J Sci Comput (2007) 33: 155–182
DOI 10.1007/s10915-007-9147-7

A Posteriori Error Estimates of Recovery Type for Distributed Convex Optimal Control Problems

Ruo Li · Wenbin Liu · Ningning Yan

Received: 24 January 2007 / Accepted: 25 June 2007 / Published online: 16 August 2007

Abstract In this paper, we derive a posteriori error estimates of recovery type, and present the superconvergence analysis for the finite element approximation of distributed convex optimal control problems. We provide a posteriori error estimates of recovery type for both the control and the state approximation, which are generally equivalent. Under some stronger assumptions, they are further shown to be asymptotically exact. Such estimates, which are apparently not available in the literature, can be used to construct adaptive finite element approximation schemes and as a reliability bound for the control problems. Numerical results demonstrating our theoretical results are also presented in this paper.

Keywords Distributed optimal control problems · Finite element approximation · Superconvergence · A posteriori error estimator

1 Introduction

The finite element approximation of optimal control problems has been extensively studied in the literature. It is impossible to even give a very brief review here. Some work relevant to this paper can be found in, for instance, [6, 7, 17, 18, 20] and [21]. Recently adaptive finite element method has been found to be able to save substantial computational work in computing optimal control. Adaptive finite element method ensures a higher density of nodes in

R. Li (✉)
School of Mathematical Sciences, Peking University, Beijing, 100871, China
e-mail: rli@math.pku.edu.cn

W. Liu
KBS & Institute of Mathematics and Statistics, The University of Kent, Canterbury, CT2 7NF, England
e-mail: W.B.Liu@ukc.ac.uk

N. Yan
Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China
e-mail: ynn@amss.ac.cn

a certain area of the given domain, where the solution is more difficult to approximate, using an a posteriori error estimator to guide mesh refinement. The decision of whether further refinement of the meshes is necessary is based on the estimate of the discretization error. If further refinement is to be performed, then the error estimator is used as a guide to show how the refinement might be accomplished most efficiently.

It has been shown in [2, 3, 9], and [13] that one of the central issues in applying the adaptive finite element method to control problems is to obtain appropriate error indicators for the mesh adaptivity. Some widely used error indicators are based on heuristic approaches or the state approximation error, and have been found to be inefficient in adaptive finite element approximation of optimal control, see [3] and [9]. Only recently have a posteriori error estimates of residual type been derived for finite element approximation of control or objective functional in some optimal control problems. Furthermore promising initial numerical results have been showed there. In [3], a posteriori error estimates were derived for the approximation error of objective functional of *unconstrained* optimal control problems. However in many control applications there are inequality constraints for the control. Furthermore it is sometimes more interesting to obtain error estimators directly for the control approximation error. In [13–15], and [16], a posteriori error estimates of residual type were obtained for the control approximation error of some optimal control problems with convex control constraints. Another important issue is that different adaptive meshes should be used for the control and the states, as generally they are of very different natures of singularities, see [8].

It is well-known that error indicators based on residual estimates are in general far less sharper than those based on recovery techniques. A posteriori error estimators of recovery type, like the Z-Z error estimator, have been widely used in engineering computations. However it is not straightforward to extend the gradient recovery techniques to derive a posteriori error estimators of recovery type for the optimal control, where for example, the control satisfies coupled algebraic variational inequalities, and normally only has lower regularity.

The problem that we are interested in is the distributed convex optimal control problem (it will be described in the next section). It is the purpose of this paper to derive a posteriori error estimates of recovery type for both the control and the state approximation.

The plan of the paper is as follows. In Sect. 2, we shall give a brief review on the finite element method and then construct the finite element approximation for the distributed convex optimal control problem. In Sect. 3, we derive a posteriori error estimators of recovery type. It is shown that for rather general meshes and solutions, the a posteriori error estimators of recovery type provide upper and lower error bounds for the finite element approximation of the control problem. In Sect. 4, the superconvergence analysis is carried out. Based on this analysis, it is shown that the a recovery type a posteriori error estimator is not only equivalent, but also asymptotically exact under some strong conditions. Some numerical examples demonstrating our theoretical results are shown in Sect. 5.

