INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 6, Number 2, Pages 335–353

L^2 NORM EQUIVALENT A POSTERIORI ERROR ESTIMATE FOR A CONSTRAINED OPTIMAL CONTROL PROBLEM

LIANG GE, WENBIN LIU, AND DANPING YANG

Abstract. Adaptive finite element approximation for a constrained optimal control problem is studied. A posteriori error estimators equivalent to the L^2 norm of the approximation error are derived both for the state and the control approximation, which are particularly suitable for an adaptive multi-mesh finite element scheme and applications where L^2 error is more important. The error estimators are then implemented and tested with promising numerical results.

Key Words. convex optimal control problem, adaptive finite element method, L^2 norm equivalent a posteriori error estimate, multi-meshes.

1. Introduction

There has been so extensive research on developing adaptive finite element algorithms for PDEs in the scientific literature that it is simply impossible to give even a very brief review here. Recently, there has been intensive research in adaptive finite element methods for optimal control problems, see, for example, [2, 3, 4, 6, 8, 9, 12, 15, 16, 18]. Particularly a posteriori error estimates equivalent to the energy norm of the approximation error were derived for several types of optimal control problems. Furthermore it has been found that for constrained control problems, different adaptive meshes are often needed for the control and the states, see [10]. Using different adaptive meshes for the control and the state allows very coarse meshes to be used in solving the state and co-state equations. Thus much computational work can be saved since one of the major computational loads in computing optimal control is to solve the state and co-state equations repeatedly. This will be also seen from our numerical experiments in Section 4.

Although a posteriori error estimates equivalent to the H^1 norm of the approximation error (to be called H^1 norm equivalent a posteriori error estimates) have been derived for several elliptic optimal control problems, see [8, 9, 10], both for the control constraints of obstacle types and integral types, there seems no existing work on L^2 norm equivalent a posteriori error estimates, which are equivalent to the L^2 norm of the approximation error, although some upper bounds were derived using the L^2 norm for the control constraint of an obstacle type, see [10, 16]. It does not seem a trivial problem whether and how some lower bounds can be derived via the L^2 norm, although it seemed possible to adapt the existing duality techniques to derive upper bounds. In many engineering applications, one cares more about averaging values of the control and the states. In these cases, it seems to be more natural to use the L^2-norm of the approximation error as the stopping criteria in

Received by the editors June 19, 2008.
2000 *Mathematics Subject Classification.* 35R35, 49J40, 60G40.
This research was supported by the National Basic Research Program of P. R. China under the grant 2005CB321703.

computations. Thus L^2 equivalent error indicators seem to be quite useful. Error indicators based on the L^2 norm error bounds tend to produce less over-reinterment in such cases.

The purpose of this article is to investigate indicators that are equivalent to the L^2 norm of the approximation error for a constrained optimal control problem, where the control constraint is of an integral type. This control problem was studied in [8, 9], where H^1 norm equivalent a posteriori error estimates were derived. We derived L^2 norm equivalent a posteriori error estimators, which allow different meshes to be used for the stats and the control. Then we performed some numerical tests to confirm the effectiveness of the error estimators.

The plan of the paper is as follows. In Section 2, we will construct the finite element approximation for the distributed optimal control problem. In Section 3, the a posteriori error estimators equivalent in the L^2 norm are derived for the control problem. Finally numerical test results are presented in Section 4.

2. Optimal control problem and its finite element approximation

Let Ω be a bounded open set in R^d $(1 \leq d \leq 3)$ with the Lipschitz boundary $\partial\Omega$. We adopt the standard notation $W^{1,q}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{W^{1,q}(\Omega)}$ and seminorm $|\cdot|_{W^{1,q}(\Omega)}$ for $1 \leq q \leq \infty$. We set $W_0^{1,q}(\Omega) \equiv \{w \in W^{1,q}(\Omega) : w|_{\partial\Omega} = 0\}$ and denote $W^{1,2}(\Omega)$ $(W_0^{1,2}(\Omega))$ by $H^1(\Omega)$ $(H_0^1(\Omega))$. In the rest of the paper, we will take the state space $V = H_0^1(\Omega)$ and the control space $U = L^2(\Omega)$. Other cases can be considered similarly. Let the observation space $Y = L^2(\Omega)$. We investigate the following distributed convex optimal control problem:

$$\min_{u\in K} \frac{1}{2}\int_\Omega (y-y_d)^2 + \frac{1}{2}\int_\Omega u^2, \tag{2.1}$$

$$-\Delta y = f + u \quad \text{in } \Omega\,, \quad y|_{\partial\Omega} = 0,$$

where $K = \{v \mid v \in L^2(\Omega),\ \int_\Omega v \geq 0\}$ is a closed convex set. We first give a weak formula for the state equation. Let

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

and

$$(u,v) = \int_\Omega uv, \quad \forall\ u, v \in L^2(\Omega).$$

It follows that

$$\alpha\|y\|_V^2 \leq a(y,y), \quad |a(y,w)| \leq M\|y\|_V\|w\|_V, \qquad \forall\ y, w \in V, \tag{2.2}$$

where $0 < \alpha \leq M < \infty$ are positive constants. With these notions the standard weak formula for the state equation reads: *find* $y \in V$ *such that*

$$a(y,w) = (f+u, w), \qquad \forall\ w \in H_0^1(\Omega). \tag{2.3}$$

Then the mentioned-above control problem can be restated as follows:

$$(\text{OCP}): \quad \begin{cases} \min\limits_{u\in K}\Big\{\dfrac{1}{2}\displaystyle\int_\Omega (y-y_d)^2 + \dfrac{1}{2}\int_\Omega u^2\Big\}, \\ a(y,w) = (f+u,w), \quad \forall\ w \in V. \end{cases} \tag{2.4}$$

It follows from [14] that the control problem (OCP) has a unique solution (y,u). Furthermore a pair (y,u) is the solution of (OCP) iff there is a co-state $p \in V$ such that the triplet (y, p, u) satisfies the following optimality conditions:

$$
\text{(2.5)} \qquad \text{(OCP-OPT)}: \quad \begin{cases} a(y, w) = (f + u, w), \quad \forall\, w \in V, \\ \\ a(q, p) = (y - y_d, q), \quad \forall\, q \in V, \\ \\ (u + p, v - u) \geq 0, \quad \forall\, v \in K. \end{cases}
$$

Let us consider the finite element approximation of the control problem (OCP). Here we consider only the conforming d-simplex elements. For simplicity, we assume that Ω is a polygonal domain. Let T^h be a triangulation Ω into disjoint regular d-simplices τ, so that $\bar{\Omega} = \bigcup_{\tau \in T^h} \bar{\tau}$. Each element has at most one face on $\partial\Omega$, and $\bar{\tau}$ and $\bar{\tau}'$ have either only one common vertex or a whole edge or face if τ and $\tau' \in T^h$.

Associated with T^h is a finite dimensional subspace S^h of $C(\bar{\Omega})$, such that $\chi|_\tau$ are polynomials of m-degree $(m \geq 1)$ for each $\chi \in S^h$ and $\tau \in T^h$. Let $V^h = S^h \cap H_0^1(\Omega)$. It is easy to see that $V^h \subset V$.

Let T_U^h be another triangulation of Ω into disjoint regular d-simplices τ_U, so that $\bar{\Omega} = \bigcup_{\tau_U \in T_U^h} \bar{\tau}_U$. Assume that $\bar{\tau}_U$ and $\bar{\tau}_U'$ have either only one common vertex or a whole face or are disjoint if τ_U and $\tau_U' \in T_U^h$.

