Journal of Computational Mathematics
Vol.27, No.4, 2009, 525–542.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2009.27.4.017

A POSTERIORI ERROR ESTIMATE OF OPTIMAL CONTROL PROBLEM OF PDE WITH INTEGRAL CONSTRAINT FOR STATE*

Lei Yuan
Department of Mathematics, Shandong University, Jinan 250100, China
Email: yuanlei.cn@gmail.com
Danping Yang
Department of Mathematics, East China Normal University, Shanghai 200241, China
Email: dpyang@math.ecnu.edu.cn

Abstract

In this paper, we study adaptive finite element discretization schemes for an optimal control problem governed by elliptic PDE with an integral constraint for the state. We derive the equivalent a posteriori error estimator for the finite element approximation, which particularly suits adaptive multi-meshes to capture different singularities of the control and the state. Numerical examples are presented to demonstrate the efficiency of a posteriori error estimator and to confirm the theoretical results.

Mathematics subject classification: 49J20, 65N30.
Key words: State-constrained optimal control problem, Adaptive finite element method, A posteriori error estimate.

1. Introduction

Finite element approximations of optimal control problems have been extensively studied in the literatures, most of which focused on control-constrained problems. In recent years, many studies have been carried out to examine finite element approximations of optimal control problems with state constraints. Many authors have studied the specific case of point-wise state constraints, where the constraints have the forms of $K = \{y : y \geqslant \varphi\}$ or $K = \{y \leqslant \varphi\}$, see, e.g., [3,6,7,9]. In engineering applications, one may care more about how to control the average value or L^2-norm of the state variable. So there exist many state-constraints of average types, such as integral constraint $K = \{y : \alpha \leqslant \int_\Omega y \leqslant \beta\}$, L^2-norm constraint $K = \{y : \int_\Omega y^2 \leqslant \gamma\}$ and so on. In this work we study a posteriori error estimates of the finite element approximation of an optimal control problem with the integral constraint for the state. It has been recently found that suitable adaptive meshes can greatly improve computational efficiency of the finite element approximation of the optimal control, see, e.g., [2, 13, 16–19]. Furthermore it has been observed that multi-meshes are often useful in computing optimal controls, see, e.g., [10, 15]. Using different adaptive meshes for the control and the state allows to use very coarse meshes in solving the state equation and the co-state equation. Thus much computational time can be saved since one of the major computational loads in computing optimal control is to solve the state and co-state equations repeatedly.

The purpose of this work is to investigate adaptive multi-mesh finite element method for a state-constrained optimal control problem with the integral constraint. We derive an equivalent

* Received June 26, 2008 / Revised version received August 26, 2008 / Accepted October 20, 2008 /

a posteriori error estimator for this control problem and then present numerical results to confirm the effectiveness of the error estimator.

The plan of the paper is as follows. In Section 2, we construct the finite element approximation for the distributed optimal control problem with the state-constraint. In Section 3, an equivalent a posteriori error estimator is derived for the control problem. Finally numerical test results are presented in Section 4.

2. Model Problem and Its Finite Element Approximation

Let Ω be a bounded domain in $\mathbb{R}^d$, $1 \leqslant d \leqslant 3$, with the Lipschitz boundary $\partial\Omega$. Throughout the paper we use the standard notations for the Sobolev spaces, norms and seminorms. We denote the L^2-inner products in $L^2(\Omega)$ and $(L^2(\Omega))^d$ by

$$(v,w) = \int_\Omega vw, \quad \forall\, v,w \in L^2(\Omega)$$

and

$$(\mathbf{w},\mathbf{v}) = \int_\Omega \mathbf{w}\cdot\mathbf{v}, \quad \forall\, \mathbf{v},\mathbf{w} \in (L^2(\Omega))^d.$$

For a nonnegative integer m, $H^m(\Omega)$ denotes the usual Sobolev spaces and $H_0^1(\Omega) = \{v \in H^1(\Omega):\ v = 0 \text{ on } \partial\Omega\}$.

2.1. Optimal control problem and its optimality condition

Introduce function spaces $U = L^2(\Omega)$, $V = H_0^1(\Omega)$, and $W = \{v \in H_0^1(\Omega):\ \Delta v \in L^2(\Omega)\}$. Obviously, W is a Hilbert space with the norm:

$$\|v\|_W = \left(\ \|v\|^2_{H^1(\Omega)} + \|\Delta v\|^2_{L^2(\Omega)}\ \right)^{1/2}.$$

The constraint set is defined as follows:

$$K = \Big\{ v \in W:\ \int_\Omega v \geqslant \gamma \Big\}, \tag{2.1}$$

where γ is a given constant.

We will investigate the distributed convex state-constrained optimal control problem (OCP):

$$(OCP) \qquad \begin{cases} \min \quad \mathcal{J}(u,y) = \dfrac{1}{2}\displaystyle\int_\Omega (y-y_d)^2 + \dfrac{\alpha}{2}\displaystyle\int_\Omega u^2, \\ s.t. \quad -\Delta y = u + f \ \ in\ \Omega, \quad y = 0 \ \ on\ \partial\Omega, \quad y \in K, \end{cases} \tag{2.2}$$

where $u \in U$ is the control and $y \in K$ is the state, $y_d \in L^2(\Omega)$ is the observation, $f \in L^2(\Omega)$ is a given function, and α is a given positive constant.

In [20],it is proved that the state-constrained optimal control problem (OCP) has a unique solution $(u,y) \in U \times V$. Further, the pair (u,y) is the solution of (OCP) if and only if there exists a unique pair $(p,\lambda) \in V \times \mathbb{R}^1_-$ where $\mathbb{R}^1_- = \{c \in \mathbb{R}^1;\ c \leqslant 0\}$, such that (u,y,p,λ) satisfies the following optimality conditions $(OCP\text{-}OPT)$:

$$(OCP\text{-}OPT): \qquad \begin{cases} (\nabla y, \nabla w) = (u+f, w), & \forall\, w \in H_0^1(\Omega); \\ (\nabla p, \nabla q) = (y - y_d, q) + (\lambda, q), & \forall\, q \in H_0^1(\Omega); \\ (\lambda, w - y) \leqslant 0, & \forall\, w \in K; \\ p + \alpha u = 0. \end{cases} \tag{2.3}$$

2.2. Finite element approximation

Let us consider the finite element approximation of the optimal control problem (OCP). We consider only the conforming n-simplex elements, which are widely used in engineering applications.

Assume that Ω is a polygonal domain. Let $T^h = \bigcup \tau$ be a quasi-regular triangulation of Ω with maximum mesh size $h := max_{\tau\in T^h}\{diam(\tau)\}$ and let $T^h_U = \bigcup \tau_U$ be another partitioning of Ω with maximum mesh size $h_U := \max_{\tau_U\in T^h_U}\{diam(\tau_U)\}$, in which each element has at most one face on $\partial\Omega$, and $\overline{\tau}$ and $\overline{\tau}'$ (or $\overline{\tau}_U$ and $\overline{\tau}'_U$) have either only one common vertex or a whole edge in 2-d case or face in 3-d case if $\overline{\tau}$ and $\overline{\tau}' \in T^h$ (or $\overline{\tau}_U$ and $\overline{\tau}'_U \in T^h_U$).

Associated with T^h is a finite dimensional subspace $V^h := \{w_h \in C(\Omega) : w_h|_\tau$ are polynomials of degree not exceeding r $(r \geqslant 1)$ for each $\tau \in T^h\} \cap H^1_0(\Omega)$. Associated with T^h_U is another finite dimensional subspace $U^h := \{v_h \in L^2(\Omega) : v_h|_{\tau_U}$ are polynomials of degree not exceeding m $(m \geqslant 0)$ for each $\tau_U \in T^h_U\} \subset U$.

