Lu et al. *Boundary Value Problems* (2016) 2016:59
DOI 10.1186/s13661-016-0562-2

Boundary Value Problems
a SpringerOpen Journal

RESEARCH

Open Access

Residual-based *a posteriori* error estimates for *hp* finite element solutions of semilinear Neumann boundary optimal control problems

Zuliang Lu[1,2*], Shuhua Zhang[2], Chuanjuan Hou[3] and Hongyan Liu[4]

*Correspondence: zulianglux@126.com
[1]Key Laboratory for Nonlinear Science and System Structure, Chongqing Three Gorges University, Chongqing, 404000, P.R. China
[2]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* finite element approximation of semilinear Neumann boundary elliptic optimal control problems. By using the *hp* finite element approximation for both the state and the co-state and the *hp* discontinuous Galerkin finite element approximation for the control, we derive *a posteriori* error bounds in L^2-H^1 norms for the Neumann boundary optimal control problems governed by semilinear elliptic equations. We also give L^2-L^2 *a posteriori* error estimates for the optimal control problems. Such estimates, which are apparently not available in the literature, can be used to construct reliable adaptive finite element approximations for the semilinear Neumann boundary optimal control problems.

MSC: 49J20; 65N30

Keywords: residual-based *a posteriori* error estimates; semilinear Neumann boundary elliptic optimal control problems; *hp* finite element methods; *hp* discontinuous Galerkin finite element methods

1 Introduction

In this paper, we study residual-based *a posteriori* error estimates for the *hp* finite element approximation of semilinear Neumann boundary optimal control problems. We consider the following semilinear elliptic optimal control problems:

$$\min_{u\in K\subset U}\{g(y)+j(u)\}, \tag{1.1}$$

$$-\operatorname{div}(A\nabla y)+\phi(y)=f, \quad \text{in } \Omega, \tag{1.2}$$

$$(A\nabla y)\cdot n|_{\partial\Omega}=u+z_0, \tag{1.3}$$

where the bounded open set $\Omega\subset\mathbb{R}^2$ is a convex polygon with the boundary $\partial\Omega$, $K=\{u\in U=L^2(\partial\Omega):\int_{\partial\Omega}u\,dx\geq 0\}$, $f\in L^2(\Omega)$, $z_0\in L^2(\partial\Omega)$, n is the outward normal on $\partial\Omega$. For $1\leq p<\infty$ and m any nonnegative integer let $W^{m,p}(\Omega)=\{v\in L^p(\Omega); D^\alpha v\in L^p(\Omega) \text{ if } |\alpha|\leq m\}$ denote the Sobolev spaces endowed with the norm $\|v\|_{m,p}^p=\sum_{|\alpha|\leq m}\|D^\alpha v\|_{L^p(\Omega)}^p$, and the

Springer

semi-norm $|v|_{m,p}^p = \sum_{|\alpha|=m} \|D^\alpha v\|_{L^p(\Omega)}^p$. 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 $\|\cdot\|_m = \|\cdot\|_{m,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$. Furthermore, we assume that 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 $\mathbf{X} \in \mathbb{R}^2$, $\mathbf{X}^t A\mathbf{X} \geq c\|\mathbf{X}\|_{\mathbb{R}^2}^2$. The function $\phi(\cdot) \in W^{1,\infty}(-R,R)$ for any $R > 0$, $\phi'(y) \in L^2(\Omega)$ for any $y \in H^1(\Omega)$, and $\phi' \geq 0$. Let g and j be strictly convex functions which are continuously differentiable on the space $L^2(\partial\Omega)$, and K be a closed convex set in the control space U. We further assume that $j(u) \to +\infty$ as $\|u\|_U \to \infty$ and $g'(\cdot)$ is a locally Lipschitz continuous function.

Optimal control problems have attracted substantial interest in recent years due to their applications in aero-hydrodynamics, atmospheric, hydraulic pollution problems, combustion, exploration and extraction of oil and gas resources, and engineering. They must be solved successfully with efficient numerical methods. Among these numerical methods, finite element methods are a successful choice for solving the optimal control problems. 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 [3–12], but there are very less published results for optimal control problems by using *hp* finite element methods. Recently, the adaptive finite element methods have been investigated extensively and became one of the most popular methods in scientific computation. In [13], the authors studied *a posteriori* error estimates for adaptive finite element discretizations of boundary control problems. *A posteriori* error estimates and adaptive finite element approximations for parameter estimation problems have been obtained in [14, 15]. There are three main versions in adaptive finite element approximation, *i.e.*, the *p*-version, the *h*-version, and the *hp*-version. The *p*-version of finite element methods uses a fixed mesh and improves the approximation of the solution by increasing 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, 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. There have been many theoretical studies as regards the *hp* finite element method in [16, 17]. An adaptive finite element approximation ensures a higher density of nodes in a certain area of the given domain, where the solution is more difficult to approximate, indicated by *a posteriori* error estimators. Hence it is an important approach to boost the accuracy and efficiency of finite element discretizations.

Actually, there are many *h*-versions of adaptive finite element methods for optimal control problems in [18–20]. But for a high order element such as a *hp*-version of the finite element method for optimal control problems they are very few. More recently, in [21], for the constrained optimal control problem governed by linear elliptic equations, the authors have derived *a posteriori* error estimates for the *hp* finite element solutions. Inspired by the work of [21], we consider *a posteriori* error estimates in L^2-H^1 norms and L^2-L^2 norms for *hp* finite element solutions of general semilinear Neumann boundary optimal control problems. To the best of our knowledge for optimal control problems, these *a posteriori* error estimates for the general semilinear boundary optimal control problems are new.

The paper is organized as follows. In Section 2, we discuss the *hp* finite element approximation for the semilinear Neumann boundary optimal control problems. In Section 3, we derive both L^2-H^1 *a posteriori* upper error bounds for the error estimates of the control, the state, and the co-state. Then we also obtain sharper *a posteriori* error estimates for the control approximation and error estimates in the L^2 norm for the state and co-state on the boundary. Finally, we give a conclusion and some possible future work in Section 4.

2 Finite element methods of boundary optimal control

In this section, we study the *hp* finite element approximation of semilinear convex optimal control problems where the control appears in the Neumann boundary conditions. To consider the *hp* finite element approximation of the semilinear boundary optimal control problems, we have to give a weak formula for the state equation. Let the state space be $V = H^1(\Omega)$ and $H = L^2(\Omega)$. Let

$$a(y,w) = \int_\Omega (A\nabla y)\cdot \nabla w\,dx, \quad \forall y,w \in V,$$

$$(f_1,f_2) = \int_\Omega f_1 f_2\,dx, \quad \forall (f_1,f_2) \in H \times H,$$

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

It follows from the assumptions on A that there are constants c and $C > 0$ such that

$$a(v,v) \geq c\|v\|_V^2, \qquad \bigl|a(v,w)\bigr| \leq C\|v\|_V\|w\|_V, \quad \forall v,w \in V. \tag{2.1}$$

Then the standard weak formula for the state equation reads as follows: find $y(u) \in V$ such that

$$a\bigl(y(u),w\bigr) + \bigl(\phi\bigl(y(u)\bigr),w\bigr) = (f,w) + (u+z_0,w)_U, \quad \forall w \in V. \tag{2.2}$$

Therefore, the above semilinear Neumann boundary optimal control problems can be restated as follows:

$$\min_{u\in K\subset U}\bigl\{g(y) + j(u)\bigr\}, \tag{2.3}$$

$$a(y,w) + \bigl(\phi(y),w\bigr) = (f,w) + (u+z_0,w)_U, \quad \forall w \in V. \tag{2.4}$$

It is well known (see [20]) that the boundary optimal control problems (2.3)-(2.4) has a solution (y,u) and that if a pair (y,u) is the solution of (2.3)-(2.4), then there is a co-state $p \in V$ such that the triplet (y,p,u) satisfies the following optimality conditions:

$$a(y,w) + \bigl(\phi(y),w\bigr) = (f,w) + (u+z_0,w)_U, \quad \forall w \in V, \tag{2.5}$$

$$a(q,p) + \bigl(\phi'(y)p,q\bigr) = \bigl(g'(y),q\bigr), \quad \forall q \in V, \tag{2.6}$$

$$\bigl(j'(u) + p, v-u\bigr)_U \geq 0, \quad \forall v \in K \subset U. \tag{2.7}$$

