SIAM J. CONTROL OPTIM.
Vol. 41, No. 5, pp. 1321–1349
© 2002 Society for Industrial and Applied Mathematics

ADAPTIVE FINITE ELEMENT APPROXIMATION FOR DISTRIBUTED ELLIPTIC OPTIMAL CONTROL PROBLEMS*

RUO LI†, WENBIN LIU‡, HEPING MA§, AND TAO TANG¶

Abstract. In this paper, sharp a posteriori error estimators are derived for a class of distributed elliptic optimal control problems. These error estimators are shown to be useful in adaptive finite element approximation for the optimal control problems and are implemented in the adaptive approach. Our numerical results indicate that the sharp error estimators work satisfactorily in guiding the mesh adjustments and can save substantial computational work.

Key words. mesh adaptivity, optimal control, a posteriori error estimate, finite element method

AMS subject classifications. 49J20, 65N30

PII. S0363012901389342

1. Introduction.

Finite element approximation of optimal control problems has long been an important topic in engineering design work and has been extensively studied in the literature. There have been extensive theoretical and numerical studies for finite element approximation of various optimal control problems; see [2, 12, 13, 15, 20, 23, 37, 44]. For instance, for the optimal control problems governed by some linear elliptic or parabolic state equations, a priori error estimates of the finite element approximation were established long ago; see, for example, [12, 13, 15, 20, 23, 37]. Furthermore, a priori error estimates were established for the finite element approximation of some important flow control problems in [17] and [11]. A priori error estimates have also been obtained for a class of state constrained control problems in [43], though the state equation is assumed to be linear. In [29], this assumption has been removed by reformulating the control problem as an abstract optimization problem in some Banach spaces and then applying nonsmooth analysis. In fact, the state equation there can be a variational inequality.

In recent years, the adaptive finite element method has been extensively investigated. Adaptive finite element approximation is among the most important means to boost the accuracy and efficiency of finite element discretizations. It ensures a higher density of nodes in a certain area of the given domain, where the solution is more difficult to approximate. At the heart of any adaptive finite element method is an a posteriori error estimator or indicator. The literature in this area is extensive. Some of the techniques directly relevant to our work can be found in [1, 5, 6, 7, 28, 32, 34, 42, 47]. It is our belief that adaptive finite element enhancement is one of the future directions to pursue in developing sophisticated numerical methods for optimal design problems.

*Received by the editors May 14, 2001; accepted for publication (in revised form) April 28, 2002; published electronically December 11, 2002. This work was supported in part by Hong Kong Baptist University, Hong Kong Research Grants Council, and the British EPSRC.

http://www.siam.org/journals/sicon/41-5/38934.html

†School of Mathematical Science, Peking University, Beijing 100871, People's Republic of China.

‡CBS & Institute of Mathematics and Statistics, The University of Kent, Canterbury, CT2 7NF England (W.B.Liu@ukc.ac.uk).

§Department of Mathematics, Shanghai University, Shanghai 200436, People's Republic of China (hpma@guomai.sh.cn).

¶Department of Mathematics, The Hong Kong Baptist University, Kowloon Tong, Hong Kong (ttang@math.hkbu.edu.hk).

Although adaptive finite element approximation is widely used in numerical simulations, it has not yet been *fully* utilized in optimal control. Initial attempts in this aspect have been reported only recently for some design problems; see, e.g., [3, 4, 38, 41]. However, a posteriori error indicators of a heuristic nature are widely used in most applications. For instance, in some existing work on adaptive finite element approximation of optimal design, the mesh refinement is guided by a posteriori error estimators *solely* from the state equation (or the displacement) for a fixed control (or design). Thus error information from approximation of the control (design) is not utilized. Although these methods may work well in some particular applications, they cannot be applied confidently in general. It is unlikely that the potential power of adaptive finite element approximation has been fully utilized due to the lack of more sophisticated a posteriori error indicators.

Very recently, some error estimators of residual type were developed in [8, 9, 30, 31, 33]. These error estimators are based on a posteriori estimation of the discretization error for the state *and* the control (design). When there is no constraint in a control problem, normally the optimality conditions consist of coupled partial differential equations only. Consequently, one may be able to write down the dual system of the *whole* optimality conditions and then apply the weighted a posteriori error estimation technique to obtain a posteriori estimators for the *objective functional* approximation error of the control problem; see [8, 9]. In many applications (like parameter estimation), it is more interesting to obtain a posteriori error estimators for the control approximation error [22]. Furthermore, there frequently exist some constraints for the control in applications. In such cases, the optimality conditions often contain a variational inequality and then have some very different properties. Thus it does not always seem to be possible to apply the techniques used in [8, 9] to constrained control problems.

In our work, constrained cases are studied via residual estimation using the norms of energy type. A posteriori error estimators are derived for quite general constrained control problems governed by the elliptic equations (see [30, 31, 33]) with upper error bounds. However, these error estimators have yet to be applied to adaptive finite element methods. Indeed, numerical experiments indicated that these estimators tend to over-refine the computational meshes. Thus the resulting computational meshes may not be efficient in reducing approximation errors. It seems that one has to derive sharper error estimators in order to obtain more efficient meshes. This seems to be possible at least for a class of control problems, which are frequently met in applications. More details on these will be given in section 3.

In this paper, we consider the convex optimal control problem

$$\begin{cases} \min\limits_{u\in K}\{g(y)+h(u)\}, \\ \\ -\mathrm{div}(A\nabla y) = f + Bu \quad \text{in } \Omega,\ y|_{\partial\Omega} = 0, \end{cases} \tag{1.1}$$

where g and h are given convex functionals, K is a closed convex set, and B is a continuous linear operator. The details will be specified later. The main objective of this work is to derive sharp a posteriori error estimators for some frequently met optimal control problems. A number of new techniques have to be introduced in order to obtain such estimators. Our numerical tests indicate that these improved error estimators indeed lead to efficient computational meshes.

The paper is organized as follows: In section 2, we describe the finite element approximation for the convex optimal control problem (1.1). In section 3, we derive

error estimates for the problem with an obstacle constraint. Both upper bounds and lower bounds are established with attention on their equivalence. In section 4, numerical experiments will be carried out, with particular attention to testing the influence of various indicators on the mesh construction.

2. The elliptic optimal control problem and finite element approximation. In this section, we describe the elliptic optimal control problem and its finite element approximation. Let Ω and Ω_U be two bounded open sets in $\mathbf{R}^n (n \leq 3)$ with the Lipschitz boundaries $\partial\Omega$ and $\partial\Omega_U$. We denote by $C^0(\bar{\Omega})$ the space of continuous functions on $\bar{\Omega}$. We adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{m,q,\Omega}$ and seminorm $|\cdot|_{m,q,\Omega}$(see (1.2) of [16, p. 2]). For $q = 2$, we denote $W^{m,2}(\Omega)$ by $H^m(\Omega)$ with norm $\|\cdot\|_{m,\Omega} := \|\cdot\|_{m,2,\Omega}$ and seminorm $|\cdot|_{m,\Omega} := |\cdot|_{m,2,\Omega}$. We set $H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$. In addition, c or C denotes a general positive constant independent of h.

In the rest of the paper, we shall take the state space $Y = H_0^1(\Omega)$, the control space $U = L^2(\Omega_U)$ with the inner product $(\cdot,\cdot)_U$, and $H = L^2(\Omega)$ with the inner product $(\cdot,\cdot)$. We wish to study the finite element approximation of the distributed elliptic convex optimal control problem (1.1). Assume that g and h are convex functionals which are continuously differentiable on $H = L^2(\Omega)$ and $U = L^2(\Omega_U)$, respectively, and h is further strictly convex. Suppose that K is a closed convex set in the control space U, $f \in L^2(\Omega)$, B is a continuous linear operator from U to $H \subset Y'$ (the dual space of Y), and

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

such that there is a constant $c > 0$ satisfying, for any vector $\xi \in \mathbf{R}^n$,

$$(A\xi)\cdot\xi \geq c|\xi|^2.$$

We further assume that $h(u) \to +\infty$ as $\|u\|_{0,\Omega_U} \to \infty$, the functional $g(\cdot)$ is bounded below, and

$$|(g'(v) - g'(w), q)| \leq C\|v - w\|_{1,\Omega}\|q\|_{1,\Omega} \qquad \forall v, w, q \in Y. \tag{2.1}$$

To consider the finite element approximation of the above optimal control problem, here we give it a weak formula

$$\text{(CCP)} \begin{cases} \min\limits_{u\in K}\{g(y) + h(u)\}, \\ \\ a(y, w) = (f + Bu, w) \quad \forall w \in Y = H_0^1(\Omega), \end{cases} \tag{2.2}$$

where

$$a(v, w) = \int_\Omega (A\nabla v)\cdot\nabla w \quad \forall v, w \in H^1(\Omega),$$
$$(f, w) = \int_\Omega fw \quad \forall f, w \in L^2(\Omega).$$

Under these assumptions, the control problem (CCP) has a unique solution (y, u), and a pair (y, u) is the solution of (CCP) if and only if there is a costate $p \in Y$ such

that the triplet (y, p, u) satisfies the following optimality conditions (see [27]):

$$\text{(CCP-OPT)}\quad \begin{cases} a(y,w) = (f + Bu, w) \quad \forall w \in Y = H_0^1(\Omega), \\ a(q,p) = (g'(y), p) \quad \forall q \in Y = H_0^1(\Omega), \\ (h'(u) + B^*p, v - u)_U \geq 0 \quad \forall v \in K \subset U = L^2(\Omega_U), \end{cases} \tag{2.3}$$

where B^* is the adjoint operator of B and g' and h' are the derivatives of g and h. Here g' and h' have been viewed as functions in $H = L^2(\Omega)$ and $U = L^2(\Omega_U)$, respectively, using the well-known representation theorem in a Hilbert space.

Let us consider the finite element approximation of the above control problem. For ease of exposition, we consider only n-simplex, conforming Lagrange elements. Also, we assume that Ω and Ω_U are polygonal. Let T^h be a partitioning of Ω into disjoint open regular n-simplices τ so that $\bar{\Omega} = \cup_{\tau \in T^h} \bar{\tau}$. Each element has at most one face on $\partial\Omega$, and the adjoining elements $\bar{\tau}$ and $\bar{\tau}'$ have either only one common vertex or a whole edge or a whole face if τ and $\tau' \in T^h$. Let h_τ denote the diameter of the element τ in T^h. Associated with T^h is a finite dimensional subspace S^h of $C^0(\bar{\Omega})$ such that $v_h|_\tau$ are polynomials of k-order ($k \geq 1$) for all $v_h \in S^h$ and $\tau \in T^h$. Denote $\{P_i\}$ $(i = 1, 2, \ldots, J)$ the vertex set associated with T^h. Let $Y^h = V_0^h := S^h \cap Y$.

Similarly, we have a regular partitioning of Ω_U, and we use the following corresponding notation: T_U^h, τ_U, h_{τ_U} and P_i^U $(i = 1, 2, \ldots, J_U)$. Associated with T_U^h is another finite dimensional subspace W_U^h of $L^2(\Omega_U)$ such that $v_h|_{\tau_U}$ are polynomials of k-order ($k \geq 0$) for all $v_h \in W_U^h$ and $\tau_U \in T_U^h$. Note here that there is no requirement for the continuity or boundary conditions. Let $U^h = W_U^h \subset U = L^2(\Omega_U)$.

Due to the limited regularity of the optimal control u (at most in $H^1(\Omega_U)$ in general), here we will consider only the piecewise constant space for the control approximation, while higher-order finite spaces may be used for the state and costate.

Then a possible finite element approximation of (CCP) is as follows:

$$(\text{CCP})^h \quad \begin{cases} \min\limits_{u_h \in K^h} \{g(y_h) + h(u_h)\}, \\ a(y_h, w_h) = (f + Bu_h, w_h) \quad \forall w_h \in Y^h, \end{cases} \tag{2.4}$$

where K^h is a closed convex set in U^h such that there are $v_h \in K^h$ converging to an element $v \in K$ in U. It follows that the control problem $(\text{CCP})^h$ has a unique solution (y_h, u_h) and that a pair $(y_h, u_h) \in Y^h \times U^h$ is the solution of $(\text{CCP})^h$ if and only if there is a costate $p_h \in Y^h$ such that the triplet (y_h, p_h, u_h) satisfies the following optimality conditions:

$$(\text{CCP-OPT})^h \quad \begin{cases} a(y_h, w_h) = (f + Bu_h, w_h) \quad \forall w_h \in Y^h \subset Y = H_0^1(\Omega), \\ a(q_h, p_h) = (g'(y_h), q_h) \quad \forall q_h \in Y^h \subset Y = H_0^1(\Omega), \\ (h'(u_h) + B^*p_h, v_h - u_h)_U \geq 0 \quad \forall v_h \in K^h \subset U = L^2(\Omega_U). \end{cases} \tag{2.5}$$

It follows that (y_h, p_h, u_h) is uniformly bounded in $Y \times Y \times U$. This is because $g(y_h) + h(u_h)$ is uniformly bounded due to the above assumption on K^h. Thus $\|u_h\|_U$ is also uniformly bounded. Then it follows from (2.5) and (2.1) that $\|y_h\|_Y$ and $\|p_h\|_Y$ are uniformly bounded.

The finite element approximation solution must be solved by using some mathematical programming algorithms such as the conjugate gradient method, the interior point method, and the SQP algorithms. This is a very active research area and is too large to be reviewed here even very briefly. Some of the recent progress in this area has been summarized in [14].