Associated with T_U^h is another finite dimensional subspace U^h of $L^2(\Omega)$, such that $\chi|_{\tau_U}$ are polynomials of r-degree $(r \geq 0)$ for each $\chi \in W_U^h$ and $\tau_U \in T_U^h$. An optimal control of a constrained problem normally has lower regularity so that we shall use discontinuous base functions to approximate the control. Hence there is no requirement for continuity of the functions in U^h.

Let h_τ (h_{τ_U}) denote the maximum diameter of the element τ (τ_U) in T^h (T_U^h). Define the discrete constraint set as

$$
\text{(2.6)} \qquad K^h = \{u_h \in U^h : \int_\Omega u_h \geq 0\}.
$$

Then a possible finite element approximation of (OCP), which will be labeled as $(\text{OCP})^h$, reads:

$$
\text{(2.7)} \qquad (\text{OCP})^h: \quad \begin{cases} \min\limits_{u_h \in K^h} \left\{ \frac{1}{2} \int_\Omega (y_h - y_d)^2 + \frac{1}{2} \int_\Omega u_h^2 \right\}, \\ \\ a(y_h, w_h) = (f + u_h, w_h), \qquad \forall\, w_h \in V^h. \end{cases}
$$

Define

$$
J(u) = \frac{1}{2} \int_\Omega (y(u) - y_d)^2 + \frac{1}{2} \int_\Omega u^2, \quad J_h(u_h) = \frac{1}{2} \int_\Omega (y_h(u_h) - y_d)^2 + \frac{1}{2} \int_\Omega u_h^2.
$$

where $y_h(u_h) \in V^h$ is given by

$$
a(y_h(u_h), w_h) = (u_h, w_h), \quad \forall\, w_h \in V^h.
$$

Then the reduced problems of (2.4) and (2.7) read: $u \in K$ such that

$$
\text{(2.8)} \qquad J(u) = \min_{v \in K} \{J(v)\},
$$

and $u_h \in K^h$ such that

$$
\text{(2.9)} \qquad J_h(u_h) = \min_{v_h \in K^h} \{J_h(v_h)\},
$$

respectively. Since this is a linear control problem, the reduced objective function is convex. Furthermore $J(\cdot)$ is uniformly convex in the sense that there is a $c > 0$, independent of h, such that

$$(J'(u) - J'(v), u - v) \geq c\|u - v\|^2_{L^2(\Omega)}, \tag{2.10}$$

where $u, v \in U$. It follows that the control problem $(\text{OCP})^h$ has a unique solution (y_h, u_h). Furthermore, a pair $(y_h, u_h) \in V^h \times U^h$ is the solution of $(\text{OCP})^h$ iff there is a co-state $p_h \in V^h$ such that the triplet (y_h, p_h, u_h) satisfies the following optimality conditions, which shall be labeled as $(\text{OCP-OPT})^h$:

$$(\text{OCP-OPT})^h : \quad \begin{cases} a(y_h, w_h) = (f + u_h, w_h), \quad \forall\, w_h \in V^h, \\ a(q_h, p_h) = (y_h - y_d, q_h), \quad \forall\, q_h \in V^h, \\ (u_h + p_h, v_h - u_h) \geq 0, \quad \forall\, v_h \in K^h. \end{cases} \tag{2.11}$$

For the last variational inequality in (2.11) we have the following conclusion, which was proved in [8, 9] but for the readers' convenience we include it here:

Lemma 2.1. *Assume p_h is known in the variational inequality of (2.11). The solution of the variational inequality in (2.11) is*

$$u_h = P_h\big(-p_h + \max\{0, \overline{p_h}\}\big), \qquad \overline{p_h} = \frac{\int_\Omega p_h}{\int_\Omega 1}, \tag{2.12}$$

where P_h is the L^2-projection from $L^2(\Omega)$ to U^h.

Proof. The proof is divided into two steps. We will prove $u_h \in K^h$ at the first step, and then prove u_h is the solution of the variational inequality at the second step.

Step 1. Let P_h be the L^2-projection from $L^2(\Omega)$ to U^h. For any $v \in U$, we have

$$\int_\Omega (P_h v - v)\phi = 0, \quad \forall\, \phi \in U^h.$$

Since $\phi \equiv 1 \in U^h$ such that

$$\int_\Omega \big[P_h(-p_h + \max\{0, \overline{p_h}\}) - (-p_h + \max\{0, \overline{p_h}\})\big] = 0,$$

hence

$$\int_\Omega u_h = \int_\Omega (-p_h + \max\{0, \overline{p_h}\}) = -\int_\Omega p_h + \int_\Omega \max\{0, \overline{p_h}\} \geq 0.$$

Thus $u_h \in K^h$.

Step 2. Noting that for each $v_h \in K^h$,

$$\int_\Omega (u_h + p_h)(v_h - u_h) = \int_\Omega \big[P_h(-p_h + \max\{0, \overline{p_h}\}) - (-p_h + \max\{0, \overline{p_h}\}) + \max\{0, \overline{p_h}\}\big](v_h - u_h) = \int_\Omega \max\{0, \overline{p_h}\}(v_h - u_h).$$

we see that if $\overline{p_h} \leq 0$ then

$$\int_\Omega (u_h + p_h)(v_h - u_h) = 0,$$

and that if $\overline{p_h} > 0$ then

$$\int_\Omega (u_h + p_h)(v_h - u_h) \geq 0,$$

since $\int_\Omega u_h = \int_\Omega (-p_h + \max\{0, \overline{p_h}\}) = 0$ and $\int_\Omega v_h \geq 0$.

Therefore it has been shown that u_h is the solution of the variational inequality in (2.11). ■

3. L^2 norm equivalent a posteriori error estimators

Adaptive finite element approximation has been found very useful in computing optimal control, as mentioned in the introduction. It uses an a posteriori error indicator to guide the mesh refinement procedure. 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 approximate. In this section, we will derive an upper bound estimate of error in L^2-norm, and then show that it also is a lower bound. In addition, c or C denotes a general positive constant independent of h and h_U.

3.1. Upper bound estimate in L^2 norm. The following theorem is one of our main results, which gives an upper bound for the approximation error in L^2 norm.

Theorem 3.1. *Assume that Ω is a convex domain. Let (y, u) and (y_h, u_h) be the solutions of* (OCP) *and* $(\mathrm{OCP})^h$*, and p and p_h be the solutions of the co-state equations* (2.5) *and* (2.11) *respectively. Then,*

$$\|u - u_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{L^2(\Omega)} + \|p - p_h\|^2_{L^2(\Omega)} \leq C \sum_{i=1}^{3} \eta_i^2, \tag{3.1}$$

where the error indicators η_1, η_2 and η_3 are given by

$$\begin{aligned} \eta_1^2 &= \sum_{\tau_U} \int_{\tau_U} (-P_h p_h + p_h)^2, \\ \eta_2^2 &= \sum_{\tau \in T^h} h_\tau^4 \int_\tau (y_h - y_d + div(\nabla p_h))^2 + \sum_{l \cap \partial\Omega = \emptyset} \int_l h_l^3 [(\nabla p_h \cdot n)]^2, \\ \eta_3^2 &= \sum_{\tau \in T^h} h_\tau^4 \int_\tau (f + u_h + div(\nabla y_h))^2 + \sum_{l \cap \partial\Omega = \emptyset} \int_l h_l^3 [(\nabla y_h \cdot n)]^2, \end{aligned} \tag{3.2}$$

where P_h is the L^2-projection from $L^2(\Omega)$ to U^h, l is a face of an element τ, $[\nabla p_h \cdot n)]$ and $[(\nabla y_h \cdot n)]$ are the normal derivative jumps over the interior face l, defined by

$$\begin{aligned} [(\nabla p_h \cdot n)]_l &= (\nabla p_h|_{\tau_l^1} - \nabla p_h|_{\tau_l^2}) \cdot n, \\ [(\nabla y_h \cdot n)]_l &= (\nabla y_h|_{\tau_l^1} - \nabla y_h|_{\tau_l^2}) \cdot n, \end{aligned} \tag{3.3}$$

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

In order to prove Theorem 3.1, we need the following important lemmas.