Now, we consider the *hp* finite element approximation for the boundary optimal control problem. We consider the triangulation $\mathcal{T}$ of the set $\Omega \subset \mathbb{R}^2$ which is a collection of elements $\tau \in \mathcal{T}$ (τ is a triangle); associated with each element τ is an affine element map $F_\tau : \hat{\tau} \to \tau$, where the reference element is the reference triangle $T = \{(x,y) \in \mathbb{R}^2 : 0 < x < 1, 0 < y < \min(x, 1-x)\}$. We consider the triangulation $\mathcal{T}$ which satisfies the standard conditions defined in [22]. We write $h_\tau = \operatorname{diam} \tau$. Assume that the triangulation $\mathcal{T}$ is γ-shape regular, *i.e.*,

$$h_\tau^{-1} \|F'_\tau\|_{L^\infty(\hat{\tau})} + h_\tau \|(F'_\tau)^{-1}\|_{L^\infty(\hat{\tau})} \le \gamma. \tag{2.8}$$

This implies (see [22]) that there exists a constant $C > 0$ that depends solely on γ such that

$$C^{-1} h_\tau \le h_{\tau'} \le C h_\tau, \quad \tau, \tau' \in \mathcal{T} \text{ with } \bar{\tau} \cap \bar{\tau}' \ne \emptyset, \tag{2.9}$$

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 further assume the triangulation $\mathcal{T}$ satisfies the relation between the patch and the reference patch. Let $\mathcal{T}_U$ be a partition of $\partial\Omega$ into disjoint regular 1-simplices s, so that $\partial\Omega = \bigcup_{s \in \mathcal{T}_U} \bar{s}$. Associated with every s is an affine map $F_s : \hat{s} \to s$, where $\hat{s} = [-1, 1]$. Assume that $\bar{s}$ and $\bar{s}'$ have either only one common vertex or are disjoint if s and $s' \in \mathcal{T}_U$.

For each element $\tau \in \mathcal{T}$, we denote $\mathcal{E}(\tau)$ the set of edges of τ and by $\mathcal{N}(\tau)$ 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 $s \in \mathcal{T}_U$, we choose a polynomial degree vector $\mathbf{p}_2 = (p_s)_{s \in \mathcal{T}_U}$ ($p_s \in \mathbb{N}$). $\mathcal{N}(\mathcal{T})$ 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 ($V \in \mathcal{N}(\mathcal{T})$, $e \in \mathcal{E}(\mathcal{T})$):

$$\mathcal{N}(e) = \{V \in \mathcal{N}(\mathcal{T}) : V \in \bar{e}\}, \qquad w_V = \{x \in \Omega : x \in \bar{\tau} \text{ and } \bar{\tau} \cap \{V\} \ne \emptyset\}^0,$$
$$w_e^1 = \bigcup_{V \in \mathcal{N}(e)} w_V, \qquad w_\tau^1 = \bigcup_{V \in \mathcal{N}(\tau)} w_V, \qquad p_e = \max\{p_\tau : e \in \mathcal{E}(\tau)\},$$

where χ^0 denotes the interior of the set χ. We denote by h_e (h_s) the length of the edge e (s). Additionally, c or C denotes a general positive constant independent of h_τ, p_τ, h_e, p_e, h_s, and p_s.

Next, we define the hp-FEM space $S^{\mathbf{p}_1}(\mathcal{T}) \subset H^1(\Omega)$ and the hp-DGFEM space $U^{\mathbf{p}_2}(\mathcal{T}_U) \subset L^2(\partial\Omega)$ by

$$S^{\mathbf{p}_1}(\mathcal{T}) = \{v \in C(\Omega) : v|_\tau \circ F_\tau \in P_{p_\tau}(\hat{\tau})\},$$
$$U^{\mathbf{p}_2}(\mathcal{T}_U) = \{v \in L^2(\partial\Omega) : v|_s \circ F_s \in P_{p_s}(\hat{s})\},$$

where $P_{p_\tau}(\hat{\tau}) := \operatorname{span}\{x^i y^j : 0 \le i + j \le p_\tau\}$, $P_{p_s}(\hat{s}) := \operatorname{span}\{x^i : 0 \le i \le p_s\}$. We assume that the polynomial degree vector $\mathbf{p}_1$ satisfies

$$\gamma^{-1} p_\tau \le p_{\tau'} \le \gamma p_\tau, \quad \tau, \tau' \in \mathcal{T} \text{ with } \bar{\tau} \cap \bar{\tau}' \ne \emptyset. \tag{2.10}$$

Let $K_{hp} = K \cap U^{\mathbf{p}_2}(\mathcal{T}_U)$ and $V_{hp} = S^{\mathbf{p}_1}(\mathcal{T})$, then for the finite element approximation of (2.3)-(2.4):

$$\min_{u_{hp} \in K_{hp}} \{g(y_{hp}) + j(u_{hp})\}, \tag{2.11}$$

$$a(y_{hp}, w_{hp}) + \big(\phi(y_{hp}), w\big) = (f, w_{hp}) + (u_{hp} + z_0, w_{hp})_U, \quad \forall w_{hp} \in V_{hp}. \tag{2.12}$$

It is well known that the boundary optimal control problem (2.11)-(2.12) has a solution (y_{hp}, u_{hp}) and that if a pair $(y_{hp}, u_{hp}) \in V_{hp} \times K_{hp}$ is the solution of (2.11)-(2.12), 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:

$$a(y_{hp}, w_{hp}) + \big(\phi(y_{hp}), w_{hp}\big) = (f, w_{hp}) + (u_{hp} + z_0, w_{hp})_U, \quad \forall w_{hp} \in V_{hp} \subset V, \tag{2.13}$$

$$a(q_{hp}, p_{hp}) + \big(\phi'(y_{hp})p_{hp}, q_{hp}\big) = \big(g'(y_{hp}), q_{hp}\big), \quad \forall q_{hp} \in V_{hp} \subset V, \tag{2.14}$$

$$\big(j'(u_{hp}) + p_{hp}, v_{hp} - u_{hp}\big)_U \geq 0, \quad \forall v_{hp} \in K_{hp} \subset U. \tag{2.15}$$

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

Lemma 2.1 *There exist a constant $C > 0$ independent of v, h_s, and p_s and a mapping $\pi_{p_s}^{h_s}$: $H^1(s) \to P_{p_s}(s)$ such that $\forall v \in H^1(s)$, $s \in \mathcal{T}_U$ the following inequality is valid*:

$$\big\|v - \pi_{p_s}^{h_s} v\big\|_{L^2(s)} \leq C\frac{h_s}{p_s}|v|_{H^1(s)},$$

where $P_{p_s}(s) := \operatorname{span}\{x^i y^j : 0 \leq i + j \leq p_s\}$.