3. Sharp a posteriori error estimators. Deriving a posteriori error estimators for the finite element approximation of the control problem (CCP) is not an easy task since the triplet (y, p, u) is the solution of the coupled system (CCP-OPT). Although there is much work on a priori error estimates for finite element approximation of optimal control problems, as seen in section 1, there are substantial differences between a priori error estimates and a posteriori error estimates for such control problems. Only very recently, some a posteriori error estimators have been derived in the literature. For the control problem (CCP), for instance, the following error estimators have been derived in [31] and [35], assuming that $(h'(u_h) + B^*p_h)|_{\tau_U} \in H^1(\tau_U)$ for any $\tau_U \in T_U^h$:

$$\|u_h - u\|_{0,\Omega_U}^2 + \|y_h - y\|_{1,\Omega}^2 + \|p_h - p\|_{1,\Omega}^2 \le C\left(\bar{\eta}_1^2 + \sum_{i=2}^{5} \eta_i^2\right) = C\bar{\eta}^2, \tag{3.1}$$

where

$$\begin{aligned}
\bar{\eta}_1^2 &= \sum_{\tau_U \in T_U^h} h_{\tau_U}^2 \|\nabla(h'(u_h) + B^*p_h)\|_{0,\tau_U}^2, \\
\eta_2^2 &= \sum_{\tau \in T^h} h_\tau^2 \int_\tau (f + Bu_h + \operatorname{div}(A\nabla y_h))^2, \\
\eta_3^2 &= \sum_{l \in \partial T^h} h_l \int_l [(A\nabla y_h) \cdot \mathbf{n}]^2, \\
\eta_4^2 &= \sum_{\tau \in T^h} h_\tau^2 \int_\tau (g'(y_h) + \operatorname{div}(A^*\nabla p_h))^2, \\
\eta_5^2 &= \sum_{l \in \partial T^h} h_l \int_l [(A^*\nabla p_h) \cdot \mathbf{n}]^2,
\end{aligned} \tag{3.2}$$

where h_l is the diameter of the face l, and the A-normal derivative jump over the interior face l is defined by

$$[(A\nabla v_h) \cdot \mathbf{n}]|_l = ((A\nabla v_h)|_{\partial\tau_l^1} - (A\nabla v_h)|_{\partial\tau_l^2}) \cdot \mathbf{n},$$

with $\mathbf{n}$ being the unit outer normal vector of τ_l^1 on $l = \bar{\tau}_l^1 \cap \bar{\tau}_l^1$. The A^*-normal derivative jump is similarly defined for the transposed matrix A^* of A.

However, major improvements on these error estimators are much needed in order that they can be used to guide mesh adaptivity efficiently in solving the optimal control problem numerically. For example, it does not seem that they are *always* sharp for the constrained cases, and this can be seen from Figures 4.4 and 4.6 in Example 4.1 of section 4, where it is clear that $|u - u_h|$ has a very different profile (the left of Figure 4.4) from that of $\bar{\eta}_1$ (or $\bar{\eta}$) (the left of Figure 4.6). Consequently, the mesh refinement adjustment schemes based on $\bar{\eta}$ may be inefficient. In Example 4.1, the state and costate are very smooth, but the optimal control u has the gradient

jumps across the free boundary, which is the boundary of the zero set $\{x : u(x) = 0\}$, as seen in Figure 4.1. This causes large control approximation errors along the free boundary, as seen from the left of Figures 4.3 and 4.4. Thus an efficient computational mesh for the control should have a higher density of nodes around the free boundary, as those in Figure 4.2. However, the mesh adjustment guided by $\bar{\eta}_1$ did not achieve this goal well, as seen from Figure 4.5. In fact, the resulting mesh even produced a larger approximation error for the control than the uniform mesh of the same size. A sharp error estimator will lead to much more efficient computational meshes, as seen in Figure 4.2.

In this section, we study sharp error estimates for finite element approximation of the convex control problem (CCP). It follows that $\bar{\eta}$ consists of three parts: The part $\bar{\eta}_1$ is contributed from the approximation error of the variational inequality, and $\eta_2^2+\eta_3^2$, $\eta_4^2+\eta_5^2$ result from the approximation error of the state and costate equations. It is well known that $\eta_2^2+\eta_3^2$ and $\eta_4^2+\eta_5^2$ are sharp error estimators for the state and costate equations. Therefore, the key to our purpose is to improve $\bar{\eta}_1^2$. However, it is difficult to derive improved estimates without knowing explicit structures of the control constraint sets K and K^h; the methods and techniques to be developed will depend heavily on these details. Here we derive a posteriori error estimators with both upper bounds and lower bounds for a class of convex sets K of obstacle type, which are most frequently met in real applications. We achieved this by exploring the special structure of the constraint sets. The ideas are applicable to some other control problems, e.g., the boundary control problems.

We shall first consider the constraint of a single obstacle

$$K = \{v \in U : v \geq \phi\}, \qquad K^h = U^h \cap K,$$

and then we will extend the results to more general cases.

We define the coincidence set (contact set) Ω_U^- and the noncoincidence set (noncontact set) Ω_U^+ as follows:

$$\Omega_U^- = \{x \in \Omega_U : u(x) = \phi(x)\}, \quad \Omega_U^+ = \{x \in \Omega_U : u(x) > \phi(x)\}.$$

It can be seen that the inequality in (2.3) is equivalent to the following:

$$h'(u) + B^*p \geq 0, \quad u \geq \phi, \quad (h'(u) + B^*p)(u - \phi) = 0, \quad \text{a.e. in } \Omega_U. \tag{3.3}$$

We shall show that the quantity $(h'(u_h) + B^*p_h)|_{\Omega_U^-}$ can be mostly removed from the error indicator $\bar{\eta}$ in this case, which enables us to obtain sharp error estimates. To make the presentation of our approach clearer and less technical, we shall first derive sharp error estimators containing an a priori quantity and then approximate it using an a posteriori quantity so that the estimators can be easily applied in numerical computations. Let us note that some approximations of a priori quantities are also used in [9].

In the following, we assume that there is a constant $c > 0$ such that

$$(h'(v) - h'(w), v - w)_U \geq c\|v - w\|_{0,\Omega_U}^2 \qquad \forall v, w \in U. \tag{3.4}$$

3.1. Upper error bounds. We first consider the case of a constant obstacle $\phi(x) \equiv \phi_0$. We define

$$\Omega_h^+ = \{\cup\bar{\tau}_U \ : \ \tau_U \subset \Omega_U^+, \ \tau_U \in T_U^h\}, \qquad \Omega_h^- = \{\cup\bar{\tau}_U \ : \ \tau_U \subset \Omega_U^-, \ \tau_U \in T_U^h\},$$
$$\Omega_h^b = \Omega_U \backslash (\Omega_h^+ \cup \Omega_h^-), \qquad \Omega_h^{+b} = \Omega_h^+ \cup \Omega_h^b, \qquad \Omega_h^{-b} = \Omega_h^- \cup \Omega_h^b,$$

and denote by χ_Q the characteristic function of Q. Let ∂T^h be the set consisting of all of the faces l of any $\tau \in T^h$ such that l is not on $\partial\Omega$. Let h_l be the diameter of the face l. We need the following lemmas in deriving residual-type a posteriori error estimates.

LEMMA 3.1 (see [10]). *Let $\pi_h : C^0(\bar{\Omega}) \to S^h$ be the standard Lagrange interpolation operator such that*

$$\pi_h v := \sum_i v(\mathbf{a}_i)\varphi_i,$$

where $\mathbf{a}_i$ are the nodes on $\bar{\Omega}$ and φ_i are the corresponding shape functions. Then, for $m = 0, 1$ and $n/2 < q \le \infty$,

$$\|v - \pi_h v\|_{m,q,\tau} \le C h_\tau^{2-m} |v|_{2,q,\tau} \quad \forall v \in W^{2,q}(\Omega), \tag{3.5}$$

where the constant C depends only on Ω and the minimum angle of the simplices in T^h.

LEMMA 3.2 (see [21]). *For all $v \in W^{1,q}(\Omega), 1 \le q \le \infty$,*

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

We need another operator $\hat{\pi}_h$: the local averaging interpolation operator defined in [42], which can be applied to functions not necessarily continuous, preserves the homogeneous boundary conditions and is stable in the $W^{1,q}$-norm. The full definition of $\hat{\pi}_h$ is rather long. Thus the readers are referred to [42]. Fortunately, we need only to use one of its properties, which is stated in the following lemma.

LEMMA 3.3. *Let $\hat{\pi}_h : W^{1,q}(\Omega) \to S^h$ be the local averaging interpolation operator defined in* (2.13) *of* [42]. *For $m = 0, 1$ and $1 \le q \le \infty$,*

$$|v - \hat{\pi}_h v|_{m,q,\tau} \le C \sum_{\bar{\tau}' \cap \bar{\tau} \ne \emptyset} h_\tau^{1-m} |v|_{1,q,\tau'} \qquad \forall v \in W^{1,q}(\Omega). \tag{3.7}$$

LEMMA 3.4. *Let $\pi_h^a : L^1(\Omega_U) \to W_U^h$ be the integral averaging operator such that*

$$(\pi_h^a v)|_{\tau_U} := \frac{1}{|\tau_U|} \int_{\tau_U} v \qquad \forall \tau_U \in T_U^h.$$

Then, for $m = 0, 1$ and $1 \le q \le \infty$,

$$\|v - \pi_h^a v\|_{0,q,\tau_U} \le C h_{\tau_U}^m |v|_{m,q,\tau_U} \qquad \forall v \in W^{m,q}(\Omega_U). \tag{3.8}$$

Proof. The result is trivial for $m = 0$. For $m = 1$, we note that $\pi_h^a v|_{\tau_U} = v|_{\tau_U}$ if v is a constant on τ_U. Thus (3.8) can be proved by the standard techniques in the finite element method [10]. □

We first give some upper bounds for $u - u_h$ in the L^2-norm and for $y - y_h$, $p - p_h$ in the H^1-norm. We shall use the following inequality:

$$|(Bv, w)| = |(v, B^* w)_U| \le C\|v\|_{0,\Omega_U}\|w\|_{1,\Omega} \qquad \forall v \in U, w \in Y, \tag{3.9}$$

which is held from our assumptions on the operator B.

THEOREM 3.1. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of* (2.3) *and* (2.5), *respectively. Let the obstacle ϕ be a constant ϕ_0. Assume that conditions* (3.4), (2.1), *and* (3.9) *hold, and $(h'(u_h) + B^* p_h)|_{\tau_U} \in H^1(\tau_U)$ for any $\tau_U \in T_U^h$. Then*

$$\|u_h - u\|_{0,\Omega_U}^2 + \|y_h - y\|_{1,\Omega}^2 + \|p_h - p\|_{1,\Omega}^2 \le C \sum_{i=1}^5 \eta_i^2, \tag{3.10}$$

where

$$\eta_1^2 = \sum_{\tau_U \in T_U^h} h_{\tau_U}^2 \|\nabla(h'(u_h) + B^* p_h)\chi_{\Omega_h^{+b}}\|_{0,\tau_U}^2, \tag{3.11}$$

$$\eta_2^2 = \sum_{\tau \in T^h} h_\tau^2 \int_\tau (f + Bu_h + \operatorname{div}(A\nabla y_h))^2,$$

$$\eta_3^2 = \sum_{l \in \partial T^h} h_l \int_l [(A\nabla y_h)\cdot \mathbf{n}]^2,$$

$$\eta_4^2 = \sum_{\tau \in T^h} h_\tau^2 \int_\tau (g'(y_h) + \operatorname{div}(A^*\nabla p_h))^2,$$

$$\eta_5^2 = \sum_{l \in \partial T^h} h_l \int_l [(A^*\nabla p_h)\cdot \mathbf{n}]^2.$$

Proof. We first estimate the error $\|u - u_h\|_{0,\Omega_U}^2$. It follows from the assumption (3.4) and the inequalities (2.3) and (2.5) that, for any $v_h \in K^h$,

$$\begin{aligned} & c\|u - u_h\|_{0,\Omega_U}^2 \\ &\le (h'(u), u - u_h)_U - (h'(u_h), u - u_h)_U \\ &\le (-B^* p, u - u_h)_U - (h'(u_h), u - u_h)_U + (h'(u_h) + B^* p_h, v_h - u_h)_U \\ &= (h'(u_h) + B^* p_h, v_h - u)_U + (B^*(p_h - p), u - u_h)_U. \end{aligned} \tag{3.12}$$

We introduce y_{u_h} and p_{u_h}, defined by

$$a(y_{u_h}, w) = (f + Bu_h, w) \qquad \forall w \in Y, \tag{3.13}$$

$$a(q, p_{u_h}) = (g'(y_{u_h}), q) \qquad \forall q \in Y. \tag{3.14}$$

It follows from (2.3), (3.13), and (3.14) that

$$a(y_{u_h} - y, w) = (B(u_h - u), w) \qquad \forall w \in Y, \tag{3.15}$$

$$a(q, p_{u_h} - p) = (g'(y_{u_h}) - g'(y), q) \qquad \forall q \in Y. \tag{3.16}$$

Taking $w = p_{u_h} - p$ in (3.15) and $q = y_{u_h} - y$ in (3.16), we have, due to the convexity of g,

$$(B(u_h - u), p_{u_h} - p) = (g'(y_{u_h}) - g'(y), y_{u_h} - y) \ge 0.$$

Using (3.12) together with (3.9) gives

$$\begin{aligned} & c\|u - u_h\|_{0,\Omega_U}^2 \\ &\le (h'(u_h) + B^* p_h, v_h - u)_U + (B^*(p_h - p_{u_h}), u - u_h)_U - (p_{u_h} - p, B(u_h - u)) \\ &\le \sum_{\tau_U \in T_U^h} (h'(u_h) + B^* p_h, v_h - u)_{\tau_U} + C\|p_h - p_{u_h}\|_{1,\Omega}^2 + \frac{c}{2}\|u - u_h\|_{0,\Omega_U}^2. \end{aligned}$$

