INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 8, Number 1, Pages 48–69

A POSTERIORI ERROR ESTIMATES OF *hp*-FEM FOR OPTIMAL CONTROL PROBLEMS

WEI GONG, WENBIN LIU, AND NINGNING YAN

Abstract. In this paper, we investigate a posteriori error estimates of the *hp*-finite element method for a distributed convex optimal control problem governed by the elliptic partial differential equations. A family of weighted a posteriori error estimators of residual type are formulated. Both reliability and efficiency of the estimators are analyzed.

Key Words. *hp*-finite element method, optimal control problem, a posteriori error estimates

1. Introduction

Finite element approximation plays an important role in the numerical methods of optimal control problems. There have been extensive theoretical and numerical studies for the finite element approximation of various optimal control problems. However the literature is too huge to even give a very brief review here. In recent years, the adaptive finite element method has been extensively investigated. Adaptive finite element approximation is among the most important means to boost the accuracy and efficiency of the finite element discretizations. It ensures a higher density of nodes in certain areas of the given domain, where the solution is more difficult to approximate, using an a posteriori error indicator. We acknowledge the pioneering work due to Babuška and Rheinboldt [4]. Further references can be found in the monographs [2], [5], [31], and the references cited therein.

In the recent years, adaptive finite elements for optimal control has become a focus of research interests. There have appeared many research papers on the adaptivity of various optimal control problems. For example, [6] studied the adaptive finite element method for optimal control problems via a goal-orientated approach, while a posteriori error estimates of residual type were derived for convex distributed optimal control problems governed by the elliptic and the parabolic equations in [18], [22]-[24], and for boundary control problems in [21].

To authors' knowledge, the papers discussing the adaptive finite element methods for optimal control problems are all related to low order FEM, i.e. h-FEM. In the adaptive h-FEM, the adaptivity is performed by mesh refinement guided by a posteriori error estimators. There are also many high order methods, such as spectral element methods, the p-version and the hp-version finite element methods, which have been applied to many practical problems. Using the local refinement of the meshes where the solution is singular and applying higher order polynomials where the solution is smooth, the adaptive hp-version finite element method can

Received by the editors October 13, 2009 and, in revised form, April 26, 2010.
2000 *Mathematics Subject Classification.* 49J20, 65N30.
This research was supported by the National Basic Research Program under the Grant 2005CB321701 and the National Natural Science Foundation of China under the Grant 10771211.

achieve very high computation efficiency. There have been some extensive investigations of adaptive hp-FEM for the elliptic partial differential equations (see, e.g., [3], [7], [10], [15], [26], [28] and [29]).

It seems to be very suitable to apply the hp finite element method to approximate optimal control problem, see [20]. The main objective of this paper is to establish a posteriori error estimates for the hp-version finite element approximation of a model optimal control problem governed by the elliptic partial differential equation, which were not available before and can be used to guide the hp-adaptivity process. In this paper, we proved the upper and lower bounds of the a posteriori error estimates, although there is a gap of order p^2 between the lower and the upper bounds due to the existing gap in the a posteriori estimates for the hp-adaptive finite element approximation of the elliptic equations. We also formulate a family of a posteriori error estimators given by weighted residuals on the elements and the edges. In our work, we used some techniques that have been used for a posteriori error estimates of the h-version FEM for optimal control problems (see, e.g., [22]-[24] for more details). We also used the weighted techniques and some estimates of the hp-interpolation of Clément-type proposed in [26], where a posteriori error estimates were obtained for hp-FEM of the elliptic partial differential equations. In comparison with the hp a posteriori error estimates for the elliptic equations, the main difference here is how to handle the variational inequality in the optimality conditions as in the h-version adaptive finite element method for optimal control. Besides different interpolators and interpolation results that now have to be used, the variational inequality is further different from that for the control constraint of obstacle type in the literature, due to the different control constraint set studied in this paper. While the existing techniques for the constraints of obstacle type can be modified for deriving the upper bounds, the techniques are very different to derive the lower bound here. These are studied in Lemma 5.2 where we use the inverse inequality to estimate the lower bound for the control, and this approach has not been used before.

The paper is organized as follows: In Section 2, we introduce the model problem and its weak formulation, and give the hp-finite element spaces and the hp-finite element approximation of the control problem. In Section 3, some technical lemmas are introduced, which are used for the later a posteriori error analysis. In Section 4, an a posteriori error estimator for the control problems is provided based on the local residual technique. It is shown that the a posteriori error estimator is an upper bound of the error. In Section 5, it is proved that the a posteriori error estimator provided in Section 4 is also a lower bound of the error, although there is a gap of order p^2 between the lower and the upper bounds. In the last section, a family of weighted a posteriori error estimators are presented as the extension of the analysis of Sections 4 and 5 using the weight function technique introduced in [26].

2. The model problem and hp-FEM approximation

Let $\Omega(\Omega_U)$ be a bounded domain in R^2 with the Lipschitz boundary $\partial\Omega(\partial\Omega_U)$. In this paper we adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with norm $||\cdot||_{W^{m,q}(\Omega)}$ and seminorm $|\cdot|_{W^{m,q}(\Omega)}$. We set $W_0^{m,q}(\Omega) \equiv \{w \in W^{m,q}(\Omega) : w|_{\partial\Omega} = 0\}$. We denote $W^{m,2}(\Omega)(W_0^{m,2}(\Omega))$ by $H^m(\Omega)(H_0^m(\Omega))$. In addition, c or C denotes a general positive constant independent of h and p.

In this paper, we will investigate the following distributed convex optimal control problems:

$$\min_{u\in K}\left\{g(y)+\frac{\delta}{2}\int_{\Omega_U}u^2\right\}, \tag{1}$$

subject to

$$\begin{aligned} -\mathrm{div}(A\nabla y) &= f+Bu \quad \text{in} \quad \Omega \\ y &= 0 \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{2}$$

where $g(\cdot)$ is a convex functional, which is bounded below and continuously differentiable on the observation space $L^2(\Omega)$, and δ is a positive constant. We take the state space $V = H_0^1(\Omega)$, the control space $U = L^2(\Omega_U)$, the observation space $Y = L^2(\Omega)$, and $H = L^2(\Omega)$. Let B be a linear continuous operator from U to H, and K be a closed set in the control space U. In this paper, we set

$$K = \{u \in L^2(\Omega_U) : \int_{\Omega_U} u \geqslant 0\}.$$

To consider the finite element approximation of the optimal control problem (1)-(2), we first give a weak formula for the state equation. Let $f \in L^2(\Omega)$, and

$$A(\cdot) = (a_{i,j}(\cdot))_{n\times n} \in (W^{1,\infty}(\Omega))^{n\times n}$$

satisfying

$$X^tAX \geqslant c||X||^2_{R^2} \quad \forall X \in R^2.$$

For simplicity, we suppose in this paper that A is constant matrix. Let

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

$$(f_1,f_2) = \int_\Omega f_1f_2 \ \ \forall\ (f_1,f_2)\in H\times H,$$

$$(u,v)_U = \int_{\Omega_U} uv \ \ \forall\ (u,v)\in U\times U.$$

It follows from the above assumptions on A that there exists constants $c, C > 0$ such that $\forall\ y, w \in V$

$$a(y,y) \geqslant c||y||_V^2, \quad |a(y,w)| \leqslant C||y||_V||w||_V.$$

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

$$a(y(u),w) = (f+Bu,w) \ \ \forall\ w\in H_0^1(\Omega). \tag{3}$$

Therefore, the control problem (1)-(2) can be restated as

$$\min_{u\in K\subset U}\left\{g(y)+\frac{\delta}{2}\int_{\Omega_U}u^2\right\}, \tag{4}$$

subject to

$$a(y(u),w) = (f+Bu,w) \quad \forall w\in V = H_0^1(\Omega). \tag{5}$$

It is well known that the control problem (4)-(5) has a solution (y^*, u^*), and that a pair (y^*, u^*) is the solution of (4)-(5) if and only if there is a co-state $\lambda^* \in V$ such that the triplet (y^*, λ^*, u^*) satisfies the following optimality conditions:

$$a(y^*,w) = (f+Bu^*,w) \qquad \forall\ w\in V = H_0^1(\Omega), \tag{6}$$

$$a(q,\lambda^*) = (g'(y^*),q) \qquad \forall\ q\in V = H_0^1(\Omega), \tag{7}$$

$$(\delta u^* + B^*\lambda^*, v-u^*)_U \geqslant 0 \quad \forall\ v\in K\subset U = L^2(\Omega_U), \tag{8}$$

where B^* is the adjoint operator of B. Just for simplicity of presentation, we still denote (y^*, λ^*, u^*) by (y, λ, u) if there is no confusion.

Next, let us consider the hp-finite element approximation of above optimal control problem. Let $\Omega^h(\Omega_U^h)$ be a polygonal approximation to $\Omega(\Omega_U)$ with the boundary $\partial\Omega^h(\partial\Omega_U^h)$. For simplicity, we assume that $\Omega^h = \Omega$ $(\Omega_U^h = \Omega_U)$. Let $\mathcal{T} = \{\tau\}$ $(\mathcal{T}_U = \{\tau_U\})$ be a local quasi-uniform partitioning of $\Omega^h(\Omega_U^h)$ into disjoint regular element $\tau(\tau_U)$, let $\mathcal{E}(\mathcal{T})$ denote all edges, and $\mathcal{E}_0(\mathcal{T})$ denote all edges that do not lie on the boundary $\partial\Omega$. Each element $\tau(\tau_U)$ can be the image of either the reference square $\hat{\tau} = S = (0,1)^2$ or the reference triangle $\hat{\tau} = T = \{(x,y) :\ 0 < x < 1,\ 0 < y < \min\{x, 1-x\}\}$ under an affine map $F_\tau : \hat{\tau} \to \tau$. We assume that $h_\tau(h_{\tau_U})$ denote the maximum diameter of the element $\tau(\tau_U)$ and the triangulation is $\gamma-$shape regular, i.e:

$$h_\tau^{-1}||F_\tau^{'}|| + h_\tau||(F_\tau^{'})^{-1}|| \leqslant \gamma. \tag{9}$$

This implies that element sizes of neighboring elements are comparable.

For γ shape regular meshes $\mathcal{T}(\mathcal{T}_U)$ on a domain $\Omega(\Omega_U)$, we associate each element $\tau \in \mathcal{T}$ $(\tau_U \in \mathcal{T}_U)$ with a polynomial degree $p_\tau \in \mathbb{N}_0$ $(p_{\tau_U} \in \mathbb{N}_0)$ and define the degree vector $\mathbf{p} = \{p_\tau :\ \tau \in \mathcal{T}\}$ $(\mathbf{p_U} = \{p_{\tau_U} :\ \tau_U \in \mathcal{T}_U\})$. Then we can define the hp-version finite element spaces $S^{\mathbf{p}}(\mathcal{T}, \Omega)$, $S^{\mathbf{p_U}}(\mathcal{T}_U, \Omega_U)$, $S_0^{\mathbf{p}}(\mathcal{T}, \Omega)$ as following

$$\begin{aligned}
&S^{\mathbf{p}}(\mathcal{T}, \Omega) := \{u \in H^1(\Omega) :\quad u|_\tau \circ F_\tau \in \Pi_{p_\tau}(\hat{\tau})\},\\
&S^{\mathbf{p_U}}(\mathcal{T}_U, \Omega_U) := \{u \in L^2(\Omega_U) :\quad u|_{\tau_U} \circ F_{\tau_U} \in \Pi_{p_{\tau_U}}(\hat{\tau}_U)\},\\
&S_0^{\mathbf{p}}(\mathcal{T}, \Omega) := S^{\mathbf{p}}(\mathcal{T}, \Omega) \cap H_0^1(\Omega),
\end{aligned}$$

$$\Pi_p(\hat{\tau}) := \begin{cases} \mathcal{Q}_p & \text{if} \quad \hat{\tau} = S \\ \mathcal{P}_p & \text{if} \quad \hat{\tau} = T \end{cases},$$

where S and T are the reference square and the reference triangle defined above, $\mathcal{P}_p$ and $\mathcal{Q}_p$ denote the polynomial spaces with polynomials of total degree no more than p and polynomials of degree no more than p in each variable, respectively.