Let Ω and Ω_U be two bounded open sets in R^2 with Lipschitz boundaries $\partial\Omega$ and $\partial\Omega_U$. In this paper we adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with the norm $\|\cdot\|_{m,q,\Omega}$ and the seminorm $|\cdot|_{m,q,\Omega}$. We set $W_0^{m,q}$ the subspace of $W^{m,q}$ with vanished traces upto order $m-1$. We denote $W^{m,2}(\Omega)$ $(W_0^{m,2}(\Omega))$ by $H^m(\Omega)(H_0^m(\Omega))$ with the norm $\|\cdot\|_{m,\Omega}$ and the seminorm $|\cdot|_{m,\Omega}$. In addition, c or C denotes a general positive constant independent of h.

2 Finite Element Approximation of Optimal Control Problems

In this section, we study the finite element approximation of distributed convex optimal control problems. In the rest of the paper, we shall take the state space $V = H_0^1(\Omega)$, the control space $U = L^2(\Omega_U)$, and $H = L^2(\Omega)$ to fix the idea. Other cases can be treated similarly. Let the observation space $Y = L^2(\Omega)$. Let B be a linear continuous operator from U to H. Let g be a strictly convex functional which is continuously differentiable on the observation space $L^2(\Omega)$. Let

$$K = \{v \in U : v \geq 0\}.$$

We further assume that the functional $g(\cdot)$ is bounded below.

We are interested in the optimal control problem

$$\begin{aligned} &\min_{u\in K}\left\{g(y) + \frac{\alpha}{2}\|u\|^2_{0,\Omega_U}\right\},\\ &-\mathrm{div}(A\nabla y) = f + Bu \quad \text{in } \Omega, \qquad y|_{\partial\Omega} = 0, \end{aligned} \tag{2.1}$$

where α is a positive constant, $f \in L^2(\Omega)$, g is a continuously differentiable convex functional on H, and $A(\cdot) = (a_{i,j}(\cdot))_{n\times n} \in R^{n\times n}$ is a symmetric positive definite matrix. To consider the finite element approximation of the above optimal control problem, we have to give a weak formula for the state equation. Let

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

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

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

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

$$a(y, y) \geq c\|y\|^2_V, \qquad |a(y, w)| \leq C\|y\|_V\|w\|_V. \tag{2.2}$$

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

$$a(y(u), w) = (f + Bu, w) \quad \forall w \in H^1_0(\Omega). \tag{2.3}$$

Therefore, the above control problem can be restated as the following, which we shall label (OCP):

$$\min_{u\in K}\left\{g(y) + \frac{\alpha}{2}\|u\|^2_{0,\Omega_U}\right\}, \tag{2.4}$$

$$a(y(u), w) = (f + Bu, w) \quad \forall w \in V = H^1_0(\Omega). \tag{2.5}$$

It is well known (see, for example, [12]) that the control problem (OCP) has a unique solution (y, u), and that the 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, which we shall label (OCP-OPT):

$$a(y, w) = (f + Bu, w) \quad \forall w \in V = H^1_0(\Omega), \tag{2.6}$$

$$a(q, p) = (g'(y), q) \quad \forall q \in V = H^1_0(\Omega), \tag{2.7}$$

$$(\alpha u + B^* p, v - u)_U \geq 0 \quad \forall v \in K \subset U = L^2(\Omega_U), \tag{2.8}$$

where B^* is the adjoint operator of B.

Also we note that for any $(y,u) \in V \times U$, $g'(y)$ is in $Y = Y' = L^2(\Omega)$. Therefore, it can be viewed as functions in $Y = L^2(\Omega)$ from the well-known representation theorem in a Hilbert space.