Now take $v_h = \pi_h^a u \in K^h$ defined in Lemma 3.4. Then we have

$$\begin{aligned} (h'(u_h) + B^* p_h, v_h - u)_{\tau_U} &= ((I - \pi_h^a)(h'(u_h) + B^* p_h), (\pi_h^a - I)(u - u_h))_{\tau_U} \\ &\le Ch_{\tau_U}\|\nabla(h'(u_h) + B^* p_h)\|_{0,\tau_U}\|u - u_h\|_{0,\tau_U} \\ &\le Ch_{\tau_U}^2\|\nabla(h'(u_h) + B^* p_h)\|_{0,\tau_U}^2 + \frac{c}{4}\|u - u_h\|_{0,\tau_U}^2. \end{aligned}$$

Noting that $(v_h - u)|_{\tau_U} = (\pi_h^a - I)u|_{\tau_U} = 0$ for any $\tau_U \in \Omega_U \backslash \Omega_h^{+b}$, we obtain

(3.17)
$$\begin{aligned}\|u - u_h\|_{0,\Omega_U}^2 &\leq C \sum_{\tau_U \in \Omega_h^{+b}} h_{\tau_U}^2 \|\nabla(h'(u_h) + B^* p_h)\|_{0,\tau_U}^2 + C\|p_h - p_{u_h}\|_{1,\Omega_U}^2 \\ &= C\eta_1^2 + C\|p_h - p_{u_h}\|_{1,\Omega_U}^2.\end{aligned}$$

The second step is to estimate the error $\|p_{u_h} - p_h\|_{1,\Omega}$. Let $e_p = p_{u_h} - p_h$. Then it follows from $(2.4)_2$, (3.14), and (2.1) that

$$\begin{aligned}& c\|p_{u_h} - p_h\|_{1,\Omega}^2 \leq a(e_p, p_{u_h}) - a(e_p, p_h) \\ = \;& (g'(y_{u_h}), e_p) - a(e_p - \hat{\pi}_h e_p, p_h) - (g'(y_h), \hat{\pi}_h e_p) \\ = \;& \sum_{\tau \in T^h} \int_\tau (g'(y_h) + \operatorname{div}(A^* \nabla p_h))(e_p - \hat{\pi}_h e_p) \\ & - \sum_{l \in \partial T^h} \int_l [(A^* \nabla p_h) \cdot \mathbf{n}](e_p - \hat{\pi}_h e_p) + (g'(y_{u_h}) - g'(y_h), e_p) \\ \leq \;& C \sum_{\tau \in T^h} h_\tau^2 \int_\tau (g'(y_h) + \operatorname{div}(A^* \nabla p_h))^2 + C \sum_{l \in \partial T^h} h_l \int_l [(A^* \nabla p_h) \cdot \mathbf{n}]^2 \\ & + C\|y_{u_h} - y_h\|_{1,\Omega}^2 + \frac{c}{2}\|e_p\|_{1,\Omega}^2,\end{aligned}$$

where we have used Lemma 3.3 to obtain

$$\|e_p - \hat{\pi}_h e_p\|_{0,\tau} \leq C h_\tau \left(\sum_{\bar{\tau}' \cap \bar{\tau} \neq \emptyset} |e_p|_{1,\tau'}^2 \right)^{1/2} \tag{3.18}$$

and Lemmas 3.2 and 3.3 to have, assuming $l \subset \bar{\tau}$,

$$\begin{aligned}\|e_p - \hat{\pi}_h e_p\|_{0,l} &\leq C(h_\tau^{-1/2}\|e_p - \hat{\pi}_h e_p\|_{0,\tau} + h_\tau^{1/2}|e_p - \hat{\pi}_h e_p|_{1,\tau}) \\ &\leq C h_\tau^{1/2} \left(\sum_{\bar{\tau}' \cap \bar{\tau} \neq \emptyset} |e_p|_{1,\tau'}^2 \right)^{1/2}.\end{aligned} \tag{3.19}$$

Thus we have

$$\|p_{u_h} - p_h\|_{1,\Omega}^2 \leq C(\hat{\eta}_4^2 + \hat{\eta}_5^2) + C\|y_{u_h} - y_h\|_{1,\Omega}^2. \tag{3.20}$$

The third step is thus to estimate the error $\|y_{u_h} - y_h\|_{1,\Omega}$. Let $e_y = y_{u_h} - y_h$, and let $\hat{\pi}_h$ be the interpolator in Lemma 3.3. It can be seen that $a(e_y, \hat{\pi}_h e_y) = 0$ due to the Galerkin orthogonality $a(e_y, w_h) = 0 \; \forall w_h \in Y^h$ from $(2.5)_1$ and (3.13). Then it follows from (2.5), (3.13), (3.6), and (3.7) that

$$
\begin{aligned}
c\|y_{u_h} - y_h\|_{1,\Omega}^2 &\le a(e_y, e_y) = a(e_y, e_y - \hat{\pi}_h e_y) \\
&= \sum_{\tau \in T^h} \int_\tau (f + Bu_h + \operatorname{div}(A\nabla y_h))(e_y - \hat{\pi}_h e_y) \\
&\qquad - \sum_{l \in \partial T^h} \int_l [(A\nabla y_h)\cdot \mathbf{n}](e_y - \hat{\pi}_h e_y) \\
&\le C \sum_{\tau \in T^h} h_\tau^2 \int_\tau (f + Bu_h + \operatorname{div}(A\nabla y_h))^2 \\
&\qquad + C \sum_{l \in \partial T^h} h_l \int_l [(A\nabla y_h)\cdot \mathbf{n}]^2 + \frac{c}{2}\|e_y\|_{1,\Omega}^2,
\end{aligned}
$$

where we have bounded $\|e_y - \hat{\pi}_h e_y\|_{0,\tau}$ and $\|e_y - \hat{\pi}_h e_y\|_{0,l}$ as in (3.18) and (3.19). Thus we have

$$
\|y_{u_h} - y_h\|_{1,\Omega}^2 \le C(\hat{\eta}_2^2 + \hat{\eta}_3^2). \tag{3.21}
$$

Finally, by noting that, from (3.15), (3.16), (3.9), and (2.1), we have

$$
\|y_{u_h} - y\|_{1,\Omega} \le C\|u_h - u\|_{0,\Omega_U}, \tag{3.22}
$$

$$
\|p_{u_h} - p\|_{1,\Omega} \le C\|y_{u_h} - y\|_{1,\Omega} \le C\|u_h - u\|_{0,\Omega_U}, \tag{3.23}
$$

we combine (3.17), (3.20), and (3.21) to obtain

$$
\begin{aligned}
& \|u - u_h\|_{0,\Omega_U}^2 + \|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{1,\Omega}^2 \\
\le\ & \|u - u_h\|_{0,\Omega_U}^2 + 2\left(\|y - y_{u_h}\|_{1,\Omega}^2 + \|p - p_{u_h}\|_{1,\Omega}^2\right) \\
& + 2\left(\|y_{u_h} - y_h\|_{1,\Omega}^2 + \|p_{u_h} - p_h\|_{1,\Omega}^2\right) \\
\le\ & C\|u - u_h\|_{0,\Omega_U}^2 + 2\left(\|y_{u_h} - y_h\|_{1,\Omega}^2 + \|p_{u_h} - p_h\|_{1,\Omega}^2\right) \le\ C\sum_{i=1}^5 \eta_i^2.
\end{aligned}
$$

Therefore, the proof is completed. □

In many applications, we are mostly interested in computing the values of the state and the control. In such cases, it is more useful to bound the errors in the L^2-norm to derive sharper estimators, which are given in the following theorem. We shall use the following condition:

$$
|(Bv, w)| = |(v, B^* w)_U| \le C\|v\|_{0,\Omega_U}\|w\|_{0,\Omega} \qquad \forall v \in U, w \in Y, \tag{3.24}
$$

which is held from our assumptions. We shall assume the following condition:

$$
|(g'(v) - g'(w), q)| \le C\|v - w\|_{0,\Omega}\|q\|_{2,\Omega} \qquad \forall v, w \in Y,\ q \in H^2(\Omega). \tag{3.25}
$$

THEOREM 3.2. *Assume that all of the conditions of Theorem* 3.1 *and* (3.25) *are satisfied except that* (3.9) *is replaced with* (3.24). *Assume that* Ω *is convex. Then*

$$
\|u - u_h\|_{0,\Omega_U}^2 + \|y - y_h\|_{0,\Omega}^2 + \|p - p_h\|_{0,\Omega}^2 \le C\left(\eta_1^2 + \sum_{i=2}^5 \hat{\eta}_i^2\right),
$$

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

$$\hat{\eta}_2^2 = \sum_{\tau\in T^h} h_\tau^4 \int_\tau (f + Bu_h + \operatorname{div}(A\nabla y_h))^2,$$

$$\hat{\eta}_3^2 = \sum_{l\in\partial T^h} h_l^3 \int_l [(A\nabla y_h)\cdot \mathbf{n}]^2,$$

$$\hat{\eta}_4^2 = \sum_{\tau\in T^h} h_\tau^4 \int_\tau (g'(y_h) + \operatorname{div}(A^*\nabla p_h))^2,$$

$$\hat{\eta}_5^2 = \sum_{l\in\partial T^h} h_l^3 \int_l [(A^*\nabla p_h)\cdot \mathbf{n}]^2.$$

Proof. Again, we first estimate the error $\|u-u_h\|^2_{0,\Omega_U}$. By the same argument as in the proof of Theorem 3.1 but using (3.24), we have

$$\|u-u_h\|^2_{0,\Omega_U} \le C\eta_1^2 + C\|p_h - p_{u_h}\|^2_{0,\Omega}. \tag{3.26}$$

To estimate $\|p_h - p_{u_h}\|^2_{0,\Omega}$, we use the dual technique. Consider the following auxiliary problems: Find $\xi \in H_0^1(\Omega)$ and $\zeta \in H_0^1(\Omega)$ such that

$$a(w,\xi) = (f_1, w) \quad \forall w \in Y, \tag{3.27}$$

$$a(\zeta, q) = (f_2, q) \quad \forall q \in Y. \tag{3.28}$$

It follows from the well-known regularity results that

$$\|\xi\|_{2,\Omega} \le C\|f_1\|_{0,\Omega}, \quad \|\zeta\|_{2,\Omega} \le C\|f_2\|_{0,\Omega}.$$

Let $f_2 = p_{u_h} - p_h$ in (3.28) and denote by $\pi_h : C^0(\bar{\Omega}) \to Y^h$ the standard Lagrange interpolation operator. It follows from $(2.5)_2$ and (3.14) that

$$\begin{aligned}
\|p_{u_h} - p_h\|^2_{0,\Omega} &= (f_2, p_h(u_h) - p_h) = a(\zeta, p_{u_h}) - a(\zeta, p_h)\\
&= (g'(y_{u_h}), \zeta) - a(\zeta - \pi_h\zeta, p_h) - (g'(y_h), \pi_h\zeta)\\
&= \sum_{\tau\in T^h}\int_\tau \operatorname{div}(A^*\nabla p_h)(\zeta - \pi_h\zeta) - \sum_{l\in\partial T^h}\int_l [(A^*\nabla p_h)\cdot\mathbf{n}](\zeta - \pi_h\zeta)\\
&\qquad + (g'(y_{u_h}), \zeta) - (g'(y_h), \pi_h\zeta)\\
&= \sum_{\tau\in T^h}\int_\tau (g'(y_h) + \operatorname{div}(A^*\nabla p_h))(\zeta - \pi_h\zeta)\\
&\quad - \sum_{l\in\partial T^h}\int_l [(A^*\nabla p_h)\cdot\mathbf{n}](\zeta - \pi_h\zeta) + (g'(y_{u_h}) - g'(y_h), \zeta).
\end{aligned}$$

By using Lemmas 3.1 and 3.2,

$$\|\zeta - \pi_h\zeta\|_{0,\tau} \le Ch_\tau^2|\zeta|_{2,\tau}, \tag{3.29}$$

$$\|\zeta - \pi_h\zeta\|_{0,l} \le C(h_\tau^{-1/2}\|\zeta - \pi_h\zeta\|_{0,\tau} + h_\tau^{1/2}|\zeta - \pi_h\zeta|_{1,\tau}) \le Ch_\tau^{3/2}|\zeta|_{2,\tau}, \tag{3.30}$$

where $l \subset \bar{\tau}$. Then it follows from (3.25) that

$$\begin{aligned}\|p_{u_h} - p_h\|_{0,\Omega}^2 &\le C \sum_{\tau\in T^h} h_\tau^2 \|g'(y_h) + \operatorname{div}(A^*\nabla p_h)\|_{0,\tau}\|\zeta\|_{2,\tau} \\ &\quad + C \sum_{l\in\partial T^h} h_l^{3/2} \left(\int_l [(A^*\nabla p_n)\cdot \mathbf{n}]^2\right)^{1/2} \|\zeta\|_{2,\tau} + C\|y_{u_h} - y_h\|_{0,\Omega}\|\zeta\|_{2,\Omega} \\ &\le C \sum_{\tau\in T^h} h_\tau^4 \int_\tau (g'(y_h) + \operatorname{div}(A^*\nabla p_h))^2 + C \sum_{l\in\partial T^h} h_l^3 \int_l [(A^*\nabla p_n)\cdot \mathbf{n}]^2 \\ &\quad + C\|y_{u_h} - y_h\|_{0,\Omega}^2 + \frac{1}{2}\|f_2\|_{0,\Omega}^2.\end{aligned}$$