Note that associated with $\mathcal{T}$ is a finite dimensional subspace $S^{\mathbf{p}}(\mathcal{T}, \Omega)$ of $C(\overline{\Omega}^h)$ and associated with $\mathcal{T}_U$ is a finite dimensional subspace $S^{\mathbf{p_U}}(\mathcal{T}_U, \Omega_U)$ of $L^2(\Omega_U^h)$. As to polynomial degree distribution $\mathbf{p}(\mathbf{p_U})$, similar to (9), we assume that the polynomial degrees of neighboring elements are comparable, i.e., there is a constant $\gamma > 0$ such that

$$\gamma^{-1}(p_\tau + 1) \leqslant p_{\tau'} + 1 \leqslant \gamma(p_\tau + 1) \quad \forall \tau, \tau' \in \mathcal{T},\ \bar{\tau} \cap \bar{\tau}' \neq \emptyset, \tag{10}$$

and

$$\gamma^{-1}(p_{\tau_U} + 1) \leqslant p_{\tau_U'} + 1 \leqslant \gamma(p_{\tau_U} + 1) \quad \forall \tau_U, \tau_U' \in \mathcal{T}_U,\ \bar{\tau}_U \cap \bar{\tau}_U' \neq \emptyset. \tag{11}$$

Let $K^{h,\mathbf{p_U}} = K \cap S^{\mathbf{p_U}}(\mathcal{T}_U, \Omega_U)$ be the finite element space for the control, and $S_0^{\mathbf{p}}(\mathcal{T}, \Omega)$ be the finite element space for the state and costate. Then the hp-FEM approximation to (4)-(5) reads

$$\min_{u_{hp} \in K^{h,\mathbf{p_U}}} \left\{ g(y_{hp}) + \frac{\delta}{2} \int_{\Omega_U} u_{hp}^2 \right\}, \tag{12}$$

subject to

$$a(y_{hp}, w) = (f + Bu_{hp}, w) \quad \forall w \in S_0^{\mathbf{p}}(\mathcal{T}, \Omega). \tag{13}$$

Again, it can be proved that (12)-(13) are equivalent to the discrete optimality conditions: Find $(y_{hp}, \lambda_{hp}, u_{hp}) \in S_0^{\mathbf{p}}(\mathcal{T}, \Omega) \times S_0^{\mathbf{p}}(\mathcal{T}, \Omega) \times K^{h,\mathbf{p_U}}$ such that

$$a(y_{hp}, w) = (f + Bu_{hp}, w) \quad \forall\ w \in S_0^{\mathbf{p}}(\mathcal{T}, \Omega), \tag{14}$$

$$a(q, \lambda_{hp}) = (g'(y_{hp}), q) \quad \forall\ q \in S_0^{\mathbf{p}}(\mathcal{T}, \Omega), \tag{15}$$

$$(\delta u_{hp} + B^*\lambda_{hp}, v - u_{hp})_U \geqslant 0 \qquad \forall\ v \in K^{h,\mathbf{p_U}}. \tag{16}$$

3. Some preliminary lemmas

We first derive a relationship between the control u and the costate λ. For the control set considered in this paper:

$$K = \{u \in L^2(\Omega_U) : \int_{\Omega_U} u \geqslant 0\},$$

it is a matter of calculation to show that the control u and the costate λ have the following relationship (see, e.g., [12]):

$$u = \frac{1}{\delta}\max\{0, \hat{\lambda}\} - \frac{1}{\delta}B^*\lambda, \tag{17}$$

where $\hat{\lambda} = \frac{\int_{\Omega_U} B^*\lambda}{\int_{\Omega_U} 1}$.

Let $y(u)$ and $y_{hp}(u_{hp})$ be the solutions of (6) and (14), respectively. Denote

$$J(u) = g(y(u)) + \frac{\delta}{2}\int_{\Omega_U} u^2, \quad J_{hp}(u_{hp}) = g(y_{hp}(u_{hp})) + \frac{\delta}{2}\int_{\Omega_U} u_{hp}^2.$$

Then the reduced problems of (4) and (12) read

$$\min_{u \in K}\{J(u)\}, \tag{18}$$

and

$$\min_{u_{hp} \in K^{h,\mathbf{p_U}}}\{J_{hp}(u_{hp})\}. \tag{19}$$

Moreover, we have that

$$\begin{aligned}(J'(u), v)_U &= (\delta u + B^*\lambda, v)_U,\\ (J'_{hp}(u_{hp}), v)_U &= (\delta u_{hp} + B^*\lambda_{hp}, v)_U,\\ (J'(u_{hp}), v)_U &= (\delta u_{hp} + B^*\lambda(u_{hp}), v)_U,\end{aligned}$$

where $\lambda(u_{hp})$ is the solution of the following auxiliary equation:

$$a(y(u_{hp}), w) = (f + Bu_{hp}, w) \quad \forall\ w \in V = H_0^1(\Omega), \tag{20}$$

$$a(q, \lambda(u_{hp})) = (g'(y(u_{hp})), q) \quad \forall\ q \in V = H_0^1(\Omega). \tag{21}$$

For the functional $J(\cdot)$ defined above, we show in the following lemma that it is uniformly convex if g is convex.

Lemma 3.1. Let J be defined above and g be convex. Then we have the following uniformly convex property: $\forall w,\ v \in U$

$$(J'(w) - J'(v), w - v)_U \geqslant \delta\|w - v\|^2_{L^2(\Omega_U)}. \tag{22}$$

Proof. Note that

$$J'(w)(w - v) = (\delta w, w - v)_U + (B^*\lambda(w), w - v)_U.$$

It follows from (6) and (7) that

$$\begin{aligned}(B^*\lambda(w), w - v)_U &= (\lambda(w), B(w - v))\\ &= a(y(w) - y(v), \lambda(w))\\ &= (g'(y(w)), y(w) - y(v)).\end{aligned}$$

Similarly

$$(B^*\lambda(v), w-v)_U = (g'(y(v)), y(w)-y(v)).$$

Note that g is convex. Then for all $w,\ v\in U$,

$$\begin{aligned}(J'(w)-J'(v), w-v)_U &= (\delta w-\delta v, w-v)_U + (B^*\lambda(w)-B^*\lambda(v), w-v)_U\\ &= \delta(w-v, w-v)_U + (g'(y(w))-g'(y(v)), y(w)-y(v))\\ &\geqslant \delta(w-v, w-v)_U + 0 = \delta\|w-v\|^2_{L^2(\Omega_U)}.\end{aligned}$$

This completes the proof. □

In the followings, we state two lemmas, which generalize the well-known Clément-type interpolation operators studied in [9] and [30] to the hp context. The readers can refer to [25] for the details. The two operators both provide piecewise polynomial approximations for H^1 functions. However, the second operator can preserve the homogenous boundary conditions.

Lemma 3.2. (Clément type quasi-interpolation). Let $\mathcal{T}_U$ be a γ-shape regular triangulation (see (9)) of a domain $\Omega_U\subset R^2$ and let $\mathbf{p_U}$ be a polynomial degree distribution which is comparable (see (10)). Then there exists a bounded linear operator $\Pi: L^1(\Omega_U)\to S^{\mathbf{p_U}}(\mathcal{T}_U,\Omega_U)$, and there exists a constant $C>0$, which depends only on γ, such that for every $u\in H^1(\Omega_U)$ and all elements $\tau_U\in\mathcal{T}_U$ and all edges $e_U\in\mathcal{E}(\mathcal{T}_U)$,

$$\|u-\Pi u\|_{L^2(\tau_U)} + \frac{h_{\tau_U}}{p_{\tau_U}}\|\nabla(u-\Pi u)\|_{L^2(\tau_U)} \leqslant C\frac{h_{\tau_U}}{p_{\tau_U}}\|\nabla u\|_{L^2(\omega_{\tau_U})}, \tag{23}$$

$$\|u-\Pi u\|_{L^2(e_U)} \leqslant C\sqrt{\frac{h_{e_U}}{p_{e_U}}}\|\nabla u\|_{L^2(\omega_{e_U})}, \tag{24}$$

where h_{e_U} is the length of the edge e_U and p_{e_U} =max(p_{τ_1}, p_{τ_2}), where $\tau_1,\ \tau_2$ are elements sharing the edge e_U, $\omega_{\tau_U},\ \omega_{e_U}$ are patches covering τ_U and e_U with a few layers, respectively. We refer to [25] for more details on ω_{τ_U} and ω_{e_U}.

Lemma 3.3. (Scott-Zhang type quasi-interpolation). Let $\mathcal{T}$ be a γ-shape regular triangulation (see (9)) of a domain $\Omega\subset R^2$ and let $\mathbf{p}$ be a polynomial degree distribution which is comparable (see (10)). Then there exists a linear operator $I: H^1_0(\Omega)\to S^{\mathbf{p}}_0(\mathcal{T},\Omega)$, and there exists a constant $C>0$, which depends only on γ, such that for every $u\in H^1_0(\Omega)$ and all elements $\tau\in\mathcal{T}$ and all edges $e\in\mathcal{E}(\mathcal{T})$,

$$\|u-Iu\|_{L^2(\tau)} + \frac{h_\tau}{p_\tau}\|\nabla(u-Iu)\|_{L^2(\tau)} \leqslant C\frac{h_\tau}{p_\tau}\|\nabla u\|_{L^2(\omega_\tau)}, \tag{25}$$

$$\|u-Iu\|_{L^2(e)} \leqslant C\sqrt{\frac{h_e}{p_e}}\|\nabla u\|_{L^2(\omega_e)}, \tag{26}$$

again, where h_e is the length of the edge e and p_e =max(p_{τ_1}, p_{τ_2}), where $\tau_1,\ \tau_2$ are elements sharing the edge e, $\omega_\tau,\ \omega_e$ are patches covering τ and e with a few layers, respectively.

Analysis of the hp a posteriori error estimators requires polynomial inverse estimates in weighted Sobolev spaces. Under this consideration, the weight functions: $\Phi_{\hat\tau}(x)$:=dist$(x,\partial\hat\tau)$ on the reference element $\hat\tau$ should be introduced (see [26] for more details). For an arbitrary element $\tau\in\mathcal{T}$, set $\Phi_\tau = c_\tau\Phi_{\hat\tau}\circ F_\tau^{-1}$, where c_τ is a scaling factor which is chosen such that $\int_\tau\Phi_\tau$dxdy=meas(τ). Similarly, one can define the weight function $\Phi_{\hat e}(x) := x(1-x)$ on the reference interval $\hat e=(0,1)$. For an interior edge e, the weight function Φ_e is then defined by $\Phi_e = c_e\Phi_{\hat e}\circ F_\tau^{-1}$, where c_e is chosen such that $\int_e \Phi_e ds$=meas(e). For the above weighted Sobolev spaces, we have the following lemmas (see [26] for more details).

Lemma 3.4. Let $\hat{\tau}$ be the reference square S or the reference triangle T defined in Section 2, let the weight function $\Phi_{\hat{\tau}}$ be defined above. Let $\gamma, \beta \in R$ satisfying $-1 < \gamma < \beta$ and $\delta \in [0,1]$. Then for all polynomial $\psi_p \in \mathcal{Q}_p(\mathcal{P}_p)$,

$$\int_{\hat{\tau}} \Phi_{\hat{\tau}} |\nabla \psi_p|^2 dxdy \leqslant C_1 p^2 \int_{\hat{\tau}} |\psi_p|^2 dxdy, \tag{27}$$

$$\int_{\hat{\tau}} (\Phi_{\hat{\tau}})^{\gamma} \psi_p^2 dxdy \leqslant C_2 p^{2(\beta-\gamma)} \int_{\hat{\tau}} (\Phi_{\hat{\tau}})^{\beta} \psi_p^2 dxdy, \tag{28}$$

$$\int_{\hat{\tau}} (\Phi_{\hat{\tau}})^{2\delta} |\nabla \psi_p|^2 dxdy \leqslant C_3 p^{2(2-\delta)} \int_{\hat{\tau}} (\Phi_{\hat{\tau}})^{\delta} \psi_p^2 dxdy, \tag{29}$$

where C_i, $i = 1, 2, 3$, are constants, C_2 is dependent on β and γ, and C_3 is dependent on δ.