In this paper, we only consider the simplest finite element spaces, i.e., $r = 1$ for V^h and $m = 0$ for U^h, which means that the piecewise linear conforming elements are used to approximate the state and co-state, and the piecewise constant elements are used to treat the control. We introduce the subspace $W^h \subset V^h$ of the form

$$W^h = \big\{w_h \in V^h;\ \exists z_h \in U^h \ \ such\ that \ \ (\nabla w_h, \nabla v_h) = (z_h, v_h), \ \ \forall\ v_h \in V^h\big\},$$

and define the discrete constraint set

$$K^h := \Big\{w_h \in W^h : \int_\Omega w_h \geqslant \gamma\Big\}. \tag{2.4}$$

The finite element approximation of the optimal control problem (OCP) reads:

$$(OCP)^h \qquad \begin{cases} \min \quad \mathcal{J}_h(u_h, y_h) = \dfrac{1}{2}\displaystyle\int_\Omega (y_h - y_d)^2 + \dfrac{\alpha}{2}\int_\Omega u_h^2, \\ s.t. \quad (\nabla y_h, \nabla w_h) = (u_h + f, w_h), \quad \forall\ w_h \in V^h, \quad y_h \in K^h. \end{cases} \tag{2.5}$$

It is proved in [20] that the finite element approximation $(OCP)^h$ has a unique solution $(u_h, y_h) \in U^h \times V^h$. Further, the pair (u_h, y_h) is the solution of $(OCP)^h$ if and only if there exists a unique pair $(p_h, \lambda_h) \in V^h \times \mathbb{R}^1_-$, such that $(u_h, y_h, p_h, \lambda_h)$ satisfies the following discrete optimality conditions $(OCP\text{-}OPT)^h$:

$$(OCP\text{-}OPT)^h: \qquad \begin{cases} (\nabla y_h, \nabla w_h) = (u_h + f, w_h), & \forall\ w_h \in V^h; \\ (\nabla p_h, \nabla q_h) = (y_h - y_d, q_h) + (\lambda_h, q_h), & \forall\ q_h \in V^h; \\ (\lambda_h, w_h - y_h) \leqslant 0, & \forall\ w_h \in K^h; \\ \alpha u_h = -\mathcal{P}_h(p_h). \end{cases} \tag{2.6}$$

where $\mathcal{P}_h$ is the L^2-projection from U onto its subspace U^h such that

$$(\mathcal{P}_h w, v_h) = (w, v_h), \qquad w \in U, \quad \forall\ v_h \in U^h. \tag{2.7}$$

Since U^h is the finite element space of piece-constant functions, we have

$$\mathcal{P}_h w|_{\tau_U} = \frac{1}{|\tau_U|}\int_{\tau_U} w, \ \ \text{in } \tau_U, \qquad \forall\ w \in U, \quad \forall\ \tau_U \in T^h_U.$$

In the following sections, we will give the a posteriori error estimator of the finite element approximation $(OCP\text{-}OPT)^h$.

3. Equivalent a Posteriori Error Estimator

Adaptive finite element approximations have been found very useful in computing optimal control problems, as mentioned in the introduction. They use a posteriori error indicators to guide mesh refinement procedures. An adaptive finite element approximation refines only the area where the error indicator is larger, so that a higher density of nodes is distributed over the area where the solution is difficult to be approximated. In this section, we will derive residual-type a posteriori error estimator for our control problem.

Introduce a posteriori error estimator η defined as follows

$$\eta^2 = \eta_U^2 + \eta_V^2, \tag{3.1}$$

where

$$\eta_U^2 = \eta_u^2, \quad \eta_V^2 = \eta_y^2 + \eta_p^2. \tag{3.2}$$

A posteriori error estimators η_U and η_V, with respect to the meshes T_U^h and T^h, are defined by

$$\eta_u^2 := \sum_{\tau_U \in T_U^h} \eta_u^2|_{\tau_U} \quad \text{with} \quad \eta_u^2|_{\tau_U} = \int_{\tau_U} (\mathcal{P}_h p_h - p_h)^2, \tag{3.3}$$

where $\mathcal{P}_h$ is the L^2-projection described in (2.7),

$$\eta_y^2 = \sum_{\tau \in T^h} \eta_y^2|_\tau + \sum_{l \cap \partial\Omega = \emptyset} \eta_y^2|_l, \tag{3.4}$$

$$\text{with} \quad \eta_y^2|_\tau = h_\tau^2 \int_\tau (u_h + f)^2, \quad \eta_y^2|_l = h_l \int_l [\nabla y_h \cdot n]^2, \tag{3.5}$$

and

$$\eta_p^2 = \sum_{\tau \in T^h} \eta_p^2|_\tau + \sum_{l \cap \partial\Omega = \emptyset} \eta_p^2|_l \tag{3.6}$$

$$\text{with} \quad \eta_p^2|_\tau = h_\tau^2 \int_\tau (y_h - y_d + \lambda_h)^2, \quad \eta_p^2|_l = h_l \int_l [\nabla p_h \cdot n]^2. \tag{3.7}$$

Here l is a face of an element τ, n is the unit normal vector on $l := \overline{\tau}_l^1 \cap \overline{\tau}_l^2$ outwards τ_l^1, and h_l is the maximum diameter of the face l. $[\nabla y_h \cdot n]$ and $[\nabla p_h \cdot n]$ are the jumps of the normal derivatives across the interior face l, defined by

$$[\nabla p_h \cdot n]_l := (\nabla p_h|_{\tau_l^1} - \nabla p_h|_{\tau_l^2}) \cdot n,$$
$$[\nabla y_h \cdot n]_l := (\nabla y_h|_{\tau_l^1} - \nabla p_h|_{\tau_l^2}) \cdot n.$$

The element residuals are weighted element-wise L^2 norms residuals with respect to the strong form of the state and co-state equation respectively. The face residuals are weighted L^2 norms of the jumps $[\nabla y_h \cdot n]$ and $[\nabla p_h \cdot n]$ of the normal derivatives across the interior faces.

3.1. Upper bound estimate

In this subsection, ϵ and C denote some general positive constants independent of grid parameters h or h_U. First we show that a posteriori error estimator η is **Reliable** in Theorem 3.1.

Theorem 3.1. *Let* (u, y, p, λ) *and* $(u_h, y_h, p_h, \lambda_h)$ *be the solutions of optimality conditions* (*OCP-OPT*) *and* $(OCP\text{-}OPT)^h$, *respectively. Then the following estimate*

$$\|u - u_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{H^1(\Omega)} + \|p - p_h\|^2_{H^1(\Omega)} + |\lambda - \lambda_h|^2 \leqslant C\eta^2 \tag{3.8}$$

holds, where η *is defined in* (3.1)-(3.2)

The proof of Theorem 3.1 is given step by step. As usual, we recall some useful lemmas.

Lemma 3.1. ([8]) *Let* π_h *be the standard Lagrange interpolation operator. For* $r = 0$ *or* 1, $1 < q \leqslant \infty$ *and* $v \in W^{2,q}(\Omega)$,

$$|v - \pi_h v|_{W^{r,q}(\Omega)} \leqslant Ch^{2-r}|v|_{W^{2,q}(\Omega)}. \tag{3.9}$$

Lemma 3.2. *Let* $\hat{\pi}_h$ *be the average interpolation operator defined in* [21]. *For* $r = 0$ *or* 1, $1 < q \leqslant \infty$ *and* $v \in W^{1,q}(\Omega)$,

$$|v - \hat{\pi}_h v|_{W^{r,q}(\tau)} \leqslant \sum_{\bar{\tau}' \cap \bar{\tau} \neq \emptyset} Ch_\tau^{1-r}|v|_{W^{1,q}(\tau')}. \tag{3.10}$$

Lemma 3.3. ([11]) *For each* $v \in W^{1,q}(\Omega), 1 \leqslant q < \infty$,

$$\|v\|_{W^{0,q}(\partial\tau)} \leqslant C\Big\{h_\tau^{-\frac{1}{q}}\|v\|_{W^{0,q}(\tau)} + h_\tau^{1-\frac{1}{q}}|v|_{W^{1,q}(\tau)}\Big\}. \tag{3.11}$$

Lemma 3.4. *Let* τ *be one element in* T^h. *For any given constant* A, *there exists the piecewise polynomial* $\omega_\tau \in H^1_0(\tau)$ *satisfying*

$$\int_\tau A\omega_\tau = h_\tau^2\int_\tau A^2, \text{ such that } \quad |\omega_\tau|^2_{m,\tau} \leqslant Ch_\tau^{2(1-m)+2}\int_\tau A^2, \qquad m = 0, 1. \tag{3.12}$$

Let τ_l^1 *and* τ_l^2 *be two elements in* T^h *sharing the edge* $l := \bar{\tau}_l^1 \cap \bar{\tau}_l^2$. *For any given constant* B, *there exists the piecewise polynomial* $\omega_l \in H^1_0(\tau_l^1 \cup \tau_l^2)$ *satisfying*

$$\int_l B\omega_l = h_l\int_l B^2, \text{ such that } \quad |\omega_l|^2_{m,\tau_l^1\cup\tau_l^2} \leqslant Ch_l^{2(1-m)+1}\int_l B^2, \qquad m = 0, 1. \tag{3.13}$$