Therefore, we have

$$\|p_{u_h} - p_h\|_{0,\Omega}^2 \le C(\hat{\eta}_4^2 + \hat{\eta}_5^2) + C\|y_{u_h} - y_h\|_{0,\Omega}^2. \tag{3.31}$$

The second step is again to estimate $\|y_{u_h} - y_h\|_{0,\Omega}^2$. Similarly, letting $f_1 = y_{u_h} - y_h$ in (3.27) gives

$$\begin{aligned}\|y_{u_h} - y_h\|_{0,\Omega}^2 = (f, y_{u_h} - y_h) &= a(y_{u_h} - y_h, \xi) = a(y_{u_h} - y_h, \xi - \pi_h\xi) \\ &= \sum_{\tau\in T^h} \int_\tau (f + Bu_h + \operatorname{div}(A\nabla y_h))(\xi - \pi_h\xi) \\ &\quad - \sum_{l\in\partial T^h} \int_l [(A\nabla y_h)\cdot \mathbf{n}](\xi - \pi_h\xi) \\ &\le C \sum_{\tau\in T^h} h_\tau^2 \int_\tau \|f + Bu_h + \operatorname{div}(A\nabla y_h)\|_{0,\tau}\|\xi\|_{2,\tau} \\ &\quad + C \sum_{l\in\partial T^h} h_l^{3/2} \left(\int_l [(A\nabla y_h)\cdot \mathbf{n}]^2\right)^{1/2} \|\xi\|_{2,\tau} \\ &\le C \sum_{\tau\in T^h} h_\tau^4 \int_\tau (f + Bu_h + \operatorname{div}(A\nabla y_h))^2 \\ &\quad + C \sum_{l\in\partial T^h} h_l^3 \int_l [(A\nabla y_h)\cdot \mathbf{n}]^2 + \frac{1}{2}\|f_1\|_{0,\Omega}^2,\end{aligned}$$

where we have estimated $\|\xi - \pi_h\xi\|_{0,\tau}$ and $\|\xi - \pi_h\xi\|_{0,l}$ as in (3.29) and (3.30). The above result leads to

$$\|y_{u_h} - y_h\|_{0,\Omega}^2 \le C(\hat{\eta}_2^2 + \hat{\eta}_3^2). \tag{3.32}$$

Then it follows from (3.26), (3.31), and (3.32) that

$$\|u - u_h\|_{0,\Omega_U}^2 \le C\left(\eta_1^2 + \sum_{i=2}^5 \hat{\eta}_i^2\right). \tag{3.33}$$

Finally, we estimate $\|y_h - y\|_{0,\Omega}$ and $\|p_h - p\|_{0,\Omega}$. It follows from (3.15), (3.16), and (2.1) that

$$\begin{aligned}\|y_h - y\|_{0,\Omega} &\le \|y_h - y_{u_h}\|_{0,\Omega} + \|y_{u_h} - y\|_{0,\Omega} \\ &\le \|y_h - y_{u_h}\|_{0,\Omega} + C\|u_h - u\|_{0,\Omega_U}\end{aligned}$$

and

$$\begin{aligned}\|p_h - p\|_{0,\Omega} &\le \|p_h - p_{u_h}\|_{0,\Omega} + \|p_{u_h} - p\|_{0,\Omega} \\ &\le \|p_h - p_{u_h}\|_{0,\Omega} + C\|y_{u_h} - y\|_{1,\Omega} \\ &\le \|p_h - p_{u_h}\|_{0,\Omega} + C\|u_h - u\|_{0,\Omega_U}.\end{aligned}$$

The above results, together with (3.31)–(3.33), yield

$$\|y_h - y\|_{0,\Omega}^2 + \|p_h - p\|_{0,\Omega}^2 \le C\left(\eta_1^2 + \sum_{i=2}^{5} \hat{\eta}_i^2\right). \tag{3.34}$$

Hence the proof is completed by combining (3.33) and (3.34). □

3.2. Lower error bounds. In this subsection, we wish to demonstrate that the error estimates obtained above are quite sharp by establishing lower error bounds for the finite element approximation. We start with the following lemma about the bubble functions, the proof of which can be found in [1, 45].

LEMMA 3.5. *Let $\tau \in T^h$. Let τ_l^1, τ_l^2 be two elements in T^h with a common edge (face) $l = \bar{\tau}_l^1 \cap \bar{\tau}_l^2$. For any constants B_τ and D_l, there exist polynomials $w_\tau \in H_0^1(\tau)$ and $w_l \in H_0^1(\tau_l^1 \cup \tau_l^2)$ such that, for $m = 0, 1$,*

$$\int_\tau B_\tau w_\tau = h_\tau^2 \int_\tau B_\tau^2, \qquad |w_\tau|_{m,\tau}^2 \le C h_\tau^{2(1-m)+2} \int_\tau B_\tau^2,$$

$$\int_l D_l w_l = h_l \int_l D_l^2, \qquad |w_l|_{m,\tau_l^1 \cup \tau_l^2}^2 \le C h_l^{2(1-m)+1} \int_l D_l^2.$$

For ease of exposition, we assume that A is a constant matrix and Y^h is the piecewise linear finite element space. We also assume that there exists an integer $k \ge 0$ independent of h such that, for any $\tau_U \in T_U^h$, $(h'(u_h) + B^* p_h)|_{\tau_U}$ is a polynomial of k-order on τ_U. This assumption is needed to apply the inverse property in our proof below, and it may impose an implicit relationship between the meshes for the state and the control. We further assume that

$$\|h'(v) - h'(w)\|_{0,\Omega_U} \le C\|v - w\|_{0,\Omega_U} \qquad \forall v, w \in Y. \tag{3.35}$$

THEOREM 3.3. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of (2.3) and (2.5), respectively. Assume that A is a constant matrix, Y^h is the piecewise linear finite element space, $f \in L^2(\Omega)$, $\phi \equiv \phi_0$, $(h'(u_h) + B^* p_h)|_{\tau_U}$ is a polynomial of k-order on τ_U for any $\tau_U \in T_U^h$ with $k \ge 0$ independent of h, and the conditions (2.1), (3.24), and (3.35) hold. Then there exists a constant C depending on the matrix A and those constants in (2.1), (3.24), (3.35), and Lemma 3.5 such that*

$$\begin{aligned}\sum_{i=1}^{5} \eta_i^2 \quad \le \quad & C(\|u - u_h\|_{0,\Omega_U}^2 + \|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{1,\Omega}^2) \\ & + C \sum_{\tau \in T^h} h_\tau^2 (\|F - \bar{F}\|_{0,\tau}^2 + \|G - \bar{G}\|_{0,\tau}^2) \\ & + C \sum_{\tau_U \in T_U^h} h_{\tau_U}^2 \|\nabla (h'(u_h) + B^* p_h) \chi_{\Omega_h^b}\|_{0,\tau_U}^2,\end{aligned}$$

where η_i $(1 \le i \le 5)$ *are defined in Theorem* 3.1, $F = f + Bu_h$, $G = g'(y_h)$, $\bar{F}|_\tau = \int_\tau F/|\tau|$, *and* $\bar{G}|_\tau = \int_\tau G/|\tau|$.

Proof. From the optimality conditions (2.3), we deduce that $(h'(u)+B^*p)|_{\Omega_U^+} = 0$. It follows from the inverse property [10], (3.35), and (3.24) that

$$\begin{aligned}
\eta_1^2 &= \sum_{\tau_U \in T_U^h} h_{\tau_U}^2 (\|\nabla(h'(u_h) + B^*p_h)\chi_{\Omega_h^+}\|_{0,\tau_U}^2 + \|\nabla(h'(u_h) + B^*p_h)\chi_{\Omega_h^b}\|_{0,\tau_U}^2) \\
&\le C\|h'(u_h) + B^*p_h - h'(u) - B^*p\|_{0,\Omega_h^+}^2 + C \sum_{\tau_U \in T_U^h} h_{\tau_U}^2 \|\nabla(h'(u_h) + B^*p_h)\chi_{\Omega_h^b}\|_{0,\tau_U}^2 \\
&\le C(\|u - u_h\|_{0,\Omega_U}^2 + \|p - p_h\|_{1,\Omega}^2) + C \sum_{\tau_U \in T_U^h} h_{\tau_U}^2 \|\nabla(h'(u_h) + B^*p_h)\chi_{\Omega_h^b}\|_{0,\tau_U}^2.
\end{aligned}$$

To bound η_2^2, let w_τ be the bubble function as in Lemma 3.5 with $B_\tau = \bar{F}|_\tau$. It follows from (2.5) and (3.13) that

$$\begin{aligned}
\eta_2^2 &= \sum_{\tau \in T^h} h_\tau^2 \int_\tau F^2 \le 2 \sum_{\tau \in T^h} h_\tau^2 \int_\tau \{\bar{F}^2 + (F - \bar{F})^2\} \\
&= 2 \sum_{\tau \in T^h} \int_\tau \{w_\tau F + w_\tau(\bar{F} - F) + h_\tau^2 (F - \bar{F})^2\} \\
&= 2 \sum_{\tau \in T^h} \int_\tau (A\nabla(y_{u_h} - y_h)) \cdot \nabla w_\tau + 2 \sum_{\tau \in T^h} \int_\tau \{w_\tau(\bar{F} - F) + h_\tau^2 (F - \bar{F})^2\} \\
&\le C \sum_{\tau \in T^h} |y_{u_h} - y_h|_{1,\tau}^2 + \delta \sum_{\tau \in T^h} (|w_\tau|_{1,\tau}^2 + h_\tau^{-2}\|w_\tau\|_{0,\tau}^2) + C \sum_{\tau \in T^h} h_\tau^2 \int_\tau (F - \bar{F})^2 \\
&\le C(|y_{u_h} - y|_{1,\Omega}^2 + |y - y_h|_{1,\Omega}^2) + C\delta\eta_2^2 + C \sum_{\tau \in T^h} h_\tau^2 \int_\tau (F - \bar{F})^2.
\end{aligned}$$

Then it follows from this inequality and (3.22) that

$$\eta_2^2 \le C(\|u - u_h\|_{0,\Omega_U}^2 + \|y - y_h\|_{1,\Omega}^2) + C \sum_{\tau \in T^h} h_\tau^2 \int_\tau (F - \bar{F})^2. \tag{3.36}$$

To estimate η_3, we define the bubble function w_l as in Lemma 3.5 with $D_l = [(A\nabla y_h) \cdot \mathbf{n}]|_l$. By (3.13),

$$\begin{aligned}
\eta_3^2 &= \sum_{l \in \partial T^h} h_l \int_l D_l^2 = \sum_{l \in \partial T^h} \int_l w_l [(A\nabla y_h) \cdot \mathbf{n}] = \sum_{l \in \partial T^h} \int_{\tau_l^1 \cup \tau_l^2} (A\nabla y_h) \cdot \nabla w_l \\
&= \sum_{l \in \partial T^h} \int_{\tau_l^1 \cup \tau_l^2} (A\nabla(y_h - y_{u_h})) \cdot \nabla w_l + \sum_{l \in \partial T^h} \int_{\tau_l^1 \cup \tau_l^2} (f + Bu_h) w_l \\
&\le C \sum_{\tau \in T^h} |y_{u_h} - y_h|_{1,\tau}^2 + \delta \sum_{l \in \partial T^h} (|w_l|_{1,\tau_l^1 \cap \tau_l^2}^2 + h_l^{-2}\|w_l\|_{0,\tau_l^1 \cap \tau_l^2}^2) + C\eta_2^2 \\
&\le C(|y_{u_h} - y|_{1,\Omega}^2 + |y - y_h|_{1,\Omega}^2) + C\delta\eta_3^2 + C\eta_2^2.
\end{aligned}$$

It follows from the above inequality, (3.22), and (3.36) that

$$\eta_3^2 \le C\left(\|u - u_h\|_{0,\Omega_U}^2 + \|y - y_h\|_{1,\Omega}^2\right) + C \sum_{\tau \in T^h} h_\tau^2 \int_\tau (F - \bar{F})^2.$$

For η_4, let w_τ be set as in Lemma 3.5 with $B_\tau = \bar{G}|_\tau$. It follows from (3.14), (2.1), (3.23), and (3.22) that

$$\begin{aligned}
\eta_4^2 &= \sum_{\tau\in T^h} h_\tau^2 \int_\tau G^2 \le 2\sum_{\tau\in T^h} h_\tau^2 \int_\tau \{\bar{G}^2 + (G-\bar{G})^2\} \\
&= 2\sum_{\tau\in T^h} \int_\tau \{w_\tau G + w_\tau(\bar{G}-G) + h_\tau^2 (G-\bar{G})^2\} \\
&= 2\sum_{\tau\in T^h} \int_\tau \{(A\nabla w_\tau)\cdot\nabla(p_{u_h}-p_h) + w_\tau(g'(y_h)-g'(y_{u_h})) \\
&\qquad\qquad + w_\tau(\bar{G}-G) + h_\tau^2(G-\bar{G})^2\} \\
&\le C|p_{u_h}-p_h|_{1,\Omega}^2 + \delta\sum_{\tau\in T^h}(|w_\tau|_{1,\tau}^2 + h_\tau^{-2}\|w_\tau\|_{0,\tau}^2) \\
&\qquad\qquad + C\|y_h - y_{u_h}\|_{1,\Omega}^2 + C\sum_{\tau\in T^h} h_\tau^2 \int_\tau (G-\bar{G})^2 \\
&\le C\left(\|u-u_h\|_{0,\Omega_U}^2 + \|p-p_h\|_{1,\Omega}^2 + \|y-y_h\|_{1,\Omega}^2\right) + C\delta\eta_4^2 + C\sum_{\tau\in T^h} h_\tau^2 \int_\tau (G-\bar{G})^2.
\end{aligned}$$