Lemma 3.5. Let $\hat{\tau}$ be the reference square S or the reference triangle T defined in Section 2, $\alpha \in (1/2, 1]$. Set $\hat{e} = (0,1) \times \{0\}$. Then there exists a constant $C_\alpha > 0$, which is dependent on α, such that the followings hold. For every univariate polynomial $\psi \in \mathcal{P}_p$ and every $\epsilon \in (0,1]$ there exists an extension $v_{\hat{e}} \in H^1(\hat{\tau})$ such that

(i) $v_{\hat{e}}|_{\hat{e}} = \psi \cdot \Phi_{\hat{e}}^{\alpha}$ and $v_{\hat{e}}|_{\partial\hat{\tau}\setminus\hat{e}} = 0$;

(ii) $\|v_{\hat{e}}\|^2_{L^2(\hat{\tau})} \leqslant C_\alpha \epsilon \|\psi \Phi_{\hat{e}}^{\alpha/2}\|^2_{L^2(\hat{e})}$;

(iii) $\|\nabla v_{\hat{e}}\|^2_{L^2(\hat{\tau})} \leqslant C_\alpha (\epsilon p^{2(2-\alpha)} + \epsilon^{-1}) \|\psi \Phi_{\hat{e}}^{\alpha/2}\|^2_{L^2(\hat{e})}$;

where $\Phi_{\hat{e}}$ is the weight function defined above, and $\hat{\tau}$ is the reference element such that $\hat{e} \subset \partial\hat{\tau}$.

4. A posteriori upper error estimates

In this section, we will derive upper a posteriori error estimates of residual type. We first define the following notations:

$$\xi^2 = \sum_{\tau\in\mathcal{T}} \xi_\tau^2 = \sum_{\tau\in\mathcal{T}} (\xi^2_{B_\tau} + \xi^2_{E_\tau}), \tag{30}$$

$$\eta^2 = \sum_{\tau\in\mathcal{T}} \eta_\tau^2 = \sum_{\tau\in\mathcal{T}} (\eta^2_{B_\tau} + \eta^2_{E_\tau}), \tag{31}$$

$$\zeta^2 = \sum_{\tau_U\in\mathcal{T}_U} \zeta^2_{\tau_U}, \tag{32}$$

where

$$\xi^2_{B_\tau} := \frac{h_\tau^2}{p_\tau^2} \|g'(y_{hp}) + div(A^*\nabla\lambda_{hp})\|^2_{L^2(\tau)}, \tag{33}$$

$$\xi^2_{E_\tau} := \sum_{e\subset\partial\tau\cap\Omega} \frac{h_e}{2p_e} \|[A^*\nabla\lambda_{hp}\cdot n_e]\|^2_{L^2(e)}, \tag{34}$$

$$\eta^2_{B_\tau} := \frac{h_\tau^2}{p_\tau^2} \|f + Bu_{hp} + div(A\nabla y_{hp})\|^2_{L^2(\tau)}, \tag{35}$$

$$\eta^2_{E_\tau} := \sum_{e\subset\partial\tau\cap\Omega} \frac{h_e}{2p_e} \|[A\nabla y_{hp}\cdot n_e]\|^2_{L^2(e)}, \tag{36}$$

$$\zeta^2_{\tau_U} := \frac{h^2_{\tau_U}}{p^2_{\tau_U}} \|\nabla(\delta u_{hp} + B^*\lambda_{hp})\|^2_{L^2(\tau_U)}, \tag{37}$$

where we denote the jump of v across the edges by $[v]$, and n_e is the unit outer normal on e. The sum in (34) and (36) extends over all the edges of τ that do not lie on the boundary $\partial\Omega$. Furthermore, h_e is the length of the edge e and p_e =max(p_{τ_1}, p_{τ_2}), where $\tau_1,\ \tau_2$ are the two elements sharing a edge e.

Using the above definitions, we have the following a posteriori error estimates.

Lemma 4.1. Let (y, λ, u) and $(y_{hp}, \lambda_{hp}, u_{hp})$ be the solutions of (6)-(8) and (14)-(16), respectively. Assume that $K^{h,p_U} \subset K$ and $g(\cdot)$ is convex. Then we have

$$\|u - u_{hp}\|^2_{L^2(\Omega_U)} \leqslant C(\zeta^2 + \|\lambda_{hp} - \lambda(u_{hp})\|^2_{L^2(\Omega)}), \tag{38}$$

where ζ is defined by (32) and (37), λ_{hp} and $\lambda(u_{hp})$ are the solutions of equations (15) and (21), respectively.

Proof. It follows from Lemma 3.1 and (8), (16) that

$$\begin{aligned} c\|u - u_{hp}\|^2_{L^2(\Omega_U)} &\leqslant (J'(u), u - u_{hp})_U - (J'(u_{hp}), u - u_{hp})_U \\ &\leqslant -(J'(u_{hp}), u - u_{hp})_U \\ &= (J'_{hp}(u_{hp}), u_{hp} - u)_U + (J'_{hp}(u_{hp}) - J'(u_{hp}), u - u_{hp})_U \\ &\leqslant \inf_{v \in K^{h,p_U}} (J'_{hp}(u_{hp}), v - u)_U + (J'_{hp}(u_{hp}) - J'(u_{hp}), u - u_{hp})_U. \end{aligned} \tag{39}$$

Note that

$$(J'_{hp}(u_{hp}), v - u)_U = (\delta u_{hp} + B^*\lambda_{hp}, v - u)_U.$$

Let $v = \pi_{p_U} u$ be the L^2-projection of u in $S^{\mathbf{p_U}}(\mathcal{T}_U, \Omega_U)$, and Π be the Clément-type interpolator defined in Lemma 3.2. Then we have

$$\begin{aligned} & (\delta u_{hp} + B^*\lambda_{hp}, \pi_{p_U} u - u)_U \\ = & \sum_{\tau_U} \int_{\tau_U} (\delta u_{hp} + B^*\lambda_{hp})(\pi_{p_U} u - u) \\ = & \sum_{\tau_U} \int_{\tau_U} (\delta u_{hp} + B^*\lambda_{hp} - \Pi(\delta u_{hp} + B^*\lambda_{hp}))(\pi_{p_U} u - u). \end{aligned}$$

Since $\pi_{p_U} u_{hp} = u_{hp}$, it follows from Lemma 3.2 that

$$\begin{aligned} & (J'_{hp}(u_{hp}), \pi_{p_U} u - u)_U \\ & = \sum_{\tau_U} \int_{\tau_U} (\delta u_{hp} + B^*\lambda_{hp} - \Pi(\delta u_{hp} + B^*\lambda_{hp}))(\pi_{p_U}(u - u_{hp}) - (u - u_{hp})) \\ & \leqslant C \sum_{\tau_U} \|\delta u_{hp} + B^*\lambda_{hp} - \Pi(\delta u_{hp} + B^*\lambda_{hp})\|_{L^2(\tau_U)} \\ & \|\pi_{p_U}(u - u_{hp}) - (u - u_{hp})\|_{L^2(\tau_U)} \\ & \leqslant C \sum_{\tau_U} \frac{h_{\tau_U}}{p_{\tau_U}} \|\nabla(\delta u_{hp} + B^*\lambda_{hp})\|_{L^2(\omega_{\tau_U})} \|u - u_{hp}\|_{L^2(\tau_U)} \\ & \leqslant C(\sigma) \sum_{\tau_U} \frac{h^2_{\tau_U}}{p^2_{\tau_U}} \|\nabla(\delta u_{hp} + B^*\lambda_{hp})\|^2_{L^2(\tau_U)} + C\sigma\|u - u_{hp}\|^2_{L^2(\Omega_U)}, \end{aligned} \tag{40}$$

where σ is an arbitrary small positive number. Moreover, it is easy to show that

$$\begin{aligned} (J'_{hp}(u_{hp}) - J'(u_{hp}), u - u_{hp})_U &= (\delta u_{hp} + B^*\lambda_{hp}, u - u_{hp})_U \\ & \quad -(\delta u_{hp} + B^*\lambda(u_{hp}), u - u_{hp})_U \\ &= (B^*(\lambda_{hp} - \lambda(u_{hp})), u - u_{hp})_U \\ &\leqslant C(\sigma)\|\lambda_{hp} - \lambda(u_{hp})\|^2_{L^2(\Omega)} + C\sigma\|u - u_{hp}\|^2_{L^2(\Omega_U)}. \end{aligned} \tag{41}$$

Note that $\pi_{p_U} u \in K^{h,\mathbf{p_U}}$. Then (38) follows from (39)-(41) by setting σ to be small enough. $\square$

It follows from Lemma 4.1 that in order to obtain a posteriori error estimates we only need to estimate $\|\lambda_{hp} - \lambda(u_{hp})\|^2_{L^2(\Omega)}$.

Theorem 4.2. Let (y, λ, u) and $(y_{hp}, \lambda_{hp}, u_{hp})$ be the solutions of (6)-(8) and (14)-(16). Assume that all the conditions in Lemma 4.1 are valid. Moreover suppose $g(\cdot)$ is Lipschitz continuous. Then there exists a $C > 0$ independent of h and p such that

$$\|u - u_{hp}\|^2_{L^2(\Omega_U)} + \|y - y_{hp}\|^2_{H^1(\Omega)} + \|\lambda - \lambda_{hp}\|^2_{H^1(\Omega)} \leqslant C(\zeta^2 + \xi^2 + \eta^2), \tag{42}$$

where ζ, ξ and η are defined by (30)-(37).

Proof. Let $e^\lambda = \lambda(u_{hp}) - \lambda_{hp}$, and $e^\lambda_I = Ie^\lambda$, where $I : H^1_0(\Omega) \to S^{\mathbf{p}}_0(\mathcal{T}, \Omega)$ is the Scott-Zhang type interpolator defined in Lemma 3.3. Note that $e^\lambda_I \in S^{\mathbf{p}}_0(\mathcal{T}, \Omega)$. Applying the standard residual techniques (see, e.g., [31]) and using equations (15) and (21), we have

$$\begin{aligned}
& c\|\lambda(u_{hp}) - \lambda_{hp}\|^2_{H^1(\Omega)} \\
\leqslant\ & a(e^\lambda, e^\lambda) = (g'(y(u_{hp})), e^\lambda) - a(e^\lambda, \lambda_{hp}) \\
=\ & (g'(y(u_{hp})), e^\lambda) - a(e^\lambda - e^\lambda_I, \lambda_{hp}) - (g'(y_{hp}), e^\lambda_I) \\
=\ & \sum_{\tau\in\mathcal{T}} \int_\tau (g'(y_{hp}) + div(A^*\nabla\lambda_{hp}))(e^\lambda - e^\lambda_I) \\
& - \sum_{e\in\mathcal{E}_0(\mathcal{T})} \int_e [A^*\nabla\lambda_{hp}\cdot n_e](e^\lambda - e^\lambda_I) + (g'(y(u_{hp})) - g'(y_{hp}), e^\lambda) \\
\leqslant\ & C\sum_{\tau\in\mathcal{T}} \|g'(y_{hp}) + div(A^*\nabla\lambda_{hp})\|_{L^2(\tau)}\|e^\lambda - e^\lambda_I\|_{L^2(\tau)} \\
& + C\sum_{e\in\mathcal{E}_0(\mathcal{T})} \|[A^*\nabla\lambda_{hp}\cdot n_e]\|_{L^2(e)}\|e^\lambda - e^\lambda_I\|_{L^2(e)} \\
& + C\|g'(y(u_{hp})) - g'(y_{hp})\|_{L^2(\Omega)}\|e^\lambda\|_{L^2(\Omega)}.
\end{aligned} \tag{43}$$

Then it follows from Lemma 3.3 that

$$\begin{aligned}
& c\|\lambda(u_{hp}) - \lambda_{hp}\|^2_{H^1(\Omega)} \\
\leqslant\ & C\sum_{\tau\in\mathcal{T}} \frac{h_\tau}{p_\tau}\|g'(y_{hp}) + div(A^*\nabla\lambda_{hp})\|_{L^2(\tau)}\|\nabla e^\lambda\|_{L^2(\omega_\tau)} \\
& + C\sum_{e\in\mathcal{E}_0(\mathcal{T})} \frac{\sqrt{h_e}}{\sqrt{p_e}}\|[A^*\nabla\lambda_{hp}\cdot n_e]\|_{L^2(e)}\|\nabla e^\lambda\|_{L^2(\omega_e)} \\
& + C\|y(u_{hp}) - y_{hp}\|_{L^2(\Omega)}\|e^\lambda\|_{L^2(\Omega)} \\
\leqslant\ & C(\sigma)\sum_{\tau\in\mathcal{T}}(\xi^2_{B_\tau} + \xi^2_{E_\tau}) + C(\sigma)\|y(u_{hp}) - y_{hp}\|^2_{L^2(\Omega)} + \sigma\|e^\lambda\|^2_{H^1(\Omega)}.
\end{aligned} \tag{44}$$

