\xi\|^2$, $\xi \in \mathbb{R}^2$. The function $\phi(\cdot) \in W^{2,\infty}(-R,R)$ for any $R > 0$, $\phi'(y) \in L^2(\Omega)$ for any $y \in H^1(\Omega)$, and $\phi'(y) \geq 0$.

Let $a(v,w) = \int_\Omega (A\nabla v) \cdot \nabla w\,dx$, $\forall v,w \in V$, $(f_1,f_2) = \int_\Omega f_1 f_2\,dx$, $\forall f_1,f_2 \in H$, $(v,w)_U = \int_{\Omega_U} vw\,dx$, $\forall v,w \in U$. It follows from the assumptions on A that there are constants c and $C > 0$ such that

$$a(v,v) \geq c\|v\|^2_{H^1(\Omega)}, \qquad \bigl|a(v,w)\bigr| \leq C|v|^2_{H^1(\Omega)}|w|^2_{H^1(\Omega)}, \quad \forall v,w \in Y.$$

Then a weak formula of the convex nonlinear parabolic optimal control problems (2.1)-(2.4) reads

$$\min_{u(t)\in K} \left\{ \frac{1}{2} \int_0^T \bigl(\|y - y_d\|^2_{L^2(\Omega)} + \|u\|^2_{L^2(\Omega_U)} \bigr)\,dt \right\}, \tag{2.5}$$

where $y \in W$, $u \in X$, $u(t) \in K$, subject to

$$(y_t, w) + a(y,w) + \bigl(\phi(y), w\bigr) = (f + Bu, w), \quad \forall w \in Y, t \in (0,T], \tag{2.6}$$

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

It is well known (see, *e.g.*, [21]) that the optimal control problem (2.5)-(2.7) has at least a solution (y,u), and if that a pair (y,u) is the solution of (2.5)-(2.7), then there is a co-state $p \in W$ such that the triplet (y,p,u) satisfies the following optimality conditions:

$$(y_t, w) + a(y,w) + \bigl(\phi(y), w\bigr) = (f + Bu, w), \quad \forall w \in Y, y(0) = y_0(x), \tag{2.8}$$

$$-(p_t, w) + a(q,p) + \bigl(\phi'(y)p, q\bigr) = (y - y_d, q), \quad \forall q \in Y, p(T) = 0, \tag{2.9}$$

$$\int_0^T \bigl(u + B^*p, v - u\bigr)_U\,dt \geq 0, \quad \forall v \in K, \tag{2.10}$$

where B^* is the adjoint operators of B, and $(\cdot,\cdot)_U$ is the inner product of U, which will be simply written as $(\cdot,\cdot)$ in the rest of the paper when no confusion is caused.

Due to the special structure of the control constraint set K, we can derive a relationship between the control variable and the co-state variable of (2.8)-(2.10) in the following lemma.

Lemma 2.1 *Let* (y,p,u) *be the solution of* (2.8)-(2.10). *Then we have*

$$u = \max\bigl\{0, \overline{B^*p}\bigr\} - B^*p,$$

where $\overline{B^*p} = \frac{\int_0^T \int_{\Omega_U} B^*p\,dx\,dt}{\int_0^T \int_{\Omega_U} 1\,dx\,dt}$ *denotes the integral average on* $\Omega_U \times [0,T]$ *of the function* B^*p.

Now, we consider the *hp* finite element approximation for the nonlinear parabolic optimal control problems. We assume that Ω and Ω_U are polygonal. We consider the triangulation $\mathcal{T}$ of the set $\Omega \subset \mathbb{R}^2$, which is a collection of elements $\tau \in \mathcal{T}$ associated with each element τ, and an affine element map $F_\tau : \widehat{\tau} \to \tau$, where the reference element $\widehat{\tau}$ is the reference triangle

$$\bigl\{(x,y) \in \mathbb{R}^2 : 0 < x < 1, 0 < y < \min(x, 1 - x)\bigr\}.$$

We consider the triangulation $\mathcal{T}$ which satisfies the standard conditions defined in [16] and write $h_\tau = \operatorname{diam}\tau$. Additionally we assume that triangulation $\mathcal{T}$ is γ-shape regular, *i.e.*,

$$h_\tau^{-1}\|F'_\tau\|_{L^\infty(\widehat{\tau})} + h_\tau\|(F'_\tau)^{-1}\|_{L^\infty(\widehat{\tau})} \le \gamma. \tag{2.11}$$

This implies that there exists a constant $C > 0$ that depends solely on γ such that

$$C^{-1}h_\tau \le h_{\tau'} \le Ch_\tau, \quad \tau, \tau' \in \mathcal{T} \text{ with } \overline{\tau} \cap \overline{\tau}' \neq \emptyset, \tag{2.12}$$

and there exists a constant $M \in \mathbb{N}$ that depends solely on γ such that no more than M elements share a common vertex. We assume that the triangulation $\mathcal{T}_U$ of Ω_U which is a collection of elements $\tau_U \in \mathcal{T}_U$, is γ-shape regular which satisfies the standard conditions as $\mathcal{T}$. Associated with each element τ_U is an affine element map $F_{\tau_U} : \widehat{\tau} \to \tau_U$. We further assume the triangulation $\mathcal{T}$ satisfies the relation between the patch and the reference patch in [16].

For each element $\tau \in \mathcal{T}$, we denote $\mathcal{E}(\tau)$ the set of edges of τ and $\mathcal{N}(\tau)$ the set of vertices of τ, and choose a polynomial degree $p_\tau \in \mathbb{N}$ and collect these numbers in the polynomial degree vector $\mathbf{p}_1 = (p_\tau)_{\tau\in\mathcal{T}}$. Similarly, for each element $\tau_U \in \mathcal{T}_U$, we choose a polynomial degree vector $\mathbf{p}_2 = (p_{\tau_U})_{\tau_U\in\mathcal{T}_U}$ ($p_{\tau_U} \in \mathbb{N}$). $\mathcal{N}(\mathcal{T})$ denotes the set of all vertices of $\mathcal{T}$, $\mathcal{E}(\mathcal{T})$ denotes the set of all edges. Additionally, we introduce the following notation ($V \in \mathcal{N}(\mathcal{T})$, $e \in \mathcal{E}(\mathcal{T})$):

$$\begin{aligned}
&\mathcal{N}(e) = \{V \in \mathcal{N}(\mathcal{T}) : V \in \overline{e}\},\\
&w_V = \{x \in \Omega : x \in \overline{\tau} \text{ and } \overline{\tau} \cap \{V\} \neq \emptyset\}^0,\\
&w_e^1 = \bigcup_{V\in\mathcal{N}(e)} w_V, \qquad w_\tau^1 = \bigcup_{V\in\mathcal{N}(\tau)} w_V,\\
&h_{\tau_U} = \operatorname{diam}\tau_U, \qquad p_e = \max\{p_\tau : e \in \mathcal{E}(\tau)\},
\end{aligned}$$

where χ^0 denotes the interior of the set χ. Finally, we denote by h_e the length of the edge e.