Lemma 3.1. [5] *Let π_h be the standard Lagrange interpolation operator. For $m = 0$ or 1, $q > \frac{d}{2}$ and each $v \in W^{2,q}(\Omega)$,*

$$|v - \pi_h v|_{W^{m,q}(\Omega)} \leq C h^{2-m} |v|_{W^{2,q}(\Omega)}. \tag{3.4}$$

Lemma 3.2. *Let $\hat{\pi}_h$ be the average interpolation operator defined in* [17]*. For $m = 0$ or 1, $1 \leq q \leq \infty$ and each $v \in W^{1,q}(\Omega)$,*

$$|v - \hat{\pi}_h v|_{W^{m,q}(\tau)} \leq \sum_{\bar{\tau}' \cap \bar{\tau} \neq \emptyset} C h_\tau^{1-m} |v|_{W^{1,q}(\tau')}. \tag{3.5}$$

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

$$\|v\|_{W^{0,q}(\partial\tau)} \le C\Big(h_\tau^{-\frac{1}{q}}\|v\|_{W^{0,q}(\tau)} + h_\tau^{1-\frac{1}{q}}|v|_{W^{1,q}(\tau)}\Big). \tag{3.6}$$

Let $J(\cdot)$ and $J_h(\cdot)$ be as before, and $p(u_h)$ be the solution of the following auxiliary equation:

$$\begin{cases} a(y(u_h), w) = (f + u_h, w), \quad \forall\ w \in V, \\ a(q, p(u_h)) = (y(u_h) - y_d, q), \quad \forall\ q \in V. \end{cases} \tag{3.7}$$

It can be shown that

$$(J'(u), v) = (u + p, v), \quad \forall\ v \in U,$$

$$(J'(u_h), v) = (u_h + p(u_h), v), \quad \forall\ v \in U$$

and

$$(J_h'(u_h), v_h) = (u_h + p_h, v_h), \quad \forall\ v_h \in U^h.$$

Then we need to prove two lemmas. In Lemma 3.4, we derive the error bound for $u - u_h$.

Lemma 3.4. *Let u and u_h be the solutions of (2.8) and (2.9) respectively. Then,*

$$\|u - u_h\|^2_{L^2(\Omega)} \le C\big\{\eta_1^2 + \|p_h - p(u_h)\|^2_{L^2(\Omega)}\big\}, \tag{3.8}$$

where p_h and $p(u_h)$ are the solutions of equations (2.11) and (3.7) respectively.

Proof. It follows from (2.10) that

$$\begin{aligned} c\|u - u_h\|_{L^2(\Omega)} &\le (J'(u), u - u_h) - (J'(u_h), u - u_h) \\ &\le -(J'(u_h), u - u_h) = (J_h'(u_h), u_h - u) + (J_h'(u_h) - J'(u_h), u - u_h) \\ &= \inf_{v_h \in K^h}(J_h'(u_h), v_h - u) + (J_h'(u_h) - J'(u_h), u - u_h). \end{aligned} \tag{3.9}$$

It is clear that

$$(J_h'(u_h), v_h - u) = (u_h + p_h, v_h - u). \tag{3.10}$$

Noting that $\int_\Omega v \ge 0$ and $\int_\Omega P_h v - v = 0$ for all $v \in K$, we have

$$\int_\Omega P_h v \ge 0,$$

which means $P_h v \in K^h$. So we can take $v_h = P_h u$ in equality (3.10). Then from Lemma 2.1, we have

$$(u_h + p_h, P_h u - u) = \sum_{\tau_U}\int_{\tau_U}(-P_h p_h + \max(0, \overline{p_h}) + p_h)(P_h u - u).$$

Since $\int_{\tau_U}(P_h u - u) = 0$, hence

$$(u_h + p_h, P_h u - u) = \sum_{\tau_U}\int_{\tau_U}(-P_h p_h + p_h)(P_h u - u).$$

Noting that $P_h u_h = u_h$, we have

$$\begin{aligned}
&\sum_{\tau}\int_{\tau_U}(-P_hp_h+p_h)(P_hu-u)\\
&\quad=\sum_{\tau_U}\int_{\tau_U}(-P_hp_h+p_h)(P_h(u-u_h)-(u-u_h))\\
&\quad\leq C\sum_{\tau_U}\int_{\tau_U}(-P_hp_h+p_h)^2+\frac{c}{3}\|u-u_h\|^2_{L^2(\Omega)},
\end{aligned} \tag{3.11}$$

which leads to

$$\|u-u_h\|^2_{L^2(\Omega)}\leq C\sum_{\tau_U}\int_{\tau_U}(-P_hp_h+p_h)^2+(J_h'(u_h)-J'(u_h),u-u_h). \tag{3.12}$$

It follows from the formulas of J' and J_h' that

$$\begin{aligned}
(J_h'(u_h)-J'(u_h),u-u_h)&=(p_h-p(u_h),u-u_h)\\
&\leq C\|p_h-p(u_h)\|^2_{L^2(\Omega)}+\frac{c}{3}\|u_h-u\|^2_{L^2(\Omega)}.
\end{aligned} \tag{3.13}$$

Substituting (3.13) into (3.12) results in (3.8). ∎

Lemma 3.5. *Assume that Ω is a convex domain. Let (y_h,p_h) and $(y(u_h),p(u_h))$ be the solutions of (2.11) and (3.7) respectively. Then,*

$$\|y_h-y(u_h)\|^2_{L^2(\Omega)}+\|p_h-p(u_h)\|^2_{L^2(\Omega)}\leq C(\eta_2^2+\eta_3^2). \tag{3.14}$$

Proof. We need an a priori regular estimate for the following auxiliary problem:

$$-div(\nabla\xi)=g\ \ in\ \Omega,\quad \xi|_{\partial\Omega}=0. \tag{3.15}$$

It is well-known that $\xi\in H^2(\Omega)\bigcap H^1_0(\Omega)$ such that

$$\|\xi\|_{H^2(\Omega)}\leq C\|g\|_{L^2(\Omega)}.$$

Firstly, we estimate $\|y_h-y(u_h)\|_{L^2(\Omega)}$. Let ξ be the solution of (3.15) with $g=y(u_h)-y_h$ and ξ_I be the average interpolation of ξ defined in Lemma 2.1. Applying the well known residual technique (see, e.g., [5] and [17]), from equations (2.11), (3.7) and Lemmas 3.1 and 3.2, we derive that

$$\begin{aligned}
&\|y(u_h)-y_h\|^2_{L^2(\Omega)}=(y(u_h)-y_h,-div(\nabla\xi))\\
&\quad=(\nabla(y(u_h)-y_h),\nabla(\xi-\xi_I))+(\nabla(y(u_h)-y_h),\nabla\xi_I)\\
&\quad=\sum_{\tau\in T^h}\int_\tau(f+u_h+div(\nabla y_h))(\xi-\xi_I)\\
&\qquad-\sum_{\tau\in T^h}\int_{\partial\tau}(\nabla y_h\cdot n)(\xi-\xi_I)ds\\
&\quad\leq C\Big(\sum_{\tau\in T^h}h_\tau^4\int_\tau(f+u_h+div(\nabla y_h))^2\\
&\qquad\quad+\sum_{l\cap\partial\Omega=\emptyset}h_l^3\int_l[\nabla y_h\cdot n]^2\Big)^{1/2}\|\xi\|_{H^2(\Omega)}\\
&\quad\leq C\Big(\sum_{\tau\in T^h}h_\tau^4\int_\tau(f+u_h+div(\nabla y_h))^2\\
&\qquad+\sum_{l\cap\partial\Omega=\emptyset}h_l^3\int_l[\nabla y_h\cdot n]^2\Big)+\frac{1}{2}\|y(u_h)-y_h\|^2_{L^2(\Omega)}.
\end{aligned}$$