Proof. Let $\tau \in T^h$ and $\{\lambda_i\}_{i=1}^3$ be the areal coordinates on the element τ. Construct

$$\omega_\tau = \alpha_\tau\lambda_1\lambda_2\lambda_3, \text{ where } \alpha_\tau = \frac{h_\tau^2\int_\tau A}{\int_\tau \lambda_1\lambda_2\lambda_3}.$$

It is clear that $\omega_\tau \in H^1_0(\tau)$. Let τ_l^1 and τ_l^2 be two elements in T^h sharing the edge $l := \bar{\tau}_l^1 \cap \bar{\tau}_l^2$ with end points $\{P_l^i\}_{i=1}^2$, and λ_i are areal coordinates with respect to P_l^i for $i = 1, 2$. Construct

$$\omega_l = \beta_l\lambda_1\lambda_2, \text{ where } \beta_l = \frac{h_l\int_l B}{\int_l \lambda_1\lambda_2}.$$

It is also clear that $\omega_l \in H^1_0(\bar{\tau}^1 \cup \bar{\tau}^2)$. One can check that all the conclusions in this lemma hold. For the detailed proof, we refer to [23]. □

We introduce an auxiliary system defined as: *seeking* $(y(u_h), p(u_h)) \in V \times V$ *such that*

$$(\nabla y(u_h), \nabla w) = (u_h + f, w), \qquad \forall\, w \in H^1_0(\Omega), \tag{3.14a}$$

$$(\nabla q, \nabla p(u_h)) = (y(u_h) - y_d + \lambda_h, q), \quad \forall\, q \in H^1_0(\Omega). \tag{3.14b}$$

Lemma 3.5. *Let (u, y, p, λ) and $(u_h, y_h, p_h, \lambda_h)$ be the solutions of optimality conditions (OCP-OPT) and (OCP-OPT)^h respectively. If $p(u_h)$ is the solution of the auxiliary system (3.14b), then there holds the following estimate:*

$$\|p - p(u_h)\|_{H^1(\Omega)} \leqslant C\Big\{\|u - u_h\|_{L^2(\Omega)} + \|p_h - p(u_h)\|_{L^2(\Omega)} + \|p_h - \mathcal{P}_h p_h\|_{L^2(\Omega)}\Big\}. \tag{3.15}$$

Proof. Let $\bar{v}$ be the average value of v over Ω given by

$$\bar{v} = \frac{1}{|\Omega|}\int_\Omega v. \tag{3.16}$$

Choose $\varphi \in C_0^\infty(\Omega)$ such that $\bar{\varphi} = 1$ and $\|\varphi\|_{H^1(\Omega)} \leqslant C$. Taking $\widetilde{C} = \bar{p} - \bar{p}(u_h)$ and noting $\widetilde{C}\varphi \in C_0^\infty(\Omega) \subset H_0^1(\Omega)$, we obtain

$$(\nabla(p - p(u_h) - \widetilde{C}\varphi), \nabla(p - p(u_h))) = (y - y(u_h) + \lambda - \lambda_h, p - p(u_h) - \widetilde{C}\varphi).$$

Since $\int_\Omega (p - p(u_h) - \widetilde{C}\varphi) = 0$ and $\lambda - \lambda_h$ is a constant, we have

$$(\nabla(p - p(u_h)), \nabla(p - p(u_h))) = \widetilde{C}(\nabla\varphi, \nabla(p - p(u_h))) + (y - y(u_h), p - p(u_h) - \widetilde{C}\varphi).$$

Therefore, we have

$$\begin{aligned}
&\|\nabla(p - p(u_h))\|^2_{L^2(\Omega)}\\
&\leqslant \frac{1}{2}\|\nabla(p - p(u_h))\|^2_{L^2(\Omega)} + C\Big[\widetilde{C}^2\|\varphi\|^2_{H^1(\Omega)} + \|y - y(u_h)\|^2_{L^2(\Omega)} + \|p - p(u_h)\|^2_{L^2(\Omega)}\Big],
\end{aligned}$$

which implies that

$$\|\nabla(p - p(u_h))\|^2_{L^2(\Omega)} \leqslant C\Big\{\widetilde{C}^2 + \|u - u_h\|^2_{L^2(\Omega)} + \|p - p(u_h)\|^2_{L^2(\Omega)}\Big\}. \tag{3.17}$$

Since Ω is a bounded domain, we have

$$\begin{aligned}
|\widetilde{C}| &= \frac{1}{|\Omega|}\Big|\int_\Omega (p - p(u_h))\Big|\\
&\leqslant \frac{1}{|\Omega|}\Big(\Big|\int_\Omega (-\alpha u + \alpha u_h)\Big| + \Big|\int_\Omega p_h - p(u_h)\Big| + \Big|\int_\Omega (\mathcal{P}_h p_h - p_h)\Big| \Big)\\
&\leqslant C\Big\{\|u - u_h\|_{L^2(\Omega)} + \|p_h - p(u_h)\|_{L^2(\Omega)} + \|\mathcal{P}_h p_h - p_h\|_{L^2(\Omega)}\Big\}.
\end{aligned} \tag{3.18}$$

Then the result (3.15) follows directly from (3.17) and (3.18). □

Lemma 3.6. *Let (u, y, p, λ) and $(u_h, y_h, p_h, \lambda_h)$ be the solutions of optimality conditions (OCP-OPT) and (OCP-OPT)^h, respectively. Then there holds the following estimate:*

$$|\lambda - \lambda_h| \leqslant C\Big\{\|u - u_h\|_{L^2(\Omega)} + \|p_h - p(u_h)\|_{L^2(\Omega)} + \|p_h - \mathcal{P}_h p_h\|_{L^2(\Omega)}\Big\}. \tag{3.19}$$

Proof. Obviously, it follows from (2.3) and (3.14b) that $\lambda - \lambda_h$ satisfies the following equation

$$(\lambda - \lambda_h, q) = (\nabla(p - p(u_h)), \nabla q) - (y(u_h) - y, q), \qquad \forall\, q \in H_0^1(\Omega),$$

which implies that

$$|\lambda - \lambda_h| \leqslant C\Big(\|p - p(u_h)\|_{H^1(\Omega)} + \|y(u_h) - y\|_{L^2(\Omega)}\Big). \tag{3.20}$$

Since $\|y(u_h) - y\|_{L^2(\Omega)} \leqslant C\|u - u_h\|_{L^2(\Omega)}$, (3.19) follows directly from(3.20) and Lemma 3.5. □

Lemma 3.7. *Let (u, y, p, λ) and $(u_h, y_h, p_h, \lambda_h)$ be the solutions of the optimality conditions (OCP-OPT) and (OCP-OPT)^h, respectively. There holds the following error estimate:*

$$\|u-u_h\|_{L^2(\Omega)} \leqslant C\Big\{\|y_h - y(u_h)\|_{L^2(\Omega)} + \|p_h - p(u_h)\|_{L^2(\Omega)} + \|p_h - \mathcal{P}_h p_h\|_{L^2(\Omega)}\Big\}. \quad (3.21)$$

Proof. By using (2.3), we have

$$\begin{aligned}
\mathcal{J}'(u)(u-u_h) &= (y-y_d, y'(u)(u-u_h)) + (\alpha u, u-u_h)\\
&= (y-y_d+\lambda, y'(u)(u-u_h)) + (\alpha u, u-u_h) - (\lambda, y'(u)(u-u_h))\\
&= (u-u_h, p) + (\alpha u, u-u_h) - (\lambda, y(u)-y(u_h))\\
&= -(\lambda, y(u)-y(u_h)).
\end{aligned}$$

Similarly,

$$\mathcal{J}'(u_h)(u-u_h) = (p(u_h)+\alpha u_h, u-u_h) - (\lambda_h, y-y(u_h)).$$