Thus

$$\eta_4^2 \le C\left(\|u-u_h\|_{0,\Omega_U}^2 + \|y-y_h\|_{1,\Omega}^2 + \|p-p_h\|_{1,\Omega}^2\right) + C\sum_{\tau\in T^h} h_\tau^2 (G-\bar{G})^2. \tag{3.37}$$

To estimate η_5, we set w_l as in Lemma 3.5 with $D_l = [(A^*\nabla p_h)\cdot\mathbf{n}]|_l$. It follows from (3.14), (2.1), (3.23), and (3.22) that

$$\begin{aligned}
\eta_5^2 &= \sum_{l\in\partial T^h} h_l \int_l D_l^2 = \sum_{l\in\partial T^h}\int_l w_l[(A^*\nabla p_h)\cdot\mathbf{n}] = \sum_{l\in\partial T^h}\int_{\tau_l^1\cup\tau_l^2} (A^*\nabla p_h)\cdot\nabla w_l \\
&= \sum_{l\in\partial T^h}\int_{\tau_l^1\cup\tau_l^2} (A\nabla w_l)\cdot\nabla(p_h - p_{u_h}) + \sum_{l\in\partial T^h}\int_{\tau_l^1\cup\tau_l^2} g'(y_{u_h})w_l \\
&\le C|p_h - p_{u_h}|_{1,\Omega}^2 + \delta\sum_{l\in\partial T^h}(|w_l|_{1,\tau_l^1\cap\tau_l^2}^2 + h_l^{-2}\|w_l\|_{0,\tau_l^1\cap\tau_l^2}^2) \\
&\quad + C\|y_{u_h} - y_h\|_{1,\Omega}^2 + C\eta_4^2 \\
&\le C(\|u-u_h\|_{0,\Omega_U}^2 + \|p-p_h\|_{1,\Omega}^2 + \|y-y_h\|_{1,\Omega}^2) + C\delta\eta_5^2 + C\eta_4^2.
\end{aligned}$$

This inequality, combined with (3.37), implies

$$\eta_5^2 \le C(\|u-u_h\|_{0,\Omega_U}^2 + \|y-y_h\|_{1,\Omega}^2 + \|p-p_h\|_{1,\Omega}^2) + C\sum_{\tau\in T^h} h_\tau^2 (G-\bar{G})^2.$$

Thus we proved the desirable result. □

We believe that the error estimator $\eta_1^2 + \sum_{i=2}^5 \hat{\eta}_i^2$ in Theorem 3.2 is also sharp, though we are unable to establish any lower error bound for it. As a matter of fact, to our best knowledge, there exist no lower a posteriori error bounds in the L^2-norm in the literature or for any control problem.

3.3. Sharp a posteriori error estimators. In the above section, we have shown the following error bounds:

(3.38)
$$c\left(\sum_{i=1}^{5}\eta_i^2 \quad -\sum_{\tau\in T^h} h_\tau^2\int_\tau\{(F-\bar F)^2+(G-\bar G)^2\}-\sum_{\tau\in\Omega_h^b} h_\tau^2\|\nabla(h'(u_h)+B^*p_h)\|_{0,\tau}^2\right)$$
$$\le \|u-u_h\|_{0,\Omega_U}^2+\|y-y_h\|_{1,\Omega}^2+\|p-p_h\|_{1,\Omega}^2\le C\left(\sum_{i=1}^{5}\eta_i^2\right),$$

provided that the conditions of Theorems 3.1 and 3.3 hold. We note that, if the free boundary $\partial\Omega_U^+$ is regular, for instance, if the free boundary consists of a finite number of smooth surfaces or if the total area of the free boundary is finite, then meas (Ω_h^b) is of the order h as $h\to 0$. Thus the second and third terms of the left side are of higher order as $h\to 0$ if the data are regular. Take the following typical quadratic control as one example: let $\Omega=\Omega_U$; let U^h and Y^h be the piecewise constant and linear spaces, respectively; let $Bu=u$, $f\in H^1(\Omega)$, $h(u)=\int_\Omega u^2$, and $g(y)=\int_\Omega(y-y_0)^2$ with $y_0\in H^1(\Omega)$. Then one has

$$\int_\Omega (F-\bar F)^2+(G-\bar G)^2\le Ch^2(|f|_{H^1}^2+|y_0|_{H^1}^2)+C\int_\Omega(Bu_h-\overline{Bu_h})^2,$$

$$\sum_{\tau\in\Omega_h^b}\|\nabla(h'(u_h)+B^*p_h)\|_{0,\tau}^2\le C\left(\int_{\Omega_h^b}|\nabla B^*p|^2+\|p_h-p\|_{H^1(\Omega)}^2\right).$$

Thus it follows that the second and third terms of the left side of (3.38) are not needed in computations. It can be seen that the above observation still holds even if f, y_0 are only piecewise smooth. For more general objective functionals, one can proceed as in Remark 3.4. Therefore, (3.38) gives equivalent a posteriori error estimates in the global sense and thus shows that the estimator $\sum_{i=1}^5\eta_i^2$ is in general quite sharp.

An obvious problem is that the characteristic function $\chi_{\Omega_h^{+b}}$ is not a posteriori in the sense that we usually do not know the position of the free boundary. Nevertheless, one can substitute it with some a posteriori quantities, thus obtaining some a posteriori error indicators, which can then be used in the adaptive finite element method.

One possible idea is to approximate $\chi_{\Omega_h^{+b}}$ by the finite element solution, as suggested in [24] and [32]. The basic idea is to approximate the characteristic function with the a posteriori quantity $\chi^h_{\Omega_U^+}$. For $\alpha>0$, let

$$\chi^h_{\Omega_U^+}=\frac{u_h-\phi_0}{h^\alpha+u_h-\phi_0}.$$

Thus, in computing, we replace η_1^2 by

$$\widetilde{\eta}_1^2=\sum_{\tau\in T_U^h}h_{\tau_U}^2\|\nabla(h'(u_h)+B^*p_h)\chi^h_{\Omega_U^+}\|_{0,\tau_U}^2.$$

In the following, we investigate the possible errors caused by this replacement. To this end, we separate Ω_U into three parts:

$$\Omega_h^-,\quad \Omega_U^{\alpha/2}:=\{x\in\Omega_U\ :\ u_h(x)<\phi_0+h^{\alpha/2},\ u(x)>\phi_0\},\ \text{and}\ \Omega_h^{+b}\backslash\Omega_U^{\alpha/2}.$$

Then, for $\tau \in \Omega_h^-$, we have

$$\|\chi^h_{\Omega_U^+} - \chi_{\Omega_h^{+b}}\|_{0,\infty,\tau} = \left\| \frac{u_h - u}{h^\alpha + u_h - u} \right\|_{0,\infty,\tau} \le \min\{1, h^{-\alpha}\|u_h - u\|_{0,\infty,\tau}\},$$

and, for $\tau \in \Omega_h^{+b} \backslash \Omega_U^{\alpha/2}$,

$$\|\chi^h_{\Omega_U^+} - \chi_{\Omega_h^{+b}}\|_{0,\infty,\tau} = \left\| \frac{h^\alpha}{h^\alpha + u_h - \phi_0} \right\|_{0,\infty,\tau} \le h^{\alpha/2}.$$

Therefore, if the error $\|u_h - u\|_{0,\infty,\tau}$ for $\tau \subset \Omega_h^-$ (where $u \equiv \phi_0$) is of the order h^β with $\beta > \alpha$, then the difference caused by the replacement is a high-order small quantity locally for all $\tau \in \Omega_h^- \cup (\Omega_h^{+b} \backslash \Omega_U^{\alpha/2})$. The size of the remaining domain $\Omega_U^{\alpha/2}$ depends on the error $\|u_h - u\|_{0,\infty}$. It can be shown (see, e.g., [32]) that meas $(\Omega_U^{\alpha/2}) \to 0$, as long as $\|u_h - u\|_{0,\infty,\Omega} \to 0$, as $h \to 0$. Thus $\chi^h_{\Omega_U^+}$ is a good approximator to $\chi_{\Omega_h^{+b}}$, and this is confirmed in our numerical tests; see section 4.

Remark 3.1. Generally speaking, for the problem considered here, the costate p is more regular than the solution u. Therefore, we may use p_h instead of u_h to approximate the characteristic function. It can be seen from (2.3) that

$$u = \max\{-(h')^{-1}(B^*p), \phi_0\}.$$

Thus, for example, we can define

$$\widetilde{\chi}^h_{\Omega_U^+} = \frac{\widetilde{u}_h - \phi_0}{h^\alpha + \widetilde{u}_h - \phi_0},$$

where $\widetilde{u}_h = \max\{-(h')^{-1}(B^*p_h), \phi_0\}$. Similarly, we can show

$$\|\widetilde{\chi}^h_{\Omega_U^+} - \chi_{\Omega_h^{+b}}\|_{0,\infty,\tau} \le \begin{cases} \min\{1, h^{-\alpha}\|(h')^{-1}(B^*p_h) - (h')^{-1}(B^*p)\|_{0,\infty,\tau}\} & \forall \tau \in \Omega_h^-, \\ h^{\alpha/2} \qquad \forall \tau \in \Omega_h^{+b} \backslash \Omega_U^{\alpha/2}. \end{cases}$$

3.4. Nonconstant obstacles. If the constraint ϕ_0 is a function $\phi(x)$, one could introduce $u^*(x) = u(x) - \phi(x)$. Then the triplet (y, p, u^*) satisfies the following optimality conditions:

$$\text{(3.39)} \qquad \begin{cases} a(y, w) = (f^* + Bu^*, w) \quad \forall w \in Y = H_0^1(\Omega), \\ a(q, p) = (g'(y), p) \quad \forall q \in Y = H_0^1(\Omega), \\ ((h^*)'(u^*) + B^*p, v - u)_U \ge 0 \quad \forall v \in K \subset U = L^2(\Omega_U), \end{cases}$$

where $f^* = f + B\phi, h^*(v) = h(v + \phi)$, and $K = \{v \in U : v \ge 0\}$. Thus the problem is reduced to the case of (2.3) with $\phi_0 = 0$.

However, this strategy, although simpler, may affect the efficiency of the resulting error estimators. Let us try to explain this: the inactive data $\phi|_{\Omega_U^-}$ on the noncoincidence set does not affect the solution of (CCP) and thus is not expected to play a major role in a sharp error estimator. However, with the transformation $u - \phi$, this data may be brought into the resulting error estimators through f^*. Thus we will directly consider the error $u - u_h$ rather than $u^* - u_h^*$. Let

$$\text{(3.40)} \quad K = \{v \in U : v \ge \phi \text{ a.e. in } \Omega_U\}, \qquad K^h = \{v_h \in U^h : v_h \ge \phi^h \text{ a.e. in } \Omega_U\},$$

where $\phi^h \in U^h$ is an approximation of ϕ. Here we take $\phi^h = \pi_h^a \phi$. It should be noticed that $K^h \not\subset K$ in general.

THEOREM 3.4. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of (2.3) and (2.5), respectively. Assume that all of the conditions of Theorem 3.1 and (3.35) hold and K^h is defined as in (3.40) with $\phi \in L^2(\Omega_U)$ and $\phi^h = \pi_h^a \phi$. Then*

$$\|u_h - u\|_{0,\Omega_U}^2 + \|y_h - y\|_{1,\Omega}^2 + \|p_h - p\|_{1,\Omega}^2 \le C \sum_{i=1}^{6} \eta_i^2, \tag{3.41}$$

where η_i $(i = 1$–$5)$ are defined in Theorem 3.1 and

$$\eta_6^2 = \sum_{\tau_U \in T_U^h} \|(\phi^h - \phi)\chi_{\Omega_h^{-b}}\|_{0,\tau_U}^2.$$

Proof. We will give only the details for the estimation of $\|u - u_h\|_{0,\Omega_U}^2$. The other terms can be estimated similarly as in Theorem 3.1. It should be emphasized that here one cannot take $v = u_h$ in (2.3) since $u_h \ge \phi$ may not be true. It follows from (3.3) that

$$h'(u) + B^* p \ge 0, \quad (h'(u) + B^* p)\chi_{\Omega_U^+} = 0. \tag{3.42}$$

Then it follows from the assumption (3.4), the inequality (2.5), and (3.42) that, for any $v_h \in K^h$,

$$\begin{aligned}
&c\|u - u_h\|_{0,\Omega_U}^2 \\
&\le (h'(u), u - u_h)_U - (h'(u_h), u - u_h)_U + (h'(u_h) + B^* p_h, v_h - u_h)_U \\
&= (h'(u_h) + B^* p_h, v_h - u)_U + (B^*(p_h - p), u - u_h)_U + (h'(u) + B^* p, u - u_h)_U \\
&= ((h'(u_h) + B^* p_h)\chi_{\Omega_h^{+b}}, v_h - u)_U + (B^*(p_h - p), u - u_h)_U \\
&\quad + ((h'(u_h) + B^* p_h - (h'(u) + B^* p))\chi_{\Omega_h^-}, v_h - u)_U \\
&\quad + ((h'(u) + B^* p)\chi_{\Omega_h^-}, v_h - u_h)_U + ((h'(u) + B^* p)\chi_{\Omega_U^- \setminus \Omega_h^-}, u - u_h)_U \\
&:= \sum_{i=1}^{5} I_i.
\end{aligned} \tag{3.43}$$

Take $v_h = \pi_h^a u$. Then I_1 and I_2 can be estimated as in the proof of Theorem 3.1 such that

$$I_1 + I_2 \le C(\eta_1^2 + \|p_h - p_{u_h}\|_{1,\Omega}^2) + \delta \|u - u_h\|_{0,\Omega_U}^2,$$

where δ is a small positive constant. It follows from (3.35), (3.24), and (3.23) that

$$\begin{aligned}
I_3 &\le \sum_{\tau_U \in T_U^h} (\|h'(u_h) - h'(u)\|_{0,\tau_U} + \|B^*(p_h - p)\|_{0,\tau_U}) \|(\phi - \phi^h)\chi_{\Omega_h^-}\|_{0,\tau_U} \\
&\le \delta(\|u - u_h\|_{0,\Omega_U}^2 + \|p_h - p_{u_h}\|_{1,\Omega}^2) + C\eta_6^2.
\end{aligned}$$

We note that $I_4 \le 0$ due to (3.42) and the fact that $(v_h - u_h)|_{\Omega_h^-} = (\phi^h - u_h)|_{\Omega_h^-} \le 0$.