Next, we define the *hp*-finite element space $S^{\mathbf{p}_1}(\mathcal{T}) \subset H^1(\Omega)$ and the *hp*-discontinuous Galerkin finite element space $U^{\mathbf{p}_2}(\mathcal{T}_U) \subset L^2(\Omega_U)$ by

$$\begin{aligned}
&S^{\mathbf{p}_1}(\mathcal{T}) = \{v \in C(\Omega) : v|_\tau \circ F_\tau \in \Pi_{p_\tau}(\widehat{\tau})\},\\
&U^{\mathbf{p}_2}(\mathcal{T}_U) = \{v \in L^2(\Omega_U) : v|_{\tau_U} \circ F_{\tau_U} \in \Pi_{p_{\tau_U}}(\widehat{\tau})\},
\end{aligned}$$

where we set

$$\Pi_k(\widehat{\tau}) = \begin{cases} P_k = \operatorname{span}\{x^iy^j : 0 \le i+j \le k\}, & \text{if } \widehat{\tau} = T,\\ Q_k = \operatorname{span}\{x^iy^j : 0 \le i,j \le k\}, & \text{if } \widehat{\tau} = S.\end{cases}$$

We also assume that the polynomial degree vector $\mathbf{p}_1$ satisfies

$$\gamma^{-1}p_\tau \le p_{\tau'} \le \gamma p_\tau, \quad \tau, \tau' \in \mathcal{T} \text{ with } \overline{\tau} \cap \overline{\tau}' \neq \emptyset. \tag{2.13}$$

Then we can introduce the finite dimensional spaces $K_{hp} = K \cap \mathcal{U}^{\mathbf{p}_2}(\mathcal{T}_U)$, $V_{hp} = V \cap \mathcal{S}^{\mathbf{p}_1}(\mathcal{T})$.

The semidiscrete *hp* finite element approximation of (2.1)-(2.4) is as follows:

$$\min_{u_{hp}\in K_{hp}} \left\{ \frac{1}{2} \int_0^T \left(\|y_{hp} - y_d\|^2_{L^2(\Omega)} + \|u_{hp}\|^2_{L^2(\Omega_U)} \right) dt \right\}, \tag{2.14}$$

$$\left(\frac{\partial y_{hp}}{\partial t}, w_{hp} \right) + a(y_{hp}, w_{hp}) + \left(\phi(y_{hp}), w\right) = (f + Bu_{hp}, w_{hp}), \quad \forall w_{hp} \in V_{hp}, \tag{2.15}$$

$$y_{hp}(x,0) = y_0^{hp}(x), \quad x \in \Omega, \tag{2.16}$$

where $y_{hp} \in H^1(0,T;V_{hp})$ and $y_0^{hp} \in V_{hp}$ is an finite element approximation of y_0.

It follows that the optimal control problems (2.14)-(2.16) has at least a solution (y_{hp}, u_{hp}) and if that a pair (y_{hp}, u_{hp}) is the solution of (2.14)-(2.16), then there is a co-state $p_{hp} \in V_{hp}$ such that the triplet (y_{hp}, p_{hp}, u_{hp}) satisfies the following optimality conditions:

$$\left(\frac{\partial y_{hp}}{\partial t}, w_{hp} \right) + a(y_{hp}, w_{hp}) + \left(\phi(y_{hp}), w\right) = (f + Bu_{hp}, w_{hp}), \quad \forall w_{hp} \in V_{hp}, \tag{2.17}$$

$$y_{hp}(x,0) = y_0^{hp}(x), \quad x \in \Omega,$$

$$-\left(\frac{\partial p_{hp}}{\partial t}, q_{hp} \right) + a(q_{hp}, p_{hp}) + \left(\phi'(y_{hp})p_{hp}, q_{hp}\right) = (y_{hp} - y_d, q_{hp}), \quad \forall w_{hp} \in V_{hp}, \tag{2.18}$$

$$p_{hp}(x,T) = 0, \quad x \in \Omega,$$

$$\left(u_{hp} + B^* p_{hp}, v_{hp} - u_{hp}\right)_U \geq 0, \quad \forall v_{hp} \in K_{hp}. \tag{2.19}$$

Furthermore, we consider the fully discrete finite element approximation for the above semidiscrete problems by using the backward Euler scheme. Let $0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T$, $k_i = t_i - t_{i-1}$, $i = 1,2,\ldots,M$, $k = \max_{1\le i\le M}\{k_i\}$.

For $i = 1,2,\ldots,M$, construct the *hp* finite element approximation spaces $V_{hp}^i \subset H_0^1(\Omega)$ (similar to V_{hp}) on the ith time step. Similarly, we construct the *hp* finite element approximation spaces $K_{hp}^i \subset L^2(\Omega_U)$ (similar to K_{hp}) on the ith time step. The fully discrete *hp* finite element approximation scheme (2.17)-(2.19) is to find $(y_{hp}^i, u_{hp}^i) \in V_{hp}^i \times K_{hp}^i$, $i = 1,2,\ldots,M$, such that

$$\min_{u_{hp}^i\in K_{hp}^i} \left\{ \sum_{i=1}^M \left(\frac{1}{2} \left\| y_{hp}^i - y_d(x,t_i) \right\|^2_{L^2(\Omega)} + \frac{1}{2} \left\| u_{hp}^i \right\|^2_{L^2(\Omega_U)} \right) \right\}, \tag{2.20}$$

$$\begin{aligned} &\left(\frac{y_{hp}^i - y_{hp}^{i-1}}{k_i}, w_{hp} \right) + a\left(y_{hp}^i, w_{hp}\right) + \left(\phi\left(y_{hp}^i\right), w_{hp}\right) \\ &\quad = \left(f(x,t_i) + Bu_{hp}^i, w_{hp}\right), \quad \forall w_{hp} \in V_{hp}^i, \end{aligned} \tag{2.21}$$

$$y_{hp}^0(x) = y_0^{hp}(x), \quad x \in \Omega. \tag{2.22}$$

It follows that the optimal control problem (2.20)-(2.22) has at least a solution (Y_{hp}^i, U_{hp}^i), and if a pair $(Y_{hp}^i, U_{hp}^i) \in V_{hp}^i \times K_{hp}^i$ is the solution of (2.20)-(2.22), then there is a co-state $P_{hp}^{i-1} \in V_{hp}^i$, such that the triplet $(Y_{hp}^i, P_{hp}^{i-1}, U_{hp}^i) \in V_{hp}^i \times V_{hp}^i \times K_{hp}^i$, satisfies the following optimality conditions:

$$\left(\frac{Y_{hp}^i - Y_{hp}^{i-1}}{k_i}, w_{hp}\right) + a\big(Y_{hp}^i, w_{hp}\big) + \big(\phi\big(Y_{hp}^i\big), w_{hp}\big) = \big(f(x,t_i) + BU_{hp}^i, w_{hp}\big), \tag{2.23}$$

$$\forall w_{hp} \in V_{hp}^i \subset V = H_0^1(\Omega), \quad i = 1,2,\ldots,M, \qquad Y_{hp}^0(x) = y_0^{hp}(x), \quad x \in \Omega,$$

$$\left(\frac{P_{hp}^{i-1} - P_{hp}^i}{k_i}, q_{hp}\right) + a\big(q_{hp}, P_{hp}^{i-1}\big) + \big(\phi'\big(Y_{hp}^{i-1}\big)P_{hp}^{i-1}, q_{hp}\big) = \big(Y_{hp}^i - y_d(x,t_i), q_{hp}\big), \tag{2.24}$$

$$\forall q_{hp} \in V_{hp}^i \subset V = H_0^1(\Omega), \quad i = M,\ldots,2,1, \qquad P_{hp}^M(x) = 0, \quad x \in \Omega,$$

$$\big(U_{hp}^i + B^* P_{hp}^{i-1}, v_{hp} - U_{hp}^i\big)_U \geq 0, \quad \forall v_{hp} \in K_{hp}^i, i = 1,2,\ldots,M. \tag{2.25}$$