Proof It follows easily from Proposition A.2 in [22] and the scaling argument. □

Lemma 2.2 [22] *Let $\mathbf{p}_1$ be an arbitrary polynomial degree distribution satisfies* (2.10). *Then there exists a linear operator $E_1 : H^1(\Omega) \to S^{\mathbf{p}_1}(\mathcal{T})$, and there exists a constant $C > 0$ depending solely on γ such that for every $v \in H^1(\Omega)$ and all elements $\tau \in \mathcal{T}$ and all edges $e \in \mathcal{E}(\mathcal{T})$,*

$$\|v - E_1 v\|_{L^2(\tau)} + \frac{h_\tau}{p_\tau}\big\|\nabla(v - E_1 v)\big\|_{L^2(\tau)} \leq C\frac{h_\tau}{p_\tau}\|\nabla v\|_{L^2(w_\tau^1)}, \tag{2.16}$$

$$\|v - E_1 v\|_{L^2(e)} \leq C\left(\frac{h_e}{p_e}\right)^{\frac{1}{2}}\|\nabla v\|_{L^2(w_e^1)}. \tag{2.17}$$

Lemma 2.3 *Let $\mathbf{p}_1$ be an arbitrary polynomial degree distribution satisfying* (2.10) *and $p_\tau \geq 2$, $\forall \tau \in \mathcal{T}$. Then there exists a bounded linear operator $E_2 : H^2(\Omega) \to S^{\mathbf{p}_1}(\mathcal{T})$, and there exists a constant $C > 0$ that depends solely on γ such that for every $v \in H^2(\Omega)$ and all elements $\tau \in \mathcal{T}$ and all edges $e \in \mathcal{E}(\mathcal{T})$,*

$$\|v - E_2 v\|_{L^2(\tau)} + \frac{h_\tau}{p_\tau}\big\|\nabla(v - E_2 v)\big\|_{L^2(\tau)} \leq C\left(\frac{h_\tau}{p_\tau}\right)^2|v|_{H^2(w_\tau^1)}, \tag{2.18}$$

$$\|v - E_2 v\|_{L^2(e)} \leq C\left(\frac{h_e}{p_e}\right)^{\frac{3}{2}}|v|_{H^2(w_e^1)}. \tag{2.19}$$

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

$$\phi(\varphi) - \phi(\rho) = -\tilde{\phi}'(\varphi)(\rho - \varphi) = -\phi'(\rho)(\rho - \varphi) + \tilde{\phi}''(\varphi)(\rho - \varphi)^2, \tag{2.20}$$

where

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

are bounded functions in $\bar{\Omega}$ [23].

3 Residual-based *a posteriori* error estimators

In this section, we discuss residual-based *a posteriori* error estimates for the semilinear Neumann boundary optimal control problems. First of all, we use the L^2 norm for estimating the control approximation error on the boundary, and the H^1 norm for the state and co-state approximation error on the domain. For simplicity of presentation, let

$$J(u) = g\big(y(u)\big) + j(u), \qquad J_{hp}(u_{hp}) = g\big(y(u_{hp})\big) + j(u_{hp}).$$

Then the optimal control problems of (2.3) and (2.11) read

$$\min_{u\in K}\big\{J(u)\big\} \tag{3.1}$$

and

$$\min_{u_{hp}\in K_{hp}}\big\{J_{hp}(u_{hp})\big\}. \tag{3.2}$$

It can be shown that

$$\big(J'(u), v\big)_U = \big(j'(u) + p, v\big)_U,$$
$$\big(J'_{hp}(u_{hp}), v\big)_U = \big(j'(u_{hp}) + p_{hp}, v\big)_U,$$
$$\big(J'(u_{hp}), v\big)_U = \big(j'(u_{hp}) + p(u_{hp}), v\big)_U,$$

where $p(u_{hp})$ is the solution of the auxiliary equations:

$$a\big(y(u_{hp}), w\big) + \big(\phi\big(y(u_{hp})\big), w\big) = (f, w) + (u_{hp} + z_0, w)_U, \quad \forall w \in V, \tag{3.3}$$
$$a\big(q, p(u_{hp})\big) + \big(\phi'\big(y(u_{hp})\big)p(u_{hp}), q\big) = \big(g'\big(y(u_{hp})\big), q\big), \quad \forall q \in V. \tag{3.4}$$

In order to estimate the control u, we introduce the $L^2(\partial\Omega)$-projection of u into $U^{\mathbf{p}_2}(\mathcal{T}_U)$, *i.e.*, let $P_{hp}u \in U^{\mathbf{p}_2}(\mathcal{T}_U)$ be the function defined by

$$(u - P_{hp}u, w_{hp})_U = 0, \quad \forall w_{hp} \in U^{\mathbf{p}_2}(\mathcal{T}_U). \tag{3.5}$$

Theorem 3.1 *Let* (y, u) *and* (y_{hp}, u_{hp}) *be the solutions of* (2.3)-(2.4) *and* (2.11)-(2.12). *Let* p *and* p_{hp} *be the solutions of the co-state equations* (2.6) *and* (2.14), *respectively. Assume that*

$$\big(J'(u) - J'(v), u - v\big)_U \ge c\|u - v\|^2_{L^2(\Omega_U)}, \quad \forall u, v \in U. \tag{3.6}$$

Moreover, we assume $j'(u_{hp}) + p_{hp} \in H^1(\Omega)$. *Then we have*

$$\|u - u_{hp}\|_{L^2(\partial\Omega)}^2 \le C\eta_1^2 + C\big\|p_{hp} - p(u_{hp})\big\|_{L^2(\partial\Omega)}^2, \tag{3.7}$$

where

$$\eta_1^2 = \sum_s \frac{h_s^2}{p_s^2}\big|j'(u_{hp}) + p_{hp}\big|_{H^1(s)}^2,$$

and $p(u_{hp})$ *is the solution of the system* (3.3)-(3.4).

Proof It follows from (2.7), (2.15), and (3.6) that

$$\begin{aligned}
c\|u - u_{hp}\|_{L^2(\partial\Omega)}^2
&\le \big(J'(u), u - u_{hp}\big)_U - \big(J'(u_{hp}), u - u_{hp}\big)_U \\
&\le -\big(J'(u_{hp}), u - u_{hp}\big)_U \\
&\le -\big(J'(u_{hp}), u - u_{hp}\big)_U + \big(j'(u_{hp}) + p_{hp}, v_{hp} - u_{hp}\big)_U \\
&= \big(J'_{hp}(u_{hp}), u_{hp} - u\big)_U + \big(J'_{hp}(u_{hp}) - J'(u_{hp}), u - u_{hp}\big)_U \\
&\quad + \big(j'(u_{hp}) + p_{hp}, v_{hp} - u_{hp}\big)_U \\
&= \big(j'(u_{hp}) + p_{hp}, u_{hp} - u\big)_U + \big(p_{hp} - p(u_{hp}), u - u_{hp}\big)_U \\
&\quad + \big(j'(u_{hp}) + p_{hp}, v_{hp} - u_{hp}\big)_U \\
&= \big(j'(u_{hp}) + p_{hp}, v_{hp} - u\big)_U + \big(p_{hp} - p(u_{hp}), u - u_{hp}\big)_U \\
&\le \big(j'(u_{hp}) + p_{hp}, v_{hp} - u\big)_U + C\big\|p_{hp} - p(u_{hp})\big\|_{L^2(\partial\Omega)}^2 \\
&\quad + \frac{c}{4}\|u - u_{hp}\|_{L^2(\partial\Omega)}^2.
\end{aligned} \tag{3.8}$$

Setting $w_{hp} = 1$ in (3.5), we have $\int_{\partial\Omega} P_{hp}u = \int_{\partial\Omega} u \ge 0$. Thus, we have $P_{hp}u \in K_{hp}$. Let $v_{hp} = P_{hp}u \in K_{hp}$. It follows from (3.5) and Lemma 2.1 that

$$\begin{aligned}
&\big(j'(u_{hp}) + p_{hp}, v_{hp} - u\big)_U \\
&\quad = \big(j'(u_{hp}) + p_{hp}, P_{hp}u - u\big)_U \\
&\quad = \sum_s \big(j'(u_{hp}) + p_{hp} - \pi_{p_s}^{h_s}\big(j'(u_{hp}) + p_{hp}\big), P_{hp}u - u\big)_s \\
&\quad \le C\sum_s \frac{h_s}{p_s}\big|j'(u_{hp}) + p_{hp}\big|_{H^1(s)}\|P_{hp}u - u\|_{L^2(s)} \\
&\quad \le C\sum_s \frac{h_s^2}{p_s^2}\big|j'(u_{hp}) + p_{hp}\big|_{H^1(s)}^2 + \frac{c}{4}\|P_{hp}u - u\|_{L^2(\partial\Omega)}^2 \\
&\quad \le C\sum_s \frac{h_s^2}{p_s^2}\big|j'(u_{hp}) + p_{hp}\big|_{H^1(s)}^2 + \frac{c}{4}\|u - u_{hp}\|_{L^2(\partial\Omega)}^2.
\end{aligned} \tag{3.9}$$

By using (3.8) and (3.9), we have

$$\|u-u_{hp}\|^2_{L^2(\partial\Omega)} \le C\sum_s \frac{h_s^2}{p_s^2}\big|j'(u_{hp})+p_{hp}\big|^2_{H^1(s)} + C\big\|p_{hp}-p(u_{hp})\big\|^2_{L^2(\partial\Omega)}. \tag{3.10}$$

Then (3.7) follows from (3.10). □

In the following theorem we estimate $\|p_{hp}-p(u_{hp})\|^2_{H^1(\Omega)}$ and then obtain the desired *hp a posteriori* error estimates.