Finally, to estimate I_5, we use $u|_{\Omega_U^-} = \phi|_{\Omega_U^-}$, (3.42) and $u_h \geq \phi^h$ to get

$$\begin{aligned}
&((h'(u) + B^*p)\chi_{\Omega_U^- \setminus \Omega_h^-}, u - u_h)_{\tau_U} = ((h'(u) + B^*p)\chi_{\Omega_U^- \setminus \Omega_h^-}, \phi - u_h)_{\tau_U} \\
&\leq ((h'(u) + B^*p)\chi_{\Omega_U^- \setminus \Omega_h^-}, \phi - \phi^h)_{\tau_U} = (h'(u) + B^*p, (\phi - \phi^h)\chi_{\Omega_h^b})_{\tau_U} \\
&= (h'(u) + B^*p - (h'(u_h) + B^*p_h), (\phi - \phi^h)\chi_{\Omega_h^b})_{\tau_U} \\
&\quad + ((I - \pi_h^a)(h'(u_h) + B^*p_h), (\phi - \phi^h)\chi_{\Omega_h^b})_{\tau_U}.
\end{aligned}$$

Thus

$$I_5 \leq \delta(\|u - u_h\|_{0,\Omega_U}^2 + \|p_h - p_{u_h}\|_{1,\Omega}^2) + C(\eta_1^2 + \eta_6^2).$$

The rest of the proof is the same as that in Theorem 3.1. □

Remark 3.2. We can approximate the characteristic functions $\chi_{\Omega_h^{+b}}$ and $\chi_{\Omega_h^{-b}}$ by

$$\chi_{\Omega_h^{+b}}^h = \frac{u_h - \phi^h}{h^{\alpha^+} + u_h - \phi^h}, \qquad \chi_{\Omega_h^{-b}}^h = \frac{h^{\alpha^-}}{h^{\alpha^-} + u_h - \phi^h},$$

where α^+ and α^- are positive parameters.

3.5. Double obstacles. We now consider the control problem with the double obstacles: $\phi_1(x) < \phi_2(x)$. Let

$$\begin{aligned}
&K = \{v \in U : \phi_1 \leq v \leq \phi_2 \text{ a.e. in } \Omega_U\}, \\
&K^h = \{v_h \in U^h : \phi_1^h \leq v_h \leq \phi_2^h \text{ a.e. in } \Omega_U\},
\end{aligned} \tag{3.44}$$

where $\phi_i^h \in U^h$ is an approximation of ϕ_i $(i = 1, 2)$. We assume that $\phi_i^h = \pi_h^a \phi_i$ $(i = 1, 2)$. To generalize the ideas used in Theorem 3.4 to this case, we define

$$\begin{aligned}
&\Omega_{\phi_i}^- = \{x \in \Omega_U : u(x) = \phi_i(x), \}, && \Omega_\phi^- = \Omega_{\phi_1}^- \cup \Omega_{\phi_2}^-, && \Omega_\phi^+ = \Omega_U \setminus \Omega_\phi^-, \\
&\Omega_{\phi_i,h}^- = \{\cup \bar{\tau}_U : \tau_U \subset \Omega_{\phi_i}^-, \tau_U \in T_U^h\}, && \Omega_{\phi,h}^- = \Omega_{\phi_1,h}^- \cup \Omega_{\phi_2,h}^-, && \Omega_{\phi,h}^{+b} = \Omega_U \setminus \Omega_{\phi,h}^-, \\
&\Omega_{\phi_i,h}^{-b} = \{\cup \bar{\tau}_U : \bar{\tau}_U \cap \Omega_{\phi_i,h}^- \neq \emptyset, \tau_U \in T_U^h\}.
\end{aligned}$$

THEOREM 3.5. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of* (2.3) *and* (2.5)*, respectively. Assume that all of the conditions of Theorem* 3.4 *hold and K and K^h are defined as in* (3.44) *with $\phi_i \in L^2(\Omega_U)$ and $\phi_i^h = \pi_h^a \phi_i$ $(i = 1, 2)$. Then*

$$\|u_h - u\|_{0,\Omega_U}^2 + \|y_h - y\|_{1,\Omega}^2 + \|p_h - p\|_{1,\Omega}^2 \leq C \sum_{i=1}^{6} \eta_i^2, \tag{3.45}$$

where η_i $(i = 2$–$5)$ are defined in Theorem 3.1 *and*

$$\begin{aligned}
\eta_1^2 &= \sum_{\tau_U \in T_U^h} h_{\tau_U}^2 \|\nabla(h'(u_h) + B^*p_h)\chi_{\Omega_{\phi,h}^{+b}}\|_{0,\tau_U}^2, \\
\eta_6^2 &= \sum_{\tau_U \in T_U^h} \sum_{i=1,2} \|(\phi_i^h - \phi_i)\chi_{\Omega_{\phi_i,h}^{-b}}\|_{0,\tau_U}^2.
\end{aligned}$$

Proof. Again, we give only the details for estimation of the error $\|u-u_h\|^2_{0,\Omega_U}$. In this case, we have

$$(h'(u)+B^*p)\chi_{\Omega^-_{\phi_1}} \ge 0, \quad (h'(u)+B^*p)\chi_{\Omega^-_{\phi_2}} \le 0, \quad (h'(u)+B^*p)\chi_{\Omega^+_{\phi}} = 0. \tag{3.46}$$

As in (3.43), it follows from the assumption (3.4), the inequality (2.5), and (3.46) that, for any $v_h \in K^h$,

$$\begin{aligned}
&c\|u-u_h\|^2_{0,\Omega_U} \\
&\le (h'(u),u-u_h)_U - (h'(u_h),u-u_h)_U + (h'(u_h)+B^*p_h, v_h-u_h)_U \\
&= (h'(u_h)+B^*p_h, v_h-u)_U + (B^*(p_h-p),u-u_h)_U + (h'(u)+B^*p,u-u_h)_U \\
&= ((h'(u_h)+B^*p_h)\chi_{\Omega^{+b}_{\phi,h}}, v_h-u)_U + (B^*(p_h-p),u-u_h)_U \\
&\quad +((h'(u_h)+B^*p_h-(h'(u)+B^*p))\chi_{\Omega^-_{\phi,h}}, v_h-u)_U \\
&\quad +((h'(u)+B^*p)\chi_{\Omega^-_{\phi,h}}, v_h-u_h)_U + ((h'(u)+B^*p)\chi_{\Omega^-_{\phi}\setminus\Omega^-_{\phi,h}}, u-u_h)_U \\
&:= \sum_{i=1}^5 J_i.
\end{aligned} \tag{3.47}$$

It is easy to see that for $1 \le i \le 3$, J_i can be estimated as I_i. Thanks to (3.46), we still have $J_4 \le 0$. Also, J_5 can be treated similarly to I_5. For instance, let us consider the case that $\tau_U \subset (\Omega^{-b}_{\phi_2,h}\setminus\Omega^-_{\phi_2,h})$. Assume that $\Omega^{-b}_{\phi_2,h} \cap \Omega^-_{\phi_1} = \emptyset$ for simplicity. We then have, from $u|_{\Omega^-_{\phi_2}} = \phi_2|_{\Omega^-_{\phi_2}}$, (3.46), and $u_h \le \phi_2^h$, that

$$\begin{aligned}
&((h'(u)+B^*p)\chi_{\Omega^{-b}_{\phi_2,h}\setminus\Omega^-_{\phi_2,h}}, u-u_h)_{\tau_U} \\
&\quad = ((h'(u)+B^*p)\chi_{\Omega^{-b}_{\phi_2,h}\setminus\Omega^-_{\phi_2,h}}, \phi_2-u_h)_{\tau_U} \\
&\qquad \le ((h'(u)+B^*p)\chi_{\Omega^{-b}_{\phi_2,h}\setminus\Omega^-_{\phi_2,h}}, \phi_2-\phi_2^h)_{\tau_U} \\
&\quad = (h'(u)+B^*p-(h'(u_h)+B^*p_h), \phi_2-\phi_2^h)_{\tau_U} \\
&\qquad + ((I-\pi_h^a)(h'(u_h)+B^*p_h), \phi_2-\phi_2^h)_{\tau_U}.
\end{aligned}$$

The rest of the proof is the same as that of Theorem 3.4. ∎

Remark 3.3. In computing, we may approximate the characteristic functions $\chi_{\Omega^{+b}_h}$ and $\chi_{\Omega^{-b}_{\phi_i,h}}$ by

$$\chi^h_{\Omega^{+b}_h} = \frac{(u_h-\phi_1^h)(\phi_2^h-u_h)}{h^{\alpha^+}+(u_h-\phi_1^h)(\phi_2^h-u_h)}, \qquad \chi^h_{\Omega^{-b}_{\phi_i,h}} = \frac{h^{\alpha^-}}{h^{\alpha^-}+|u_h-\phi_i^h|},$$

where α^+ and α^- are positive parameters.

Remark 3.4. It is clear that the uniform monotonicity conditions and Lipschitz continuity (2.1), (3.4), (3.25), (3.35), assumed in the proofs of Theorems 3.1–3.3, are needed to hold only in a neighborhood of the true solutions. This observation is useful in some applications involving a nonquadratic objective functional like $g(y) = \int_\Omega (y-y_0)^4$.

For (2.1) and (3.25), let us assume that $g(y) = \int_\Omega j(y)$, where j is twice continuously differentiable on R^1, to just fix the idea. Then it follows from the Sobolev embedding result $H^1(\Omega) \to L^\beta(\Omega)$ ($\beta < \infty$ if $n = 2$, and $\beta = 6$ if $n = 3$) that we have, using the Hölder inequality,

$$\begin{aligned} |(g'(v) - g'(w), q)| &\le \|j''(z)\|_{0,\beta^*,\Omega}\|v - w\|_{0,\beta,\Omega}\|q\|_{0,\beta,\Omega} \\ &\le C\|j''(z)\|_{0,\beta^*,\Omega}\|v - w\|_{1,\Omega}\|q\|_{1,\Omega}, \end{aligned}$$

where $z = \theta v + (1-\theta)w$ with $\theta \in [0,1]$, $\beta^* = (1 - 2/\beta)^{-1}$ for any $\beta > 2$ if $n = 2$ and $\beta^* = 3/2$ if $n = 3$.

Also, by using the embedding result $H^2(\Omega) \to L^\infty(\Omega)$ for $n \le 3$, we have

$$|(g'(v) - g'(w), q)| \le \|j''(z)\|_{0,\Omega}\|v - w\|_{0,\Omega}\|q\|_{2,\Omega}.$$

For example, if $g(y) = \int_\Omega (y - y_0)^4$ with $y_0 \in L^4(\Omega)$, we have by $H^1(\Omega) \to L^{2\beta^*}(\Omega)$ that

$$\begin{aligned} |(g'(v) - g'(w), q)| &\le C(\|v^2\|_{0,\beta^*,\Omega} + \|w^2\|_{0,\beta^*,\Omega} + \|y_0^2\|_{0,\beta^*,\Omega})\|v - w\|_{1,\Omega}\|q\|_{1,\Omega}, \\ &\le C(\|v\|_{1,\Omega}^2 + \|w\|_{1,\Omega}^2 + \|y_0\|_{L^4(\Omega)}^2)\|v - w\|_{1,\Omega}\|q\|_{1,\Omega} \end{aligned}$$

and

$$\begin{aligned} |(g'(v) - g'(w), q)| &\le C(\|v^2\|_{0,\Omega} + \|w^2\|_{0,\Omega} + \|y_0^2\|_{0,\Omega})\|v - w\|_{0,\Omega}\|q\|_{2,\Omega}, \\ &\le C(\|v\|_{1,\Omega}^2 + \|w\|_{1,\Omega}^2 + \|y_0\|_{L^4(\Omega)}^2)\|v - w\|_{0,\Omega}\|q\|_{2,\Omega}. \end{aligned}$$

Thus (2.1) and (3.25) hold as long as v, w are in a bounded set of Y. One can discuss (3.4) and (3.35) similarly.

It follows from the proofs of Theorems 3.4–3.5 that Theorem 3.2 can also be generalized to the nonconstant or double obstacle cases in the same way. Thus one can just use $\eta_1^2 + \eta_6^2$ as the error indicator in adaptive finite element methods if only the values of the control and state are important in an application.

4. Numerical experiments. In this section, we carry out some numerical experiments to demonstrate possible applications of the error estimators obtained in section 3. In most control problems, the optimal control is often of prime interest. Thus it is important to develop mesh refinement schemes that efficiently reduce the error $\|u - u_h\|$. In practice, there are four major types of adaptive finite element methods—namely, the h-method (mesh refinement), the p-method (order enrichment), the r-method (mesh redistribution), and the hp-method. A posteriori error estimators can be used as error indicators to guide the mesh refinement in adaptive finite element methods. For our numerical tests, using an adaptive mesh redistribution (AMR) method is advantageous since it can keep the number of the total nodes unchanged while adjusting the distribution of the nodes.