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

$$Y_{hp}|_{(t_{i-1},t_i]} = \big((t_i - t)Y_{hp}^{i-1} + (t - t_{i-1})Y_{hp}^i\big)/k_i,$$

$$P_{hp}|_{(t_{i-1},t_i]} = \big((t_i - t)P_{hp}^{i-1} + (t - t_{i-1})P_{hp}^i\big)/k_i,$$

$$U_{hp}|_{(t_{i-1},t_i]} = U_{hp}^i.$$

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 (2.23)-(2.25) can be restated as

$$\left(\frac{\partial Y_{hp}}{\partial t}, w_{hp}\right) + a(\hat{Y}_{hp}, w_{hp}) + \big(\phi(\hat{Y}_{hp}), w_{hp}\big) = (\hat{f} + BU_{hp}, w_{hp}), \tag{2.26}$$

$$\forall w_{hp} \in V_{hp}^i \subset V = H_0^1(\Omega), \quad t \in (t_{i-1}, t_i], i = 1,2,\ldots,M,$$

$$Y_{hp}(x,0) = y_0^{hp}(x), \quad x \in \Omega,$$

$$-\left(\frac{\partial P_{hp}}{\partial t}, q_{hp}\right) + a(q_{hp}, \tilde{P}_{hp}) + \big(\phi'(\tilde{Y}_{hp})\tilde{P}_{hp}, q_{hp}\big) = (\hat{Y}_{hp} - \hat{y}_d, q_{hp}), \tag{2.27}$$

$$\forall q_{hp} \in V_{hp}^i \subset V = H_0^1(\Omega), \quad t \in (t_{i-1}, t_i], i = M,\ldots,2,1,$$

$$P_{hp}(x,T) = 0, \quad x \in \Omega,$$

$$\big(U_{hp} + B^*\tilde{P}_{hp}, v_{hp} - U_{hp}\big)_U \geq 0, \tag{2.28}$$

$$\forall v_{hp} \in K_{hp}^i, \quad t \in (t_{i-1}, t_i], i = 1,2,\ldots,M.$$

The following lemmas [10, 16, 19] are important in deriving *a posteriori* error estimates of residual type.

Lemma 2.2 *There exist a constant $C > 0$ independent of v, h_{τ_U}, and p_{τ_U} and a mapping $\pi_{p_{\tau_U}}^{h_{\tau_U}} : H^1(\tau_U) \to \mathcal{P}_{p_{\tau_U}}(\tau_U)$ such that $\forall v \in H^1(\tau_U)$, $\tau_U \in \mathcal{T}_U$ the following inequality is valid*:

$$\big\| v - \pi_{p_{\tau_U}}^{h_{\tau_U}} \big\|_{L^2(\tau_U)} \leq C\frac{h_{\tau_U}}{p_{\tau_U}} \|v\|_{H^1(\tau_U)},$$

where we will write $v \in \mathcal{P}_{p_{\tau_U}}(\tau_U)$ if the following satisfied: *$v|_{\tau_U} \circ F_{\tau_U} \in P_{p_{\tau_U}}(\widehat{\tau})$ if τ_U is a triangle*; *$v|_{\tau_U} \circ F_{\tau_U} \in Q_{p_{\tau_U}}(\widehat{\tau})$ if τ_U is a parallelogram.*

Lemma 2.3 *Let* $\mathbf{p}_1$ *be an arbitrary polynomial degree distribution satisfies* (2.13). *Then there exists a linear operator* $E_1 : H_0^1(\Omega) \to S^{\mathbf{p}_1}(\mathcal{T}) \cap H_0^1(\Omega)$, *and there exists a constant* $C > 0$ *depending solely on* γ *such that for every* $v \in H_0^1(\Omega)$ *and all elements* $\tau \in \mathcal{T}$ *and all edges* $e \in \mathcal{E}(\mathcal{T})$:

$$\|v - E_1 v\|_{L^2(\tau)} + \frac{h_\tau}{p_\tau}\big\|\nabla(v - E_1 v)\big\|_{L^2(\tau)} \le C\frac{h_\tau}{p_\tau}\|\nabla v\|_{L^2(w_\tau^1)},$$

$$\|v - E_1 v\|_{L^2(e)} \le C\left(\frac{h_e}{p_e}\right)^{\frac{1}{2}}\|\nabla v\|_{L^2(w_e^1)}.$$

Lemma 2.4 *Let* $\mathbf{p}_1$ *be an arbitrary polynomial degree distribution satisfying* (2.13) *and* $p_\tau \ge 2$, $\forall \tau \in \mathcal{T}$. *Then there exists a bounded linear operator* $E_2 : H_0^1(\Omega) \cap H^2(\Omega) \to S^{\mathbf{p}_1}(\mathcal{T}) \cap H_0^1(\Omega)$, *and there exists a constant* $C > 0$ *that depends solely on* γ *such that for every* $v \in H_0^1(\Omega) \cap H^2(\Omega)$ *and all elements* $\tau \in \mathcal{T}$ *and all edges* $e \in \mathcal{E}(\mathcal{T})$:

$$\|v - E_2 v\|_{L^2(\tau)} + \frac{h_\tau}{p_\tau}\big\|\nabla(v - E_2 v)\big\|_{L^2(\tau)} \le C\left(\frac{h_\tau}{p_\tau}\right)^2 |v|_{H^2(w_\tau^1)},$$

$$\|v - E_2 v\|_{L^2(e)} \le C\left(\frac{h_e}{p_e}\right)^{\frac{3}{2}} |v|_{H^2(w_e^1)}.$$

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

$$\phi(\varphi) - \phi(\rho) = -\tilde{\phi}'(\varphi)(\rho - \varphi) = -\phi'(\rho)(\rho - \varphi) + \tilde{\phi}''(\varphi)(\rho - \varphi)^2, \tag{2.29}$$

where

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

are bounded functions in $\bar{\Omega}$ [22].

3 *A posteriori* error estimates

In this section, we shall derive some *a posteriori* error estimates for the *hp* finite element approximation of the optimal control problems governed by nonlinear parabolic equations.

Let

$$S(u) = \frac{1}{2}\big(\|y - y_d\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega_U)}^2\big), \tag{3.1}$$

$$S_{hp}(U_{hp}) = \frac{1}{2}\big(\|Y_{hp} - y_d\|_{L^2(\Omega)}^2 + \|U_{hp}\|_{L^2(\Omega_U)}^2\big). \tag{3.2}$$

It can be shown [19] that

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

$$\big(S'(U_{hp}), v\big) = \big(U_{hp} + B^* p(U_{hp}), v\big), \tag{3.4}$$

$$\big(S_{hp}'(U_{hp}), v\big) = \big(U_{hp} + B^* \tilde{P}_{hp}, v\big). \tag{3.5}$$

It is clear that S and S_{hp} are well defined and continuous on K and K_{hp}. Also the functional S_{hp} can be naturally extended on K. Then (2.5) and (2.14) can be represented as

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

and

$$\min_{U_{hp}\in K_{hp}}\left\{\int_0^T S_{hp}(U_{hp})\,dt\right\}. \tag{3.7}$$

In many application, $S(\cdot)$ 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, there is a $c > 0$, independent of h, such that

$$\int_0^T \bigl(S'(u) - S'(U_{hp}), u - U_{hp}\bigr)_U\,dt \ge c\|u - U_{hp}\|^2_{L^2(J;L^2(\Omega))}. \tag{3.8}$$

The following theorem is the first step to derive *a posteriori* error estimates.