Let us consider the finite element approximation of the control problem (OCP). Here we consider only n-simplex elements, as they are among the most widely used ones. Also we consider only conforming finite elements.

Let Ω^h be a polygonal approximation to Ω with a boundary $\partial\Omega^h$. For simplicity, we assume that $\Omega^h = \Omega$ in this paper. Let T^h be a partitioning of Ω^h into disjoint regular triangular element τ, so that $\bar{\Omega}^h = \bigcup_{\tau \in T^h} \bar{\tau}$. Each element has at most one face on $\partial\Omega^h$, 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 $C(\bar{\Omega}^h)$, such that $\chi|_\tau$ are polynomials of m-order ($m \geq 1$) $\forall \chi \in S^h$ 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 a partitioning of Ω_U^h into disjoint regular triangular element τ_U, so that $\bar{\Omega}_U^h = \bigcup_{\tau_U \in T_U^h} \bar{\tau}_U$. Assume that $\Omega_U^h = \Omega_U$, and $\bar{\tau}_U$ 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 W_U^h of $L^2(\Omega_U^h)$, such that $\chi|_{\tau_U}$ are polynomials of m-order ($m \geq 0$) $\forall \chi \in W_U^h$ and $\tau_U \in T_U^h$. Here there is no requirement for the continuity. Let $U^h = W_U^h$.

In this paper, we will only consider the simplest finite element spaces, i.e., $m = 1$ for V^h and $m = 0$ for U^h. Let h_τ (h_{τ_U}) denote the maximum diameter of the element τ (τ_U) in T^h (T_U^h), let $h = \max_{\tau \in T^h}\{h_\tau\}$, $h_U = \max_{\tau_U \in T_U^h}\{h_{\tau_U}\}$. Note that the order of the finite element space for U^h ($m = 0$) is lower than the one for V^h ($m = 1$). The size of the element in T_U^h is smaller than the one in T^h generally. Therefore, we assume that $h_U/h \leq C$ in this paper.

Then a possible finite element approximation of (OCP), which we shall label $(OCP)^h$, is as follows:

$$\min_{u_h \in K^h \subset U^h} \left\{ g(y_h) + \frac{\alpha}{2}\|u_h\|_{0,\Omega}^2 \right\}, \tag{2.9}$$

$$a(y_h, w_h) = (f + Bu_h, w_h) \quad \forall w_h \in V^h, \tag{2.10}$$

where K^h is a closed convex set in U^h. This is a finite dimensional optimization problem and may be solved by existing mathematical programming methods such as the steepest descent method, conjugate gradient method, trust domain method.

It follows that the control problem $(OCP)^h$ has a unique solution (y_h, u_h) and that the pair $(y_h, u_h) \in V^h \times K^h$ is the solution of $(OCP)^h$ 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 we shall label $(OCP-OPT)^h$:

$$a(y_h, w_h) = (f + Bu_h, w_h) \quad \forall w_h \in V^h \subset V = H_0^1(\Omega), \tag{2.11}$$

$$a(q_h, p_h) = (g'(y_h), q_h) \quad \forall q_h \in V^h \subset V = H_0^1(\Omega), \tag{2.12}$$

$$(\alpha u_h + B^* p_h, v_h - u_h)_U \geq 0 \quad \forall v_h \in K^h \subset U = L^2(\Omega_U). \tag{2.13}$$

It is well known that for the problem (2.6–2.8) and its finite element approximation (2.11–2.13), the following error estimate holds:

$$\|u - u_h\|_{0,\Omega} + \|y - y_h\|_{1,\Omega} + \|p - p_h\|_{1,\Omega} \leq C(h + h_U), \tag{2.14}$$

if $y, p \in H^2(\Omega)$, $u \in H^1(\Omega_U)$, where (y, p, u) and (y_h, p_h, u_h) are the solutions of (2.6–2.8) and (2.11–2.13), respectively.