Hence,

$$\|y(u_h) - y_h\|^2_{L^2(\Omega)} \le C\eta_3^2. \tag{3.16}$$

Next we estimate $\|p(u_h) - p_h\|_{L^2(\Omega)}$. Let ξ be the solution of (3.15) with $g = p(u_h) - p_h$, and $\xi_I = \pi_h \xi \in V^h$ be the standard Lagrange interpolation of ξ. Then

$$\begin{aligned}
\|p(u_h) - p_h\|^2_{L^2(\Omega)} &= (p(u_h) - p_h, -div(\nabla\xi)) \\
&= (\nabla(p(u_h) - p_h), \nabla(\xi - \xi_I)) + (\nabla(p(u_h) - p_h), \nabla\xi_I) \\
&= \sum_{\tau\in T^h} \int_\tau (y(u_h) - y_d + div(\nabla p_h))(\xi - \xi_I) \\
&\quad - \sum_{l\cap\partial\Omega=\emptyset} \int_l [\nabla p_h \cdot n](\xi - \xi_I)ds + (y(u_h) - y_h, \xi_I) \\
&= \sum_{\tau\in T^h} \int_\tau (y_h - y_d + div(\nabla p_h))(\xi - \xi_I) \\
&\quad - \sum_{l\cap\partial\Omega=\emptyset} \int_l [\nabla p_h \cdot n](\xi - \xi_I)ds + (y(u_h) - y_h, \xi) \\
&\le C\Big(\sum_{\tau\in T^h} h_\tau^4 \int_\tau (y_h - y_d + div(\nabla p_h))^2 + \sum_{l\cap\partial\Omega=\emptyset} h_l^3 \int_l [\nabla p_h \cdot n]^2\Big)^{1/2} \\
&\quad \cdot \Big(\sum_{\tau\in T^h} h_\tau^{-4} \int_\tau |\xi - \xi_I|^2 + \sum_{l\cap\partial\Omega=\emptyset} h_\tau^{-3} \int_l |\xi - \xi_I|^2\Big)^{1/2} \\
&\quad + C\|y_h - y(u_h)\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} \\
&\le C\Big\{\sum_{\tau\in T^h} h_\tau^4 \int_\tau (y_h - y_d + div(\nabla p_h))^2 + \sum_{l\cap\partial\Omega=\emptyset} h_l^3 \int_l [\nabla p_h \cdot n]^2 \\
&\quad + \|y_h - y(u_h)\|^2_{L^2(\Omega)}\Big\} + \frac{1}{2}\|p(u_h) - p_h\|^2_{L^2(\Omega)}.
\end{aligned}$$

Thus

$$\begin{aligned}
\|p(u_h) - p_h\|^2_{L^2(\Omega)} \le C\Big\{&\sum_{\tau\in T^h} h_\tau^4 \int_\tau (y_h - y_d + div(\nabla p_h))^2 \\
&+ \sum_{l\cap\partial\Omega=\emptyset} h_l^3 \int_l [\nabla p_h \cdot n]^2 + \|y_h - y(u_h)\|^2_{L^2(\Omega)}\Big\}.
\end{aligned} \tag{3.17}$$

As a result, it follows from (3.17) and (3.16) that

$$\|p_h - p(u_h)\|^2_{L^2(\Omega)} \le C\Big\{\eta_2^2 + \eta_3^2\Big\}. \tag{3.18}$$

The proof of Lemma 3.5 is completed. ■

Now we can prove Theorem 3.1.

Proof of Theorem 3.1. It follows from Lemmas 3.4 and 3.5 that

$$\|u - u_h\|^2_{L^2(\Omega)} + \|p_h - p(u_h)\|^2_{L^2(\Omega)} + \|y_h - y(u_h)\|^2_{L^2(\Omega)} \le C\sum_{i=1}^3 \eta_i^2. \tag{3.19}$$

Noting that

$$\|y_h - y\|_{L^2(\Omega)} \le \|y_h - y(u_h)\|_{L^2(\Omega)} + \|y(u_h) - y\|_{L^2(\Omega)}, \tag{3.20}$$

$$\|p_h - p\|_{L^2(\Omega)} \le \|p_h - p(u_h)\|_{L^2(\Omega)} + \|p(u_h) - p\|_{L^2(\Omega)}, \tag{3.21}$$

and

$$\|p - p(u_h)\|^2_{L^2(\Omega)} + \|y - y(u_h)\|^2_{L^2(\Omega)} \le C\|u - u_h\|^2_{L^2(\Omega)}, \tag{3.22}$$

we then derive (3.1) from (3.19)-(3.22). ■

3.2. Lower bound estimate in L^2 norm. Now we are in the position of deriving a posteriori lower bounds. That is to show the derived estimators in the above theorem are in fact equivalent in the sense that there are two constants $C \ge c > 0$ such that

$$c\sum_{i=1}^{3}\eta_i^2 - c\epsilon^2 \le \|u - u_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{L^2(\Omega)} + \|p - p_h\|^2_{L^2(\Omega)} \le C\sum_{i=1}^{3}\eta_i^2,$$

where ϵ is of higher order. The following theorem confirms this result.

Theorem 3.2. *Let (y, u) and (y_h, u_h) be the solutions of* (OCP) *and* $(\text{OCP})^h$*, and p and p_h be the solutions of the co-state equations* (2.5) *and* (2.11) *respectively. Then,*

$$\sum_{i=1}^{3}\eta_i^2 \le C\Big\{\|u - u_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{L^2(\Omega)} + \|p - p_h\|^2_{L^2(\Omega)} + \epsilon^2\Big\}, \tag{3.23}$$

where η_i, $i = 1, 2, 3$, are defined in Theorem 3.1 and

$$\epsilon^2 = \epsilon_2^2 + \epsilon_3^2$$

associated with

$$\epsilon_2^2 = \sum_{\tau\in T^h}\int_\tau h_\tau^4(y_h - y_d - (\overline{y_h}|_\tau - \overline{y_d}|_\tau))^2 \tag{3.24}$$

and

$$\epsilon_3^2 = \sum_{\tau\in T^h}\int_\tau h_\tau^4(f - \bar f|_\tau)^2, \tag{3.25}$$

where $\bar v|_\tau$ is the integral average value of v on the element τ such that $\bar v|_\tau = \dfrac{\int_\tau v}{\int_\tau 1}$.

It is clear that ϵ is of higher order. The proof of Theorem 3.2 is completed by the following lemmas.

Lemma 3.6. *Let (y, u) and (y_h, u_h) be the solutions of* (OCP) *and* $(\text{OCP})^h$*, and p and p_h be the solutions of the co-state equations* (2.5) *and* (2.11) *respectively. Then,*

$$\eta_1^2 \le C\Big\{\|u - u_h\|^2_{L^2(\Omega)} + \|p - p_h\|^2_{L^2(\Omega)}\Big\}, \tag{3.26}$$

where η_1 is defined in Theorem 3.1.