Hence we have

$$\begin{aligned}
&\|y-y(u_h)\|^2_{L^2(\Omega)} + \|u-u_h\|^2_{L^2(\Omega)}\\
&= \mathcal{J}'(u)(u-u_h) - \mathcal{J}'(u_h)(u-u_h)\\
&= -(p(u_h)+\alpha u_h, u-u_h) - (\lambda-\lambda_h, y-y(u_h))\\
&= (p_h - p(u_h), u-u_h) - (\lambda-\lambda_h, y-y_h) - (\lambda-\lambda_h, y_h-y(u_h)) + (\mathcal{P}_h p_h - p_h, u-u_h)\\
&\leqslant (p_h - p(u_h), u-u_h) - (\lambda-\lambda_h, y_h-y(u_h)) + (\mathcal{P}_h p_h - p_h, u-u_h)\\
&\leqslant \epsilon\Big\{\|u-u_h\|^2_{L^2(\Omega)} + |\lambda-\lambda_h|^2\Big\}\\
&\quad + C(\epsilon)\Big\{\|p_h-p(u_h)\|^2_{L^2(\Omega)} + \|y_h-y(u_h)\|^2_{L^2(\Omega)} + \|\mathcal{P}_h p_h - p_h\|^2_{L^2(\Omega)}\Big\}\\
&\leqslant C\epsilon\Big\{\|u-u_h\|^2_{L^2(\Omega)} + \|p_h-p(u_h)\|^2_{L^2(\Omega)}\Big\}\\
&\quad + C(\epsilon)\Big\{\|p_h-p(u_h)\|^2_{L^2(\Omega)} + \|y_h-y(u_h)\|^2_{L^2(\Omega)} + \|\mathcal{P}_h p_h - p_h\|^2_{L^2(\Omega)}\Big\},
\end{aligned}$$

which implies that

$$\|u-u_h\|_{L^2(\Omega)} \leqslant C\Big\{\|y_h - y(u_h)\|_{L^2(\Omega)} + \|p_h - p(u_h)\|_{L^2(\Omega)} + \|p_h - \mathcal{P}_h p_h\|_{L^2(\Omega)}\Big\}. \quad (3.22)$$

Here we have used $(\lambda-\lambda_h, y-y_h) = (\lambda-\lambda_h)(\bar{y}-\bar{y}_h)|\Omega| \geqslant 0$ due to the following facts

$$\lambda(\bar{y}-\bar{y}_h)|\Omega| = \begin{cases} 0, & \bar{y}|\Omega| > \gamma \quad \text{such that } \lambda = 0;\\ \lambda(\gamma - \bar{y}_h|\Omega|) \geqslant 0, & \bar{y}|\Omega| = \gamma \quad \text{such that } \lambda \leqslant 0 \ \textit{and} \ \gamma - \bar{y}_h|\Omega| \leqslant 0; \end{cases}$$

and

$$-\lambda_h(\bar{y}-\bar{y}_h)|\Omega| = \begin{cases} 0, & \bar{y}_h|\Omega| > \gamma \quad \text{such that } \lambda_h = 0;\\ -\lambda_h(\bar{y}|\Omega| - \gamma) \geqslant 0, & \bar{y}_h|\Omega| = \gamma \quad \text{such that } \lambda_h \leqslant 0 \ \textit{and} \ \bar{y}|\Omega| - \gamma \geqslant 0. \end{cases}$$

The proof of Lemma 3.7 is then complete. □

Summing up results of the above lemmas, we have the following lemma.

Lemma 3.8. *Let (u,y,p,λ) and (u_h,y_h,p_h,λ_h) be the solutions of the optimality conditions (OCP-OPT) and $(OCP\text{-}OPT)^h$ respectively. If $(p(u_h),y(u_h))$ is the solution of the auxiliary system,* (3.14*a*)-(3.14*b*)*, then there holds the following estimate:*

$$\begin{aligned}&\|u-u_h\|_{L^2(\Omega)}+\|y-y_h\|_{H^1(\Omega)}+\|p-p_h\|_{H^1(\Omega)}+|\lambda-\lambda_h|\\&\leqslant C\Big\{\|y_h-y(u_h)\|_{H^1(\Omega)}+\|p_h-p(u_h)\|_{H^1(\Omega)}+\|p_h-\mathcal{P}_hp_h\|_{L^2(\Omega)}\Big\}.\end{aligned}\tag{3.23}$$

Proof. It is clear that

$$\begin{aligned}&\|y-y_h\|_{H^1(\Omega)}+\|p-p_h\|_{H^1(\Omega)}\\&\leqslant\|y-y(u_h)\|_{H^1(\Omega)}+\|y(u_h)-y_h\|_{H^1(\Omega)}+\|p-p(u_h)\|_{H^1(\Omega)}+\|p(u_h)-p_h\|_{H^1(\Omega)}\\&\leqslant C\Big\{\|u-u_h\|_{L^2(\Omega)}+\|y(u_h)-y_h\|_{H^1(\Omega)}+\|p-p(u_h)\|_{H^1(\Omega)}+\|p(u_h)-p_h\|_{H^1(\Omega)}\Big\}.\end{aligned}$$

Applying Lemmas 3.5-3.7 to this inequality leads to (3.23). □

Now, we can prove Theorem 3.1.

Proof of Theorem 3.1. Noting $\|p_h-\mathcal{P}_hp_h\|_{L^2(\Omega)}=\eta_u$, we only need to estimate $\|p_h-p(u_h)\|_{H^1(\Omega)}$ and $\|y_h-y(u_h)\|_{H^1(\Omega)}$.

Firstly we need some notations. Let $E^p=p_h-p(u_h)$ and $E^p_I=\hat{\pi}_hE^p$, where $\hat{\pi}_h$ is the average interpolator defined in Lemma 3.2. Applying the standard residual technique, we have

$$\begin{aligned}&c\|E^p\|^2_{H^1(\Omega)}\leqslant(\nabla E^p,\nabla E^p)\\&=(\nabla E^p,\nabla(E^p-E^p_I))+(y(u_h)-y_h,E^p_I)\\&=\sum_{\tau\in T^h}\int_\tau-\Delta(p(u_h)-p_h)(E^p-E^p_I)-\sum_{\tau\in T^h}\int_l(\nabla p_h\cdot n)(E^p-E^p_I)+(y(u_h)-y_h,E^p_I)\\&=\sum_{\tau\in T^h}\int_\tau(y(u_h)-y_d+\lambda_h)(E^p-E^p_I)-\sum_{l\cap\partial\Omega=\emptyset}\int_l[\nabla p_h\cdot n](E^p-E^p_I)+(y(u_h)-y_h,E^p_I)\\&=\sum_{\tau\in T^h}\int_\tau(y_h-y_d+\lambda_h)(E^p-E^p_I)-\sum_{l\cap\partial\Omega=\emptyset}\int_l[\nabla p_h\cdot n](E^p-E^p_I)+(y(u_h)-y_h,E^p)\\&\leqslant C(\epsilon)\sum_{\tau\in T^h}h^2_\tau\int_\tau(y_h-y_d+\lambda_h)^2+C(\epsilon)\sum_{l\cap\partial\Omega=\emptyset}h_l\int_l[\nabla p_h\cdot n]^2+C(\epsilon)\|y(u_h)-y_h\|^2_{L^2(\Omega)}\\&\quad+\epsilon\|E^p\|^2_{L^2(\Omega)}+\epsilon\sum_{\tau\in T^h}h^{-2}_\tau\int_\tau(E^p-E^p_I)^2+\epsilon\sum_{\tau\in T^h}\int_\tau(\nabla(E^p-E^p_I))^2\\&\leqslant C\sum_{\tau\in T^h}h^2_\tau\int_\tau(y_h-y_d+\lambda_h)^2+C\sum_{l\cap\partial\Omega=\emptyset}h_l\int_l[\nabla p_h\cdot n]^2\\&\quad+C\|y(u_h)-y_h\|^2_{L^2(\Omega)}+C\epsilon\|E^p\|^2_{H^1(\Omega)},\end{aligned}$$

where we have used results in Lemmas 3.1-3.3. Hence we obtain

$$\begin{aligned}&\|E^p\|^2_{H^1(\Omega)}\\&\leqslant C\Big\{\sum_{\tau\in T^h}h^2_\tau\int_\tau(y_h-y_d+\lambda_h)^2+\sum_{l\cap\partial\Omega=\emptyset}h_l\int_l[\nabla p_h\cdot n]^2+\|y(u_h)-y_h\|^2_{L^2(\Omega)}\Big\}.\end{aligned}\tag{3.24}$$

Similarly, letting $E^y = y_h - y(u_h)$ and $E_I^y = \hat{\pi}_h E^y$, we have

$$\begin{aligned}
& c\|E^y\|^2_{H^1(\Omega)} \\
& \leqslant (\nabla E^y, \nabla E^y) = (\nabla E^y, \nabla(E^y - E_I^y)) \\
& = \sum_{\tau\in T^h} \int_\tau -\Delta(y(u_h) - y_h))(E^y - E_I^y) - \sum_{\tau\in T^h} \int_l (\nabla y_h \cdot n)(E^y - E_I^y) \\
& = \sum_{\tau\in T^h} \int_\tau (u_h + f)(E^y - E_I^y) - \sum_{l\cap\partial\Omega=\emptyset} \int_l [\nabla y_h \cdot n](E^y - E_I^y) \\
& \leqslant C(\epsilon) \sum_{\tau\in T^h} h_\tau^2 \int_\tau (u_h + f)^2 + C(\epsilon) \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [\nabla y_h \cdot n]^2 \\
& \quad + C\epsilon \sum_{\tau\in T^h} h_\tau^{-2} \int_\tau |E^y - E_I^y|^2 + C\epsilon \sum_{\tau\in T^h} \int_\tau |\nabla(E^y - E_I^y)|^2 \\
& \leqslant C \sum_{\tau\in T^h} h_\tau^2 \int_\tau (u_h + f)^2 + C \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [\nabla y_h \cdot n]^2 + C\epsilon\|E^y\|^2_{H^1(\Omega)},
\end{aligned}$$

where we have used Lemmas 3.1-3.3. Hence

$$\|E^y\|^2_{H^1(\Omega)} \leqslant C \sum_{\tau\in T^h} h_\tau^2 \int_\tau (u_h + f)^2 + C \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [\nabla y_h \cdot n]^2. \tag{3.25}$$