Theorem 3.2 *Let* (y,p,u) *and* (y_{hp},p_{hp},u_{hp}) *be the solutions of* (2.5)-(2.7) *and* (2.13)-(2.15), *respectively. Assume that all the conditions in Theorem* 3.1 *hold. Then we have*

$$\|u-u_{hp}\|^2_{L^2(\partial\Omega)} + \|y-y_{hp}\|^2_{H^1(\Omega)} + \|p-p_{hp}\|^2_{H^1(\Omega)} \le C\sum_{i=1}^{7}\eta_i^2, \tag{3.11}$$

where η_1^2 *is defined in Theorem* 3.1 *and*

$$\begin{aligned}
\eta_2^2 &= \sum_\tau \int_\tau \frac{h_\tau^2}{p_\tau^2}\big(\operatorname{div}(A\nabla p_{hp}) - \phi'(y_{hp})p_{hp}\big)^2,\\
\eta_3^2 &= \sum_{e\subset\partial\Omega} \int_e \frac{h_e}{p_e}\big(A\nabla p_{hp}\cdot n - g'(y_{hp})\big)^2,\\
\eta_4^2 &= \sum_{e\cap\partial\Omega=\emptyset} \int_e \frac{h_e}{p_e}\big[(A\nabla p_{hp}\cdot n)\big]^2,\\
\eta_5^2 &= \sum_\tau \int_\tau \frac{h_\tau^2}{p_\tau^2}\big(f+\operatorname{div}(A\nabla y_{hp}) - \phi(y_{hp})\big)^2,\\
\eta_6^2 &= \sum_{e\subset\partial\Omega} \int_e \frac{h_e}{p_e}(A\nabla y_{hp}\cdot n - u_{hp} - z_0)^2,\\
\eta_7^2 &= \sum_{e\cap\partial\Omega=\emptyset} \int_e \frac{h_e}{p_e}\big[(A\nabla y_{hp}\cdot n)\big]^2.
\end{aligned}$$

Proof Let $e_p = p_{hp} - p(u_{hp})$ and E_1 be the linear operator defined in Lemma 2.2, we have

$$\begin{aligned}
c\|e_p\|^2_{H^1(\Omega)} &\le a(e_p,e_p) + \big(\phi'\big(y(u_{hp})\big)e_p, e_p\big)\\
&= a(e_p - E_1e_p, e_p) + a(E_1e_p, e_p) + \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), e_p\big)\\
&\quad - \big(\big(\phi'(y_{hp}) - \phi'\big(y(u_{hp})\big)\big)p_{hp}, e_p\big)\\
&= a(e_p - E_1e_p, e_p) + \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), e_p - E_1e_p\big)\\
&\quad + a(E_1e_p, e_p) + \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), E_1e_p\big)\\
&\quad - \big(\big(\phi'(y_{hp}) - \phi'\big(y(u_{hp})\big)\big)p_{hp}, e_p\big)\\
&= \sum_\tau \int_\tau A\nabla\big(p_{hp} - p(u_{hp})\big)\cdot\nabla(e_p - E_1e_p)\\
&\quad + \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), e_p - E_1e_p\big)\\
&\quad + \big(g'(y_{hp}) - g'\big(y(u_{hp})\big), E_1e_p\big) - \big(\big(\phi'(y_{hp}) - \phi'\big(y(u_{hp})\big)\big)p_{hp}, e_p\big).
\end{aligned} \tag{3.12}$$

It follows from (2.1), (2.14), (2.16)-(2.17), and (3.12) that

$$\begin{aligned}
c\|e_p\|_{H^1(\Omega)}^2 &\le \sum_\tau \int_\tau \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)(e_p - E_1 e_p) - \big(g'\big(y(u_{hp})\big), e_p - E_1 e_p\big) \\
&\quad + \sum_\tau \int_{\partial\tau} (A\nabla p_{hp}) \cdot n(e_p - E_1 e_p) \\
&\quad + \big(g'(y_{hp}) - g'\big(y(u_{hp})\big), E_1 e_p\big) - \big(\big(\phi'(y_{hp}) - \phi'\big(y(u_{hp})\big)\big)p_{hp}, e_p\big) \\
&= \sum_\tau \int_\tau \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)(e_p - E_1 e_p) + \big(g'(y_{hp}) - g'\big(y(u_{hp})\big), e_p\big) \\
&\quad + \sum_\tau \int_{\partial\tau} (A\nabla p_{hp}) \cdot n(e_p - E_1 e_p) - \big(\big(\phi'(y_{hp}) - \phi'\big(y(u_{hp})\big)\big)p_{hp}, e_p\big) \\
&\quad + \sum_{e\cap\partial\Omega=\emptyset} \int_e \big[(A\nabla p_{hp} \cdot n)\big](e_p - E_1 e_p) \\
&\quad + \sum_{e\subset\partial\Omega} \int_e \big(A\nabla p_{hp} \cdot n - g'(y_{hp})\big)(e_p - E_1 e_p) \\
&\le C\sum_\tau \frac{h_\tau^2}{p_\tau^2} \int_\tau \big(\operatorname{div}(A\nabla p_{hp}) - \phi'(y_{hp})p_{hp}\big)^2 + C\sum_{e\cap\partial\Omega=\emptyset} \frac{h_e}{p_e} \int_e \big[(A\nabla p_{hp} \cdot n)\big]^2 \\
&\quad + C\sum_{e\subset\partial\Omega} \frac{h_e}{p_e} \int_e \big(A\nabla p_{hp} \cdot n - g'(y_{hp})\big)^2 \\
&\quad + C\big\|\phi'(y_{hp}) - \phi'\big(y(u_{hp})\big)\big\|_{L^2(\partial\Omega)}^2 \\
&\quad + C\big\|g'(y_{hp}) - g'\big(y(u_{hp})\big)\big\|_{L^2(\partial\Omega)}^2 + \frac{c}{2}\|e_p\|_{H^1(\Omega)}^2.
\end{aligned} \tag{3.13}$$

Therefore, noting that g' is locally Lipschitz continuous, we have

$$\big\|p_{hp} - p(u_{hp})\big\|_{H^1(\Omega)}^2 \le C\sum_{i=2}^4 \eta_i^2 + C\big\|y_{hp} - y(u_{hp})\big\|_{L^2(\partial\Omega)}^2. \tag{3.14}$$

Similarly, it can be proved that

$$\big\|y_{hp} - y(u_{hp})\big\|_{H^1(\Omega)}^2 \le C\sum_{i=5}^7 \eta_i^2. \tag{3.15}$$

It follows from (3.14), (3.15), and the trace theorem that

$$\begin{aligned}
\big\|p_{hp} - p(u_{hp})\big\|_{H^1(\Omega)}^2 &\le C\sum_{i=2}^4 \eta_i^2 + C\big\|y_{hp} - y(u_{hp})\big\|_{L^2(\partial\Omega)}^2 \\
&\le C\sum_{i=2}^4 \eta_i^2 + C\big\|y_{hp} - y(u_{hp})\big\|_{H^1(\Omega)}^2 \\
&\le C\sum_{i=2}^7 \eta_i^2.
\end{aligned} \tag{3.16}$$