3 Recovery Type a Posteriori Error Estimates

In order to discuss the recovery type a posteriori error estimate, let us construct the recovery operator R_h. Let $R_h v \in S^h$ be a continuous piecewise linear function (without zero boundary constraint). The values of $R_h v$ on the nodes are defined by least-squares argument on an element patches surrounding the nodes, similar as Z-Z patch recovery (see, e.g., [25, 26]), as follows. Let z be a node, $\omega_z = \bigcup_{\tau_U \in T_U^h, z \in \bar{\tau}_U} \tau_U$, V_z be the linear function space on ω_z. Set $R_h v(z) = \sigma_z(z)$ where

$$E(\sigma_z) = \min_{w \in V_z} E(w),$$

and

$$E(w) = \sum_{\tau_U \subset \omega_z} \left(\int_{\tau_U} w - \int_{\tau_U} v \right)^2.$$

When $z \in \partial\Omega_U$, we should add a few extra neighbor elements to ω_z such that ω_z contains more than three elements. For the regular mesh and the suitable choice of ω_z, we can conclude that for any $v \in L^2(\Omega)$, $R_h v$ exists. Moreover, for any domain $D \subset \Omega$, $R_h v = v$ on D if v is a linear function on $\hat{D}$, where $\hat{D} = \{\bigcup \tau_U : \bar{\tau}_U \cap \bar{D} \neq \emptyset\}$.

Remark 3.1 Generally, $R_h v(z)$, $z = (x_0, y_0)$, can be calculated as follows. Let $\omega_z = \sum_{i=1}^m \tau_U^i$, the three nodes of τ_U^i are (x_0, y_0), (x_i, y_i) and (x_{i+1}, y_{i+1}), $i = 1, 2, \ldots, m-1$, the three nodes of τ_U^m are (x_0, y_0), (x_m, y_m) and (x_1, y_1). Then, $R_h v(z) = a + bx_0 + dy_0$, where (a, b, d) is the solution of the following linear system:

$$\begin{bmatrix} \sum_{i=1}^m |k_i|^2 & \frac{1}{3}\sum_{i=1}^m |k_i|^2 \times (x_0 + x_i + x_{i+1}) & \frac{1}{3}\sum_{i=1}^m |k_i|^2 \times (y_0 + y_i + y_{i+1}) \\ \frac{1}{3}\sum_{i=1}^m |k_i|^2 \times (x_0 + x_i + x_{i+1}) & \frac{1}{9}\sum_{i=1}^m |k_i|^2 \times (x_0 + x_i + x_{i+1})^2 & \frac{1}{9}\sum_{i=1}^m |k_i|^2 \times (x_0 + x_i + x_{i+1}) \times (y_0 + y_i + y_{i+1}) \\ \frac{1}{3}\sum_{i=1}^m |k_i|^2 \times (y_0 + y_i + y_{i+1}) & \frac{1}{9}\sum_{i=1}^m |k_i|^2 \times (x_0 + x_i + x_{i+1}) \times (y_0 + y_i + y_{i+1}) & \frac{1}{9}\sum_{i=1}^m |k_i|^2 \times (y_0 + y_i + y_{i+1})^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ d \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^m |k_i| \int_{\tau_U^i} v \\ \frac{1}{3}\sum_{i=1}^m |k_i| \int_{\tau_U^i} v \times (x_0 + x_i + x_{i+1}) \\ \frac{1}{3}\sum_{i=1}^m |k_i| \int_{\tau_U^i} v \times (y_0 + y_i + y_{i+1}) \end{bmatrix},$$

where $|k_i|$ is the area of the element τ_U^i, and $(x_{m+1}, y_{m+1}) = (x_1, y_1)$.