Setting $\sigma = \frac{c}{2}$, we have

$$\|\lambda(u_{hp}) - \lambda_{hp}\|^2_{H^1(\Omega)} \leqslant C\sum_{\tau\in\mathcal{T}}(\xi^2_{B_\tau} + \xi^2_{E_\tau}) + C\|y(u_{hp}) - y_{hp}\|^2_{L^2(\Omega)}. \tag{45}$$

Similarly, let $e^y = y(u_{hp}) - y_{hp}$ and e_I^y be the Scott-Zhang interpolation of e^y defined in Lemma 3.3. We have

$$\begin{aligned}
c\|y(u_{hp}) - y_{hp}\|^2_{H^1(\Omega)} &\leqslant a(e^y, e^y) = a(e^y, e^y - e_I^y) \\
&= \sum_{\tau\in\mathcal{T}} \int_\tau (f + Bu_{hp} + div(A\nabla y_{hp}))(e^y - e_I^y) \\
&\quad - \sum_{e\in\mathcal{E}_0(\mathcal{T})} \int_e [A\nabla y_{hp}\cdot n_e](e^y - e_I^y) \\
&\leqslant C \sum_{\tau\in\mathcal{T}} \frac{h_\tau}{p_\tau} \|f + Bu_{hp} + div(A\nabla y_{hp})\|_{L^2(\tau)} \|\nabla e^y\|_{L^2(\omega_\tau)} \\
&\quad + C \sum_{e\in\mathcal{E}_0(\mathcal{T})} \frac{\sqrt{h_e}}{\sqrt{p_e}} \|[A\nabla y_{hp}\cdot n_e]\|_{L^2(e)} \|\nabla e^y\|_{L^2(\omega_e)} \\
&\leqslant C(\sigma) \sum_{\tau\in\mathcal{T}} (\eta^2_{B_\tau} + \eta^2_{E_\tau}) + \sigma\|e^y\|^2_{H^1(\Omega)},
\end{aligned}$$

which implies

$$\|y(u_{hp}) - y_{hp}\|^2_{H^1(\Omega)} \leqslant C \sum_{\tau\in\mathcal{T}} (\eta^2_{B_\tau} + \eta^2_{E_\tau}). \tag{46}$$

Hence, it follows from (45), (46) and Lemma 4.1 that

$$\|u-u_{hp}\|^2_{L^2(\Omega_U)} + \|y(u_{hp}) - y_{hp}\|^2_{H^1(\Omega)} + \|\lambda(u_{hp}) - \lambda_{hp}\|^2_{H^1(\Omega)} \leqslant C(\zeta^2 + \xi^2 + \eta^2). \tag{47}$$

Note that

$$\|y - y_{hp}\|_{H^1(\Omega)} \leqslant \|y(u_{hp}) - y_{hp}\|_{H^1(\Omega)} + \|y - y(u_{hp})\|_{H^1(\Omega)}, \tag{48}$$

$$\|\lambda - \lambda_{hp}\|_{H^1(\Omega)} \leqslant \|\lambda(u_{hp}) - \lambda_{hp}\|_{H^1(\Omega)} + \|\lambda - \lambda(u_{hp})\|_{H^1(\Omega)}, \tag{49}$$

and

$$\|y - y(u_{hp})\|^2_{H^1(\Omega)} + \|\lambda - \lambda(u_{hp})\|^2_{H^1(\Omega)} \leqslant C\|u - u_{hp}\|^2_{L^2(\Omega_U)}. \tag{50}$$

Therefore, (42) follows from (47)-(50). □

5. A posteriori lower error estimates

In this section, we discuss lower a posteriori bounds, i.e., the efficiency of the error estimates provided in Theorem 4.2. As pointed in [26], we cannot obtain equivalent a posteriori error estimates for hp-FEM of optimal control problems using the current techniques. Thus there exists a gap of the order p^2 between the lower and upper bounds in our estimates as in the case of the hp-adaptive finite element approximation of elliptic equations.

Firstly, we estimate the residual ξ_{B_τ} and η_{B_τ}.

Lemma 5.1. Let (y, λ, u) and $(y_{hp}, \lambda_{hp}, u_{hp})$ be the solutions of (6)-(8) and (10)-(12) respectively. Let ξ_{B_τ} and η_{B_τ} be defined by (33) and (35), respectively.

Assume that $g(\cdot)$ is Lipschitz continuous. Then

$$\begin{aligned}\xi^2_{B_\tau} &\leqslant Cp_\tau^2\|\lambda-\lambda_{hp}\|^2_{H^1(\tau)}+Cp_\tau^{2\alpha}\frac{h_\tau^2}{p_\tau^2}\Big(\|y-y_{hp}\|^2_{L^2(\tau)}\\ &\quad +\|\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp})\|^2_{L^2(\tau)}\Big),\end{aligned}\tag{51}$$

$$\begin{aligned}\eta^2_{B_\tau} &\leqslant Cp_\tau^2\|y-y_{hp}\|^2_{H^1(\tau)}+Cp_\tau^{2\alpha}\frac{h_\tau^2}{p_\tau^2}\Big(\|B(u-u_{hp})\|^2_{L^2(\tau)}\\ &\quad +\|\pi_{p_\tau}f-f\|^2_{L^2(\tau)}+\|\pi_{p_\tau}(Bu)-Bu\|^2_{L^2(\tau)}\Big),\end{aligned}\tag{52}$$

where π_{p_τ} is the L^2-project operator on the space of polynomials of degree p_τ on the element τ, $\frac{1}{2}<\alpha\leqslant 1$, and the constant C depends on α.

Proof. Let Φ_τ be the weight function defined before Lemma 3.4 in Section 3. Define $w_\tau=(\pi_{p_\tau}g'(y_{hp})+div(A^*\nabla\lambda_{hp}))\Phi_\tau^\alpha$, $\frac{1}{2}<\alpha\leqslant 1$. Using the trivial extension by zero on $\Omega\setminus\tau$, we have

$$\begin{aligned}&\|w_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|^2_{L^2(\tau)}\\ =\;& \int_\tau(\pi_{p_\tau}(g'(y_{hp}))+div(A^*\nabla\lambda_{hp}))w_\tau dxdy\\ =\;& \int_\tau(g'(y)+div(A^*\nabla\lambda_{hp}))w_\tau dxdy+\int_\tau(g'(y_{hp})-g'(y))w_\tau dxdy\\ &+\int_\tau(\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp}))w_\tau dxdy\\ =\;& a(w_\tau,\lambda-\lambda_{hp})+\int_\tau(g'(y_{hp})-g'(y))w_\tau dxdy\\ &+\int_\tau(\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp}))w_\tau dxdy\\ \leqslant\;& C\|\lambda-\lambda_{hp}\|_{H^1(\tau)}|w_\tau|_{H^1(\tau)}+\|(g'(y_{hp})-g'(y))\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)}\|w_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)}\\ &+\|(\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp}))\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)}\|w_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)}.\end{aligned}\tag{53}$$

Then we should estimate w_τ with H^1-seminorm. Using the inverse estimates (28)-(29) with $\beta=\alpha$, $\gamma=2(\alpha-1)$ (note that we have $\gamma=2(\alpha-1)>-1$ when $\alpha>\frac{1}{2}$), $\delta=\alpha$, and a suitable transformation from the reference element $\hat\tau$ to τ, we have that

$$\begin{aligned}|w_\tau|^2_{H^1(\tau)} &\leqslant C\int_\tau\Phi_\tau^{2\alpha}|\nabla(\pi_{p_\tau}g'(y_{hp})+div(A^*\nabla\lambda_{hp}))|^2dxdy\\ &\quad +C\int_\tau(\pi_{p_\tau}g'(y_{hp})+div(A^*\nabla\lambda_{hp}))^2|\nabla\Phi_\tau^\alpha|^2dxdy\\ &\leqslant C\frac{p_\tau^{2(2-\alpha)}}{h_\tau^2}\int_\tau\Phi_\tau^\alpha(\pi_{p_\tau}g'(y_{hp})+div(A^*\nabla\lambda_{hp}))^2dxdy\\ &\quad +\frac{C}{h_\tau^2}\int_\tau\Phi_\tau^{2(\alpha-1)}(\pi_{p_\tau}g'(y_{hp})+div(A^*\nabla\lambda_{hp}))^2dxdy\\ &\leqslant Cp_\tau^{2(1-\alpha)}\frac{p_\tau^2}{h_\tau^2}\int_\tau\Phi_\tau^\alpha(\pi_{p_\tau}g'(y_{hp})+div(A^*\nabla\lambda_{hp}))^2dxdy\\ &= Cp_\tau^{2(1-\alpha)}\frac{p_\tau^2}{h_\tau^2}\|w_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|^2_{L^2(\tau)}.\end{aligned}\tag{54}$$

Therefore it follows from (53) and (54) that

$$
\begin{aligned}
&\|w_\tau \Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)} \\
\leqslant\ & C(p_\tau^{1-\alpha}\frac{p_\tau}{h_\tau}\|\lambda-\lambda_{hp}\|_{H^1(\tau)} + \|(g'(y_{hp})-g'(y))\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)} \\
& +\|(\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp}))\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)}) \qquad (55)\\
\leqslant\ & C(\frac{p_\tau^{2-\alpha}}{h_\tau}\|\lambda-\lambda_{hp}\|_{H^1(\tau)} + \|y-y_{hp}\|_{L^2(\tau)} + \|\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp})\|_{L^2(\tau)}).
\end{aligned}
$$

Furthermore, it follows from (55) and (28) with $\beta=\alpha$ and $\alpha=0$ that

$$
\begin{aligned}
&\|\pi_{p_\tau}(g'(y_{hp})) + div(A^*\nabla\lambda_{hp})\|_{L^2(\tau)} \\
&\leqslant Cp_\tau^\alpha\|(\pi_{p_\tau}(g'(y_{hp})) + div(A^*\nabla\lambda_{hp}))\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)} = Cp_\tau^\alpha\|w_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)} \\
&\leqslant Cp_\tau^\alpha(\frac{p_\tau^{2-\alpha}}{h_\tau}\|\lambda-\lambda_{hp}\|_{H^1(\tau)} + \|y-y_{hp}\|_{L^2(\tau)} + \|\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp})\|_{L^2(\tau)}).
\end{aligned}
$$

Thus

$$
\begin{aligned}
\xi_{B_\tau}^2 &= \frac{h_\tau^2}{p_\tau^2}\|g'(y_{hp}) + div(A^*\nabla\lambda_{hp})\|_{L^2(\tau)}^2 \\
&\leqslant C\frac{h_\tau^2}{p_\tau^2}\|\pi_{p_\tau}(g'(y_{hp})) + div(A^*\nabla\lambda_{hp})\|_{L^2(\tau)}^2 + C\frac{h_\tau^2}{p_\tau^2}\|\pi_{p_\tau}(g'(y_{hp}))-g'(y_{hp})\|_{L^2(\tau)}^2 \\
&\leqslant Cp_\tau^2\|\lambda-\lambda_{hp}\|_{H^1(\tau)}^2 + Cp_\tau^{2\alpha}\frac{h_\tau^2}{p_\tau^2}(\|y-y_{hp}\|_{L^2(\tau)}^2 + \|\pi_{p_\tau}g'(y_{hp})-g'(y_{hp})\|_{L^2(\tau)}^2).
\end{aligned}
$$

This proves (51).