Proof. It is easily seen that

$$\begin{aligned}\sum_{\tau_U}\int_{\tau_U}(-P_hp_h + p_h)^2 &= \sum_{\tau_U}\int_{\tau_U}(p_h - P_hp_h)(p_h - p + p - P_hp + P_hp - P_hp_h)\\ &\le \sum_{\tau_U}\int_{\tau_U}(p_h - P_hp_h)(p - P_hp) + \frac{1}{3}\sum_{\tau_U}\int_{\tau_U}(p_h - P_hp_h)^2 + C\|p_h - p\|^2_{L^2(\Omega)}.\end{aligned}$$

Since $u + p = \max(0, \overline{p}) = const$, hence

$$P_h(u + p) = u + p,$$

such that

$$\begin{aligned}
&\sum_{\tau_U}\int_{\tau_U}(p_h-P_hp_h)(p-P_hp)\\
&=\sum_{\tau_U}\int_{\tau_U}(p_h-P_hp_h)(p+u-P_h(p+u)+P_hu-u)\\
&=\sum_{\tau_U}\int_{\tau_U}(p_h-P_hp_h)(P_h(u-u_h)-(u-u_h))\\
&\leq\frac{1}{3}\sum_{\tau_U}\int_{\tau_U}(p_h-P_hp_h)^2+C\|u-u_h\|^2_{L^2(\Omega)}.
\end{aligned}$$

Therefore,

$$\eta_1^2\leq C\Big\{\|u-u_h\|^2_{L^2(\Omega)}+\|p-p_h\|^2_{L^2(\Omega)}\Big\}.$$

This is (3.26). ■

Then we prove the following lemmas by modifying the standard bubble function technique as in [1, 19].

Lemma 3.7. *Let* (y,u) *and* (y_h,u_h) *be the solutions of* (OCP) *and* (OCP)^h, *and* p *and* p_h *be the solutions of the co-state equations* (2.5) *and* (2.11) *respectively. Then,*

$$\sum_{\tau\in T^h}h_\tau^4\int_\tau(y_h-y_d+div(\nabla p_h))^2\leq C\Big\{\|p-p_h\|^2_{L^2(\Omega)}+\|y-y_h\|^2_{L^2(\Omega)}+\epsilon_2^2\Big\}.\tag{3.27}$$

Proof. Following the idea in the standard bubble function technique (see [1, 19]), we indicate with b_τ the standard third order polynomial bubble on τ scaled such that $b_\tau=\lambda_1\lambda_2\lambda_3$, where $\{\lambda_1,\lambda_2,\lambda_3\}$ denote the barycentric coordinates on τ. Let

$$w_\tau=c_1(div(\nabla p_h)+\overline{(y_h-y_d)}|_\tau)b_\tau^2,\tag{3.28}$$

where $c_1=|\tau|/\int_\tau b_\tau^2$ is a constant. From the standard scaling arguments it is not difficult to show that

$$\begin{aligned}
\|w_\tau\|^2_{L^2(\tau)}&=\int_\tau c_1^2(div(\nabla p_h)+\overline{(y_h-y_d)}|_\tau)^2b_\tau^4\\
&\leq c_1\int_\tau(div(\nabla p_h)+\overline{(y_h-y_d)}|_\tau)^2.
\end{aligned}\tag{3.29}$$

On the other hand, since $\lambda_1\lambda_2\lambda_3=0$ on $\partial\tau$, then

$$w_\tau|_{\partial\tau}=c_1(div(\nabla p_h)+\overline{y_h-y_d})(\lambda_1\lambda_2\lambda_3)^2|_{\partial\tau}=0.$$

Furthermore, we also have

$$\nabla w_\tau\big|_{\partial\tau}=2c_1(div(\nabla p_h)+\overline{y_h-y_d})(\lambda_1\lambda_2\lambda_3)\nabla(\lambda_1\lambda_2\lambda_3)\big|_{\partial\tau}=0.\tag{3.30}$$

So we know $w_\tau\in H_0^2(\tau)$. By using the inverse property of the bubble functions, (see [1, 19]), we derive

$$|w_\tau|^2_{H^2(\tau)}\leq Ch_\tau^{-4}\int_\tau|w_\tau(\tau)|^2.\tag{3.31}$$

Let $\widehat{\tau}$ be a reference element and $\widehat{\mathbf{x}}=F_\tau(\mathbf{x})=B_\tau\mathbf{x}+\mathbf{b}_\tau$ be an affine map from τ onto $\widehat{\tau}$ and set $\widehat{w}=w\circ F_\tau^{-1}(\widehat{\mathbf{x}})$. Then we have

$$|w_\tau|^2_{L^2(\tau)}=\int_{\widehat{\tau}}|\widehat{w}_\tau(\widehat{\mathbf{x}})|^2|detB^{-1}|.$$

Since $\widehat{w}_\tau \in H_0^2(\widehat{\tau})$ due to $w_\tau \in H_0^2(\tau)$, by using Poincare inequality, we have

$$\int_{\widehat{\tau}} |\widehat{w}_\tau(\widehat{\mathbf{x}})|^2 |detB^{-1}| \le C \sum_{|\alpha|=2} \int_{\widehat{\tau}} |D^\alpha \widehat{w}_\tau(\widehat{\mathbf{x}})|^2 |detB^{-1}|$$

such that

$$|w_\tau|^2_{L^2(\tau)} \le C h_\tau^4 |w_\tau|^2_{H^2(\tau)}. \tag{3.32}$$

So we obtain

$$c h_\tau^{-2} \|w_\tau\|_{L^2(\tau)} \le |w_\tau|_{H^2(\tau)} \le C h_\tau^{-2} \|w_\tau\|_{L^2(\tau)}, \quad \forall\, \tau \in T^h. \tag{3.33}$$

By use of the bubble function w_τ, (3.29) and (3.33), we have

$$\begin{aligned}
&\int_\tau h_\tau^4 (div(\nabla p_h) + \overline{(y_h - y_d)}|_\tau)^2 \\
&= \int_\tau h_\tau^4 (div(\nabla p_h) + \overline{(y_h - y_d)}|_\tau) w_\tau \\
&= \int_\tau h_\tau^4 (div(\nabla p_h) + y_h - y_d) w_\tau \\
&\quad + \int_\tau h_\tau^4 (-(y_h - y_d) + \overline{(y_h - y_d)}|_\tau) w_\tau \\
&= \int_\tau h_\tau^4 (div(\nabla p_h) + y_h - y_d - div(\nabla p) - (y - y_d)) w_\tau \\
&\quad + \int_\tau h_\tau^4 (-(y_h - y_d) + \overline{(y_h - y_d)}|_\tau) w_\tau \\
&= \int_\tau h_\tau^4 \nabla(p - p_h) \cdot \nabla w_\tau + \int_\tau h_\tau^4 (y_h - y) w_\tau \\
&\quad + \int_\tau h_\tau^4 (-(y_h - y_d) + \overline{(y_h - y_d)}|_\tau) w_\tau \\
&= \int_\tau h_\tau^4 (p_h - p) \Delta w_\tau + \int_\tau h_\tau^4 (y_h - y) w_\tau \\
&\quad + \int_\tau h_\tau^4 (-(y_h - y_d) + \overline{(y_h - y_d)}|_\tau) w_\tau \\
&\le C\Big(\|p - p_h\|^2_{L^2(\tau)} + \|y_h - y\|^2_{L^2(\tau)} \\
&\quad + \int_\tau h_\tau^4 (-(y_h - y_d) + \overline{(y_h - y_d)}|_\tau)^2 \Big)^{1/2} \\
&\quad \cdot \Big(h_\tau^4 \|w_\tau\|^2_{L^2(\tau)} + h_\tau^8 |w_\tau|^2_{H^2(\tau)} \Big)^{1/2} \\
&\le C h_\tau^2 \|w_\tau\|_{L^2(\tau)} \Big(\|p - p_h\|^2_{L^2(\tau)} + \|y_h - y\|^2_{L^2(\tau)} \\
&\quad + \int_\tau h_\tau^4 (-(y_h - y_d) + \overline{(y_h - y_d)}|_\tau)^2 \Big)^{1/2} \\
&\le C\Big(\|p - p_h\|^2_{L^2(\tau)} + \|y_h - y\|^2_{L^2(\tau)} \\
&\quad + \int_\tau h_\tau^4 (-(y_h - y_d) + \overline{(y_h - y_d)}|_\tau)^2 \Big) \\
&\quad + \frac{h_\tau^4}{2} \|div(\nabla p_h) + \overline{(y_h - y_d)}|_\tau\|^2_{L^2(\tau)},
\end{aligned}$$