Theorem 3.1 *Let* (y,u,p) *and* (Y_{hp},P_{hp},U_{hp}) *be the solutions of* (2.8)-(2.10) *and* (2.23)-(2.25), *respectively. Assume that* $(S'_{hp}(U_{hp}))|_\tau \in H^s(\tau)$, $\forall\tau \in \mathcal{T}_h$ $(s = 0,1)$, *and there is a* $v_{hp} \in K_{hp}$ *such that*

$$\bigl|\bigl(S'_{hp}(U_{hp}), v_{hp} - u\bigr)\bigr| \le C\sum_{\tau\in\mathcal{T}_h} h_\tau \bigl\|S'_{hp}(U_{hp})\bigr\|_{H^s(\tau)}\|u - U_{hp}\|^s_{L^2(\tau)}. \tag{3.9}$$

Then we have

$$\|u - U_{hp}\|^2_{L^2(J;L^2(\Omega))} \le C\eta_1^2 + C\bigl\|p(U_{hp}) - \tilde{P}_{hp}\bigr\|^2_{L^2(J;L^2(\Omega))}, \tag{3.10}$$

where

$$\eta_1^2 = \int_0^T \sum_{\tau\in\mathcal{T}_h} h_\tau^{1+s}\bigl\|U_{hp} + B^*\tilde{P}_{hp}\bigr\|^{1+s}_{H^1(\tau)}\,dt,$$

and $p(U_{hp})$ *is defined by the following equations*:

$$\left(\frac{\partial}{\partial t}y(U_{hp}), w\right) + a\bigl(y(U_{hp}), w\bigr) + \bigl(\phi\bigl(y(U_{hp})\bigr), w\bigr) = (f + BU_{hp}, w), \quad \forall w \in V, \tag{3.11}$$

$$y(U_{hp})(x,0) = y_0(x), \quad x \in \Omega,$$

$$\begin{aligned}&-\left(\frac{\partial}{\partial t}p(U_{hp}), q\right) + a\bigl(q, p(U_{hp})\bigr) + \bigl(\phi'\bigl(y(U_{hp})\bigr)p(U_{hp}), q\bigr)\\&\quad = \bigl(y(U_{hp}) - y_d, q\bigr), \quad \forall q \in V,\end{aligned} \tag{3.12}$$

$$p(U_{hp})(x,T) = 0, \quad x \in \Omega.$$

Proof It follows from (3.6) and (3.7) that $\int_0^T (S'(u), u - v)\,dt \leq 0$, $\forall v \in K$, and $\int_0^T (S'_{hp}(U_{hp}), U_{hp} - v_{hp})\,dt \leq 0$, $\forall v_{hp} \in K_{hp} \subset K$. Then it follows from the assumptions (3.8), (3.9), and the Schwartz inequality that

$$\begin{aligned} c\|u - U_{hp}\|^2_{L^2(J;L^2(\Omega))} &\leq \int_0^T \bigl(S'(u) - S'(U_{hp}), u - U_{hp}\bigr)\,dt \\ &\leq \int_0^T \bigl\{\bigl(S'_{hp}(U_{hp}), v_{hp} - u\bigr) + \bigl(S'_{hp}(U_{hp}) - S'(U_{hp}), u - U_{hp}\bigr)\bigr\}\,dt \\ &\leq C\int_0^T \biggl\{\sum_{\tau\in\mathcal{T}_h} h_\tau^{1+s}\bigl\|S'_{hp}(U_{hp})\bigr\|^{1+s}_{H^s(\tau)} + \bigl\|S'_{hp}(U_{hp}) - S'(U_{hp})\bigr\|^2_{L^2(\Omega)}\biggr\}\,dt \\ &\quad + \frac{c}{2}\|u - U_{hp}\|^2_{L^2(J;L^2(\Omega))}. \end{aligned} \tag{3.13}$$

It is not difficult to show

$$S'_{hp}(U_{hp}) = U_{hp} + B^*\tilde{P}_{hp}, \qquad S'(U_{hp}) = U_{hp} + B^*p(U_{hp}), \tag{3.14}$$

where $p(U_{hp})$ is defined in (3.11)-(3.12). Thanks to (3.14), it is easy to derive

$$\bigl\|S'_{hp}(U_{hp}) - S'(U_{hp})\bigr\|_{L^2(\Omega)} = \bigl\|p(U_{hp}) - \tilde{P}_{hp}\bigr\|_{L^2(\Omega)}. \tag{3.15}$$

Then by using the estimates (3.13) and (3.15) we can prove the requested result (3.10). □

Next, we are in the position to estimate the errors $\|Y_{hp} - y(U_{hp})\|^2_{L^2(0,T;H^1(\Omega))}$ and $\|P_{hp} - p(U_{hp})\|^2_{L^2(0,T;H^1(\Omega))}$.

Theorem 3.2 *Let* (Y_{hp}, P_{hp}, U_{hp}) *be the solutions of* (2.26)-(2.28), *let* $(y(U_{hp}), p(U_{hp}))$ *be defined by* (3.11)-(3.12). *Then we have*

$$\bigl\|Y_{hp} - y(U_{hp})\bigr\|^2_{L^2(0,T;H^1(\Omega))} + \bigl\|P_{hp} - p(U_{hp})\bigr\|^2_{L^2(0,T;H^1(\Omega))} \leq C\sum_{i=2}^{6}\eta_i^2, \tag{3.16}$$

where

$$\begin{aligned} \eta_2^2 &= \int_0^T \sum_\tau \frac{h_\tau^2}{p_\tau^2}\int_\tau \biggl(\hat{f} + BU_{hp} + \operatorname{div}(A\nabla\hat{Y}_{hp}) - \phi(\hat{Y}_{hp}) - \frac{\partial Y_{hp}}{\partial t}\biggr)^2\,dx\,dt, \\ \eta_3^2 &= \int_0^T \sum_\tau \frac{h_\tau^2}{p_\tau^2}\int_\tau \biggl(\hat{Y}_{hp} - \hat{y}_d + \operatorname{div}\bigl(A^*\nabla\tilde{P}_{hp}\bigr) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp} + \frac{\partial P_{hp}}{\partial t}\biggr)^2\,dx\,dt, \\ \eta_4^2 &= \int_0^T \sum_e \int_e \frac{h_e}{p_e}\bigl[(A\nabla\hat{Y}_{hp})\cdot n\bigr]^2\,de\,dt + \int_0^T \sum_e \int_e \frac{h_e}{p_e}\bigl[\bigl(A^*\nabla\tilde{P}_{hp}\bigr)\cdot n\bigr]^2\,de\,dt, \\ \eta_5^2 &= \int_0^T \int_\Omega \bigl|A\nabla(\hat{Y}_{hp} - Y_{hp})\bigr|^2\,dx\,dt + \int_0^T \int_\Omega \bigl|A^*\nabla(\tilde{P}_{hp} - P_{hp})\bigr|^2\,dx\,dt, \\ \eta_6^2 &= \|Y_{hp} - \tilde{Y}_{hp}\|^2_{L^2(0,T;L^2(\Omega))} + \|Y_{hp} - \hat{Y}_{hp}\|^2_{L^2(0,T;L^2(\Omega))} + \|y_d - \hat{y}_d\|^2_{L^2(0,T;L^2(\Omega))} \\ &\quad + \|f - \hat{f}\|^2_{L^2(0,T;L^2(\Omega))} + \bigl\|y_0(x) - Y_{hp}(x,0)\bigr\|^2_{L^2(\Omega)}, \end{aligned}$$