For gradient of y and p, we construct the gradient recovery operator $G_h v = (R_h v_x, R_h v_y)$, where R_h is the recovery operator defined above for the recovery of u, $v_x = \frac{\partial v}{\partial x}$ and $v_y = \frac{\partial v}{\partial y}$. It should be noted that G_h is same as the Z-Z gradient recovery (see e.g., [25, 26]) in our piecewise linear case.

Based on the recovery operator R_h and G_h defined above, we can define the recovery type a posteriori error estimator:

$$\eta^2 = \|G_h y_h - \nabla y_h\|^2_{0,\Omega} + \|G_h p_h - \nabla p_h\|^2_{0,\Omega} + \|R_h u_h - u_h\|^2_{0,\Omega_U}. \tag{3.1}$$

To analyze the estimator, let us divide the domain Ω_U into three disjoint subdomains:

$$\bar{\Omega}_U = \bar{\Omega}_U^- \cup \bar{\Omega}_U^{00} \cup \bar{\Omega}_U^{++},$$

where

$$\Omega_U^- = \{x \in \Omega_U : B^* p_h(x) \le 0\},$$
$$\Omega_U^{00} = \{x \in \Omega_U : B^* p_h(x) > 0, u_h(x) = 0\},$$
$$\Omega_U^{++} = \{x \in \Omega_U : B^* p_h(x) > 0, u_h(x) > 0\}.$$

Let

$$\Omega_U^{-+} = \Omega_U^- \cup \Omega_U^{++},$$
$$\Omega_h^+ = \{x \in \Omega_U : u_h(x) > 0\}.$$

Then, it is easy to see that

$$\Omega_U^{-+} = \Omega_h^+ \cup \Omega_U^-.$$

Moreover, set

$$e^2 = \int_{\Omega^*} (\alpha u + B^* p - \mathcal{P}_h(\alpha u + B^* p))^2, \tag{3.2}$$

where $\mathcal{P}_h$ is the L^2-project operator from $L^2(\Omega_U)$ to U^h, and

$$\Omega^* = \{x \in \Omega_U^{++} : u(x) = 0, u_h(x) > 0\}.$$

Theorem 3.1 *Let* (y, p, u) *and* (y_h, p_h, u_h) *be the solutions of the systems* (2.6–2.8) *and* (2.11–2.13), *respectively. Assume that* g' *is Lipschitz continuous, and* A *is a constant positive definite matrix. Then,*

$$|y_h - y|^2_{1,\Omega} + |p_h - p|^2_{1,\Omega} + \|u_h - u\|^2_{0,\Omega_U} + e^2 \le C\eta^2 + C(\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2), \tag{3.3}$$

where e *is defined by* (3.2), η *is defined by* (3.1),

$$\epsilon_1^2 = \int_{\Omega_U^{-+}} |B^* p_h + \alpha R_h u_h|^2,$$

$$\epsilon_2^2 = \sum_{z \in \Lambda} \int_{\omega_z} h_z^2 |f + B u_h - (\overline{f + B u_h})_z|^2,$$

$$\epsilon_3^2 = \sum_{z \in \Lambda} \int_{\omega_z} h_z^2 |g'(y_h) - \overline{g'(y_h)}_z|^2,$$

where

$$\bar{f}_z = \frac{\int_{\omega_z} f}{\int_{\omega_z} 1},$$

where ω_z *is the support of* ϕ_z, ϕ_z *is the base function on the node point* z, h_z *is the size of* ω_z, *and* Λ *is the set of all inner nodes.*

Proof Note that

$$\begin{aligned}\alpha\|u-u_h\|^2_{0,\Omega_U} &= (\alpha u, u-u_h)_U - (\alpha u_h, u-u_h)_U \\ &\le -(B^*p, u-u_h)_U - (\alpha u_h, u-u_h)_U \\ &= (B^*p_h+\alpha u_h, u_h-u)_U + (B^*(p_h-p(u_h)), u-u_h)_U \\ &\quad + (B^*(p(u_h)-p), u-u_h)_U, \end{aligned} \tag{3.4}$$

where $p(u_h)$ is the solution of the auxiliary equation:

$$a(y(u_h), w) = (f + Bu_h, w) \quad \forall w \in V = H^1_0(\Omega), \tag{3.5}$$

$$a(q, p(u_h)) = (g'(y(u_h)), q) \quad \forall q \in V = H^1_0(\Omega). \tag{3.6}$$