Similarly, we define $v_\tau = (\pi_{p_\tau}f + \pi_{p_\tau}(Bu_{hp}) + div(A\nabla y_{hp}))\Phi_\tau^\alpha$. Then again we have that

$$
\begin{aligned}
|v_\tau|_{H^1(\tau)}^2 &\leqslant Cp_\tau^{2(1-\alpha)}\frac{p_\tau^2}{h_\tau^2}\int_\tau \Phi_\tau^\alpha(\pi_{p_\tau}f + \pi_{p_\tau}(Bu_{hp}) + div(A\nabla y_{hp}))^2dxdy \\
&= Cp_\tau^{2(1-\alpha)}\frac{p_\tau^2}{h_\tau^2}\|v_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)}^2.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\|v_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)}^2 &= \int_\tau(\pi_{p_\tau}f + \pi_{p_\tau}(Bu_{hp}) + div(A\nabla y_{hp}))v_\tau dxdy \\
&= a(y-y_{hp}, v_\tau) + \int_\tau(\pi_{p_\tau}f - f)v_\tau dxdy \\
&\quad + \int_\tau(\pi_{p_\tau}(Bu_{hp}) - \pi_{p_\tau}(Bu))v_\tau dxdy + \int_\tau(\pi_{p_\tau}(Bu) - Bu)v_\tau dxdy \\
&\leqslant C\|y-y_{hp}\|_{H^1(\tau)}|v_\tau|_{H^1(\tau)} + \|(\pi_{p_\tau}f - f)\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)}\|v_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)} \\
&\quad +\|\pi_{p_\tau}(Bu_{hp} - Bu)\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)}\|v_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)} \\
&\quad +\|(\pi_{p_\tau}(Bu) - Bu)\Phi_\tau^{\frac{\alpha}{2}}\|_{L^2(\tau)}\|v_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)},
\end{aligned}
$$

and hence,

$$
\begin{aligned}
\|v_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|_{L^2(\tau)} &\leqslant C\bigg(\frac{p_\tau^{2-\alpha}}{h_\tau}\|y-y_{hp}\|_{H^1(\tau)} + (\|B(u-u_{hp})\|_{L^2(\tau)} \\
&\quad +\|\pi_{p_\tau}f - f\|_{L^2(\tau)} + \|\pi_{p_\tau}(Bu) - Bu\|_{L^2(\tau)})\bigg). \qquad (56)
\end{aligned}
$$

Using similar techniques for ξ_{B_τ}, it is deduced that

$$\begin{aligned}
\eta_{B_\tau} &= \frac{h_\tau^2}{p_\tau^2}\|f+Bu_{hp}+div(A\nabla y_{hp})\|^2_{L^2(\tau)}\\
&\leqslant C\frac{h_\tau^2}{p_\tau^2}\|\pi_{p_\tau}f+\pi_{p_\tau}(Bu_{hp})+div(A^*\nabla\lambda_{hp})\|^2_{L^2(\tau)}+C\frac{h_\tau^2}{p_\tau^2}\|f-\pi_{p_\tau}f\|^2_{L^2(\tau)}\\
&\quad +C\frac{h_\tau^2}{p_\tau^2}\|Bu_{hp}-\pi_{p_\tau}(Bu_{hp})\|^2_{L^2(\tau)}\\
&\leqslant C\frac{h_\tau^2}{p_\tau^2}p_\tau^{2\alpha}\|v_\tau\Phi_\tau^{-\frac{\alpha}{2}}\|^2_{L^2(\tau)}+C\frac{h_\tau^2}{p_\tau^2}\|f-\pi_{p_\tau}f\|^2_{L^2(\tau)}+C\frac{h_\tau^2}{p_\tau^2}\|Bu_{hp}-\pi_{p_\tau}(Bu_{hp})\|^2_{L^2(\tau)}\\
&\leqslant Cp_\tau^2\|y-y_{hp}\|^2_{H^1(\tau)}+Cp_\tau^{2\alpha}\frac{h_\tau^2}{p_\tau^2}\bigg(\|B(u-u_{hp})\|^2_{L^2(\tau)}+\|\pi_{p_\tau}f-f\|^2_{L^2(\tau)}\\
&\quad +\|\pi_{p_\tau}(Bu)-Bu\|^2_{L^2(\tau)}\bigg).
\end{aligned}$$

Thus (52) follows, and the proof of the lemma is completed. □

Similarly we can have estimation for the residual ζ_{τ_U}.

Lemma 5.2. Let (y,λ,u) and $(y_{hp},\lambda_{hp},u_{hp})$ be the solutions of (6)-(8) and (10)-(12) respectively, and $\zeta^2_{\tau_U}$ be defined by (37). Then

$$\begin{aligned}
\zeta^2_{\tau_U} &\leqslant Cp^2_{\tau_U}\bigg(\|u-u_{hp}\|^2_{L^2(\tau_U)}+\|B^*(\lambda-\lambda_{hp})\|^2_{L^2(\tau_U)}\\
&\quad +\frac{h^2_{\tau_U}}{p^2_{\tau_U}}(\|B^*(\lambda_{hp}-\lambda)\|^2_{H^1(\tau_U)}+\|\pi_{p_{\tau_U}+1}(B^*\lambda)-B^*\lambda\|^2_{H^1(\tau_U)})\bigg),
\end{aligned}\tag{57}$$

where $\pi_{p_{\tau_U}+1}$ is the L^2-project operator on the space of polynomials of degree $p_{\tau_U}+1$ on the element τ_U.

Proof. Let $\vec{w}_{\tau_U}=\nabla(\delta u_{hp}+\pi_{p_{\tau_U}+1}(B^*\lambda_{hp}))\Phi_{\tau_U}$. Note that (17) implies $\nabla(\delta u+B^*\lambda)=0$. Then using the trivial extension by zero on $\Omega_U\backslash\tau_U$, we have

$$\begin{aligned}
\|\vec{w}_{\tau_U}\Phi_{\tau_U}^{-\frac12}\|^2_{L^2(\tau_U)} &= \int_{\tau_U}\nabla(\delta u_{hp}+\pi_{p_{\tau_U}+1}(B^*\lambda_{hp}))\cdot\vec{w}_{\tau_U}dxdy\\
&= \int_{\tau_U}\nabla(\delta u_{hp}+B^*\lambda_{hp}-\delta u-B^*\lambda)\cdot\vec{w}_{\tau_U}dxdy\\
&\quad +\int_{\tau_U}\nabla(\pi_{p_{\tau_U}+1}(B^*\lambda_{hp})-B^*\lambda_{hp})\cdot\vec{w}_{\tau_U}dxdy\\
&= -\int_{\tau_U}(\delta u_{hp}+B^*\lambda_{hp}-\delta u-B^*\lambda)\cdot div\vec{w}_{\tau_U}dxdy\\
&\quad +\int_{\tau_U}\nabla(\pi_{p_{\tau_U}+1}(B^*\lambda_{hp})-B^*\lambda_{hp})\cdot\vec{w}_{\tau_U}dxdy\\
&\leqslant C(\|B^*(\lambda-\lambda_{hp})\|_{L^2(\tau_U)}+\|u-u_{hp}\|_{L^2(\tau_U)})\|div\vec{\omega}_{\tau_U}\|_{L^2(\tau_U)}\\
&\quad +C(\|\nabla(\pi_{p_{\tau_U}+1}(B^*(\lambda_{hp}-\lambda)))\Phi^{\frac12}_{\tau_U}\|_{L^2(\tau_U)}\\
&\quad +\|\nabla(\pi_{p_{\tau_U}+1}(B^*\lambda)-B^*\lambda)\Phi^{\frac12}_{\tau_U}\|_{L^2(\tau_U)}\\
&\quad +\|\nabla(B^*(\lambda-\lambda_{hp}))\Phi^{\frac12}_{\tau_U}\|_{L^2(\tau_U)})\|\vec{w}_{\tau_U}\Phi^{-\frac12}_{\tau_U}\|_{L^2(\tau_U)}.
\end{aligned}$$

Using the same technique used in Lemma 5.1, we can derive that

$$\|div\vec{\omega}_{\tau_U}\|_{L^2(\tau_U)} \leqslant C|\vec{\omega}_{\tau_U}|_{H^1(\tau_U)} \leqslant C\frac{p_{\tau_U}}{h_{\tau_U}}\|\vec{\omega}_{\tau_U}\Phi^{-\frac{1}{2}}\|_{L^2(\tau_U)},$$

and hence

$$\begin{aligned}\|\vec{w}_{\tau_U}\Phi_{\tau_U}^{-\frac{1}{2}}\|_{L^2(\tau_U)} \quad &\leqslant \quad C\frac{p_{\tau_U}}{h_{\tau_U}}(\|u-u_{hp}\|_{L^2(\tau_U)}+\|B^*(\lambda-\lambda_{hp})\|_{L^2(\tau_U)})\\ &\quad +C(\|B^*(\lambda_{hp}-\lambda)\|_{H^1(\tau_U)}+\|\pi_{p_{\tau_U}+1}(B^*\lambda)-B^*\lambda\|_{H^1(\tau_U)}). \qquad (58)\end{aligned}$$

Then it can be deduced from (58) that

$$\begin{aligned}\zeta_{\tau_U} \quad &= \quad \frac{h_{\tau_U}^2}{p_{\tau_U}^2}\|\nabla(\delta u_{hp}+B^*\lambda_{hp})\|_{L^2(\tau_U)}^2\\ &\leqslant \quad C\frac{h_{\tau_U}^2}{p_{\tau_U}^2}\|\nabla(\delta u_{hp}+\pi_{p_{\tau_U}+1}(B^*\lambda_{hp}))\|_{L^2(\tau_U)}^2\\ &\quad\quad +C\frac{h_{\tau_U}^2}{p_{\tau_U}^2}\|\nabla(B^*\lambda_{hp}-\pi_{p_{\tau_U}+1}(B^*\lambda_{hp}))\|_{L^2(\tau_U)}^2\\ &\leqslant \quad C\frac{h_{\tau_U}^2}{p_{\tau_U}^2}p_{\tau_U}^2\|\vec{w}_{\tau_U}\Phi_{\tau_U}^{-\frac{1}{2}}\|_{L^2(\tau_U)}^2+C\frac{h_{\tau_U}^2}{p_{\tau_U}^2}\|\nabla(B^*\lambda_{hp}-\pi_{p_{\tau_U}+1}(B^*\lambda_{hp}))\|_{L^2(\tau_U)}^2\\ &\leqslant \quad Cp_{\tau_U}^2(\|u-u_{hp}\|_{L^2(\tau_U)}^2+\|B^*(\lambda-\lambda_{hp})\|_{L^2(\tau_U)}^2)\\ &\quad\quad +Cp_{\tau_U}^2\frac{h_{\tau_U}^2}{p_{\tau_U}^2}(\|B^*(\lambda_{hp}-\lambda)\|_{H^1(\tau_U)}^2+\|\pi_{p_{\tau_U}+1}(B^*\lambda)-B^*\lambda\|_{H^1(\tau_U)}^2).\end{aligned}$$

This proves (57). $\square$

In order to obtain a local upper bound for the edge contribution $\xi_{E_\tau},\ \eta_{E_\tau}$, we introduce the set

$$\omega_\tau=\Big\{\cup\tau':\ \tau' \text{ and } \tau \text{ share at least one edge}\Big\}.$$

Lemma 5.3. Let (y,λ,u) and $(y_{hp},\lambda_{hp},u_{hp})$ be the solutions of (6)-(8) and (10)-(12), ξ_{E_τ} and η_{E_τ} be defined by (34) and (36), respectively. Then

$$\begin{aligned}\xi_{E_\tau}^2 \quad &\leqslant \quad C\bigg(p_\tau^{2+2\epsilon}\|\lambda-\lambda_{hp}\|_{H^1(\omega_\tau)}^2+\frac{h_\tau^2}{p_\tau^{1-4\epsilon}}\|y-y_{hp}\|_{L^2(\omega_\tau)}^2\\ &\quad\quad +\frac{h_\tau^2}{p_\tau^{1-4\epsilon}}\sum_{\tau'\subset\omega_\tau}\|\pi_{p_{\tau'}}g'(y_{hp})-g'(y_{hp})\|_{L^2(\tau')}^2\bigg), \qquad (59)\end{aligned}$$

$$\begin{aligned}\eta_{E_\tau}^2 \quad &\leqslant \quad C\bigg(p_\tau^{2+2\epsilon}\|y-y_{hp}\|_{H^1(\omega_\tau)}^2+\frac{h_\tau^2}{p_\tau^{1-4\epsilon}}\|B(u-u_{hp})\|_{L^2(\omega_\tau)}^2\\ &\quad\quad +\frac{h_\tau^2}{p_\tau^{1-4\epsilon}}\sum_{\tau'\subset\omega_\tau}(\|\pi_{p_{\tau'}}f-f\|_{L^2(\tau')}^2+\|\pi_{p_{\tau'}}(Bu)-Bu\|_{L^2(\tau')}^2)\bigg), \qquad (60)\end{aligned}$$

where π_{p_τ} is the L^2-project operator defined in Lemma 5.1, $0<\epsilon\leqslant\frac{1}{4}$ is an arbitrary small positive number, the constant C depends on ϵ.