Due to estimates (3.24) and (3.25), we obtain

$$\begin{aligned}
& \|y_h - y(u_h)\|^2_{H^1(\Omega)} + \|p_h - p(u_h)\|^2_{H^1(\Omega)} \\
= \quad & \|E^y\|^2_{H^1(\Omega)} + \|E^p\|^2_{H^1(\Omega)} \leqslant C\Big\{\eta_y^2 + \eta_p^2\Big\}.
\end{aligned} \tag{3.26}$$

Hence due to Lemma 3.8, we have obtained the upper bound (3.8). □

3.2. Lower bound estimate

In this subsection, we will show the a posteriori error estimator also provides a lower bound for the discretization errors in the state, co-state, control and multiplier; that is to say that the estimator η is **Efficient**:

$$\eta^2 - c\sigma^2 \leqslant \|u - u_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{H^1(\Omega)} + \|p - p_h\|^2_{H^1(\Omega)} + |\lambda - \lambda_h|^2,$$

where σ is a higher order term.

Theorem 3.2. *Let (u, y, p, λ) and $(u_h, y_h, p_h, \lambda_h)$ be the solutions of the optimal conditions (OCP-OPT) and $(OCP\text{-}OPT)^h$, respectively. Then there holds the estimate:*

$$\begin{aligned}
& \eta^2 - c(\sigma_1^2 + \sigma_2^2) \\
& \leqslant C\Big\{\|u - u_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{H^1(\Omega)} + \|p - p_h\|^2_{H^1(\Omega)} + |\lambda - \lambda_h|^2\Big\},
\end{aligned} \tag{3.27}$$

where η is defined in (3.1)*, σ_1 and σ_2 are defined by*

$$\begin{aligned}
\sigma_1^2 &:= \sum_{\tau\in T^h} \int_\tau h_\tau^2 (\bar{y}_h|_\tau - y_h + \bar{y}_d|_\tau - y_d)^2, \\
\sigma_2^2 &:= \sum_{\tau\in T^h} \int_\tau h_\tau^2 (\bar{u}_h|_\tau - u_h + \bar{f}|_\tau - f)^2.
\end{aligned}$$

Here $\bar{v}|_\tau$ is the averaging of v over the element τ.

The proof of Theorem 3.2 is divided into the following three lemmas.

Lemma 3.9. *There holds the following error estimate:*

$$\eta_u^2 \leqslant \|p-p_h\|^2_{L^2(\Omega)} + \|u-u_h\|^2_{L^2(\Omega)}. \tag{3.28}$$

Proof. It is clear that

$$\begin{aligned}&\sum_{\tau\in T_U^h}\int_\tau (\mathcal{P}_h p_h - p_h)^2 = \|\mathcal{P}_h p_h - p_h\|^2_{L^2(\Omega)} = \|\alpha u_h + p_h\|^2_{L^2(\Omega)}\\ &\leqslant \|p-p_h\|^2_{L^2(\Omega)} + \alpha^2\|u-u_h\|^2_{L^2(\Omega)},\end{aligned}$$

which is (3.28). □

In order to derive the lower bound, we use the standard bubble function technique (see, e.g., [23]). As in Lemma 3.4, there exist piecewise polynomials $\omega_\tau \in H_0^1(\tau)$ and $\omega_l \in H_0^1(\tau_l^1\cup\tau_l^2)$ such that

$$\begin{aligned}&\int_\tau \overline{(y_h-y_d+\lambda_h)}|_\tau \omega_\tau = h_\tau^2\int_\tau \overline{(y_h-y_d+\lambda_h)}|_\tau^2,\\ &|\omega_\tau|^2_{m,\tau} \leqslant Ch_\tau^{2(1-m)+2}\int_\tau \overline{(y_h-y_d+\lambda_h)}|_\tau^2, \qquad m=0,1,\end{aligned} \tag{3.29}$$

and

$$\begin{aligned}&\int_l [\nabla p_h\cdot n]\omega_l = h_l\int_l [\nabla p_h\cdot n]^2,\\ &|\omega_l|^2_{m,\tau_l^1\cup\tau_l^2} \leqslant Ch_l^{2(1-m)+1}\int_l [\nabla p_h\cdot n]^2, \quad m=0,1.\end{aligned} \tag{3.30}$$

We will use these bubble functions and their properties.

Lemma 3.10. *There holds the following error estimate:*

$$\eta_p^2 \leqslant C\Big\{\|p_h-p\|^2_{H^1(\Omega)} + \|y_h-y\|^2_{L^2(\Omega)} + |\lambda_h-\lambda|^2 + \sigma_1^2\Big\}. \tag{3.31}$$

Proof. By using the standard bubble function technique, which infers (3.29) and (3.30), and noting that p_h is the piecewise linear element, we have

$$\begin{aligned}&\int_\tau h_\tau^2(y_h-y_d+\lambda_h)^2\\ &\leqslant C\int_\tau h_\tau^2\big(\overline{(y_h-y_d+\lambda_h)}|_\tau\big)^2 + C\int_\tau h_\tau^2(y_h-y_d+\lambda_h-\overline{(y_h-y_d+\lambda_h)}|_\tau)^2\\ &= C\int_\tau \overline{(y_h-y_d+\lambda_h)}|_\tau\omega_\tau + C\int_\tau h_\tau^2(y_h-y_d-\overline{(y_h-y_d)}|_\tau)^2\\ &= C\Big\{\int_\tau (\Delta p_h + y_h - y_d + \lambda_h - \Delta p - y + y_d - \lambda)\omega_\tau\\ &\quad + \int_\tau ((\overline{y_h-y_d+\lambda_h})|_\tau - (y_h-y_d+\lambda_h))\omega_\tau + \int_\tau h_\tau^2(y_h-y_d-(\overline{y_h-y_d})|_\tau)^2\Big\}\\ &\leqslant -C\int_\tau \nabla(p_h-p)\nabla\omega_\tau + C\int_\tau (y_h-y+\lambda_h-\lambda)\omega_\tau\\ &\quad + C\int_\tau ((\overline{y_h-y_d})|_\tau - (y_h-y_d))\omega_\tau + C\int_\tau h_\tau^2(y_h-y_d-(\overline{y_h-y_d})|_\tau)^2.\end{aligned} \tag{3.32}$$

Applying the Cauchy-Schwartz inequality, we have

$$
\begin{aligned}
&-\int_\tau \nabla(p_h-p)\nabla\omega_\tau+\int_\tau(y_h-y+\lambda_h-\lambda)\omega_\tau\\
&\leqslant C(\epsilon)\Big\{\|\nabla(p_h-p)\|^2_{L^2(\tau)}+\|y_h-y\|^2_{L^2(\tau)}+\|\lambda_h-\lambda\|^2_{L^2(\tau)}\Big\}+\epsilon\|\omega_\tau\|^2_{H^1(\tau)}
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_\tau((\overline{y_h-y_d})|_\tau-(y_h-y_d))\omega_\tau\\
&\leqslant C(\epsilon)\int_\tau h_\tau^2((\overline{y_h-y_d})|_\tau-(y_h-y_d))^2+\epsilon h_\tau^{-2}\|\omega_\tau\|^2_{L^2(\tau)}.
\end{aligned}
$$