4.1. AMR method. The general idea behind the AMR method is to adjust meshes such that the a posteriori error estimators (the monitor functions to be called) are *equally* distributed over the computational meshes, while the total number of the nodes remains the same. Clearly, this method particularly suits our purposes of testing the efficiency of the known a posteriori error estimators.

In solving the optimal control problem (1.1), we use an iterative method to move the meshes and to redistribute the solutions on the new grid points. The procedure

for the mesh moving part is described in [24, 25, 26]. The key idea here is to use some kind of equivalent error estimators as the monitor function (or moving mesh indicator). More precisely, let $(x(\xi,\eta), y(\xi,\eta))$ be the mesh map in two dimensions. Here (ξ,η) are the computational coordinates. Let $M > 0$ be the monitor function which depends on the physical solution to be adapted. By solving the Euler–Lagrange equation

$$\nabla \cdot (M^{-1}\nabla\xi) = 0, \qquad \nabla \cdot (M^{-1}\nabla\eta) = 0, \tag{4.1}$$

a map between the physical domain Ω and the logical domain Ω_c can be computed. Typically, the map transforms a uniform mesh in the logical domain to cluster grid points at the regions of physical domain where the solutions are of greater physical interest. One of the crucial issues is what monitor functions are to be used. One popular choice in the AMR method literature is a gradient-based monitor function like $M_\tau = \sqrt{1 + |\nabla y_h|^2_\tau}$, which moves more grids to the regions of the largest solution gradients. In [24], it was shown that the gradient-based monitor functions may not be suitable for free boundary problems, and a monitor function associated with a posteriori error estimators is introduced which was found particularly useful in approximating the variational inequalities with free boundaries. In this section, we will use the same solution procedures as described in [24] to obtain the numerical solutions with moving grids, except that monitor functions will be based on the error estimators developed in this work.

4.2. Numerical tests. Our numerical example is the following type of optimal control problem:

$$\text{(OCP)} \qquad \begin{aligned} &\min \frac{1}{2}\int_\Omega (y - y_0)^2 + \frac{1}{2}\int_{\Omega_U} (u - u_0)^2 \\ &\text{s.t.} \quad \begin{cases} -\Delta y = Bu + f, \\ y|_{\partial\Omega} = y_0|_{\partial\Omega} = 0, \\ u \geq 0 \quad \text{in } \Omega_U. \end{cases} \end{aligned}$$

In our example, $\Omega_U = \Omega = [0,1] \times [0,1]$ and $B = I$. We also use the same meshes for the approximation of the state and the control. Thus $\tau_U = \tau$. Let Ω^h be a polygonal approximation to Ω with boundary $\partial\Omega^h$. Let T^h be a partitioning of Ω^h into a disjoint regular triangular τ so that $\bar{\Omega}^h = \cup_{\tau\in T^h}\bar{\tau}$. Assume that the state y is approximated in the finite element space Y^h with Φ^i as basis functions and u is approximated in U^h with Ψ^i as basis functions. Thus the problem (OCP) is discretized as the following optimization problem:

$$\begin{aligned} \min \quad & \tfrac{1}{2}\left\{(Y - Y^0)^T Q (Y - Y^0) + (U - U^0)^T M (U - U^0)\right\} \\ \text{s.t.} \quad & AY = BU + F, \\ & U \geq 0, \end{aligned} \tag{4.2}$$

with

$$Q^{ij} = \int_\Omega \Phi^i \Phi^j dx, \qquad M^{ij} = \int_\Omega \Psi^i \Psi^j dx,$$
$$A^{ij} = \int_\Omega \nabla\Phi^i \nabla\Phi^j dx, \qquad B^{ij} = \int_\Omega \Phi^i \Psi^j dx,$$
$$F_i = \int_\Omega f\Phi^i dx.$$

The finite element solution (y_h, u_h) is given by $y_h = \sum_i Y_i \Phi^i$ and $u_h = \sum_i U_i \Psi^i$, and (y_0, u_0) is approximated by $y_0^h = \sum_i Y_i^0 \Phi^i$ and $u_0^h = \sum_i U_i^0 \Psi^i$.

In solving the above optimization problem, we use a projection gradient method developed by He [19]. The projection method, though simple, is by no means the most efficient algorithm for solving our problem, but the purpose of the experiments in this section is to test the efficiency of the error indicators. The idea in [19] is the first to introduce the Lagrange multiplier P and then to set

$$H = \begin{pmatrix} Q & 0 & -A^T \\ 0 & M & B^T \\ A & -B & 0 \end{pmatrix}, \qquad x = \begin{pmatrix} Y \\ U \\ P \end{pmatrix}, \qquad c = \begin{pmatrix} QY^0 \\ MU^0 \\ F \end{pmatrix}.$$

The algorithm for solving the optimization problem (4.2) is described by the following pseudocode:

```
du = beta*(Hx + c)
e = x - max(x-du,b)
error =||e||
do  while error >= TOL
    d= beta*H^T*e
    g = d + du
    beta=beta*error/||d||
    e = e + d
    rho = error^2/||e||^2
    x = max(x - gamma \rho g,b)
    du = \beta (H x + c)
    e = x - max(x-du,b)
    error = ||e||
end do
```

We now briefly describe the solution algorithm to be used for solving the numerical examples in this section.

ALGORITHM 0

(i) Solve the optimization problem (4.2) with the above optimization code on the current mesh, and calculate the error monitor function M;
(ii) move the mesh to a new location, and update the solution on new meshes using the monitor M, as described in [25].

It is important to note from Theorem 3.2 that the error $\|u - u_h\|_{L^2(\Omega_U)}$ is largely controlled by η_1. Thus, in Algorithm 0, η_1, in (3.11) will be used to construct the

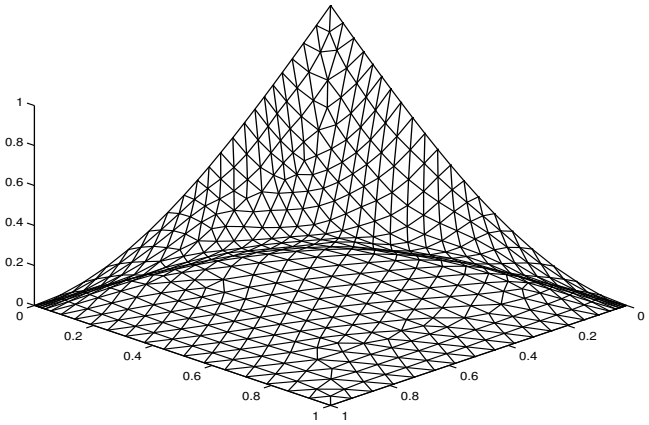

FIG. 4.1. *The surface of the solution u.*

monitor function M discussed in section 4.1,

$$M|_\tau = \sqrt{1 + \lambda \widetilde{\eta}_1^2|_\tau}, \tag{4.3}$$

where $\lambda > 0$ is a positive constant, and

$$\widetilde{\eta}_1^2|_\tau = h_\tau^2 \|\nabla(h'(u_h) + B^* p_h)\chi_{u_h}\|_{0,\tau}^2 . \tag{4.4}$$

In general, λ should be chosen such that $\lambda\|\eta_1\| >> 1$. Here we let $\lambda\|\eta_1\|_2 = 10^4$. As discussed in section 3, in our computation, we approximate the characteristic function used in η_1 by the following approximation:

$$\chi_{u_h} = \frac{u_h}{u_h + \epsilon} , \tag{4.5}$$

where $\epsilon > 0$ is a (small) positive number. In our experiments, we tried a range of values for ϵ between 0.1 and 1, and similar computational results were obtained.

Example 4.1. In this example we have

$$u_0 = 1 - \sin(\pi x_1/2) - \sin(\pi x_2/2), \quad y_0 = 0, \quad p = Z(x_1, x_2), \quad f = 4\pi^4 Z - u,$$

where $Z = \sin \pi x_1 \sin \pi x_2$. The exact solution of this problem is $y = 2\pi^2 Z, u = \max(u_0 - p, 0)$.

20×20 *nodes solution.* We first compute Example 4.1 on a 20×20 uniform mesh and then adjust the mesh by using Algorithm 0. The parameters λ and ϵ in (4.3) and (4.5) are 10^5 and 0.1, respectively. In Figure 4.1, the exact solution u is plotted. It is seen that the free boundary for this problem is just a single curve, and the maximum magnitude of the solution u is 1. The state and costate are approximated by piecewise linear elements. Both piecewise constant and piecewise linear elements are used to approximate the control in this example. In Figure 4.2, the 20×20 adaptive meshes are displayed. The control approximation errors are presented in Figures 4.3 and 4.4. It is observed that the maximum errors are distributed along the free boundary, as seen from Figures 4.3 and 4.4.

The adaptive meshes shown in Figure 4.2 are obtained by using the AMR method with the monitor function defined by (4.3). It is seen that a higher density of node

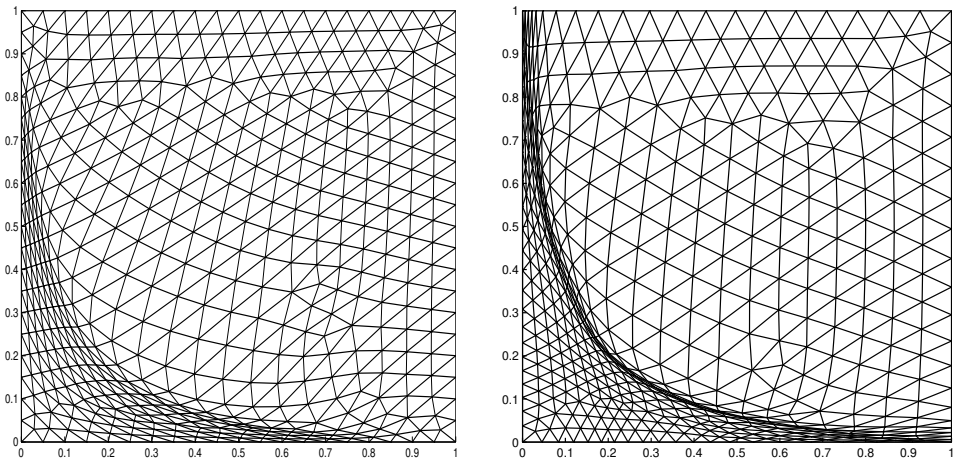

FIG. 4.2. *The adaptive mesh obtained by using piecewise constant elements (left) and piecewise linear elements (right), with* 20×20 *nodes.*

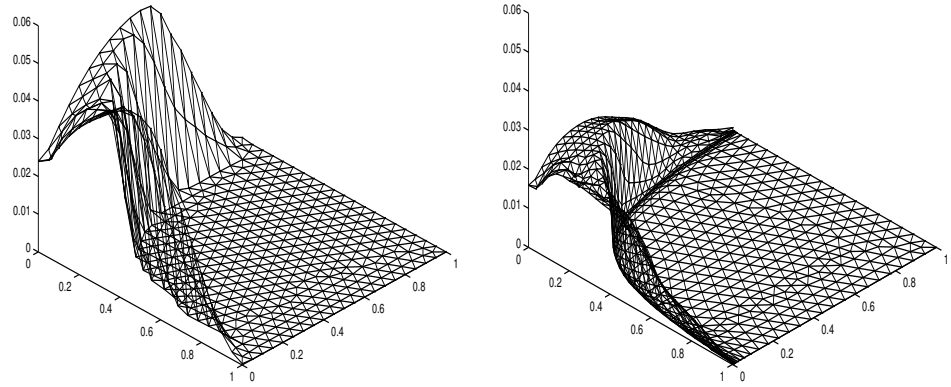

FIG. 4.3. L^2*-error* $\|u - u_h\|$ *with uniform mesh (left) and adaptive mesh (right), obtained by using piecewise constant elements with* 20×20 *nodes.*

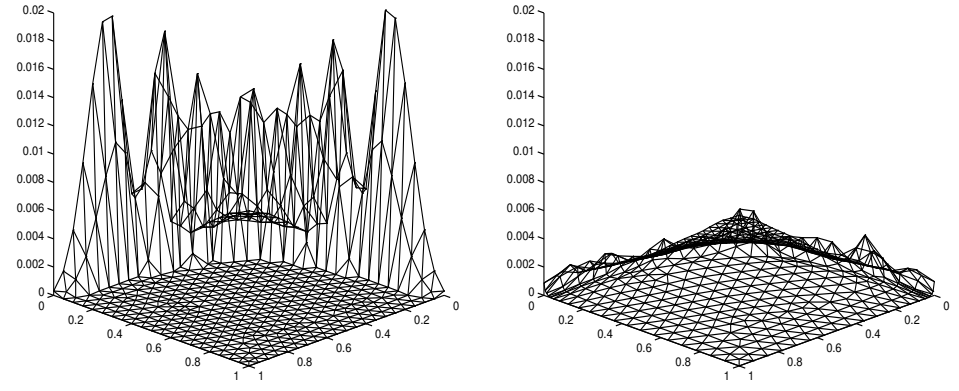

FIG. 4.4. *Same as Figure* 4.3*, except with linear elements.*

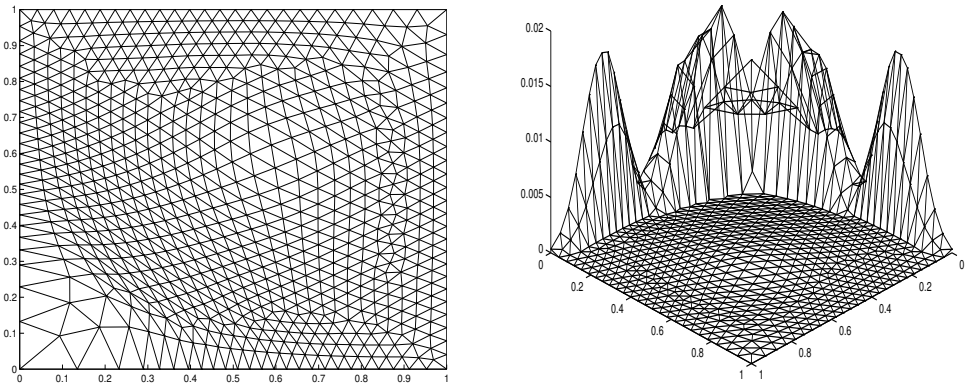

FIG. 4.5. *Example* 4.1 *with* 20×20 *nodes: Mesh (left) and error (right) obtained by using linear elements with unsharp error estimator associated with* $\bar{\eta}_1$*, as defined by* (3.2).