Proof. To obtain an upper bound for the edge contribution, we will resort to weight functions associated with the edges and a suitable extension operator. For a given element τ with the (interior) edge e, we set τ_e to be the union of all the elements sharing the edge e. We construct a function $\omega_e\in H_0^1(\tau_e)$ with $w_e|_e=[A\nabla y_{hp}\cdot n_e]\Phi_e^\beta$, $\frac{1}{2}<\beta\leqslant 1$, such that w_e and $[A\nabla y_{hp}\cdot n_e]$ will be $v_{\hat{e}}$ and ψ in Lemma 3.5

on the reference element. Noting that $\omega_e \in H_0^1(\tau_e)$, we can view ω_e as a function in $H_0^1(\Omega)$ by the trivial extension. Then we have

$$\begin{aligned}\|[A\nabla y_{hp}\cdot n_e]\Phi_e^{\frac{\beta}{2}}\|^2_{L^2(e)} &= \|\omega_e\Phi_{\tau_e}^{-\frac{\beta}{2}}\|^2_{L^2(e)} = \int_e [A\nabla(y_{hp}-y)\cdot n_e]\omega_e ds \\ &= a(y_{hp}-y,\omega_e) + \int_{\tau_e}(f+Bu+div(A\nabla y_{hp}))\omega_e dxdy \\ &\leqslant C\|y-y_{hp}\|_{H^1(\tau_e)}|\omega_e|_{H^1(\tau_e)} \\ &\quad +C\|f+Bu+div(A\nabla y_{hp})\|_{L^2(\tau_e)}\|\omega_e\|_{L^2(\tau_e)}.\end{aligned}$$

Using the affine equivalence and Lemma 3.5 with $\alpha=\beta$, we obtain the upper bounds for $|\omega_e|_{H^1(\tau_e)}$ and $\|\omega_e\|_{L^2(\tau_e)}$ in terms of $\|[A\nabla y_{hp}\cdot n_e]\Phi_e^{\frac{\beta}{2}}\|_{L^2(e)}$:

$$|\omega_e|^2_{H^1(\tau_e)} \leqslant C\frac{1}{h_\tau}(\epsilon p_\tau^{2(2-\beta)}+\epsilon^{-1})\|[A\nabla y_{hp}\cdot n_e]\Phi_e^{\frac{\beta}{2}}\|^2_{L^2(e)},$$

$$\|\omega_e\|^2_{L^2(\tau_e)} \leqslant Ch_\tau\epsilon\|[A\nabla y_{hp}\cdot n_e]\Phi_e^{\frac{\beta}{2}}\|^2_{L^2(e)},$$

where $\epsilon\in(0,1]$ is an arbitrary small positive number. Summing up, we have

$$\begin{aligned}\|[A\nabla y_{hp}\cdot n_e]\Phi_e^{\frac{\beta}{2}}\|_{L^2(e)} \leqslant C((&\frac{1}{h_\tau}(\epsilon p_\tau^{2(2-\beta)}+\epsilon^{-1}))^{\frac12}\|y-y_{hp}\|_{H^1(\tau_e)} \\ &+(h_\tau\epsilon)^{\frac12}\|f+Bu+div(A\nabla y_{hp})\|_{L^2(\tau_e)}).\end{aligned}$$

Considering (52), we sum up all the edges $e\subset\partial\tau\cap\Omega$ and then obtain that

$$\begin{aligned}&\sum_{e\subset\partial\tau\cap\Omega}\frac{h_e}{p_e}\|[A\nabla y_{hp}\cdot n_e]\Phi_e^{\frac{\beta}{2}}\|^2_{L^2(e)} \\ \leqslant\ & C\left(\frac{1}{p_\tau}(\epsilon p_\tau^{2(2-\beta)}+\epsilon^{-1})\|y-y_{hp}\|^2_{H^1(\omega_\tau)}+\epsilon p_\tau\frac{h_\tau^2}{p_\tau^2}\|f+Bu+div(A\nabla y_{hp})\|^2_{L^2(\omega_\tau)}\right) \\ \leqslant\ & C\frac{1}{p_\tau}(\epsilon p_\tau^{2(2-\beta)}+\epsilon^{-1})\|y-y_{hp}\|^2_{H^1(\omega_\tau)}+C\epsilon p_\tau\sum_{\tau'\subset\omega_\tau}\eta^2_{B_{\tau'}} \\ \leqslant\ & C\frac{1}{p_\tau}(\epsilon p_\tau^{2(2-\beta)}+\epsilon^{-1})\|y-y_{hp}\|^2_{H^1(\omega_\tau)}+C\epsilon p_\tau^3\sum_{\tau'\subset\omega_\tau}\|y-y_{hp}\|^2_{H^1(\tau')} \\ & +C\epsilon p_\tau^{1+2\alpha}\frac{h_\tau^2}{p_\tau^2}\sum_{\tau'\subset\omega_\tau}\Bigg(\|B(u-u_{hp})\|^2_{L^2(\tau')} \\ & +\|\pi_{p_{\tau'}}f-f\|^2_{L^2(\tau')}+\|\pi_{p_{\tau'}}(Bu)-Bu\|^2_{L^2(\tau')}\Bigg),\end{aligned}$$

where ϵ is an arbitrary positive numbers, and $\frac12<\alpha\leqslant 1$ is defined in Lemma 5.1. Setting $\epsilon=1/p_\tau^2$ yields that

$$\begin{aligned}&\sum_{e\subset\partial\tau\cap\Omega}\frac{h_e}{p_e}\|[A\nabla y_{hp}\cdot n_e]\Phi_e^{\frac{\beta}{2}}\|^2_{L^2(e)} \\ \leqslant\ & Cp_\tau\|y-y_{hp}\|^2_{H^1(\omega_\tau)}+Cp_\tau^{2\alpha-1}\frac{h_\tau^2}{p_\tau^2}\sum_{\tau'\subset\omega_\tau}\Bigg(\|B(u-u_{hp})\|^2_{L^2(\tau')} \\ & +\|\pi_{p_{\tau'}}f-f\|^2_{L^2(\tau')}+\|\pi_{p_{\tau'}}(Bu)-Bu\|^2_{L^2(\tau')}\Bigg).\end{aligned}$$

Using the inverse estimate for one dimension analogue to Lemma 3.4 and setting $\alpha = \beta = \frac{1}{2} + \epsilon$ with $0 < \epsilon \leqslant \frac{1}{4}$, we have that

$$\begin{aligned}
\eta_{E_\tau}^2 &= \sum_{e \subset \partial\tau \cap \Omega} \frac{h_e}{2p_e} \|[A\nabla y_{hp} \cdot n_e]\|_{L^2(e)}^2 \leqslant C p_\tau^{2\beta} \sum_{e \subset \partial\tau \cap \Omega} \frac{h_e}{p_e} \|[A\nabla y_{hp} \cdot n_e]\Phi_e^{\frac{\beta}{2}}\|_{L^2(e)}^2 \\
&\leqslant C p_\tau^{1+2\beta} \|y - y_{hp}\|_{H^1(\omega_\tau)}^2 + C p_\tau^{2\beta+2\alpha-1} \frac{h_\tau^2}{p_\tau^2} \sum_{\tau' \subset \omega_\tau} \Big(\|B(u - u_{hp})\|_{L^2(\tau')}^2 \\
&\quad + \|\pi_{p_{\tau'}} f - f\|_{L^2(\tau')}^2 + \|\pi_{p_{\tau'}}(Bu) - Bu\|_{L^2(\tau')}^2 \Big) \\
&\leqslant C \Big(p_\tau^{2+2\epsilon} \|y - y_{hp}\|_{H^1(\omega_\tau)}^2 + \frac{h_\tau^2}{p_\tau^{1-4\epsilon}} \|B(u - u_{hp})\|_{L^2(\omega_\tau)}^2 \\
&\quad + \frac{h_\tau^2}{p_\tau^{1-4\epsilon}} \sum_{\tau' \subset \omega_\tau} (\|\pi_{p_{\tau'}} f - f\|_{L^2(\tau')}^2 + \|\pi_{p_{\tau'}}(Bu) - Bu\|_{L^2(\tau')}^2) \Big).
\end{aligned}$$

This proves (60). The estimate for ξ_{E_τ} (59) can be proved similarly. □

Summing up, we can obtain the following global lower estimates using Lemmas 5.1-5.3.

Theorem 5.4. Let (y, λ, u) and $(y_{hp}, \lambda_{hp}, u_{hp})$ be the solutions of (6)-(8) and (10)-(12), respectively. Let ξ, η and ζ be defined by (30)-(37). Assume that all the conditions in Lemmas 5.1-5.3 are valid. Then we have

$$\begin{aligned}
\xi^2 + \eta^2 + \zeta^2 &\leqslant C \sum_{\tau \in \mathcal{T}} p_\tau^{2+2\epsilon} (\|\lambda - \lambda_{hp}\|_{H^1(\omega_\tau)}^2 + \|y - y_{hp}\|_{H^1(\omega_\tau)}^2 + E_1^2) \\
&\quad + C \sum_{\tau_U \in \mathcal{T}_U} p_{\tau_U}^2 (\|u - u_{hp}\|_{L^2(\tau_U)}^2 + \|B^*(\lambda - \lambda_{hp})\|_{L^2(\tau_U)}^2 + E_2^2),
\end{aligned} \tag{61}$$

where ξ, η and ζ are defined by (30)-(37), $0 < \epsilon \leqslant \frac{1}{4}$ is an arbitrary small positive number, and

$$\begin{aligned}
E_1^2 &= \sum_{\tau \in \mathcal{T}} \frac{h_\tau^2}{p_\tau^2} (\|\pi_{p_\tau} f - f\|_{L^2(\tau)}^2 + \|\pi_{p_\tau} g'(y_{hp}) - g'(y_{hp})\|_{L^2(\tau)}^2 + \|\pi_{p_\tau}(Bu) - Bu\|_{L^2(\tau)}^2), \\
E_2^2 &= \sum_{\tau_U \in \mathcal{T}_U} \frac{h_{\tau_U}^2}{p_{\tau_U}^2} \|\pi_{p_{\tau_U}+1}(B^*\lambda) - B^*\lambda\|_{H^1(\tau_U)}^2.
\end{aligned}$$

Proof. Summing up the results in Lemmas 5.1-5.3 and setting $\alpha = \frac{1}{2} + \epsilon$ in Lemma 5.1, we have that

$$
\begin{aligned}
& \xi^2 + \eta^2 + \zeta^2 \\
= \;& \sum_{\tau\in\mathcal{T}} (\xi_{B_\tau}^2 + \xi_{E_\tau}^2) + \sum_{\tau\in\mathcal{T}} (\eta_{B_\tau}^2 + \eta_{E_\tau}^2) + \sum_{\tau_U\in\mathcal{T}_U} \zeta_{\tau_U}^2 \\
\leqslant \;& C\sum_{\tau\in\mathcal{T}} \bigg(p_\tau^2 \|\lambda - \lambda_{hp}\|_{H^1(\tau)}^2 + \frac{h_\tau^2}{p_\tau^{1-2\epsilon}} (\|y - y_{hp}\|_{L^2(\tau)} + \|\pi_{p_\tau} g'(y_{hp}) - g'(y_{hp})\|_{L^2(\tau)}^2) \bigg) \\
& + C\sum_{\tau\in\mathcal{T}} \bigg(p_\tau^{2+2\epsilon} \|\lambda - \lambda_{hp}\|_{H^1(\omega_\tau)}^2 + \frac{h_\tau^2}{p_\tau^{1-4\epsilon}} \|y - y_{hp}\|_{L^2(\omega_\tau)} \\
& + \frac{h_\tau^2}{p_\tau^{1-4\epsilon}} \sum_{\tau'\subset\omega_\tau} \|\pi_{p_{\tau'}} g'(y_{hp}) - g'(y_{hp})\|_{L^2(\tau')}^2 \bigg) \\
& + C\sum_{\tau\in\mathcal{T}} \bigg(p_\tau^2 \|y - y_{hp}\|_{H^1(\tau)}^2 + \frac{h_\tau^2}{p_\tau^{1-2\epsilon}} (\|B(u - u_{hp})\|_{L^2(\tau)}^2 \\
& + \|\pi_{p_\tau} f - f\|_{L^2(\tau)}^2 + \|\pi_{p_\tau}(Bu) - Bu\|_{L^2(\tau)}^2) \bigg) \\
& + C\sum_{\tau\in\mathcal{T}} \bigg(p_\tau^{2+2\epsilon} \|y - y_{hp}\|_{H^1(\omega_\tau)}^2 + \frac{h_\tau^2}{p_\tau^{1-4\epsilon}} \|B(u - u_{hp})\|_{L^2(\omega_\tau)}^2 \\
& + \frac{h_\tau^2}{p_\tau^{1-4\epsilon}} \sum_{\tau'\subset\omega_\tau} (\|\pi_{p_{\tau'}} f - f\|_{L^2(\tau')}^2 + \|\pi_{p_{\tau'}}(Bu) - Bu\|_{L^2(\tau')}^2) \bigg) \\
& + C\sum_{\tau_U\in\mathcal{T}_U} p_{\tau_U}^2 \bigg(\|u - u_{hp}\|_{L^2(\tau_U)}^2 + \|B^*(\lambda - \lambda_{hp})\|_{L^2(\tau_U)}^2 \\
& + \frac{h_{\tau_U}^2}{p_{\tau_U}^2} (\|B^*(\lambda_{hp} - \lambda)\|_{H^1(\tau_U)}^2 + \|\pi_{p_{\tau_U}+1}(B^*\lambda) - B^*\lambda\|_{H^1(\tau_U)}^2) \bigg) \\
\leqslant \;& C\bigg(\sum_{\tau\in\mathcal{T}} p_\tau^{2+2\epsilon} (\|\lambda - \lambda_{hp}\|_{H^1(\omega_\tau)}^2 + \|y - y_{hp}\|_{H^1(\omega_\tau)}^2) \\
& + \sum_{\tau_U\in\mathcal{T}_U} p_{\tau_U}^2 (\|u - u_{hp}\|_{L^2(\tau_U)}^2 + \|B^*(\lambda - \lambda_{hp})\|_{L^2(\tau_U)}^2) \\
& + \sum_{\tau\in\mathcal{T}} p_\tau^{2+2\epsilon} \frac{h_\tau^2}{p_\tau^{3-2\epsilon}} (\|B(u - u_{hp})\|_{L^2(\omega_\tau)}^2 + \sum_{\tau'\subset\omega_\tau} (\|\pi_{p_{\tau'}} f - f\|_{L^2(\tau')}^2 \\
& + \|\pi_{p_{\tau'}}(Bu) - Bu\|_{L^2(\tau')}^2 + \|\pi_{p_{\tau'}} g'(y_{hp}) - g'(y_{hp})\|_{L^2(\tau')}^2)) \\
& + \sum_{\tau_U\in\mathcal{T}_U} p_{\tau_U}^2 \frac{h_{\tau_U}^2}{p_{\tau_U}^2} (\|B^*(\lambda_{hp} - \lambda)\|_{H^1(\tau_U)}^2 + \|\pi_{p_{\tau_U}+1}(B^*\lambda) - B^*\lambda\|_{H^1(\tau_U)}^2) \bigg).
\end{aligned}
$$