Therefore, by using the estimate of bubble function (3.29), we obtain

$$
\begin{aligned}
&\int_\tau h_\tau^2(y_h-y_d+\lambda_h)^2\\
&\leqslant C\Big\{\|p_h-p\|^2_{H^1(\tau)}+\|y_h-y\|^2_{L^2(\tau)}+\|\lambda_h-\lambda\|^2_{L^2(\tau)}\Big\}+\epsilon\|\omega_\tau\|^2_{H^1(\tau)}\\
&+C\int_\tau h_\tau^2((\overline{y_h-y_d})|_\tau-(y_h-y_d))^2+\epsilon h_\tau^{-2}\|\omega_\tau\|^2_{L^2(\tau)}\\
&\leqslant C\Big\{\|p_h-p\|^2_{H^1(\tau)}+\|y_h-y\|^2_{L^2(\tau)}+\|\lambda_h-\lambda\|^2_{L^2(\tau)}\Big\}\\
&+C\int_\tau h_\tau^2((\overline{y_h-y_d})|_\tau-(y_h-y_d))^2+\epsilon\int_\tau h_\tau^2((\overline{y_h-y_d+\lambda_h})|_\tau)^2.
\end{aligned}
$$

Hence we have the estimate

$$
\begin{aligned}
&\int_\tau h_\tau^2(y_h-y_d+\lambda_h)^2\\
&\leqslant C\Big\{\|p_h-p\|^2_{H^1(\tau)}+\|y_h-y\|^2_{L^2(\tau)}+\|\lambda_h-\lambda\|^2_{L^2(\tau)}\Big\}\\
&+C\int_\tau h_\tau^2((\overline{y_h-y_d})|_\tau-(y_h-y_d))^2.
\end{aligned}\tag{3.33}
$$

On the other hand, for $l\cap\partial\Omega\neq\emptyset$, ($l:=\overline{\tau}_l^1\cap\overline{\tau}_l^2$), by using the bubble function technique, we obtain

$$
\begin{aligned}
&\int_l h_l[\nabla p_h\cdot n]^2=\int_l[\nabla p_h\cdot n]\omega_l=\int_l[\nabla p_h\cdot n-\nabla p\cdot n]\omega_l\\
&=\int_{\partial\tau_l^1\cup\partial\tau_l^2}(\nabla p_h\cdot n-\nabla p\cdot n)\omega_l\\
&=\int_{\tau_l^1\cup\tau_l^2}(\nabla p_h-\nabla p)\nabla\omega_l+\int_{\tau_l^1\cup\tau_l^2}div(\nabla p_h-\nabla p)\omega_l\\
&=\int_{\tau_l^1\cup\tau_l^2}(\nabla p_h-\nabla p)\nabla\omega_l+\int_{\tau_l^1\cup\tau_l^2}(div\nabla p_h+y-y_d+\lambda)\omega_l\\
&=\int_{\tau_l^1\cup\tau_l^2}(\nabla p_h-\nabla p)\nabla\omega_l+\int_{\tau_l^1\cup\tau_l^2}(y_h-y_d+\lambda_h)\omega_l+\int_{\tau_l^1\cup\tau_l^2}(y-y_h+\lambda-\lambda_h)\omega_l\\
&\leqslant C(\epsilon)\Big\{\|\nabla(p_h-p)\|^2_{L^2(\tau_l^1\cup\tau_l^2)}+\|y_h-y\|^2_{L^2(\tau_l^1\cup\tau_l^2)}+\|\lambda_h-\lambda\|^2_{L^2(\tau_l^1\cup\tau_l^2)}\Big\}\\
&\quad+\epsilon\|\omega_\tau\|^2_{H^1(\tau_l^1\cup\tau_l^2)}+\epsilon h_l^{-2}\|\omega_\tau\|^2_{L^2(\tau_l^1\cup\tau_l^2)}+C(\epsilon)\int_{\tau_l^1\cup\tau_l^2}h_l^2(y_h-y_d+\lambda_h)^2\\
&\leqslant C(\epsilon)\Big\{\|\nabla(p_h-p)\|^2_{L^2(\tau_l^1\cup\tau_l^2)}+\|y_h-y\|^2_{L^2(\tau_l^1\cup\tau_l^2)}+\|\lambda_h-\lambda\|^2_{L^2(\tau_l^1\cup\tau_l^2)}\Big\}\\
&\quad+\epsilon\int_l h_l[\nabla p_h\cdot n]^2+C(\epsilon)\int_{\tau_l^1\cup\tau_l^2}h_l^2(y_h-y_d+\lambda_h)^2,
\end{aligned}\tag{3.34}
$$

where we have used the estimate (3.30). Therefore, combining (3.33) and (3.34), gives

$$\sum_{\tau\in T^h} h_\tau^2 \int_\tau (y_h - y_d + \lambda_h)^2 + C \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [\nabla p_h\cdot n]^2$$
$$\leqslant C\Big\{\|p_h - p\|_{H^1(\Omega)}^2 + \|y_h - y\|_{L^2(\Omega)}^2 + |\lambda_h - \lambda|^2 + \sum_{\tau\in T^h}\int_\tau h_\tau^2((\overline{y_h - y_d})|_\tau - (y_h - y_d))^2\Big\}.$$

This ends the proof of Lemma 3.10. □

Similarly, we have the following estimate for η_y^2.

Lemma 3.11. *There holds the following error estimate:*

$$\eta_y^2 \leqslant \Big\{\|y - y_h\|_{H^1(\Omega)}^2 + \|u - u_h\|_{L^2(\Omega)}^2 + \sigma_2^2\Big\}. \tag{3.35}$$

Remark 3.1. Actually, σ_1 and σ_2 are higher terms. Due to the Poincare inequality, we have

$$\sigma_1^2 = \sum_{\tau\in T^h}\int_\tau h_\tau^2(\bar{y}_h|_\tau - y_h + \bar{y}_d|_\tau - y_d)^2 \leqslant C\sum_{\tau\in T^h} h_\tau^4\|y_h - y_d\|_{1,\tau}^2$$
$$\leqslant Ch^4\|y_h - y_d\|_{1,\Omega}^2 = \mathcal{O}(h^4).$$

On the other hand, since $\alpha u_h = -\mathcal{P}_h p_h$ such that

$$\sum_{\tau\in T^h}\int_\tau h_\tau^2(\bar{u}_h|_\tau - u_h)^2 = \alpha^{-2}\sum_{\tau\in T^h}\int_\tau h_\tau^2(\overline{\mathcal{P}_h p_h}|_\tau - \mathcal{P}_h p_h)^2$$
$$\leqslant C\sum_{\tau\in T^h} h_\tau^2\Big[\int_\tau \overline{(\mathcal{P}_h p_h - p_h)}|_\tau^2 + \int_\tau(\bar{p}_h|_\tau - p_h)^2 + \int_\tau(\mathcal{P}_h p_h - p_h)^2\Big]$$
$$\leqslant Ch^2\big[\|\mathcal{P}_h p_h - p_h\|_{L^2(\Omega)}^2 + h^2\|p_h\|_{H^1(\Omega)}^2\big] \leqslant Ch^2(h^2 + h_U^2)\|p_h\|_{H^1(\Omega)}^2,$$

we have

$$\sigma_2^2 = \sum_{\tau\in T^h}\int_\tau h_\tau^2(\bar{u}_h|_\tau - u_h + \bar{f}|_\tau - f)^2$$
$$\leqslant 2\Big[\sum_{\tau\in T^h}\int_\tau h_\tau^2(\bar{u}_h|_\tau - u_h)^2 + \sum_{\tau\in T^h}\int_\tau h_\tau^2(\bar{f}|_\tau - f)^2\Big]$$
$$\leqslant Ch^2\big[(h^2 + h_U^2)\|p_h\|_{1,\Omega}^2 + h^2\|f\|_{1,\Omega}^2\big] = \mathcal{O}(h^4 + h_U^4).$$

Hence, we infer that σ_1 and σ_2 are the terms of the order $\mathcal{O}(h^2 + h_U^2)$. From the a priori estimate in [20], we known that $\|u - u_h\|_{L^2(\Omega)} + \|y - y_h\|_{L^2(\Omega)} + \|p - p_h\|_{L^2(\Omega)} = \mathcal{O}(h + h_U)$. Thus σ_1 and σ_2 are a small amount of higher order.

4. Numerical Experiments

In this section, we carry out some numerical experiments to demonstrate the error estimator developed in Section 3. Basically we are able to show that the derived error estimators can be effectively used in adaptive finite element approximation of the control problem in the framework of multiple adaptive meshes. We note that computational savings mainly come from two aspects: one is the adaptive meshes suitable for the singularities of the optimal control or states; the other is the possible DOFs reduction of the states for a given error of the optimal control via using multiple meshes. In all the examples, we use different meshes for the control and the states. And we use the piecewise constant elements to approximate the control, and the piecewise linear elements for the state and co-state. In solving these examples, we use a projection gradient algorithm, whose convergence has been proved in [20]. For the details of this algorithm we refer to [20]. We combine this algorithm with the standard adaptivity procedures, and only adjust the meshes after each iteration. We describe the algorithm briefly as follow:

Adaptive Process

1. Start with initial partition T^h, T_U^h and corresponding finite element space U^h, V^h.
2. Solve the discretization problem with the projection gradient method on current meshes.