Note that

$$\|p-p_{hp}\|_{H^1(\Omega)} \leq \|p-p(u_{hp})\|_{H^1(\Omega)} + \|p_{hp}-p(u_{hp})\|_{H^1(\Omega)}, \tag{3.17}$$

$$\|y-y_{hp}\|_{H^1(\Omega)} \leq \|y-y(u_{hp})\|_{H^1(\Omega)} + \|y_{hp}-y(u_{hp})\|_{H^1(\Omega)}, \tag{3.18}$$

and

$$\|y-y(u_{hp})\|_{H^1(\Omega)} \leq C\|u-u_{hp}\|_{L^2(\partial\Omega)}, \tag{3.19}$$

$$\|p-p(u_{hp})\|_{H^1(\Omega)} \leq C\|y-y(u_{hp})\|_{H^1(\Omega)} \leq C\|u-u_{hp}\|_{L^2(\partial\Omega)}. \tag{3.20}$$

Combining (3.7), (3.14)-(3.16), and (3.17)-(3.20), we derive

$$\begin{aligned}
&\|u-u_{hp}\|^2_{L^2(\partial\Omega)} + \|y-y_{hp}\|^2_{H^1(\Omega)} + \|p-p_{hp}\|^2_{H^1(\Omega)} \\
&\quad\leq \|u-u_{hp}\|^2_{L^2(\partial\Omega)} + \|y-y(u_{hp})\|^2_{H^1(\Omega)} \\
&\qquad + \|y_{hp}-y(u_{hp})\|^2_{H^1(\Omega)} + \|p-p(u_{hp})\|^2_{H^1(\Omega)} + \|p_{hp}-p(u_{hp})\|^2_{H^1(\Omega)} \\
&\quad\leq \|u-u_{hp}\|^2_{L^2(\partial\Omega)} + \|y-y(u_{hp})\|^2_{H^1(\Omega)} + \|p-p(u_{hp})\|^2_{H^1(\Omega)} \\
&\quad\leq C\sum_{i=1}^{7}\eta_i^2.
\end{aligned} \tag{3.21}$$

Then we have proved (3.11). □

Next, we shall derive sharper *a posteriori* error estimates for the control approximation and error estimates in the L^2 norm for the state and co-state on the boundary. We introduce a subset of Ω: $\Omega_d = \{\tau \in \mathcal{T} : \bar{\tau} \cap \bar{\Omega}_d^- \neq \emptyset\}$, where $\Omega_d^- = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) < d\}$ and d is a constant independent of h_τ, p_τ, h_e, and p_e. Then we have the following improved residual-based *a posteriori* error estimates.

Theorem 3.3 *Let* (y, p, u) *and* (y_{hp}, p_{hp}, u_{hp}) *be the solutions of* (2.5)-(2.7) *and* (2.13)-(2.15), *respectively. Assume that* $j'(u_{hp}) + p_{hp} \in H^1(\Omega)$ *and* $(J'(u) - J'(v), u - v)_U \geq c\|u - v\|^2_{L^2(\Omega_U)}$, $\forall u, v \in U$. *Moreover*, $p_\tau \geq 2$, $\forall \tau \in \mathcal{T}$ *and* $g'(\cdot)$ *is locally Lipschitz continuous. Then*

$$\|u-u_{hp}\|^2_{L^2(\partial\Omega)} + \|y-y_{hp}\|^2_{L^2(\partial\Omega)} + \|p-p_{hp}\|^2_{L^2(\partial\Omega)} \leq C\sum_{i=1}^{13}\kappa_i^2, \tag{3.22}$$

where

$$\kappa_1^2 = \sum_s \frac{h_s^2}{p_s^2}\left|j'(u_{hp}) + p_{hp}\right|^2_{H^1(s)},$$

$$\kappa_2^2 = \sum_{\tau\subset\Omega_d}\int_\tau \frac{h_\tau^2}{p_\tau^2}\left(\operatorname{div}(A\nabla p_{hp}) - \phi'(y_{hp})p_{hp}\right)^2,$$

$$\kappa_3^2 = \sum_{e\cap\partial\Omega=\emptyset, e\in\Omega_d}\int_e \frac{h_e}{p_e}\left[(A\nabla p_{hp}\cdot n)\right]^2,$$

$$\kappa_4^2 = \sum_{e\subset\partial\Omega} \int_e \frac{h_e}{p_e}\big(A\nabla p_{hp}\cdot n - g'(y_{hp})\big)^2,$$

$$\kappa_5^2 = \sum_{\tau} \int_\tau \frac{h_\tau^4}{p_\tau^4}\big(\operatorname{div}(A\nabla p_{hp}) - \phi'(y_{hp})p_{hp}\big)^2,$$

$$\kappa_6^2 = \sum_{e\cap\partial\Omega=\emptyset} \int_e \frac{h_e^3}{p_e^3}\big[(A\nabla p_{hp}\cdot n)\big]^2,$$

$$\kappa_7^2 = \sum_{e\subset\partial\Omega} \int_e \frac{h_e^3}{p_e^3}\big(A\nabla p_{hp}\cdot n - g'(y_{hp})\big)^2,$$

and

$$\kappa_8^2 = \sum_{\tau\subset\Omega_d} \int_\tau \frac{h_\tau^2}{p_\tau^2}\big(f + \operatorname{div}(A\nabla y_{hp}) - \phi(y_{hp})\big)^2,$$

$$\kappa_9^2 = \sum_{e\cap\partial\Omega=\emptyset, e\in\Omega_d} \int_e \frac{h_e}{p_e}\big[(A\nabla y_{hp}\cdot n)\big]^2,$$

$$\kappa_{10}^2 = \sum_{e\subset\partial\Omega} \int_e \frac{h_e}{p_e}(A\nabla y_{hp}\cdot n - u_{hp} - z_0)^2,$$

$$\kappa_{11}^2 = \sum_{\tau} \int_\tau \frac{h_\tau^4}{p_\tau^4}\big(f + \operatorname{div}(A\nabla y_{hp}) - \phi(y_{hp})\big)^2,$$

$$\kappa_{12}^2 = \sum_{e\cap\partial\Omega=\emptyset} \int_e \frac{h_e^3}{p_e^3}\big[(A\nabla y_{hp}\cdot n)\big]^2,$$

$$\kappa_{13}^2 = \sum_{e\subset\partial\Omega} \int_e \frac{h_e^3}{p_e^3}(A\nabla y_{hp}\cdot n - u_{hp} - z_0)^2.$$

Proof For the proof of this theorem, we estimate (3.22) in the following five parts, respectively.

Part I. First, we estimate $\|p_{hp} - p(u_{hp})\|_{L^2(\partial\Omega)}$. Let $e_p = p_{hp} - p(u_{hp})$ and $e_y = y_{hp} - y(u_{hp})$, then there is some $\xi \in C^\infty(\Omega_d^-)$ satisfying $\xi = 0$ on $\partial\Omega_d^-\backslash\partial\Omega$ and $\xi = 1$ on $\partial\Omega$. It follows from the trace theorem that

$$\|e_p\|_{L^2(\partial\Omega)}^2 = \|\xi e_p\|_{L^2(\partial\Omega)}^2 \le C\|\xi e_p\|_{H^1(\Omega)}^2. \tag{3.23}$$

By using the assumption of A, we have

$$\int_\Omega A\nabla(\xi e_p)\nabla(\xi e_p) = \int_\Omega A\nabla e_p\nabla\big(\xi^2 e_p\big) + \int_\Omega (e_p)^2 A\nabla\xi\nabla\xi. \tag{3.24}$$