Noting that when $0 < \epsilon \leqslant \frac{1}{4}$, we have $3 - 2\epsilon \geqslant 2$, $3 - 2\epsilon - (2 + 2\epsilon) = 1 - 4\epsilon \geqslant 0$, and

$$
\sum_{\tau\in\mathcal{T}} p_\tau^{2+2\epsilon} \frac{h_\tau^2}{p_\tau^{3-2\epsilon}} \|B(u - u_{hp})\|_{L^2(\omega_\tau)}^2 \leqslant C\|u - u_{hp}\|_{L^2(\Omega_U)}^2 \leqslant C \sum_{\tau_U\in\mathcal{T}_U} p_{\tau_U}^2 \|u - u_{hp}\|_{L^2(\tau_U)}^2,
$$

$$
\sum_{\tau_U\in\mathcal{T}_U} p_{\tau_U}^2 \frac{h_{\tau_U}^2}{p_{\tau_U}^2} \|B^*(\lambda_{hp} - \lambda)\|_{H^1(\tau_U)}^2 \leqslant C\|\lambda_{hp} - \lambda\|_{H^1(\Omega)}^2 \leqslant C \sum_{\tau\in\mathcal{T}} p_\tau^{2+2\epsilon} \|\lambda - \lambda_{hp}\|_{H^1(\omega_\tau)}^2.
$$

Then (61) is proved. □

Remark 5.5. It follows from Theorems 4.2 and 5.4 that

$$\|u-u_{hp}\|^2_{L^2(\Omega_U)} + \|y-y_{hp}\|^2_{H^1(\Omega)} + \|\lambda-\lambda_{hp}\|^2_{H^1(\Omega)} \leqslant C(\zeta^2+\xi^2+\eta^2),$$

and

$$\begin{aligned}\xi^2+\eta^2+\zeta^2 \quad &\leqslant \quad C\sum_{\tau\in\mathcal{T}} p_\tau^{2+2\epsilon}(\|\lambda-\lambda_{hp}\|^2_{H^1(\omega_\tau)} + \|y-y_{hp}\|^2_{H^1(\omega_\tau)} + E_1^2) \\ &\quad +C\sum_{\tau_U\in\mathcal{T}_U} p^2_{\tau_U}(\|u-u_{hp}\|^2_{L^2(\tau_U)} + \|B^*(\lambda-\lambda_{hp})\|^2_{L^2(\tau_U)} + E_2^2),\end{aligned}$$

where E_1 and E_2 are defined in Theorems 5.4, which are all higher order terms under some regularity conditions. Then we obtain the a posteriori error estimates with the upper and lower bounds, although there is a gap of order p^2 between the upper and lower bounds. In order to obtain equivalent a posteriori error estimates for the hp-finite element method, the Jacobi-weighted Sobolev space may should be explored (see [15]). But to our knowledge, it seems to be difficult to use the approach in constrained optimal control problems.

6. A family of a posteriori error estimates

In this section, for each $\alpha\in[0,1]$ we introduce a family of weighted local error estimators $\xi_{\alpha;\tau},\eta_{\alpha;\tau}$ for each element τ, and $\zeta_{\alpha;\tau_U}$ for each element τ_U, as in [26]. The estimators are defined as follows:

$$\xi^2_{\alpha;\tau} := \xi^2_{\alpha;B_\tau} + \xi^2_{\alpha;E_\tau}, \tag{62}$$

$$\eta^2_{\alpha;\tau} := \eta^2_{\alpha;B_\tau} + \eta^2_{\alpha;E_\tau}, \tag{63}$$

and

$$\zeta^2_{\alpha;\tau_U} := \frac{h^2_{\tau_U}}{p^2_{\tau_U}}\|\nabla(\delta u_{hp} + \pi_{p_{\tau_U}+1}(B^*\lambda_{hp}))\Phi^{\alpha/2}_{\tau_U}\|^2_{L^2(\tau_U)}, \tag{64}$$

where

$$\xi^2_{\alpha;B_\tau} := \frac{h^2_\tau}{p^2_\tau}\|(\pi_{p_\tau}g'(y_{hp}) + div(A^*\nabla\lambda_{hp}))\Phi^{\alpha/2}_\tau\|^2_{L^2(\tau)}, \tag{65}$$

$$\xi^2_{\alpha;E_\tau} := \sum_{e\subset\partial\tau\cap\Omega}\frac{h_e}{2p_e}\|[A^*\nabla\lambda_{hp}\cdot n_e]\Phi^{\alpha/2}_e\|^2_{L^2(e)}, \tag{66}$$

$$\eta^2_{\alpha;B_\tau} := \frac{h^2_\tau}{p^2_\tau}\|(\pi_{p_\tau}f + \pi_{p_\tau}(Bu_{hp}) + div(A\nabla y_{hp}))\Phi^{\alpha/2}_\tau\|^2_{L^2(\tau)}, \tag{67}$$

$$\eta^2_{\alpha;E_\tau} := \sum_{e\subset\partial\tau\cap\Omega}\frac{h_e}{2p_e}\|[A\nabla y_{hp}\cdot n_e]\Phi^{\alpha/2}_e\|^2_{L^2(e)}, \tag{68}$$

where π_{p_τ} ($\pi_{p_{\tau_U}+1}$) is the $L^2(\tau)$ ($L^2(\tau_U)$) projection operator on the space of polynomials of degree p_τ ($p_{\tau_U}+1$), Φ_τ, Φ_{τ_U} and Φ_e are the weight functions defined before Lemma 3.4 in Section 3. Comparing with the definitions of the a posteriori estimators defined by (30)-(37), it can be found that ξ_α, η_α and ζ_α are similar to ξ, η and ζ when $\alpha=0$. Only differences are that $g'(y_{hp})$, f, Bu_{hp} and $B^*\lambda_{hp}$, in ξ, η and ζ are now replaced by $\pi_{p_\tau}g'(y_{hp})$, $\pi_{p_\tau}f$, $\pi_{p_\tau}(Bu_{hp})$ and $\pi_{p_{\tau_U}+1}(B^*\lambda_{hp})$ in ξ_0, η_0 and ζ_0. As usual the global error estimators are given by the sum of their local contributions:

$$\xi^2_\alpha = \sum_{\tau\in\mathcal{T}}\xi^2_{\alpha;\tau}, \quad \eta^2_\alpha = \sum_{\tau\in\mathcal{T}}\eta^2_{\alpha;\tau}, \quad \zeta^2_\alpha = \sum_{\tau_U\in\mathcal{T}_U}\zeta^2_{\alpha;\tau_U}. \tag{69}$$

For the above weighted a posteriori error estimators, we can obtain the following a posteriori error estimates, using the weighted technique (see [26] for more details) and the approaches similar to those used in Sections 4 and 5. Here we only state the results and omit their proofs.

Theorem 6.1. Let (y, λ, u) and $(y_{hp}, \lambda_{hp}, u_{hp})$ be the solutions of (6)-(8) and (10)-(12), $\alpha \in [0,1]$. Assume that all the conditions in Theorem 4.2 are valid. Then when $h = \max\{h_\tau\}$ and $h_U = \max\{h_{\tau_U}\}$ are small enough, there exists a constant $C > 0$ independent of h and p such that

$$\begin{aligned} & \|u - u_{hp}\|^2_{L^2(\Omega_U)} + \|y - y_{hp}\|^2_{H^1(\Omega)} + \|\lambda - \lambda_{hp}\|^2_{H^1(\Omega)} \\ & \leqslant C \sum_{\tau \in \mathcal{T}} (p_\tau^{2\alpha} \xi^2_{\alpha;\tau} + p_\tau^{2\alpha} \eta^2_{\alpha;\tau}) + C \sum_{\tau_U \in \mathcal{T}_U} p_{\tau_U}^{2\alpha} \zeta^2_{\alpha;\tau_U} + CE^2, \end{aligned} \tag{70}$$

where $\xi_{\alpha;\tau}$, $\eta_{\alpha;\tau}$ and $\zeta_{\alpha;\tau_U}$ are defined by (62)-(68),

$$\begin{aligned} E^2 &= \sum_{\tau \in \mathcal{T}} \frac{h_\tau^2}{p_\tau^2} (\|\pi_{p_\tau} f - f\|^2_{L^2(\tau)} + \|\pi_{p_\tau}(Bu) - Bu\|^2_{L^2(\tau)} \\ &\quad + \|\pi_{p_\tau} g'(y_{hp}) - g'(y_{hp})\|^2_{L^2(\tau)}) + \sum_{\tau_U \in \mathcal{T}_U} \frac{h^2_{\tau_U}}{p^2_{\tau_U}} \|\pi_{p_{\tau_U}+1}(B^*\lambda) - B^*\lambda\|^2_{H^1(\tau_U)}, \end{aligned}$$

π_{p_τ} and $\pi_{p_{\tau_U}+1}$ are the $L^2(\tau)$-projection operator on the space of polynomials of degree p_τ, and the $L^2(\tau_U)$-projection operator on the space of polynomials of degree $p_{\tau_U} + 1$, respectively.