   Step 1. Seek an initial approximation $u_h^0 \in U_{ad}^h$, and $y_h^0,\ g_h,\ r_h \in V^h$ such that
   $$(\nabla y_h^0, \nabla w_h) = (u_h^0, w_h), \quad (\nabla g_h, \nabla w_h) = (1, w_h),$$
   $$(\nabla r_h, \nabla w_h) = (\mathcal{P}_h g_h, w_h), \quad \forall\ w_h \in V^h; \quad \kappa_h = \frac{1}{\rho \bar{r}_h}.$$
   Step 2. Seek $\tilde{p}_h^n \in V^h$ such that
   $$(\nabla \tilde{p}_h^n, \nabla q_h) = (y_h^n - y_d, q_h), \quad \forall\ q_h \in V^h.$$
   Step 3. Set $u_h^{n+\frac{1}{2}} = u_h^n - \rho(\alpha u_h^n + \mathcal{P}_h \tilde{p}_h^n)$. Seek $y_h^{n+\frac{1}{2}} \in V^h$ such that
   $$(\nabla y_h^{n+\frac{1}{2}}, \nabla w_h) = (u_h^{n+\frac{1}{2}}, w_h), \quad \forall\ w_h \in V^h.$$
   Step 4. Set $\lambda_h^n = \kappa_h \min\left\{\bar{y}_h^{n+\frac{1}{2}} - \gamma, 0\right\}$ and $u_h^{n+1} = u_h^{n+\frac{1}{2}} - \lambda_h^n \mathcal{P}_h g_h$.
   Step 5. Seek $y_h^{n+1} \in V^h$ such that
   $$(\nabla y_h^{n+1}, \nabla w_h) = (u_h^{n+1}, w_h), \quad \forall\ w_h \in V^h.$$
   Step 6. Stop if stopping criterion is satisfied. Otherwise set $n = n + 1$,
3. Calculate local error estimators $\eta_u|_{\tau_U}, \eta_y|_\tau, \eta_y|_l, \eta_p|_\tau$ and $\eta_p|_l$ on each element $\tau_U \in T_U^h$ and $\tau \in T^h$ by (3.3), (3.5) and (3.7), then global error estimators η_u, η_y and η_p by (3.3), (3.4) and (3.6).
4. Refinement: for mesh T_U^h, let $\bar{\eta}_U = \frac{1}{N}\eta_u$ with N being the numbers of elements in partition T_U^h. An element τ_U is marked for refinement if $\eta_u|_{\tau_U} > \theta_U \bar{\eta}_U$, where θ_U is a suitable parameter. Perform refinement and then obtain new triangulations, T_U^h, for control. The treatment of mesh T^h is similar with $\bar{\eta}_V = (\eta_y + \eta_p)/M$ with M being the numbers of elements in partition T^h. An element τ is marked for refinement if
   $$\eta_y|_\tau + \eta_p|_\tau + \sum_{l \in \partial\tau \backslash \partial\Omega} (\eta_y|_l + \eta_p|_l) > \theta_V \bar{\eta}_V,$$
   where θ_V is a suitable parameter. Perform refinement and then obtain new triangulations T^h for the state and co-state.
5. Return to (2) on new meshes to update the solutions.

We solve the following problem on $\Omega = (0,1)^2$ to confirm the theoretical results.

$$\begin{cases} \min \quad \mathcal{J}(u,y) = \dfrac{1}{2}\displaystyle\int_\Omega (y-y_d)^2 + \dfrac{1}{2}\int_\Omega u^2, \\ s.t. \quad -\Delta y = u + f, \quad in\ \Omega, \qquad y = 0, \quad on\ \partial\Omega, \qquad \displaystyle\int_\Omega y \geqslant 0. \end{cases} \tag{4.1}$$

We perform three groups of numerical experiments. In the first group of numerical experiments, we investigate the advantage of multi-meshes. In the second and the third groups of numerical experiments, we check the efficiency of our posteriori error indictor. In our examples, the iteration parameter was set $\rho = 0.8$.

Example 4.1. The first example is to solve a control problem of type (4.1) with the exact solutions as follows:

$$\begin{aligned} &p = \sin 2\pi x_1 \sin 2\pi x_2 + \frac{3}{8}\sin 2\pi x_1 \sin 4\pi x_2; \\ &u = -p, \;\; y = p, \\ &y_d = y - \Delta p - 0.4; \\ &f = -\Delta y - u, \;\; \lambda = -0.4. \end{aligned} \tag{4.2}$$

We compute Example 4.1 on a uniform mesh and an adaptive mesh, respectively. The numerical results are presented in Table 4.1, where the mesh information is displayed together with L^2 approximation errors for the control and states.

Table 4.1: Numerical results of Example 4.1

	u,y,p on uniform mesh			u,y,p on adaptive mesh		
	u	y	p	u	y	p
Number of nodes	65536	65536	65536	65519	8639	8639
L^2 error	6.14e-03	1.23e-04	1.31e-04	6.13e-03	1.04e-03	1.06e-03

In Figure 4.1, we present the adaptive mesh and the discretized solution of y.

In this example, the optimal control and states are quite smooth so there is no much difference from numbers of nodes of meshes in either cases for the control with an approximation accuracy. However from Table 4.1, it can be clearly seen that using the multi-meshes the numbers of nodes for y and p can be reduced substantially to achieve this given control accuracy (here we are mainly interested in computing the control).

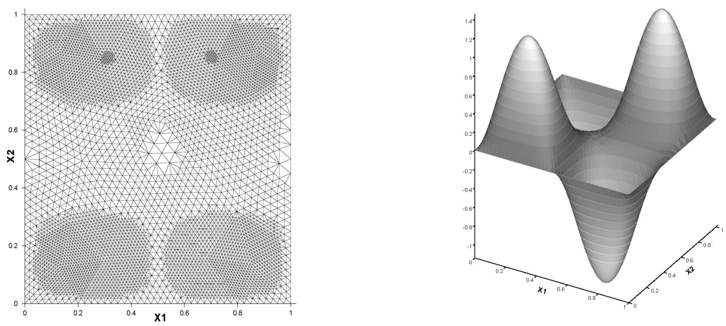

Fig. 4.1. Example 4.1: the adaptive mesh and discretized solution of y.

Since the main computational loads in solving the control problem come from repeatedly solving the state and the co-state equations, substantial computing work is thus saved.

Example 4.2. In this experiment, we show the efficiency of our posteriori error indictor. The example is to solve a control problem of type (4.1) with the exact solutions as follows:

$$\begin{aligned} &u = -p, \quad y_d = y - \Delta p - 0.4; \\ &\lambda = -0.4, \quad f = -\Delta y - u; \end{aligned}$$

$$p = \begin{cases} 7 \times 10^{10} \exp\left(\frac{1}{s(x_1, x_2)})\right), & s(x_1, x_2) < 0, \\ 0, & s(x_1, x_2) \geq 0. \end{cases}$$

$$y = \begin{cases} 8 \times 10^{10} \exp\left(\frac{1}{s(x_1, x_2)})\right), & s(x_1, x_2) < 0, \\ 0, & s(x_1, x_2) \geq 0. \end{cases}$$

where $s(x_1, x_2) = (x_1 - 0.2)^2 + (x_2 - 0.6)^2 - 0.04$.

In Figures 4.2 and 4.3, we display the adaptive meshes and the discretized solution of u and y in Example 4.2.

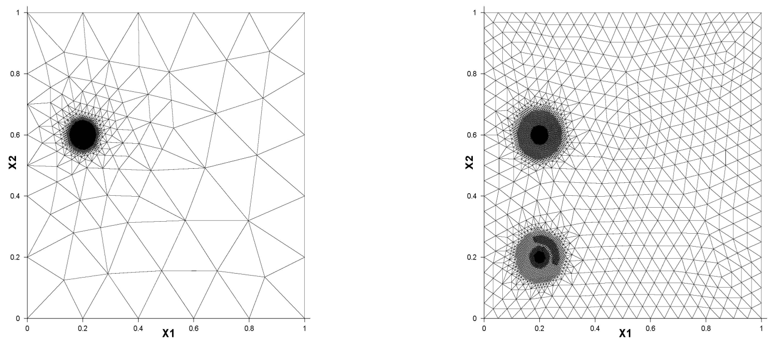

Fig. 4.2. Example 4.2: the adaptive meshes for the control (left) and for the state (right).