where $e = \bar{\tau}_e^1 \cap \bar{\tau}_e^2$, τ_e^1, τ_e^2 are two neighboring elements in $\mathcal{T}$, $[A\nabla\hat{Y}_{hp}\cdot n]_e$ and $[A^*\nabla\tilde{P}_{hp}\cdot n]_e$ are the A-normal and A^*-normal derivative jumps over the interior edge e, respectively, defined by

$$\big[(A\nabla\hat{Y}_{hp})\cdot n\big]_e = (A\nabla\hat{Y}_{hp}|_{\tau_e^1} - A\nabla\hat{Y}_{hp}|_{\tau_e^2})\cdot n,$$
$$\big[\big(A^*\nabla\tilde{P}_{hp}\big)\cdot n\big]_e = \big(A^*\nabla\tilde{P}_{hp}|_{\tau_e^1} - A^*\nabla\tilde{P}_{hp}|_{\tau_e^2}\big)\cdot n,$$

where n is the unit normal vector on $e = \bar{\tau}_e^1 \cap \bar{\tau}_e^2$ outwards τ_e^1. For later convenience, we define $[A\nabla\hat{Y}_{hp}\cdot n]_e = 0$ and $[A^*\nabla\tilde{P}_{hp}\cdot n]_e = 0$ when $e \subset \partial\Omega$.

Proof Let $r_{hp} = p(U_{hp}) - P_{hp}$ and E_1 be the linear operator defined in Lemma 2.3. Note that $(p(U_{hp}) - P_{hp})(x,T) = 0$, hence

$$\int_0^T -\left(\frac{\partial(p(U_{hp}) - P_{hp})}{\partial t}, r_{hp}\right)dt \geq 0. \tag{3.17}$$

By using the assumption of ϕ, (2.29), and (3.17), we obtain

$$\begin{aligned}
&c\|r_{hp}\|^2_{L^2(0,T;H^1(\Omega))}\\
&\quad\leq \int_0^T a\big(r_{hp}, p(U_{hp}) - P_{hp}\big)\,dt + \int_0^T \big(\phi'\big(y(U_{hp})\big)\big(p(U_{hp}) - P_{hp}\big), p(U_{hp}) - P_{hp}\big)\,dt\\
&\quad\leq \int_0^T \big(\nabla r_{hp}, A^*\nabla\big(p(U_{hp}) - P_{hp}\big)\big)\,dt - \int_0^T \left(\frac{\partial(p(U_{hp}) - P_{hp})}{\partial t}, r_{hp}\right)dt\\
&\qquad + \int_0^T \big(\phi'\big(y(U_{hp})\big)\big(p(U_{hp}) - \tilde{P}_{hp}\big), p(U_{hp}) - P_{hp}\big)\,dt\\
&\qquad + \int_0^T \big(\phi'\big(y(U_{hp})\big)(\tilde{P}_{hp} - P_{hp}), p(U_{hp}) - P_{hp}\big)\,dt\\
&\quad= \int_0^T \big(\nabla r_{hp}, A^*\nabla\big(p(U_{hp}) - \tilde{P}_{hp}\big)\big)\,dt - \int_0^T \left(\frac{\partial(p(U_{hp}) - P_{hp})}{\partial t}, r_{hp}\right)dt\\
&\qquad + \int_0^T \big(\phi'\big(y(U_{hp})\big)p(U_{hp}) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp}, p(U_{hp}) - P_{hp}\big)\,dt\\
&\qquad + \int_0^T \big(\tilde{\phi}''(\tilde{Y}_{hp})\big(\tilde{Y}_{hp} - y(U_{hp})\big)\tilde{P}_{hp}, p(U_{hp}) - P_{hp}\big)\,dt\\
&\qquad + \int_0^T \big(\phi'\big(y(U_{hp})\big)(\tilde{P}_{hp} - P_{hp}), p(U_{hp}) - P_{hp}\big)\,dt\\
&\qquad + \int_0^T \big(\nabla r_{hp}, A^*\nabla(\tilde{P}_{hp} - P_{hp})\big)\,dt.
\end{aligned} \tag{3.18}$$

Connecting (2.27), (3.12), and (3.18), we have

$$\begin{aligned}
&c\|r_{hp}\|^2_{L^2(0,T;H^1(\Omega))}\\
&\quad\leq \int_0^T \big(\nabla(r_{hp} - E_1 r_{hp}), A^*\nabla\big(p(U_{hp}) - \tilde{P}_{hp}\big)\big)\,dt\\
&\qquad - \int_0^T \left(\frac{\partial(p(U_{hp}) - P_{hp})}{\partial t}, r_{hp} - E_1 r_{hp}\right)dt
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^T \big(\phi'\big(y(U_{hp})\big)p(U_{hp}) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp}, r_{hp} - E_1 r_{hp}\big)\,dt \\
&-\int_0^T \left(\frac{\partial (p(U_{hp}) - P_{hp})}{\partial t}, E_1 r_{hp}\right)dt + \int_0^T \big(\nabla E_1 r_{hp}, A^*\nabla\big(p(U_{hp}) - \tilde{P}_{hp}\big)\big)\,dt \\
&+\int_0^T \big(\phi'\big(y(U_{hp})\big)p(U_{hp}) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp}, E_1 r_{hp}\big)\,dt \\
&+\int_0^T \big(\tilde{\phi}''(\tilde{Y}_{hp})\big(\tilde{Y}_{hp} - y(U_{hp})\big)\tilde{P}_{hp}, p(U_{hp}) - P_{hp}\big)\,dt \\
&+\int_0^T \big(\phi'\big(y(U_{hp})\big)(\tilde{P}_{hp} - P_{hp}), p(U_{hp}) - P_{hp}\big)\,dt + \int_0^T \big(\nabla r_{hp}, A^*\nabla(\tilde{P}_{hp} - P_{hp})\big)\,dt \\
=&\int_0^T \left(y(U_{hp}) - y_d + \operatorname{div}\big(A^*\nabla\tilde{P}_{hp}\big) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp} + \frac{\partial P_{hp}}{\partial t}, r_{hp} - E_1 r_{hp}\right)dt \\
&+\int_0^T \sum_e \int_e \big[\big(A^*\nabla\tilde{P}_{hp}\big)\cdot n\big](r_{hp} - E_1 r_{hp})\,de\,dt \\
&+\int_0^T \big(\tilde{\phi}''(\tilde{Y}_{hp})\big(\tilde{Y}_{hp} - y(U_{hp})\big)\tilde{P}_{hp}, p(U_{hp}) - P_{hp}\big)\,dt \\
&+\int_0^T \big(y(U_{hp}) - \hat{Y}_{hp}, E_1 r_{hp}\big)\,dt + \int_0^T (\hat{y}_d - y_d, E_1 r_{hp})\,dt \\
&+\int_0^T \big(\phi'\big(y(U_{hp})\big)(\tilde{P}_{hp} - P_{hp}), p(U_{hp}) - P_{hp}\big)\,dt \\
&+\int_0^T \big(\nabla r_{hp}, A^*\nabla(\tilde{P}_{hp} - P_{hp})\big)\,dt. \qquad (3.19)
\end{aligned}
$$