Theorem 6.2. Let (y, λ, u) and $(y_{hp}, \lambda_{hp}, u_{hp})$ be the solutions of (6)-(8) and (10)-(12), $\zeta^2_{\alpha;\tau_U}$, $\xi_{\alpha;B_\tau}$, $\xi_{\alpha;E_\tau}$, $\eta_{\alpha;B_\tau}$ and $\eta_{\alpha;E_\tau}$ be defined by (64)-(68), $\alpha \in [0,1]$, and ϵ be an arbitrary small positive number. Assume that $g(\cdot)$ is Lipschitz continuous. Then there exists a constant $C(\epsilon) > 0$, which is dependent on ϵ but independent of h, $\mathbf{p}$ and $\tau \in \mathcal{T}$ $(\tau_U \in \mathcal{T}_U)$, such that

$$\begin{aligned} \zeta^2_{\alpha;\tau_U} &\leqslant C(\epsilon)\bigg(p_{\tau_U}^{2(1-\alpha)} (\|u - u_{hp}\|^2_{L^2(\tau_U)} + \|B^*(\lambda - \lambda_{hp})\|^2_{L^2(\tau_U)}) \\ &\quad + p_{\tau_U}^{max\{1+2\epsilon-2\alpha,0\}} \frac{h^2_{\tau_U}}{p^2_{\tau_U}} (\|B^*(\lambda_{hp} - \lambda)\|^2_{H^1(\tau_U)} + \|\pi_{p_{\tau_U}+1}(B^*\lambda) - B^*\lambda\|^2_{H^1(\tau_U)}) \bigg), \end{aligned}$$

$$\begin{aligned} \xi^2_{\alpha;B_\tau} &\leqslant C(\epsilon)\bigg(p_\tau^{2(1-\alpha)} \|\lambda - \lambda_{hp}\|^2_{H^1(\tau)} + p_\tau^{max\{1+2\epsilon-2\alpha,0\}} \frac{h_\tau^2}{p_\tau^2} (\|y - y_{hp}\|^2_{L^2(\tau)} \\ &\quad + \|\pi_{p_\tau} g'(y_{hp}) - g'(y_{hp})\|^2_{L^2(\tau)}) \bigg), \end{aligned}$$

$$\begin{aligned} \eta^2_{\alpha;B_\tau} &\leqslant C(\epsilon)\bigg(p_\tau^{2(1-\alpha)} \|y - y_{hp}\|^2_{H^1(\tau)} + p_\tau^{max\{1+2\epsilon-2\alpha,0\}} \frac{h_\tau^2}{p_\tau^2} (\|B(u - u_{hp})\|^2_{L^2(\tau)} \\ &\quad + \|\pi_{p_\tau} f - f\|^2_{L^2(\tau)} + \|\pi_{p_\tau}(Bu) - Bu\|^2_{L^2(\tau)}) \bigg), \end{aligned}$$

$$\begin{aligned} \xi^2_{\alpha;E_\tau} &\leqslant C(\epsilon) p_\tau^{max\{1+2\epsilon-2\alpha,0\}} \bigg(p_\tau \|\lambda - \lambda_{hp}\|^2_{H^1(\omega_\tau)} + p_\tau^{2\epsilon} \frac{h_\tau^2}{p_\tau^2} \|y - y_{hp}\|^2_{L^2(\omega_\tau)} \\ &\quad + p_\tau^{2\epsilon} \frac{h_\tau^2}{p_\tau^2} \sum_{\tau' \subset \omega_\tau} \|\pi_{p_{\tau'}} g'(y_{hp}) - g'(y_{hp})\|^2_{L^2(\tau')} \bigg), \end{aligned}$$

$$\begin{aligned}\eta^2_{\alpha;E_\tau} \leqslant\ & C(\epsilon)p_\tau^{max\{1+2\epsilon-2\alpha,0\}}\Bigg(p_\tau\|y-y_{hp}\|^2_{H^1(\omega_\tau)}+p_\tau^{2\epsilon}\frac{h_\tau^2}{p_\tau^2}\|B(u-u_{hp})\|^2_{L^2(\omega_\tau)}\\ &+p_\tau^{2\epsilon}\frac{h_\tau^2}{p_\tau^2}\sum_{\tau'\subset\omega_\tau}(\|\pi_{p_{\tau'}}f-f\|^2_{L^2(\tau')}+\|\pi_{p_{\tau'}}(Bu)-Bu\|^2_{L^2(\tau')})\Bigg),\end{aligned}$$

where π_{p_τ} and $\pi_{p_{\tau_U}+1}$ are the L^2-project operators defined in Theorem 6.

Theorem 6.3. Let (y,λ,u) and $(y_{hp},\lambda_{hp},u_{hp})$ be the solutions of (6)-(8) and (10)-(12), respectively. Let ξ_α, η_α and ζ_α be defined by (62)-(69). Assume that all the conditions in Lemma 5.1 are all valid. Then we have

$$\begin{aligned}&\xi_\alpha^2+\eta_\alpha^2+\zeta_\alpha^2\\ (71)\leqslant\ & C(\epsilon)\sum_{\tau\in\mathcal{T}}p_\tau^{max\{1+2\epsilon-2\alpha,0\}}p_\tau(\|\lambda-\lambda_{hp}\|^2_{H^1(\omega_\tau)}+\|y-y_{hp}\|^2_{H^1(\omega_\tau)})\\ &+C(\epsilon)\sum_{\tau_U\in\mathcal{T}_U}p_{\tau_U}^{2(1-\alpha)}(\|u-u_{hp}\|^2_{L^2(\tau_U)}+\|B^*(\lambda-\lambda_{hp})\|^2_{L^2(\tau_U)})+C(\epsilon)E_\alpha^2,\end{aligned}$$

where

$$\begin{aligned}E_\alpha^2 =\ & \sum_{\tau\in\mathcal{T}}p_\tau^{max\{1+2\epsilon-2\alpha,0\}}p_\tau^{2\epsilon}\frac{h_\tau^2}{p_\tau^2}\sum_{\tau'\subset\omega_\tau}\Bigg(\|\pi_{p_{\tau'}}f-f\|^2_{L^2(\tau')}\\ &+\|\pi_{p_{\tau'}}g'(y_{hp})-g'(y_{hp})\|^2_{L^2(\tau')}+\|\pi_{p_{\tau'}}(Bu)-Bu\|^2_{L^2(\tau')}\Bigg)\\ &+\sum_{\tau_U\in\mathcal{T}_U}p_{\tau_U}^{max\{1+2\epsilon-2\alpha,0\}}\frac{h^2_{\tau_U}}{p^2_{\tau_U}}\|\pi_{p_{\tau_U}+1}(B^*\lambda)-B^*\lambda\|^2_{H^1(\tau_U)}.\end{aligned}$$

Remark 6.4. It follows from Theorems 6.1-6.3 that $\xi_0^2+\eta_0^2+\zeta_0^2$ provides an a posteriori error upper bound, while

$$\sum_{\tau\in\mathcal{T}}(\xi^2_{1;B_\tau}+\eta^2_{1;B_\tau})+\sum_{\tau\in\mathcal{T}}p_\tau^{-1}(\xi^2_{\frac12+\epsilon;E_\tau}+\eta^2_{\frac12+\epsilon;E_\tau})+\sum_{\tau_U\in\mathcal{T}_U}\zeta^2_{1;\tau_U}$$

provides an a posteriori error lower bound, if the higher order terms can be ignored. The estimator $\xi_\alpha^2+\eta_\alpha^2+\zeta_\alpha^2$ also provides an a posteriori error lower bound with the factors $p_\tau^{1+max\{1+2\epsilon-2\alpha,0\}}$ and $p_{\tau_U}^{2(1-\alpha)}$. But as pointed in Remark 5.5, there is still a gap between the upper and lower bounds.

7. Discussions

In this paper, we discussed a posteriori error estimates of the hp-finite element method for distributed convex optimal control problem governed by the elliptic partial differential equation. It is shown that the a posteriori error estimators derived in this paper provide both upper and lower bounds for the approximation errors, although the lower bound is suboptimal in the sense that there is a gap of order p^2 between the upper and lower bounds.

In this area there are many important issues that still need to be addressed. For example, studies for more complicated control problems and constraint sets are needed. Also it is interesting to explore the Jacobi-weighted Sobolev space to improve the a posteriori error estimates of the hp-finite element method for optimal control problems. Furthermore many computational issues have to be addressed. For example, adaptive refinement strategy should be investigated for efficiently implementing adaptive hp-finite element method for optimal control problems.

References

[1] M. Ainsworth and J. T. Oden, A posteriori error estimators in finite element analysis, Comput. Methods Appl. Mech. Engrg., 142(1997), pp. 1-88.

[2] M. Ainsworth and J.T. Oden, A Posteriori Error Estimation in Finite Element Analysis, Wiley Interscience, New York, 2000.

[3] M. Ainsworth and B. Senior, An adaptive refinement strategy for *hp*-finite element computations, Appl. Numer. Math., 26(1998), pp. 165-178.

[4] I. Babuška and C. Rheinboldt, Error estimates for adaptive finite element computations, SIAM J. Numer. Anal., 15(1978), pp. 736-754.

[5] I. Babuška and T. Strouboulis, The Finite Element Method and its Reliability, Oxford University Press, 2001.

[6] R. Becker, H. Kapp and R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: Basic concept, SIAM J. Control Optimization, 39(2000), pp. 113-132.

[7] C. Bernardi, Indicateurs d'erreur en h-N version des elements spectraux, Math. Modeling Numer. Anal., 30(1)(1996), pp. 1-38.

[8] P. G. Ciarlet, The finite element methods for elliptic problems, North-Holland, Amsterdam, 1978.

[9] Ph. Clément, Approximation by finite element functions using local regularization, RAIRO Anal. Numer. 9(1975), pp. 77-84.

[10] D. Dunavant and B. Szabó, A posteriori error indicators for the p-version of the finite element method, Internat. J. Numer. Methods Engrg., 19(1983), pp. 1851-1870.

[11] F. S. Falk, Approximation of a class of optimal control problems with order of convergence estimates, J. Math. Anal. Appl., 44(1973), pp. 28-47.

[12] L. Ge, W. B. Liu and D.P Yang, Adaptive finite element approximation for a constrained optimal control problem via multi-meshes, submitted.

[13] T. Geveci, On the approximation of the solution of an optimal control problem governed by an elliptic equation, RAIRO Anal. Numer., 13(1979), pp. 313-328.

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

[15] B. Q. Guo, Recent Progress in A-posteriori Error Estimation for the p-version of Finite Element Method, Recent Advances in Adaptive, eds. Z-C. Shi, Z. Chen, T.Tang and D.Yu, AMS, Comtemporary Mathematics 383 (2005), 47-61.

[16] P. Houston, D. Schötzau and T. P. Wihler, Energy norm a posteriori error estimation of *hp*-adaptive discontinuous galerkin methods for elliptic problems, Math. Models and Meth. in Appl. Sciences, 17(2007), pp. 33-62.

[17] Y. Q. Huang, R. Li, W. B. Liu and N. N. Yan, Efficient discretization for finite element approximation of constrained optimal control problems, submitted.

[18] R. Li, W. B. Liu, H. P. Ma and T. Tang, Adaptive finite element approximation of elliptic optimal controls, SIAM J. Control Optimization, 41(2002), pp. 1321-1349.

[19] J. L. Lions, Optimal control of systems governed by partial differential equations, Springer-Verlag, Berlin, 1971.

[20] W.B Liu, Adaptive multi-meshes in finite element approximation of optimal control, Contemporary Mathematics, Volume 383. pp.113-132, October 2005.

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

[22] W. B. Liu and N. N. Yan, A posteriori error analysis for convex distributed optimal control problems, Adv. Comp. Math., 15, 1-4(2001), pp. 285-309.

[23] W. B. Liu and N. N. Yan, A posteriori error estimates for optimal problems governed by Stokes equations, SIAM J. Numer. Anal., 40(2003), pp. 1850-1869.

[24] W. B. Liu and N. N. Yan, A posteriori error estimates for optimal control problems governed by parabolic equations, Numer. Math., 93(2003), pp. 497-521.

[25] J. M. Melenk, *hp*-interpolation of non-smooth functions, SIAM J. Numer. Anal., 43(2005), pp. 127-155.

[26] J. M. Melenk and B. I. Wohlmuth, On residual-based a posteriori error estimation in *hp*-FEM, Adv. Comp. Math., 15(2001), pp. 311-331.

[27] P. Neittaanmaki and D. Tiba, Optimal control of nonlinear parabolic systems. Theory, Algorithms and Applications, M. Dekker, New York, 1994.

[28] T. J. Oden, L. Demkowicz, W. Rachowicz and T. A. Westermann, Towards a universal hp finite element strategy. Part 2. A posteriori error estimation, Comput. Methods Appl. Mech. Engrg. 77(1989), pp. 113-180.

[29] C. Schwab, p- and hp-finite element methods, Oxford Univ. Press, Oxford, 1998.

[30] L. R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, Math. Comp., 54(1990), pp. 483-493.

[31] R. Verfürth, A review of a posteriori error estimation and adaptive mesh refinement, Wiley-Teubner, London, UK, 1996.

LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.

E-mail: `wgong@lsec.cc.ac.cn`

KBS & Institute of Mathematics and Statistics, The University of Kent, Canterbury, CT2 7NF England.

E-mail: `W.B.Liu@ukc.ac.uk`

LSEC, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.

E-mail: `ynn@amss.ac.cn`