Let $v^p = \xi^2 e_p$, and let E_1 be the linear operator defined in Lemma 2.2. It follows from (2.1), (2.14), (3.4), and (3.24) that

$$\begin{aligned} c\|\xi e_p\|_{H^1(\Omega)}^2 &\le a(\xi e_p, \xi e_p) = a\big(v^p, e_p\big) + \int_\Omega (e_p)^2 A\nabla\xi\nabla\xi \\ &= a\big(v^p - E_1 v^p, e_p\big) + a\big(E_1 v^p, e_p\big) + \int_\Omega (e_p)^2 A\nabla\xi\nabla\xi \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\tau} \int_{\tau} \big(A\nabla\big(p_{hp} - p(u_{hp})\big)\big) \cdot \nabla\big(v^p - E_1 v^p\big) \\
&\quad + \int_{\partial\Omega} \big(g'(y_{hp}) - g'\big(y(u_{hp})\big)\big) E_1 v^p \\
&\quad - \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), E_1 v^p\big) + \int_{\Omega} (e_p)^2 A\nabla\xi\nabla\xi \\
&= \sum_{\tau} \int_{\tau} \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)\big(v^p - E_1 v^p\big) - \big(\phi'(y_{hp})p_{hp}, v^p - E_1 v^p\big) \\
&\quad - \int_{\partial\Omega} g'\big(y(u_{hp})\big)\big(v^p - E_1 v^p\big) + \big(\phi'\big(y(u_{hp})\big)p(u_{hp}), v^p - E_1 v^p\big) \\
&\quad + \int_{\partial\Omega} \big(g'(y_{hp}) - g'\big(y(u_{hp})\big)\big) E_1 v^p - \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), E_1 v^p\big) \\
&\quad + \sum_{e\cap\partial\Omega=\emptyset} \int_{e} \big[(A\nabla p_{hp} \cdot n)\big]\big(v^p - E_1 v^p\big) \\
&\quad + \sum_{e\subset\partial\Omega} \int_{e} (A\nabla p_{hp}) \cdot n\big(v^p - E_1 v^p\big) + \int_{\Omega} (e_p)^2 A\nabla\xi\nabla\xi \\
&= \sum_{\tau} \int_{\tau} \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)\big(v^p - E_1 v^p\big) \\
&\quad - \int_{\partial\Omega} \big(g'(y_{hp}) - g'\big(y(u_{hp})\big)\big) v^p + \sum_{e\cap\partial\Omega=\emptyset} \int_{e} \big[(A\nabla p_{hp} \cdot n)\big]\big(v^p - E_1 v^p\big) \\
&\quad + \sum_{e\subset\partial\Omega} \int_{e} \big((A\nabla p_{hp}) \cdot n - g'(y_{hp})\big)\big(v^p - E_1 v^p\big) \\
&\quad - \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), v^p\big) + \int_{\Omega} (e_p)^2 A\nabla\xi\nabla\xi. \qquad (3.25)
\end{aligned}
$$

By using Lemma 2.2, we have

$$
\begin{aligned}
c\|\xi e_p\|_{H^1(\Omega)}^2 &= \sum_{\tau} \int_{\tau} \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)\big(v^p - E_1 v^p\big) \\
&\quad - \int_{\partial\Omega} \big(g'(y_{hp}) - g'\big(y(u_{hp})\big)\big) v^p + \sum_{e\cap\partial\Omega=\emptyset} \int_{e} \big[(A\nabla p_{hp} \cdot n)\big]\big(v^p - E_1 v^p\big) \\
&\quad + \sum_{e\subset\partial\Omega} \int_{e} \big((A\nabla p_{hp}) \cdot n - g'(y_{hp})\big)\big(v^p - E_1 v^p\big) \\
&\quad - \big(\phi'(y_{hp})e^p, v^p\big) - \big(\tilde{\phi}''(y_{hp})p(u_{hp})e^y, v^p\big) + \int_{\Omega} (e_p)^2 A\nabla\xi\nabla\xi \\
&\leq C \sum_{\tau\subset\Omega_d} \frac{h_\tau^2}{p_\tau^2} \big(\operatorname{div}(A\nabla p_{hp}) - \phi'(y_{hp})p_{hp}\big)^2 \\
&\quad + C \sum_{e\cap\partial\Omega=\emptyset, e\in\Omega_d} \int_{e} \frac{h_e}{p_e} \big[(A\nabla p_{hp} \cdot n)\big]^2 \\
&\quad + C \sum_{e\subset\partial\Omega} \frac{h_e}{p_e} \int_{e} \big(A\nabla p_{hp} \cdot n - g'(y_{hp})\big)^2 + C\big\|y_{hp} - y(u_{hp})\big\|_{L^2(\Omega)}^2 \\
&\quad + C\big\|y_{hp} - y(u_{hp})\big\|_{L^2(\partial\Omega)}^2 + C\|e_p\|_{L^2(\Omega)}^2 + \frac{c}{2}\big\|v^p\big\|_{H^1(\Omega)}^2,
\end{aligned}
$$

where we use the property that $\xi = 0$ on Ω/Ω_d^-. Noting that $\|v^p\|_{H^1(\Omega)}^2 \le \|\xi e_p\|_{H^1(\Omega)}^2$, we have

$$\begin{aligned}
&\big\|\xi\big(p_{hp}-p(u_{hp})\big)\big\|_{H^1(\Omega)}^2\\
&\quad\le C\sum_{i=2}^{4}\kappa_i^2 + C\big\|y_{hp}-y(u_{hp})\big\|_{L^2(\Omega)}^2\\
&\qquad + C\big\|y_{hp}-y(u_{hp})\big\|_{L^2(\partial\Omega)}^2 + C\big\|p_{hp}-p(u_{hp})\big\|_{L^2(\Omega)}^2.
\end{aligned}\tag{3.26}$$

It follows from (3.23) and (3.26) that

$$\begin{aligned}
&\big\|p_{hp}-p(u_{hp})\big\|_{L^2(\partial\Omega)}^2\\
&\quad\le C\big\|\xi\big(p_{hp}-p(u_{hp})\big)\big\|_{H^1(\Omega)}^2\\
&\quad\le C\sum_{i=2}^{4}\kappa_i^2 + C\big\|y_{hp}-y(u_{hp})\big\|_{L^2(\Omega)}^2\\
&\qquad + C\big\|y_{hp}-y(u_{hp})\big\|_{L^2(\partial\Omega)}^2 + C\big\|p_{hp}-p(u_{hp})\big\|_{L^2(\Omega)}^2.
\end{aligned}\tag{3.27}$$

Part II. Now, we estimate $\|p_{hp}-p(u_{hp})\|_{L^2(\Omega)}$. Let φ_p be the solution of the following equation:

$$a(\varphi_p, w) = (e_p, w), \quad \forall w \in V. \tag{3.28}$$

Noting that Ω is convex [24], it has been shown that

$$\|\varphi_p\|_{2,\Omega} \le C\|e_p\|_{0,\Omega}. \tag{3.29}$$

Let E_2 be the linear operator defined in Lemma 2.3. It follows from (2.1), (2.14), (3.28)-(3.29), and Lemma 2.3 that

$$\begin{aligned}
\|e_p\|_{L^2(\Omega)}^2 &= a(\varphi_p, e_p) = a(\varphi_p - E_2\varphi_p, e_p) + a(E_2\varphi_p, e_p)\\
&= \sum_\tau \int_\tau \big(A\nabla\big(p_{hp}-p(u_{hp})\big)\big)\cdot\nabla(\varphi_p - E_2\varphi_p)\\
&\quad + \big(g'(y_{hp}) - g'\big(y(u_{hp})\big), E_2\varphi_p\big) - \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), E_2\varphi_p\big)\\
&= \sum_\tau \int_\tau \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)(\varphi_p - E_2\varphi_p) - \big(\phi'(y_{hp})p_{hp}, \varphi_p - E_2\varphi_p\big)\\
&\quad - \big(g'\big(y(u_{hp})\big), \varphi_p - E_2\varphi_p\big) + \big(\phi'\big(y(u_{hp})\big)p(u_{hp}), \varphi_p - E_2\varphi_p\big)\\
&\quad + \big(g'(y_{hp}) - g'\big(y(u_{hp})\big), E_2\varphi_p\big) - \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), E_2\varphi_p\big)\\
&\quad + \sum_{e\cap\partial\Omega=\emptyset}\int_e \big[(A\nabla p_{hp}\cdot n)\big](\varphi_p - E_2\varphi_p) + \sum_{e\subset\partial\Omega}\int_e (A\nabla p_{hp})\cdot n(\varphi_p - E_2\varphi_p)\\
&= \sum_\tau \int_\tau \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)(\varphi_p - E_2\varphi_p) + \big(g'(y_{hp}) - g'\big(y(u_{hp})\big), \varphi_p\big)
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{e\cap\partial\Omega=\emptyset} \int_e \big[(A\nabla p_{hp}\cdot n)\big](\varphi_p - E_2\varphi_p) \\
&+ \sum_{e\subset\partial\Omega} \int_e \big((A\nabla p_{hp})\cdot n - g'(y_{hp})\big)(\varphi_p - E_2\varphi_p) \\
&- \big(\phi'(y_{hp})p_{hp} - \phi'\big(y(u_{hp})\big)p(u_{hp}), \varphi_p\big).
\end{aligned}
$$