which lead to

$$\int_\tau h_\tau^4(div(\nabla p_h) + \overline{(y_h - y_d)}|_\tau)^2$$
$$\leq C\Big(\|p - p_h\|^2_{L^2(\tau)} + \|y - y_h\|^2_{L^2(\tau)} + \int_\tau h_\tau^4(y_h - y_d - \overline{(y_h - y_d)}|_\tau)^2\Big)$$

such that

$$\int_\tau h_\tau^4(div(\nabla p_h) + y_h - y_d|_\tau)^2$$
$$\leq 2\Big(\int_\tau h_\tau^4(div(\nabla p_h) + \overline{(y_h - y_d)}|_\tau)^2 + \int_\tau h_\tau^4(y_h - y_d - \overline{(y_h - y_d)}|_\tau)^2\Big)$$
$$\leq C\Big(\|p - p_h\|^2_{L^2(\tau)} + \|y - y_h\|^2_{L^2(\tau)} + \int_\tau h_\tau^4(y_h - y_d - (\overline{y_h} - \overline{y_d}))^2\Big).$$

This ends the proof of lemma 3.7. ■

Lemma 3.8. *Let* (y, u) *and* (y_h, u_h) *be the solutions of* (OCP) *and* (OCP)^h, *and* p *and* p_h *be the solutions of the co-state equations* (2.5) *and* (2.11) *respectively. Then,*

$$\sum_{l\cap\partial\Omega=\emptyset} \int_l h_l^3[(\nabla p_h \cdot n)]^2 \leq C\{\|p - p_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{L^2(\Omega)} + \epsilon_2^2\}. \tag{3.34}$$

Proof. We need to introduce new bubble functions. Let b_l be the four order polynomial for l, where $l = \partial\tau_1 \cap \partial\tau_2$, see Figure 1, as follows:

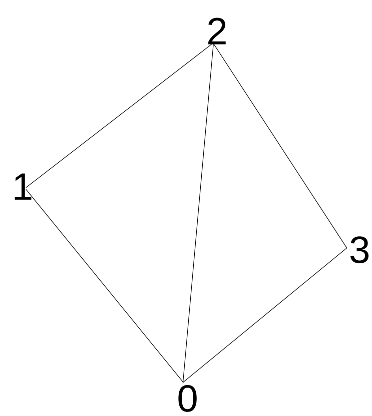

FIGURE 1. triangle $\tau_1 \cup \tau_2$

Define

$$\widehat{\lambda}_0 = \frac{\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix}}{\begin{vmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} + \begin{vmatrix} x_0 & y_0 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}$$

$$\widehat{\lambda}_2 = \frac{\begin{vmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x & y & 1 \end{vmatrix}}{\begin{vmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} + \begin{vmatrix} x_0 & y_0 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}$$

$$\widehat{\lambda}'_0 = \frac{\begin{vmatrix} x & y & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}{\begin{vmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} + \begin{vmatrix} x_0 & y_0 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}$$

$$\widehat{\lambda}'_2 = \frac{\begin{vmatrix} x_0 & y_0 & 1 \\ x & y & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}{\begin{vmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} + \begin{vmatrix} x_0 & y_0 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}$$

and

$$b_l = \begin{cases} \widehat{\lambda}_0 \widehat{\lambda}_2 \widehat{\lambda}'_0 \widehat{\lambda}'_2, & \tau_1 \cup \tau_2, \\ \\ 0, & \Omega \backslash \tau_1 \cup \tau_2. \end{cases}$$

Introduce the new bubble function:

$$w_l = c'_1 [(\nabla p_h \cdot n_l)] b_l^2, \quad c'_1 = \frac{\int_l 1}{\int_l b_l^2}. \tag{3.35}$$

Following the standard scaling arguments, it is not difficult to show that

$$\|w_l\|^2_{L^2(l)} = \int_l {c'_1}^2 [\nabla p_h \cdot n_l]^2 b_l^4 \le c'_1 \int_l [\nabla p_h \cdot n_l]^2. \tag{3.36}$$

Let $\varpi_l = \tau_1 \cup \tau_2$. Since $\widehat{\lambda}_0 \widehat{\lambda}_2 \widehat{\lambda}'_0 \widehat{\lambda}'_2 = 0$ on $\partial \varpi_l$, we have

$$w_l = \frac{\partial w_l}{\partial n_l} = \frac{\partial w_l}{\partial s_l} = 0, \quad on\ \partial \varpi_l, \tag{3.37}$$

where n_l and s_l are the unit vector normal to $\partial \varpi_l$ outwards ϖ_l and the tangent vector along $\partial \varpi_l$. So we have $w_l \in H_0^2(\varpi_l)$. Similarly, from the standard scaling arguments, it is easily shown that

$$\|w_l\|_{L^2(\tau_1 \cup \tau_2)} \le C h_l^{\frac{1}{2}} \|w_l\|_{L^2(l)} \tag{3.38}$$

and

$$c h_l^{-2} \|w_l\|_{L^2(\tau_1 \cup \tau_2)} \le \|w_l\|_{H^2(\tau_1 \cup \tau_2)} \le C h_l^{-2} \|w_l\|_{L^2(\tau_1 \cup \tau_2)}. \tag{3.39}$$

By use of the bubble function w_l, (3.36), (3.38) and (3.39), we have

$$\begin{aligned} \int_l h_l^3 [\nabla p_h \cdot n]^2 &= \int_l h_l^3 [\nabla p_h \cdot n] w_l \\ &= \int_l h_l^3 [(\nabla p_h \cdot n) - (\nabla p \cdot n)] w_l \\ &= \int_{\partial \tau_l^1 \cup \partial \tau_l^2} h_l^3 ((\nabla p_h \cdot n) - (\nabla p \cdot n)) w_l \end{aligned}$$

$$= \int_{\tau_l^1 \cup \tau_l^2} h_l^3 \nabla(p_h - p) \cdot \nabla w_l + \int_{\tau_l^1 \cup \tau_l^2} h_l^3 div(\nabla(p_h - p)) w_l$$
$$= \int_{\tau_l^1 \cup \tau_l^2} h_l^3 (p - p_h) \Delta w_l + \int_{\tau_l^1 \cup \tau_l^2} h_l^3 (div(\nabla p_h) + y_h - y_d) w_l$$
$$+ \int_{\tau_l^1 \cup \tau_l^2} h_l^3 (y - y_h) w_l$$
$$\leq C\Big(\|p_h - p\|_{L^2(\tau_1 \cup \tau_2)} + \|y_h - y\|_{L^2(\tau_1 \cup \tau_2)}$$
$$+ \|h_l^2 (div(\nabla p_h) + y_h - y_d)\|_{L^2(\tau_1 \cup \tau_2)} \Big)$$
$$\cdot \Big(h_l \|w_l\|_{L^2(\tau_l \cup \tau_2)} + h_l^3 |w_l|_{H^2(\tau_1 \cup \tau_2)} \Big)$$
$$\leq C h_l^{\frac{3}{2}} \|w_l\|_{L^2(l)} \Big(\|p_h - p\|_{L^2(\tau_1 \cup \tau_2)} + \|y_h - y\|_{L^2(\tau_1 \cup \tau_2)}$$
$$+ \|h_l^2 (div(\nabla p_h) + y_h - y_d)\|_{L^2(\tau_1 \cup \tau_2)} \Big)$$
$$\leq C\Big(\|p_h - p\|^2_{L^2(\tau_1 \cup \tau_2)} + \|y_h - y\|^2_{L^2(\tau_1 \cup \tau_2)}$$
$$+ \int_{\tau_1 \cup \tau_2} h_l^4 (div(\nabla p_h) + y_h - y_d)^2 \Big) + \frac{1}{2} \int_l h_l^3 (\nabla p_h \cdot n_l)^2.$$