Then we have

$$
\begin{aligned}
&c\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \\
&\quad\le \int_0^T \left(\hat{Y}_{hp} - \hat{y}_d + \operatorname{div}\big(A^*\nabla\tilde{P}_{hp}\big) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp} + \frac{\partial P_{hp}}{\partial t}, r_{hp} - E_1 r_{hp}\right)dt \\
&\qquad+\int_0^T \sum_e \int_e \big[\big(A^*\nabla\tilde{P}_{hp}\big)\cdot n\big](r_{hp} - E_1 r_{hp})\,de\,dt \\
&\qquad+\int_0^T \big(y(U_{hp}) - \hat{Y}_{hp}, r_{hp}\big)\,dt + \int_0^T (\hat{y}_d - y_d, r_{hp})\,dt \\
&\qquad+\int_0^T \big(\nabla r_{hp}, A^*\nabla(\tilde{P}_{hp} - P_{hp})\big)\,dt \\
&\qquad+\int_0^T \big(\tilde{\phi}''(\tilde{Y}_{hp})\big(\tilde{Y}_{hp} - y(U_{hp})\big)\tilde{P}_{hp}, p(U_{hp}) - P_{hp}\big)\,dt \\
&\qquad+\int_0^T \big(\phi'\big(y(U_{hp})\big)(\tilde{P}_{hp} - P_{hp}), p(U_{hp}) - P_{hp}\big)\,dt \\
&\quad\equiv \sum_{i=1}^{7} I_i. \qquad (3.20)
\end{aligned}
$$

It follows from Lemma 2.3 that

$$\begin{aligned}
I_1 &= \int_0^T \left(\hat{Y}_{hp} - \hat{y}_d + \operatorname{div}\bigl(A^*\nabla\tilde{P}_{hp}\bigr) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp} + \frac{\partial P_{hp}}{\partial t}, r_{hp} - E_1 r_{hp} \right) dt \\
&\leq C(\delta) \int_0^T \sum_\tau \frac{h_\tau^2}{p_\tau^2} \int_\tau \left(\hat{Y}_{hp} - \hat{y}_d + \operatorname{div}\bigl(A^*\nabla\tilde{P}_{hp}\bigr) - \phi'(\tilde{Y}_{hp})\tilde{P}_{hp} + \frac{\partial P_{hp}}{\partial t} \right)^2 dx\, dt \\
&\quad + \delta \int_0^T \|r_{hp}\|_{H^1(\Omega)}^2 \, dt \\
&= C(\delta)\eta_3^2 + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2.
\end{aligned} \tag{3.21}$$

Similarly,

$$\begin{aligned}
I_2 &= \int_0^T \sum_e \int_e \bigl[\bigl(A^*\nabla\tilde{P}_{hp}\bigr)\cdot n\bigr](r_{hp} - E_1 r_{hp})\, de\, dt \\
&\leq C(\delta) \int_0^T \sum_e \int_e \frac{h_e}{p_e} \bigl[\bigl(A^*\nabla\tilde{P}_{hp}\bigr)\cdot n\bigr]^2 de\, dt + \delta \int_0^T \|r_{hp}\|_{H^1(\Omega)}^2 \, dt \\
&\leq C(\delta)\eta_4^2 + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2.
\end{aligned} \tag{3.22}$$

For I_3-I_5, we have

$$\begin{aligned}
I_3 &= \int_0^T \bigl(y(U_{hp}) - \hat{Y}_{hp}, r_{hp}\bigr)\, dt \\
&\leq C(\delta) \int_0^T \int_\Omega \bigl|y(U_{hp}) - \hat{Y}_{hp}\bigr|^2 dx\, dt + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2 \\
&\leq C(\delta)\bigl\|y(U_{hp}) - Y_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2 + C(\delta)\|Y_{hp} - \hat{Y}_{hp}\|_{L^2(0,T;L^2(\Omega))}^2 \\
&\quad + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2 \\
&\leq C(\delta)\eta_6^2 + C(\delta)\bigl\|y(U_{hp}) - Y_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2 + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2,
\end{aligned} \tag{3.23}$$

$$\begin{aligned}
I_4 &= \int_0^T (\hat{y}_d - y_d, r_{hp})\, dt \\
&\leq C(\delta)\|y_d - \hat{y}_d\|_{L^2(0,T;L^2(\Omega))}^2 + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2 \\
&\leq C(\delta)\eta_6^2 + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2,
\end{aligned} \tag{3.24}$$

and

$$\begin{aligned}
I_5 &= \int_0^T \bigl(\nabla r_{hp}, A^*\nabla(\tilde{P}_{hp} - P_{hp})\bigr)\, dt \\
&\leq C(\delta) \int_0^T \int_\Omega \bigl|A^*\nabla(\tilde{P}_{hp} - P_{hp})\bigr|^2 dx\, dt + \delta \int_0^T \int_\Omega |\nabla r_{hp}|^2 \, dx\, dt \\
&\leq C(\delta)\eta_5^2 + \delta \bigl\|p(U_{hp}) - P_{hp}\bigr\|_{L^2(0,T;H^1(\Omega))}^2.
\end{aligned} \tag{3.25}$$

Similarly, we can obtain

$$\begin{aligned} I_6 &= \int_0^T \big(\tilde{\phi}''(\tilde{Y}_{hp})\big(\tilde{Y}_{hp} - y(U_{hp})\big)\tilde{P}_{hp}, p(U_{hp}) - P_{hp}\big)\,dt \\ &\leq C(\delta)\int_0^T \int_\Omega \big|y(U_{hp}) - \tilde{Y}_{hp}\big|^2\,dx\,dt + \delta\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \\ &\leq C(\delta)\big\|y(U_{hp}) - Y_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} + C(\delta)\|Y_{hp} - \tilde{Y}_{hp}\|^2_{L^2(0,T;L^2(\Omega))} \\ &\quad + \delta\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \\ &\leq C(\delta)\eta_6^2 + C(\delta)\big\|y(U_{hp}) - Y_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} + \delta\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \end{aligned} \tag{3.26}$$

and

$$\begin{aligned} I_7 &= \int_0^T \big(\phi'\big(y(U_{hp})\big)(\tilde{P}_{hp} - P_{hp}), p(U_{hp}) - P_{hp}\big)\,dt \\ &\leq C(\delta)\|\tilde{P}_{hp} - P_{hp}\|^2_{L^2(0,T;L^2(\Omega))} + \delta\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \\ &\leq C(\delta)\int_0^T \int_\Omega \big|A^*\nabla(\tilde{P}_{hp} - P_{hp})\big|^2\,dx\,dt + \delta\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \\ &\leq C(\delta)\eta_5^2 + \delta\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))}. \end{aligned} \tag{3.27}$$

Then let δ be small enough, from (3.20)-(3.27), we derive

$$\big\|p(U_{hp}) - P_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \leq C(\delta)\sum_{i=2}^{6}\eta_i^2 + C(\delta)\big\|y(U_{hp}) - Y_{hp}\big\|^2_{L^2(0,T;L^2(\Omega))}. \tag{3.28}$$

Furthermore, we estimate the error $\|y(U_{hp}) - Y_{hp}\|^2_{L^2(0,T;L^2(\Omega))}$. Let $e_{hp} = y(U_{hp}) - Y_{hp}$ and E_1 be the linear operator defined in Lemma 2.3. Note that

$$\begin{aligned} &\int_0^T \left(\frac{\partial(y(U_{hp}) - Y_{hp})}{\partial t}, e_{hp}\right)dt \\ &\quad = \int_0^T \int_\Omega \frac{\partial(y(U_{hp}) - Y_{hp})}{\partial t} e_{hp}\,dx\,dt \\ &\quad = \int_\Omega \big(\big(y(U_{hp}) - Y_{hp}\big)(x,T)\big)^2\,dx - \int_\Omega \big(\big(y(U_{hp}) - Y_{hp}\big)(x,0)\big)^2\,dx \\ &\qquad - \int_0^T \left(\frac{\partial(y(U_{hp}) - Y_{hp})}{\partial t}, e_{hp}\right)dt. \end{aligned}$$