By using Lemma 2.3, we have

$$
\begin{aligned}
\|e_p\|^2_{L^2(\Omega)} &= \sum_\tau \int_\tau \big(-\operatorname{div}(A\nabla p_{hp}) + \phi'(y_{hp})p_{hp}\big)(\varphi_p - E_2\varphi_p) + \big(g'(y_{hp}) - g'\big(y(u_{hp})\big), \varphi_p\big) \\
&\quad + \sum_{e\cap\partial\Omega=\emptyset} \int_e \big[(A\nabla p_{hp}\cdot n)\big](\varphi_p - E_2\varphi_p) \\
&\quad + \sum_{e\subset\partial\Omega} \int_e \big((A\nabla p_{hp})\cdot n - g'(y_{hp})\big)(\varphi_p - E_2\varphi_p) \\
&\quad - \big(\phi'(y_{hp})e^p, v^p\big) - \big(\tilde{\phi}''(y_{hp})p(u_{hp})e^y, v^p\big) \\
&\le C(\delta) \sum_\tau \frac{h_\tau^4}{p_\tau^4} \int_\tau \big(\operatorname{div}(A\nabla p_{hp}) - \phi'(y_{hp})p_{hp}\big)^2 \\
&\quad + C(\delta) \sum_{e\cap\partial\Omega=\emptyset} \frac{h_e^3}{p_e^3} \int_e \big[(A\nabla p_{hp}\cdot n)\big]^2 \\
&\quad + C(\delta) \sum_{e\subset\partial\Omega} \frac{h_e^3}{p_e^3} \int_e \big(A\nabla p_{hp}\cdot n - g'(y_{hp})\big)^2 + C(\delta)\big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\Omega)} \\
&\quad + C(\delta)\big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\partial\Omega)} + C\delta\|\xi\|^2_{H^2(\Omega)} \\
&= C(\delta)\left(\sum_{i=5}^7 \kappa_i^2 + \big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\Omega)} + \big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\partial\Omega)}\right) + C\delta\|\xi\|^2_{H^2(\Omega)}.
\end{aligned}
$$

Let δ be small enough, it follows from (3.29) that

$$
\big\|p_{hp} - p(u_{hp})\big\|^2_{L^2(\Omega)} \le C\sum_{i=5}^7 \kappa_i^2 + C\big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\Omega)} + C\big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\partial\Omega)}. \quad (3.30)
$$

Part III. Next, we estimate $\|y_{hp} - y(u_{hp})\|_{L^2(\partial\Omega)}$. Let $e_y = y_{hp} - y(u_{hp})$ and $v^y = \xi^2 e_y$, by using (2.1), (2.13), (3.3), and Lemma 2.2, then we have

$$
\begin{aligned}
c\|\xi e_y\|^2_{H^1(\Omega)} &\le a(\xi e_y, \xi e_y) = a\big(v^y, e_y\big) + \int_\Omega (e_y)^2 A\nabla\xi\nabla\xi \\
&= a\big(e_y, v^y - E_1 v^y\big) + \int_\Omega (e_y)^2 A\nabla\xi\nabla\xi \\
&= \sum_{\tau\subset\Omega_d} \int_\tau A\nabla\big(y_{hp} - y(u_{hp})\big)\cdot\big(v^y - E_1 v^y\big) + \int_\Omega (e_y)^2 A\nabla\xi\nabla\xi \\
&= \sum_{\tau\subset\Omega_d} \int_\tau \big(-\operatorname{div}(A\nabla y_{hp}) + \phi(y_{hp}) - f\big)\big(v^y - E_1 v^y\big)
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{e\cap\partial\Omega=\emptyset}\int_e \big[(A\nabla y_{hp}\cdot n)\big]\big(v^y - E_1 v^y\big) \\
&+ \int_{\partial\Omega}\big((A\nabla y_{hp})\cdot n - u_{hp} - z_0\big)\big(v^y - E_1 v^y\big) \\
&- \big(\phi(y_{hp}) - \phi\big(y(u_{hp})\big), v^y - E_1 v^y\big) + \int_\Omega (e_y)^2 A\nabla\xi\nabla\xi \\
= {} & \sum_{\tau\subset\Omega_d}\int_\tau \big(-\operatorname{div}(A\nabla y_{hp}) + \phi(y_{hp}) - f\big)\big(v^y - E_1 v^y\big) \\
&+ \sum_{e\cap\partial\Omega=\emptyset}\int_e \big[(A\nabla y_{hp}\cdot n)\big]\big(v^y - E_1 v^y\big) \\
&+ \int_{\partial\Omega}\big((A\nabla y_{hp})\cdot n - u_{hp} - z_0\big)\big(v^y - E_1 v^y\big) \\
&- \big(\tilde{\phi}'(y_{hp})\big(y_{hp} - y(u_{hp})\big), v^y - E_1 v^y\big) + \int_\Omega (e_y)^2 A\nabla\xi\nabla\xi \\
\le {} & C\sum_{\tau\subset\Omega_d}\int_\tau h_\tau^2\big(f + \operatorname{div}(A\nabla y_{hp}) - \phi(y_{hp})\big)^2 \\
&+ C\sum_{e\cap\partial\Omega=\emptyset, e\in\Omega_d}\int_e h_e\big[(A\nabla y_{hp}\cdot n)\big]^2 \\
&+ C\sum_{e\cap\partial\Omega}\int_e h_e(A\nabla y_{hp}\cdot n - u_{hp} - z_0)^2 + C\|e_y\|^2_{L^2(\Omega)} + \frac{c}{2}\|\xi e_y\|^2_{H^1(\Omega)} \\
= {} & C\sum_{i=8}^{10}\kappa_i^2 + C\big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\Omega)} + \frac{c}{2}\|\xi e_y\|^2_{H^1(\Omega)}.
\end{aligned}
\tag{3.31}
$$

Therefore, it follows from (3.31) and the trace theorem that

$$
\begin{aligned}
\big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\partial\Omega)} &\le C\big\|\xi\big(y_{hp} - y(u_{hp})\big)\big\|^2_{H^1(\Omega)} \\
&\le C\sum_{i=8}^{10}\kappa_i^2 + C\big\|y_{hp} - y(u_{hp})\big\|^2_{L^2(\Omega)}.
\end{aligned}
\tag{3.32}
$$

Part IV. Furthermore, we estimate $\|y_{hp} - y(u_{hp})\|_{L^2(\Omega)}$. Let φ_y be the solution of the equation