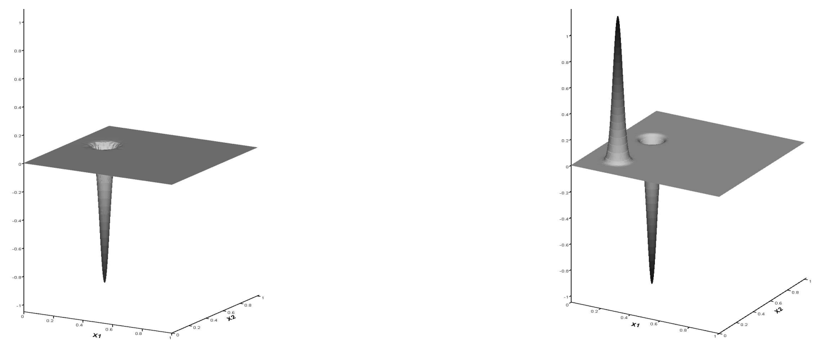

Fig. 4.3. Example 4.2: the approximation solutions of u (left) and y (right).

Likewise, we compute this problem on a uniform mesh and an adaptive mesh, respectively. Numerical results are presented in Table 4.2, where the mesh information is displayed together with L^2 approximation errors for the control and states.

Table 4.2: Numerical results of Example 4.2

	u, y, p on uniform mesh			u, y, p on adaptive mesh		
	u	y	p	u	y	p
Number of nodes	143360	143360	143360	15481	12247	12247
L^2 error	1.45e-03	1.58e-04	9.45e-05	1.14e-03	1.50e-04	8.75e-05

In this case, the control and states have different singularities. It can be clearly seen that on the multiple adaptive meshes one may use 10 times fewer degree of freedoms (DOFs) in the control and states variables to produce a given L^2 control error reduction.

Example 4.3. The final example is to solve a control problem of type (4.1) with the exact solutions:

$$
\begin{aligned}
&p = \sin \pi x_1 \sin \pi x_2, \quad u = u_0 - p, \\
&u_0 = \begin{cases} 1, & \text{if } x_1^2 < x_2; \\ 0, & \text{else}, \end{cases} \\
&y = \sin 2\pi x_1 \sin 2\pi x_2, \quad y_d = y - \Delta p - 0.4, \\
&\lambda = -0.4, \quad f = -\Delta y - u.
\end{aligned} \tag{4.3}
$$

Again, we compute this problem on a uniform mesh and an adaptive mesh, respectively. Numerical results are presented in Table 4.3, in which the mesh information is displayed together with L^2 approximation errors for the control and states.

Table 4.3: Numerical results of Example 4.3

	u, y, p on uniform mesh			u, y, p on adaptive mesh		
	u	y	p	u	y	p
Number of nodes	65536	65536	65536	4246	2407	2407
L^2 error	3.08e-02	9.82e-05	5.09e-05	2.94e-02	4.27e-03	1.85e-03

The numerical results show that the a posteriori error indicator works well. Adaptive meshes can capture the singularities of the solutions so that the number of DOFs is reduced substantially to one tenth of that used for the uniform mesh.

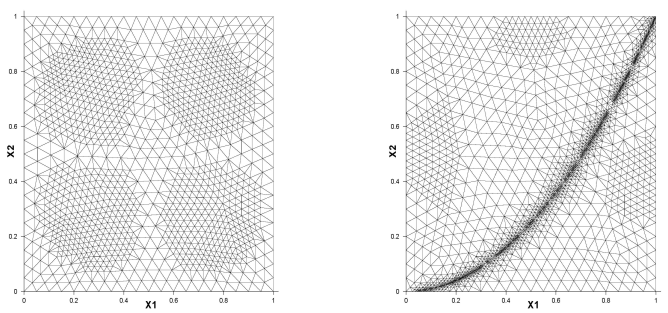

Fig. 4.4. Adaptive meshes for the state (left) and for the control (right) in Example 4.3.

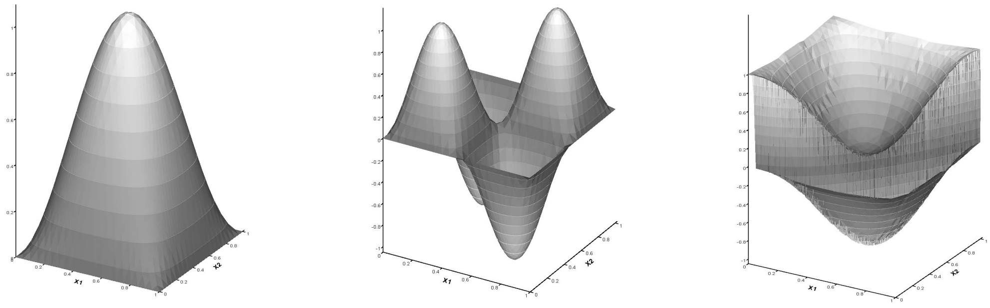

Fig. 4.5. Discretized solutions of p, y and u in Example 4.3.

In Figures 4.4-4.5, we present the adaptive meshes and the discretized solutions of u and y. It can be clearly seen from the right of Figure 4.4 that the locations of the jumps were correctly reflected in the adaptive meshes. These jumps require a large number of the DOFs to resolve and this would significantly increase the number of the DOFs for the computation should a uniform mesh were used.

Acknowledgments. The authors would like to thank referees for their useful suggestion. The first and second authors were supported by the National Basic Research Program of P. R. China under the grant 2005CB321703, the NSFC under the grants: 10441005 and 10571108, and the Research Fund for Doctoral Program of High Education by China State Education Ministry.

References

[1] V. Barbu and T. Precupanu, Convexity and Optimization in Banach Spaces, Sÿthoff & Noordhoff International Publishers, Bucharest, Roumania, 1978.

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

[3] M. Bergounioux and K. Kunisch, On the structure of Lagrange multipliers for state-constrained optimal control problems, *Syst. Control Lett.*, **48**(2003), 169-176.

[4] M. Bergounioux and D. Tiba, General optimality conditions for constrained convex control problems, *SIAM J. Control Optim.*, **34**:2(1996), 698-711.

[5] J. F. Bonnas and A. Shapiro, Perturbation Analysis of Optimization Problems, Springer-Verlag, NewYork, 2000.

[6] E. Casas, Control of an elliptic problem with pointwise state constraints, *SIAM J. Control Optim.*, **24**:4(1986), 1309-1322.

[7] E. Casas, M. Mateos and F. Tröltzsch, Necessary and sufficients optimality conditions for optimization problems in function spaces and applications to control theory, *ESAIM: Proceedinings*, **13**(2003), 18-30.

[8] P. G. Ciarlet, The Finite Element Methods for Elliptic Problems, North-Holland, Amsterdam, 1978.

[9] K. Deckelnick and M. Hinze, Convergence of a finite element approximation to a state-constrained elliptic control problem, *SIAM J. Numerical Analysis*, **45**(2007), 1937-1953.

[10] Y.Q. Huang, R. Li, W.B. Liu and N.N. Yan, Adaptive multi-mesh finite element approximation for constrained optimal control problems, Submitted to *SIAM J. Control Optim.*.

[11] A. Kufner, O. John and S. Fucik, Function Spaces, Nordhoff, Leiden, 1977.

[12] I. Lasiecka, State constrained control problems for parabolic systems: regularity of optimal solutions, *Appl. Math. Opt.*, **6**(1980), 1-29.

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

[14] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1971.

[15] W.B. Liu, Adaptive multi-meshes in finite element approximation of optimal control, *Contemporary Mathematics*, **383**(2005), 113-132.

[16] W.B. Liu and N.N. Yan, A posteriori error estimates for distributed optimal control problems, *Adv. Comput. Math.*, **15**(2001), 285-309.

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

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

[19] W.B. Liu and N.N. Yan, A posteriori error estimates for optimal control of Stokes flows, *SIAM J. Numer. Anal.*, **40**(2003), 1805-1869.

[20] W.B. Liu, D.P. Yang and L. Yuan, A priori error estimates of state-constrained optimal control of PDEs with integral constraint, Submitted.

[21] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary condition, *Math. Comput.*, **54**(1990), 483-493.

[22] R. Verfürth, A posteriori error estimators for the Stokes equations, *Numer. Math.*, **55**(1989), 309-325.

[23] R. Verfürth, A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques, Wiley-Teubner, New York-Stuttgart, 1996.

[24] Z. Zhang and N.N. Yan, Recovery type a posteriori error estmates in finite element methods, *Korean Journal Computational and Applied Mathematics*, **8**(2001), 1331-1382.