Thus

$$\int_0^T \left(\frac{\partial(y(U_{hp}) - Y_{hp})}{\partial t}, e_{hp}\right)dt + \frac{1}{2}\big\|y_0(x) - Y_{hp}(x,0)\big\|^2_{L^2(\Omega)} \geq 0. \tag{3.29}$$

From (2.29), it is easy to see that

$$\big(\phi\big(y(U_{hp})\big) - \phi(Y_{hp}), e_{hp}\big) = \big(\tilde{\phi}'\big(y(U_{hp})\big)\big(y(U_{hp}) - Y_{hp}\big), e_{hp}\big) \geq 0. \tag{3.30}$$

By using (3.29) and (3.30), we have

$$\begin{aligned}
&c\|e_{hp}\|^2_{L^2(0,T;H^1(\Omega))} \\
&\quad\leq \int_0^T a\big(y(U_{hp})-Y_{hp},e_{hp}\big)\,dt+\int_0^T \big(\phi\big(y(U_{hp})\big)-\phi(Y_{hp}),e_{hp}\big)\,dt \\
&\quad\leq \int_0^T a\big(y(U_{hp})-Y_{hp},e_{hp}\big)\,dt+\int_0^T \big(\phi\big(y(U_{hp})\big)-\phi(\hat{Y}_{hp}),e_{hp}\big)\,dt \\
&\qquad+\int_0^T \left(\frac{\partial(y(U_{hp})-Y_{hp})}{\partial t},e_{hp}\right)dt+\int_0^T \big(\phi(\hat{Y}_{hp})-\phi(Y_{hp}),e_{hp}\big)\,dt \\
&\qquad+\frac{1}{2}\big\|y_0(x)-Y_{hp}(x,0)\big\|^2_{L^2(\Omega)} \\
&\quad=\int_0^T \big(A\nabla\big(y(U_{hp})-\hat{Y}_{hp}\big),\nabla e_{hp}\big)\,dt+\int_0^T \big(\phi\big(y(U_{hp})\big)-\phi(\hat{Y}_{hp}),e_{hp}\big)\,dt \\
&\qquad+\int_0^T \left(\frac{\partial(y(U_{hp})-Y_{hp})}{\partial t},e_{hp}\right)dt+\int_0^T \big(A\nabla(\hat{Y}_{hp}-Y_{hp}),\nabla e_{hp}\big)\,dt \\
&\qquad+\int_0^T \big(\phi(\hat{Y}_{hp})-\phi(Y_{hp}),e_{hp}\big)\,dt \\
&\qquad+\frac{1}{2}\big\|y_0(x)-Y_{hp}(x,0)\big\|^2_{L^2(\Omega)}.
\end{aligned} \tag{3.31}$$

Connecting (2.26), (2.29), (3.11), and (3.31), we obtain

$$\begin{aligned}
&c\|e_{hp}\|^2_{L^2(0,T;H^1(\Omega))} \\
&\quad\leq \int_0^T \left(\hat{f}+BU_{hp}+\operatorname{div}(A\nabla\hat{Y}_{hp})-\phi(\hat{Y}_{hp})-\frac{\partial Y_{hp}}{\partial t},e_{hp}-E_1e_{hp}\right)dt \\
&\qquad+\int_0^T \sum_e\int_e\big[(A\nabla\hat{Y}_{hp})\cdot n\big](e_{hp}-E_1e_{hp})\,de\,dt+\int_0^T (f-\hat{f},e_{hp})\,dt \\
&\qquad+\int_0^T \big(A\nabla(\hat{Y}_{hp}-Y_{hp}),\nabla e_{hp}\big)\,dt+\frac{1}{2}\big\|y_0(x)-Y_{hp}(x,0)\big\|^2_{L^2(\Omega)} \\
&\qquad+\int_0^T \big(\tilde{\phi}'(\hat{Y}_{hp})(\hat{Y}_{hp}-Y_{hp}),e_{hp}\big)\,dt \\
&\quad\leq C(\delta)\int_0^T\sum_\tau\frac{h_\tau^2}{p_\tau^2}\int_\tau\left(\hat{f}+BU_{hp}+\operatorname{div}(A\nabla\hat{Y}_{hp})-\phi(\hat{Y}_{hp})-\frac{\partial Y_{hp}}{\partial t}\right)^2dx\,dt \\
&\qquad+C(\delta)\int_0^T\sum_e\int_e\frac{h_e}{p_e}\big[(A\nabla\hat{Y}_{hp})\cdot n\big]^2\,de\,dt+C(\delta)\|f-\hat{f}\|^2_{L^2(0,T;L^2(\Omega))} \\
&\qquad+C(\delta)\int_0^T\int_\Omega\big|A\nabla(\hat{Y}_{hp}-Y_{hp})\big|^2\,dx\,dt+C(\delta)\|\hat{Y}_{hp}-Y_{hp}\|^2_{L^2(0,T;H^1(\Omega))} \\
&\qquad+\frac{1}{2}\big\|y_0(x)-Y_{hp}(x,0)\big\|^2_{L^2(\Omega)}+\delta\big\|y(U_{hp})-Y_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))} \\
&\quad\leq C(\delta)\sum_{i=2}^6\eta_i^2+\delta\big\|y(U_{hp})-Y_{hp}\big\|^2_{L^2(0,T;H^1(\Omega))}.
\end{aligned} \tag{3.32}$$

Hence, let δ be small enough, we have

$$\left\| y(U_{hp}) - Y_{hp} \right\|^2_{L^2(0,T;H^1(\Omega))} \leq C(\delta) \sum_{i=1}^{6} \eta_i^2. \tag{3.33}$$

Then (3.16) follows from (3.28) and (3.33). □

Finally, collecting Theorems 3.1-3.2, we derive the following residual-based *a posteriori* error estimates.

Theorem 3.3 *Let* (y, p, u) *and* (Y_{hp}, P_{hp}, U_{hp}) *are the solutions of* (2.8)-(2.10) *and* (2.26)-(2.28) *respectively. Assume that all the conditions in Theorem* 3.1 *are valid. Then*

$$\begin{aligned} &\|y - Y_{hp}\|^2_{L^2(0,T;H^1(\Omega))} + \|p - P_{hp}\|^2_{L^2(0,T;H^1(\Omega))} + \|u - U_{hp}\|^2_{L^2(0,T;L^2(\Omega_U))} \\ &\quad \leq C \sum_{i=1}^{6} \eta_i^2, \end{aligned} \tag{3.34}$$

where η_i^2, $i = 1, \ldots, 6$ *are defined in Theorem* 3.1 *and Theorem* 3.2.