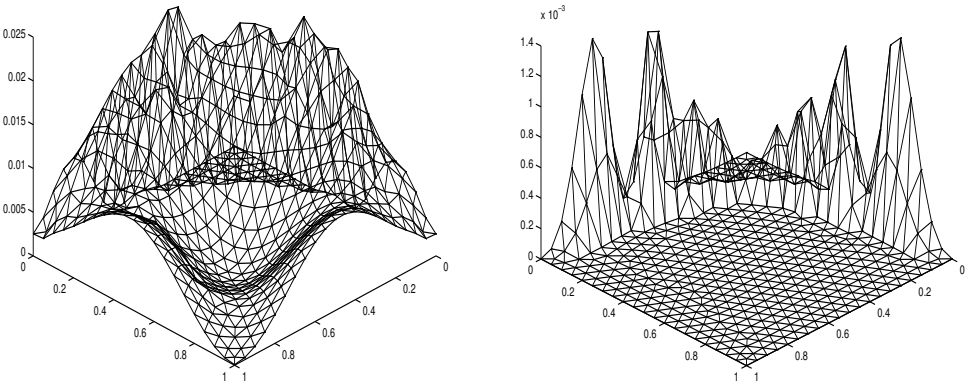

FIG. 4.6. *Profiles of unsharp estimator* $\bar{\eta}_1$ *and the sharp estimator* $\widetilde{\eta}_1$ *obtained by using the linear element with* 20×20 *nodes.*

points are now distributed along the free boundary. Furthermore, the approximation error is substantially reduced, as seen in Figures 4.3 and 4.4. In Figure 4.4, the L^2-norm of $u - u_h$ is 4.3×10^{-3} on the uniform mesh but is reduced 10 times to 4.4×10^{-4} on the adaptive mesh, while the L^2 error of the state approximation becomes slightly larger. It was found that one would need a 100×100 uniform mesh to produce such an error reduction. Thus efficient adaptive meshes can indeed save substantial computational work.

However, if we replace the estimator $\widetilde{\eta}_1$ in the monitor (4.3) with the estimator $\bar{\eta}_1$ given by (3.2), then a very different mesh is obtained; see Figure 4.5. As also seen in Figure 4.5, such a mesh is not efficient in reducing the control error; the error is virtually the same as that on the uniform mesh. The main reason is that the estimator $\bar{\eta}_1$ may not be sharp in this case. In fact, from Figure 4.6, it is clear that $\bar{\eta}_1$ and $|u - u_h|$ have very different profiles, while $\widetilde{\eta}_1$ has a profile similar to that of $|u - u_h|$.

40×40 *nodes solution.* To see the effect of mesh refinement, numerical solutions for Example 4.1 are obtained by using 40×40 linear elements. The control error distributions in this case are plotted in Figure 4.7, while the adapted mesh is plotted

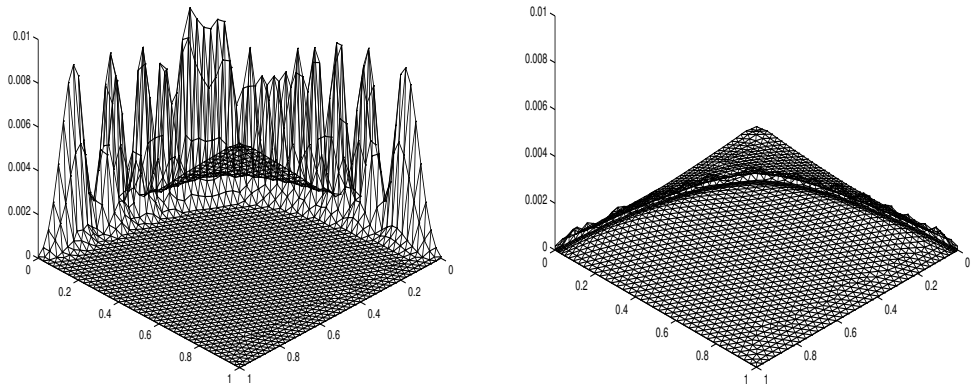

FIG. 4.7. *Same as Figure* 4.4*, except with* 40×40 *nodes.*

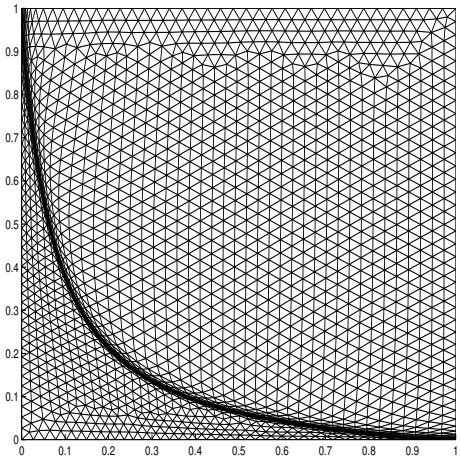

FIG. 4.8. *The adaptive mesh obtained by using piecewise linear elements with* 40×40 *nodes.*

in Figure 4.8. It is clear that the control errors are reduced with the finer mesh and with the use of the adaptive meshes.

5. Conclusion. In this work, we have derived some sharp a posteriori error indicators for the distributed elliptic optimal control problems. It is shown that the error indicators obtained can be applied in adaptive finite element computations and are found efficient in guiding mesh adjustments for our numerical examples. It is clear from the numerical experiments that the AMR methods can substantially increase the approximation accuracy. We point out that the approaches used in this work can be generalized to study other control problems.

REFERENCES

[1] M. AINSWORTH AND J. T. ODEN, *A posteriori error estimation in finite element analysis*, Comput. Methods Appl. Mech. Engrg., 142 (1997), pp. 1–88.

[2] W. ALT AND U. MACKENROTH, *Convergence of finite element approximations to state constrained convex parabolic boundary control problems*, SIAM J. Control Optim., 27 (1989), pp. 718–736.

[3] P. Alotto, P. Girdinio, O. Horigami, S. Ito, K. Iwanaga, K. Kato, T. Kuriyama, and H. Maeda, *Mesh adaption and optimization techniques in magnet design*, IEEE Trans. Magnetics, 32 (1996), pp. 1–8.
[4] N. V. Banichuk, F. J. Barthold, A. Falk, and E. Stein, *Mesh refinement for shape optimization*, Structural Optim., 9 (1995), pp. 46–51.
[5] R. E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp., 44 (1985), pp. 283–301.
[6] J. Baranger and H. E. Amri, *A posteriori error estimators in finite element approximation of quasi-Newtonian flows*, M2AN Math. Model. Numer. Anal., 25 (1991), pp. 31–47.
[7] J. W. Barrett and W. B. Liu, *Finite element approximation of some degenerate quasilinear elliptic and parabolic problems*, in Numerical Analysis (Dundee, 1993), Pitman Res. Notes Math. Ser. 303, Longman, Harlow, UK, 1994, pp. 1–16.
[8] R. Becker and H. Kapp, *Optimization in PDE models with adaptive finite element discretization*, in Proceedings of ENUMATH-97, World Scientific, Singapore, 1998, pp. 147–155.
[9] 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), pp. 113–132.
[10] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.
[11] K. Deckelnick and M. Hinze, *Error estimates in space and time for tracking-type control of the instationary Stokes system*, in Proceedings of Control and Estimation of Distributed Parameter Systems, Graz, Austria, 2001, to appear.
[12] F. S. Falk, *Approximation of a class of optimal control problems with order of convergence estimates*, J. Math. Anal. Appl., 44 (1973), pp. 28–47.
[13] D. A. French and J. T. King, *Approximation of an elliptic control problem by the finite element method*, Numer. Funct. Anal. Optim., 12 (1991), pp. 299–314.
[14] M. Hinze, *Optimal and Instantaneous Control of the Instationary Navier-Stokes Equations*, Habilitation thesis, Technische Universitat Berlin, Berlin, Germany, 2000.
[15] T. Geveci, *On the approximation of the solution of an optimal control problem governed by an elliptic equation*, RAIRO Anal. Numér., 13 (1979), pp. 313–328.
[16] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, Berlin, 1986.
[17] M. D. Gunzburger and L. S. Hou, *Finite-dimensional approximation of a class of constrained nonlinear control problems*, SIAM J. Control Optim., 34 (1996), pp.1001–1043.
[18] J. Haslinger and P. Neittaanmaki, *Finite Element Approximation for Optimal Shape Design*, John Wiley and Sons, Chichester, UK, 1989.
[19] B. S. He, *Solving a class of linear projection equations*, Numer. Math., 68 (1994), pp. 71–80.
[20] G. Knowles, *Finite element approximation of parabolic time optimal control problems*, SIAM J. Control Optim., 20 (1982), pp. 414–427.
[21] A. Kufner, O. John, and S. Fucik, *Function Spaces*, Nordhoff, Leyden, The Netherlands, 1977.
[22] K. Kunisch, W.-B. Liu, and N. N. Yan, *A posteriori error estimators for a model parameter estimation problem*, submitted to Proceedings of the European Conference on Numerical Mathematics and Advanced Applications, Ischia, Italy, 2001.
[23] I. Lasiecka, *Ritz–Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary conditions*, SIAM J. Control Optim., 22 (1984), pp. 477–500.
[24] R. Li, W.-B. Liu, and T. Tang, *Moving Mesh Method with Error-Estimator-Based Monitor and Its Application to Static Obstacle Problem*, manuscript.
[25] R. Li, T. Tang, and P.-W. Zhang, *Moving mesh methods in multiple dimensions based on harmonic maps*, J. Comput. Phys., 170 (2001), pp. 562–588.
[26] R. Li, T. Tang, and P.-W. Zhang, *A moving mesh finite element algorithm for singular problems in two and three space dimensions*, J. Comput. Phys., 177 (2002), pp. 365–393.
[27] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.
[28] W.-B. Liu and J. W. Barrett, *Quasi-norm error bounds for the finite element approximation of some degenerate quasilinear elliptic equations and variational inequalities*, RAIRO Modél. Math. Anal. Numér., 28 (1994), pp. 725–744.
[29] W.-B. Liu and D. Tiba, *Error estimates in the approximation of optimization problems governed by nonlinear operators*, Numer. Funct. Anal. Optim., 22 (2001), pp. 953–972.
[30] W.-B. Liu and N. N. Yan, *A posteriori error estimates for a model boundary optimal control problem*, J. Comput. Appl. Math., 120 (2000), pp. 159–173.

[31] W.-B. Liu and N. N. Yan, *A posteriori error analysis for control problem governed by nonlinear elliptic equations*, in Proceedings of EUNMATH99, Science Press, Singapore, 2000, pp. 146–153.
[32] W.-B. Liu and N. N. Yan, *A posteriori error estimators for a class of variational inequalities*, J. Sci. Comput., 35 (2000), pp. 361–393.
[33] W.-B. Liu and N. N. Yan, *A posteriori error estimates for convex boundary control problems*, SIAM. J. Numer. Anal., 39 (2001), pp. 73–99.
[34] W.-B. Liu and N. N. Yan, *Quasi-norm local error estimates for p-Laplacian*, SIAM. J. Numer. Anal., 39 (2001), pp. 100–127.
[35] W.-B. Liu and N. N. Yan, *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math., 15 (2001), pp. 285–309.
[36] W.-B. Liu and N. N. Yan, *A posteriori error estimates for optimal control problems governed by parabolic equations*, Numer. Math., to appear.
[37] K. Malanowski, *Convergence of approximations vs. regularity of solutions for convex, control constrained, optimal control systems*, Appl. Math. Optim., 8 (1982), pp. 69–95.
[38] K. Maute, S. Schwarz, and E. Ramm, *Adaptive topology optimization of elastoplastic structures*, Structural Optim., 15 (1998), pp. 81–91.
[39] P. Neittaanmaki and D. Tiba, *Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications*, Marcel Dekker, New York, 1994.
[40] O. Pironneau, *Optimal Shape Design for Elliptic System*, Springer-Verlag, Berlin, 1984.
[41] A. Schleupen, K. Maute, and E. Ramm, *Adaptive FE-procedures in shape optimization*, Structural and Multidisciplinary Optimization, 19 (2000), pp. 282–302.
[42] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[43] D. Tiba and F. Troltzsch, *Error estimates for the discretization of state constrained convex control problems*, Numer. Funct. Anal. Optim., 17 (1996), pp. 1005–1028.
[44] F. Troltzsch, *Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems—strong convergence of optimal control*, Appl. Math. Optim., 29 (1994), pp. 309–329.
[45] R. Verfürth, *A posteriori error estimators for the Stokes equations*, Numer. Math., 55 (1989), pp. 309–325.
[46] R. Verfürth, *The equivalence of a posteriori error estimators*, in Fast Solvers for Flow Problems (Kiel, 1994), Vieweg, Braunschweig, Germany, 1995, pp. 273–283.
[47] R. Verfürth, *A Review of Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*, Wiley-Teubner, Chichester, UK, 1996.