$$a(w, \varphi_y) = (e_y, w), \quad \forall w \in V.$$

Then we have

$$\|\varphi_y\|_{2,\Omega} \le C\|e_y\|_{0,\Omega}. \tag{3.33}$$

Similarly, we have

$$
\begin{aligned}
c\|e_y\|^2_{L^2(\Omega)} &= a(e_y, \varphi_y) = a(e_y, \varphi_y - E_2\varphi_y) \\
&= \sum_\tau \int_\tau A\nabla\big(y_{hp} - y(u_{hp})\big)\cdot\nabla(\varphi_y - E_2\varphi_y)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_\tau \int_\tau \big(-\operatorname{div}(A\nabla y_{hp}) + \phi(y_{hp}) - f\big)(\varphi_y - E_2\varphi_y) \\
&\quad + \sum_{e\cap\partial\Omega=\emptyset} \int_e \big[(A\nabla y_{hp}\cdot n)\big](\varphi_y - E_2\varphi_y) \\
&\quad + \int_{\partial\Omega} \big((A\nabla y_{hp})\cdot n - u_{hp} - z_0\big)(\varphi_y - E_2\varphi_y) \\
&\quad - \big(\phi(y_{hp}) - \phi\big(y(u_{hp})\big), \varphi_y - E_2\varphi_y\big) \\
&= \sum_\tau \int_\tau \big(-\operatorname{div}(A\nabla y_{hp}) + \phi(y_{hp}) - f\big)(\varphi_y - E_2\varphi_y) \\
&\quad + \sum_{e\cap\partial\Omega=\emptyset} \int_e \big[(A\nabla y_{hp}\cdot n)\big](\varphi_y - E_2\varphi_y) \\
&\quad + \int_{\partial\Omega} \big((A\nabla y_{hp})\cdot n - u_{hp} - z_0\big)(\varphi_y - E_2\varphi_y) \\
&\quad - \big(\tilde{\phi}'(y_{hp})\big(y_{hp} - y(u_{hp})\big), \varphi_y - E_2\varphi_y\big) \\
&\le C(\delta) \sum_\tau \frac{h_\tau^4}{p_\tau^4} \int_\tau \big(f + \operatorname{div}(A\nabla p_{hp}) - \phi'(y_{hp})p_{hp}\big)^2 \\
&\quad + C(\delta) \sum_{e\cap\partial\Omega=\emptyset} \frac{h_e^3}{p_e^3} \int_e \big[(A\nabla p_{hp}\cdot n)\big]^2 \\
&\quad + C(\delta) \sum_{e\subset\partial\Omega} \frac{h_e^3}{p_e^3} \int_e (A\nabla y_{hp}\cdot n - u_{hp} - z_0)^2 + \frac{c}{2}\|e_y\|_{L^2(\Omega)}^2 + C\delta\|\varphi_y\|_{H^2(\Omega)}^2 \\
&= C(\delta) \sum_{i=11}^{13} \kappa_i^2 + \frac{c}{2}\|e_y\|_{L^2(\Omega)}^2 + C\delta\|\varphi_y\|_{H^2(\Omega)}^2.
\end{aligned}
$$

Let δ be small enough, it follows from (3.33) that

$$
\big\|y_{hp} - y(u_{hp})\big\|_{L^2(\Omega)}^2 \le C \sum_{i=11}^{13} \kappa_i^2. \tag{3.34}
$$

It follows from (3.7), (3.27), (3.30), (3.32), and (3.34) that

$$
\|u - u_{hp}\|_{L^2(\partial\Omega)}^2 + \big\|y(u_{hp}) - y_{hp}\big\|_{L^2(\partial\Omega)}^2 + \big\|p(u_{hp}) - p_{hp}\big\|_{L^2(\partial\Omega)}^2 \le C \sum_{i=1}^{13} \kappa_i^2. \tag{3.35}
$$

Part V. Finally, it is easy to see that

$$
\|y - y_{hp}\|_{L^2(\partial\Omega)} \le \big\|y - y(u_{hp})\big\|_{L^2(\partial\Omega)} + \big\|y(u_{hp}) - y_{hp}\big\|_{L^2(\partial\Omega)}, \tag{3.36}
$$

$$
\|p - p_{hp}\|_{L^2(\partial\Omega)} \le \big\|p - p(u_{hp})\big\|_{L^2(\partial\Omega)} + \big\|p(u_{hp}) - p_{hp}\big\|_{L^2(\partial\Omega)}, \tag{3.37}
$$

and

$$
\big\|y - y(u_{hp})\big\|_{L^2(\partial\Omega)} \le C\|u - u_{hp}\|_{L^2(\partial\Omega)}, \tag{3.38}
$$

$$
\big\|p - p(u_{hp})\big\|_{L^2(\partial\Omega)} \le C\big\|y - y(u_{hp})\big\|_{L^2(\partial\Omega)} \le C\|u - u_{hp}\|_{L^2(\partial\Omega)}. \tag{3.39}
$$

It follows from (3.35) and (3.36)-(3.39) that

$$\begin{aligned}
&\|u-u_{hp}\|^2_{L^2(\partial\Omega)}+\|y-y_{hp}\|^2_{L^2(\partial\Omega)}+\|p-p_{hp}\|^2_{L^2(\partial\Omega)}\\
&\quad\le\|u-u_{hp}\|^2_{L^2(\partial\Omega)}+\big\|y-y(u_{hp})\big\|^2_{L^2(\partial\Omega)}\\
&\qquad+\big\|y_{hp}-y(u_{hp})\big\|^2_{L^2(\partial\Omega)}+\big\|p-p(u_{hp})\big\|^2_{L^2(\partial\Omega)}+\big\|p_{hp}-p(u_{hp})\big\|^2_{L^2(\partial\Omega)}\\
&\quad\le\|u-u_{hp}\|^2_{L^2(\partial\Omega)}+\big\|y-y(u_{hp})\big\|^2_{L^2(\partial\Omega)}+\big\|p-p(u_{hp})\big\|^2_{L^2(\partial\Omega)}\\
&\quad\le C\sum_{i=1}^{13}\eta_i^2. \qquad (3.40)
\end{aligned}$$

Then (3.22) follows from (3.40). □

4 Conclusion and future work

In this paper, we use 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 residual-based *a posteriori* error estimates in L^2-H^1 norms for the semilinear Neumann boundary optimal control problems. Then we also give sharper *a posteriori* error estimates for the control approximation and error estimates in the L^2 norm for the state and co-state on the boundary. To the best of our knowledge in the context of optimal control problems, these *a posteriori* error estimates for the semilinear Neumann boundary optimal control problems are new.

In future, we shall consider the *hp* finite element method for hyperbolic optimal control problems. Furthermore, we shall consider *a posteriori* error estimates and 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, SZ, CH, and HL participated in the sequence alignment and drafted the manuscript. All authors read and approved the final manuscript.

Author details

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

Acknowledgements

This work is supported by National Basic Research Program (2012CB955804), Major Research Plan of National Natural Science Foundation of China (91430108), National Science Foundation of China (11201510, 11171251), China Postdoctoral Science Foundation (2015M580197), Chongqing Research Program of Basic Research and Frontier Technology (cstc2015jcyjA20001), Ministry of education Chunhui projects (Z2015139), Major Program of Tianjin University of Finance and Economics (ZD1302) and Science and Technology Project of Wanzhou District of Chongqing (2013030050).

Received: 26 August 2015 Accepted: 15 February 2016 Published online: 02 March 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: 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)
4. 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)

5. 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)
6. 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)
7. Li, R, Liu, W, Ma, H, Tang, T: Adaptive finite element approximation for distributed elliptic optimal control problems. SIAM J. Control Optim. **41**, 1321-1349 (2002)
8. Lions, JL: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
9. Liu, W, Yan, N: *A posteriori* error estimates for convex boundary control problems. SIAM J. Numer. Anal. **39**, 73-99 (2001)
10. 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)
11. 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)
12. Lu, Z, Chen, Y, Zheng, W: *A posteriori* error estimates of lowest order Raviart-Thomas mixed finite element methods for bilinear optimal control problems. East Asian J. Appl. Math. **2**, 108-125 (2012)
13. 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)
14. 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)
15. 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)
16. Babuska, I, Suri, M: The *hp*-version of the finite element method with quasiuniform meshes. RAIRO Modél. Math. Anal. Numér. **21**, 199-238 (1987)
17. 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)
18. 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)
19. Liu, W, Yan, N: *A posteriori* error estimates for some model boundary control problems. J. Comput. Appl. Math. **120**, 159-173 (2000)
20. Liu, W, Yan, N: Adaptive Finite Element Methods for Optimal Control Governed by PDEs. Science Press, Beijing (2008)
21. 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)
22. Melenk, JM: *hp*-Interpolation of nonsmooth functions and an application to *hp-a posteriori* error estimation. SIAM J. Numer. Anal. **43**, 127-155 (2005)
23. Milner, FA: Mixed finite element methods for quasilinear second-order elliptic problems. Math. Comput. **44**, 303-320 (1985)
24. Lasyzhenskaya, O, Uraltseva, N: Linear and Quasilinear Elliptic Equations. Springer, Berlin (1972)