Hence,

$$\sum_{l \cap \partial\Omega = \emptyset} \int_l h_l^3 [(\nabla p_h \cdot n)]^2$$
$$\leq C\Big\{ \|p_h - p\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{L^2(\Omega)} + \sum_{\tau \in T^h} \int_\tau h_\tau^4 (div(\nabla p_h) + y_h - y_d)^2 \Big\}. \tag{3.40}$$

Applying (3.27) into (3.40) leads to (3.34). ■

As a corollary of Lemmas 3.7 and 3.8 , we have

Lemma 3.9. *Let* (y, u) *and* (y_h, u_h) *be the solutions of* (OCP) *and* $(\mathrm{OCP})^h$, *and* p *and* p_h *be the solutions of the co-state equations* (2.5) *and* (2.11) *respectively. Then,*

$$\eta_2^2 \leq C\Big\{ \|p - p_h\|^2_{L^2(\Omega)} + \|y - y_h\|^2_{L^2(\Omega)} + \epsilon_2^2 \Big\}. \tag{3.41}$$

Similarly, we can prove the following lower bound estimate for η_3.

Lemma 3.10. *Let* (y, u) *and* (y_h, u_h) *be the solutions of* (OCP) *and* $(\mathrm{OCP})^h$, *and* p *and* p_h *be the solutions of the co-state equations* (2.5) *and* (2.11) *respectively. Then,*

$$\eta_3^2 \leq C\Big\{ \|y - y_h\|^2_{L^2(\Omega)} + \|u - u_h\|^2_{L^2(\Omega)} + \epsilon_3^2 \Big\}, \tag{3.42}$$

where η_3 *and* ϵ_3 *are defined in Theorem 3.1.*

Then Theorem 3.2 follows from Lemmas 3.6, 3.9 and 3.10.

4. Numerical Experiments

In this section, we carry out some numerical experiments to demonstrate the a posteriori error estimators developed in Section 3. We wish to emphasize that the main purpose of this paper is to compute the optimal control effectively, not the states. In these cases, L^2 error of the control and states is more important than their H^1 error. We consider the following control problem on $\Omega = (0,1)^2$:

$$\begin{cases} \min \dfrac{1}{2}\displaystyle\int_\Omega (y-y_d)^2 dx + \dfrac{1}{2}\int_\Omega (u-u_0)^2 dx \\ s.t. \quad -\Delta y = u + f, \quad \displaystyle\int_\Omega u \geq 0. \end{cases} \tag{4.1}$$

To solve the optimal control numerically, we used the following iterations:

ALGORITHM. *Give an initial control $u_h^0 \in K^h$. Then for $k = 0, 1, 2, \cdots\cdots$, seek $(u_h^k, y_h^k, p_h^k) \in V^h \times V^h \times U^h$, in iteration, such that*

$$\begin{cases} a(y_h^k, w_h) = (f + u_h^{k-1}, w_h), \quad \forall\ w_h \in V^h, \\ a(q_h, p_h^k) = (y_h^k - y_d, q_h), \quad \forall\ q_h \in V^h, \\ (u_h^k + p_h^k, v_h - u_h^k) \geq 0, \quad \forall\ v_h \in K^h. \end{cases} \tag{4.2}$$

The last inequality was solved by the explicit formula:

$$u_h^k = -P_h p_h^k + \max\left\{0, \overline{p_h^k}\right\},$$

where P_h is the L^2-projection from $L^2(\Omega)$ to U^h.

The proof of convergence of the iteration can be found in [10].

We performed two different examples. In Example 1, the optimal control is smooth, so the adaptivity will not contribute greatly to the computation savings. But it is clear that the multi-meshes still save much computational work. In Example 2 we compare effectiveness of different error estimators. We use the indicators given in [9], which is equivalent in energy norm, and the indicators derived in Section 3.

In computing these examples, we used the software package: AFEpack, see [12] for the details.

Example 1. In the first example, the data and solutions are:

$$\begin{aligned} y|_{\partial\Omega} &= 0 \\ p &= \sin \pi x_1 \sin \pi x_2 \\ u_0 &= 0 \\ u &= \max(\bar{p}, 0) - p \\ y_d &= 0 \\ f &= 4\pi^4 p + p - \frac{4}{\pi^2} \\ y &= 2\pi^2 p + y_d \end{aligned} \tag{4.3}$$

In Example 1, the state and co-state are approximated by the piecewise linear elements, while piecewise constant elements are used to approximate the control. We compute Example 1 on a uniform mesh and an adaptive mesh respectively. Two numerical experiments with different numbers of nodes are performed. Numerical results are presented in Tables 1-2 respectively. In Tables 1-2, the mesh information is displayed with L^2 approximation errors for the control and the states.

TABLE 1. Piecewise constant element approximation for control

	u, y, p on uniform mesh			u, y, p on adaptive mesh		
	u	y	p	u	y	p
# nodes	24993	24993	24993	24991	3737	3737
L^2 error	3.112e-03	5.862e-04	5.822e-05	3.091e-03	6.274e-03	6.050e-04

TABLE 2. Piecewise constant element approximation for control

	u, y, p on uniform mesh			u, y, p on adaptive mesh		
	u	y	p	u	y	p
# nodes	131713	131713	131713	122157	6362	6362
L^2 error	1.404e-03	1.108e-04	1.087e-05	1.402e-03	3.757e-03	3.533e-04

In this example, the optimal control is quite smooth so that there was no much difference in using either the uniform or adaptive meshes to approximate the control. However the multiple meshes can still save much computational work in this case. In fact it can be clearly seen that on the multiple adaptive meshes one may use 10 times fewer degree of freedoms(DOFs) in the state variables to produce a given L^2 control error reduction. 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. It is important to note that if one used just one set of adaptive meshes, then much more degree of freedoms has to be used for solving the state and the costate.

As mentioned in Introduction, the indicators equivalent in H^1 norm may produce over-refinement when the L^2 norm of the approximation error is used as the stopping criteria, although this is not always visible. In the following example, we try to illustrate this by using the H^1 equivalent indicator derived in [9] and the L^2 equivalent indicator derived here. The control problem in Example 2 is almost identical to that of [9], except that here the costate p was multiplied by a factor of 10. Without this factor, both indicators perform similarly and lead to similar meshes both for the control and the states. With the multiplier, the difference of the performance of the two indicators can be clearly seen from the results below:

Example 2. In the second example, the data and solutions are:

$$z = \begin{cases} 0.5, & x_1 + x_2 > 1.0 \ , \\ 0, & x_1 + x_2 \le 1.0 \ , \end{cases}$$

$$\begin{aligned} p &= \beta \sin \pi x_1 \sin \pi x_2 \ , \\ u_0 &= 1.0 - \sin \frac{\pi x_1}{2} - \sin \frac{\pi x_2}{2} + z \ , \\ u &= \max(p - u_0, 0) + u_0 - p \ , \\ y_d &= 0 \ , \\ f &= 4\pi^4 p - u \ , \\ y &= 2\pi^2 p + y_d \ . \end{aligned} \tag{4.4}$$

where β is a positive constant given late. In this case, the control is discontinuous so that it has much weaker global regularity than co-state.