Proof It follows from Theorem 3.1 and Theorem 3.2 that

$$\begin{aligned} \|u - U_{hp}\|^2_{L^2(0,T;L^2(\Omega_U))} &\leq C\eta_1^2 + C\left\| \tilde{P}_{hp} - p(U_{hp}) \right\|^2_{L^2(0,T;L^2(\Omega))} \\ &\leq C\eta_1^2 + C\|\tilde{P}_{hp} - P_{hp}\|^2_{L^2(0,T;L^2(\Omega))} \\ &\quad + C\left\| P_{hp} - p(U_{hp}) \right\|^2_{L^2(0,T;L^2(\Omega))} \\ &\leq C \sum_{i=1}^{6} \eta_i^2 + C\|\tilde{P}_{hp} - P_{hp}\|^2_{L^2(0,T;L^2(\Omega))}. \end{aligned} \tag{3.35}$$

Note that A is positive definite, it follows from the Poincaré inequality that

$$\begin{aligned} \|\tilde{P}_{hp} - P_{hp}\|^2_{L^2(0,T;L^2(\Omega))} &\leq \int_0^T \int_\Omega \left| A^* \nabla (\tilde{P}_{hp} - P_{hp}) \right|^2 dx\, dt \\ &\leq C\eta_5^2. \end{aligned} \tag{3.36}$$

Then it follows from (3.35) and (3.36) that

$$\|u - U_{hp}\|^2_{L^2(0,T;L^2(\Omega_U))} \leq C \sum_{i=1}^{6} \eta_i^2. \tag{3.37}$$

Note that

$$\|y - Y_{hp}\|^2_{L^2(0,T;H^1(\Omega))} \leq \left\| y - y(U_{hp}) \right\|^2_{L^2(0,T;H^1(\Omega))} + \left\| y(U_{hp}) - Y_{hp} \right\|^2_{L^2(0,T;H^1(\Omega))}, \tag{3.38}$$

$$\|p - P_{hp}\|^2_{L^2(0,T;H^1(\Omega))} \leq \left\| p - p(U_{hp}) \right\|^2_{L^2(0,T;H^1(\Omega))} + \left\| p(U_{hp}) - P_{hp} \right\|^2_{L^2(0,T;H^1(\Omega))}, \tag{3.39}$$

and

$$\|y - y(U_{hp})\|^2_{L^2(0,T;H^1(\Omega))} \leq C\|u - U_{hp}\|^2_{L^2(0,T;L^2(\Omega_U))}, \tag{3.40}$$

$$\|p - p(U_{hp})\|^2_{L^2(0,T;H^1(\Omega))} \leq \|y - y(U_{hp})\|^2_{L^2(0,T;L^2(\Omega))} \leq C\|u - U_{hp}\|^2_{L^2(0,T;L^2(\Omega_U))}. \tag{3.41}$$

Therefore, we obtain (3.34) from (3.16) and (3.37)-(3.41). □

4 Conclusion and future work

In this paper, we present the *hp* version of the finite element approximation for the optimal control problems governed by nonlinear parabolic equations. By using the *hp* finite element approximation for both the state and the co-state variables and the *hp* discontinuous Galerkin finite element approximation for the control variable, we derive *hp* residual-based *a posteriori* error estimates for the nonlinear parabolic optimal control problems. To the best of our knowledge in the context of optimal control problems, these residual-based *a posteriori* error estimates for the nonlinear parabolic optimal control problems are new.

In future, we shall consider the *hp* version of the finite element method for hyperbolic optimal control problems. Furthermore, we shall consider *a posteriori* error estimates and the superconvergence of the *hp* finite element solutions for hyperbolic optimal control problems.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ZL and LC have participated in the sequence alignment and drafted the manuscript. HL and CH have made substantial contributions to the conception and design. All authors read and approved the final manuscript.

Author details

[1] Key Laboratory of Signal and Information Processing, Chongqing Three Gorges University, Chongqing, 404000, P.R. China. [2] Key Laboratory for Nonlinear Science and System Structure, Chongqing Three Gorges University, Chongqing, 404000, P.R. China. [3] Research Center for Mathematics and Economics, Tianjin University of Finance and Economics, Tianjin, 300222, P.R. China. [4] Chongqing Wanzhou Long Bao Middle School, Chongqing, 404001, P.R. China. [5] Huashang College, Guangdong University of Finance, Guangzhou, 511300, P.R. China.

Acknowledgements

The authors express their thanks to the referees for their helpful suggestions, which led to improvements of the presentation. This work is 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).

Received: 17 June 2015 Accepted: 1 February 2016 Published online: 13 February 2016

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. Chen, Y, Lu, Z, Huang, Y: Superconvergence of triangular Raviart-Thomas mixed finite element methods for bilinear constrained optimal control problem. Comput. Math. Appl. **66**, 1498-1513 (2013)
4. Chen, Y, Yi, N, Liu, W: A Legendre-Galerkin spectral method for optimal control problems governed by elliptic equations. SIAM J. Numer. Anal. **46**, 2254-2275 (2008)
5. Lu, Z: A residual-based posteriori error estimates for *hp* finite element solutions of general bilinear optimal control problems. J. Math. Inequal. **9**, 665-682 (2015)
6. Hoppe, RHW, Iliash, Y, Iyyunni, C, Sweilam, NH: A posteriori error estimates for adaptive finite element discretizations of boundary control problems. J. Numer. Math. **14**, 57-82 (2006)
7. Kröner, A, Vexler, B: A priori error estimates for elliptic optimal control problems with a bilinear state equation. J. Comput. Appl. Math. **230**, 781-802 (2009)

8. Kunisch, K, Liu, W, Chang, Y, Yan, N, Li, R: Adaptive finite element approximation for a class of parameter estimation problems. J. Comput. Math. **28**, 645-675 (2001)
9. Melenk, JM, Wohlmuth, B: On residual-based a posteriori error estimation in *hp*-FEM. Adv. Comput. Math. **15**, 311-331 (2001)
10. Babuska, I, Suri, M: The *hp*-version of the finite element method with quasiuniform meshes. Modél. Math. Anal. Numér. **21**, 199-238 (1987)
11. Oden, JT, Demkowicz, L, Rachowicz, W, Estermann, TAW: Toward a universal *h-p* adaptive finite element strategy, part 2. A posteriori error estimation. Comput. Methods Appl. Mech. Eng. **77**, 113-180 (1989)
12. Babuska, I, Suri, M: The *p*- and *h-p* version of the finite element method, an overview. Comput. Methods Appl. Mech. Eng. **80**, 5-26 (1990)
13. Babuska, I, Guo, B, Stephan, EP: The *h-p* version of the finite element method, an overview. Comput. Methods Appl. Mech. Eng. **80**, 319-325 (1990)
14. Babuska, I, Guo, B: Approximation properties of the *h-p* version of the finite element method. Comput. Methods Appl. Mech. Eng. **133**, 319-346 (1996)
15. Guo, B, Cao, W: An additive Schwarz method for the *hp* version of the finite element method in three dimensions. SIAM J. Numer. Anal. **35**, 632-654 (1998)
16. Melenk, JM: *hp*-interpolation of nonsmooth functions and an application to *hp*-a posteriori error estimation. SIAM J. Numer. Anal. **43**, 127-155 (2005)
17. Liu, W: Adaptive multi-meshes in finite element approximation of optimal control. Contemp. Math. **383**, 113-132 (2005)
18. Liu, W, Yan, N: A posteriori error estimates for optimal control problems governed by parabolic equations. Numer. Math. **93**, 497-521 (2003)
19. Chen, Y, Lin, Y: A posteriori error estimates for *hp* finite element solutions of convex optimal control problems. J. Comput. Appl. Math. **235**, 3435-3454 (2011)
20. Lions, JL, Magenes, E: Non Homogeneous Boundary Value Problems and Applications. Springer, Berlin (1972)
21. Lions, JL: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
22. Milner, FA: Mixed finite element methods for quasilinear second-order elliptic problems. Math. Comput. **44**, 303-320 (1985)