In Example 2, the state and co-state are approximated by the conforming piecewise linear elements, while the piecewise discontinuous linear elements are used to approximate the control. We compute Example 2 on two different adaptive meshes, one is produced by the indicators given in [9], and another is produced by the indicators given in this article. We consider two cases of $\beta = 0.5$ and $\beta = 5$. In every case, two numerical experiments with different numbers of nodes are performed. Numerical results are presented in Tables 3-4 and 5-6 respectively, in which, the mesh information is displayed with L^2 approximation errors for the control and the state.

Table 3. Piecewise linear element approximation for control with $\beta = 0.5$

	u, y, p on adaptive mesh produced by the indicators in [9]			u, y, p on adaptive mesh produced by the indicators in this article		
	u	y	p	u	y	p
# nodes	4693	1542	1542	4847	517	517
L^2 error	9.799e-03	8.562e-03	8.024e-04	9.797e-03	1.548e-02	1.523e-03

Table 4. Piecewise linear element approximation for control with $\beta = 0.5$

	u, y, p on adaptive mesh produced by the indicators in [9]			u, y, p on adaptive mesh produced by the indicators in this article		
	u	y	p	u	y	p
# nodes	16156	5597	5597	16067	1939	1939
L^2 error	4.826e-03	2.393e-03	2.252e-04	4.817e-03	3.979e-03	3.892e-04

From Tables 3-4, we see that the numbers of nodes of meshes produced by indicators given in [9] for the states is more triple of ones by the adaptive indicators given in this article. Further, we investigate the case of $\beta = 5$.

Table 5. Piecewise linear element approximation for control with $\beta = 5$

	u, y, p on adaptive mesh produced by the indicators in [9]			u, y, p on adaptive mesh produced by the indicators in this article		
	u	y	p	u	y	p
# nodes	3720	17105	17105	4159	1939	1939
L^2 error	1.420e-02	1.472e-02	1.336e-03	1.417e-02	3.973e-02	3.889e-03

TABLE 6. Piecewise linear element approximation for control with $\beta = 5$

	u, y, p on adaptive mesh produced by the indicators in [9]			u, y, p on adaptive mesh produced by the indicators in this article		
	u	y	p	u	y	p
# nodes	10562	18385	18385	11701	1939	1939
L^2 error	7.567e-03	1.446e-02	1.306e-03	7.554e-03	3.973e-02	3.889e-03

From Tables 5-6, we see that the numbers of nodes of meshes produced by indicators given in [9] for the states become almost ten-multiple of ones by the adaptive indicators given in this article. These numerical results show that the indicator derived here performs better than that in [9].

FIGURE 2. The adaptive meshes for the state and co-state with different indicators

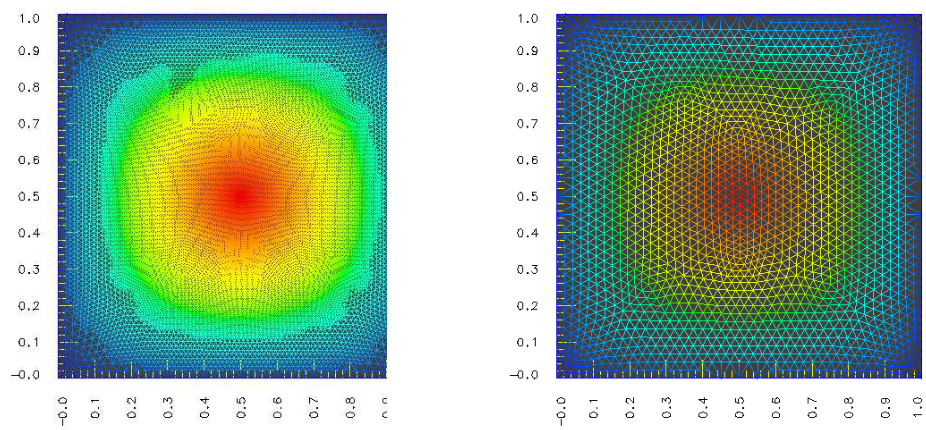

Figure 2 displays the adaptive meshes of the state and co-state in Table 5. The left mesh is produced by using the indicators derived in [9], and the right mesh is produced by using the adaptive indicators derived in this article. It was found that the adaptive meshes guided by the indicators in Section 3 can further reduce the DOFs to produce a given L^2 control error reduction as shown in Table 5 and Table 6. So we can see that the new indicator is more suitable for the application where the L^2-error is more important.

Acknowledgments

We wish to express our thanks to referees for their useful suggestions and comments.

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. BERGOUNIOUX, M. HADDOU, M. HINTERMULLER AND K. KUNISCH, *A comparison of a Moreau-Yosida based active set strategy and interior point methods for constrained optimal control problems*, SIAM J.Optim. Vol.11, pp.495-521, 2000.

[3] R. Becker and R. Rannacher, *An optimal control approach to a-posteriori error estimation*, In A. Iserles, editor, Acta Numerica 2001, pages 1-102. Cambridge University Press, 2001.
[4] R. Becker, H. Kapp and R. Rannacher, *Adaptive finite element methods for optimal control of partial differential equations: Basic concept*, SIAM J. Control Optimization, Vol. 39 (2000), pp. 113-132.
[5] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
[6] W. Dorfler , *A Convergent Adaptive Algorithm for Poisson's Equation*, SIAM J. Numer. Anal., Vol. 33 (1996), pp. 1106-1124.
[7] M. D. Gunzburger and S. L. Hou, *Finite dimensional approximation of a class of constrained nonlinear control problems*, SIAM J.Control and Optimization, Vol. 34 (1996), 1001-1043.
[8] L. Ge, WB. Liu and D.P. Yang *Equivalent a Posteriori Error Estimates for A Constrained Optimal Control Problem*, submitted.
[9] L. Ge, WB. Liu and D.P. Yang *Adaptive Finite Element Approximation for A Constrained Optimal Control Problem*, submitted.
[10] Y. Q. Huang, R. Li, WB. Liu, and N. N. Yan, *Adaptive multi-mesh finite element approximation for constrained optimal control*, Submitted to SIAM Optim.Control (first in 2004).
[11] A. Kufner, O. John and S. Fucik, *Function Spaces*, Nordhoff, Leiden, The Netherlands, 1977.
[12] R. Li *On Multi-Mesh h-Adaptive Algorithm*, JSC, Vol.24, pp.321-341, 2005.
[13] R. Li, WB. Liu, H.P. Ma and T. Tang, *Adaptive finite element approximation of elliptic optimal control*, SIAM, J. Control. Optim., vol 41, pp. 1321-1349, 2002.
[14] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.
[15] WB. Liu and N. N. Yan, *A posteriori error estimates for convex boundary control problems*, SIAM Numer. Anal., Vol.39 (2001), pp. 73-99.
[16] WB. Liu and N. N. Yan, *A posteriori error analysis for convex distributed optimal control problems*, Adv. Comp. Math., Vol.15, No.1-4 (2001), pp. 285-309.
[17] L.R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[18] A. Veeser, *Efficient and reliable a posteriori error estimators for elliptic obstacle problems*, SIAM J. Numer. Anal., Vol. 39(2001), no. 1, pp. 146-167.
[19] R. Verfurth, *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement*, Wiley-Teubner, London, UK, 1996.

Department of Mathematics, Shandong University, Jinan, Shandong, 250100, China
E-mail: geliangsdu@mail.sdu.edu.cn

KBS and IMS, University of Kent, Canterbury, CT2 7NF, England
E-mail: W.B.Liu@kent.ac.uk

Department of Mathematics, East China Normal University, Shanghai, 200241, China
E-mail: dpyang@math.ecnu.edu.cn