It is easy to show that

$$(B^*(p_h - p(u_h)), u-u_h)_U \le C\|p_h - p(u_h)\|^2_{0,\Omega} + C\delta\|u-u_h\|^2_{0,\Omega_U}, \tag{3.7}$$

and

$$(B^*(p(u_h)-p), u-u_h)_U = (g'(y(u_h)) - g'(y), y - y(u_h)) \le 0. \tag{3.8}$$

For the first term of (3.4), we have that

$$(B^*p_h+\alpha u_h, u_h-u)_U = \int_{\Omega_U^{-+}} (B^*p_h+\alpha u_h)(u_h-u) + \int_{\Omega_U^{00}} (B^*p_h+\alpha u_h)(u_h-u), \tag{3.9}$$

$$\begin{aligned}&\int_{\Omega_U^{-+}} (B^*p_h+\alpha u_h)(u_h-u) \\ &\le C\int_{\Omega_U^{-+}} (B^*p_h+\alpha u_h)^2 + C\delta\|u-u_h\|^2_{0,\Omega_U} \\ &\le C\int_{\Omega_U^{-+}} (B^*p_h+\alpha R_h u_h)^2 + C\int_{\Omega_U^{-+}} \alpha^2(u_h - R_h u_h)^2 + C\delta\|u-u_h\|^2_{0,\Omega_U} \\ &\le C\|u_h - R_h u_h\|^2_{0,\Omega_U} + C\epsilon_1^2 + C\delta\|u-u_h\|^2_{0,\Omega_U}, \end{aligned} \tag{3.10}$$

and

$$\int_{\Omega_U^{00}} (B^*p_h+\alpha u_h)(u_h-u) = -\int_{\Omega_U^{00}} B^*p_h u \le 0. \tag{3.11}$$

Therefore, it follows from (3.9–3.11) that

$$(B^*p_h+\alpha u_h, u_h-u)_U \le C\|u_h - R_h u_h\|^2_{0,\Omega_U} + C\epsilon_1^2 + C\delta\|u-u_h\|^2_{0,\Omega_U}. \tag{3.12}$$

It follows from (3.4), (3.7), (3.8) and (3.12) that

$$\|u-u_h\|^2_{0,\Omega_U} \le C\|u_h - R_h u_h\|^2_{0,\Omega_U} + C\epsilon_1^2 + C\|p_h - p(u_h)\|^2_{0,\Omega}. \tag{3.13}$$

Now let us consider the last term $\|p_h - p(u_h)\|^2_{0,\Omega}$. Firstly, let us introduce the weighted Clément-type interpolation operator π defined in [4]. For all $v \in H_0^1(\Omega)$, $\pi v \in V^h \subset H_0^1(\Omega)$ and

$$\sum_{\tau\in T^h} \|h_\tau^{-1}(v-\pi v)\|^2_{0,\tau} \le C|v|^2_{1,\Omega},$$

$$|\pi v|^2_{1,\Omega} \le C|v|^2_{1,\Omega}.$$

Furthermore, if $f \in L^2(\Omega)$,

$$\int_\Omega f(v-\pi v) \le C|v|_{1,\Omega}\left(\sum_{z\in\Lambda}\int_{\omega_z} h_z^2|f-\bar{f}_z|^2\right)^{\frac{1}{2}},$$

where $\bar{f}_z$, ω_z, h_z, Λ are dined in Theorem 3.1 (the proof can be found in, e.g., [4, 23]).

Let $e^p = p(u_h) - p_h$, $e_I^p = \pi e^p$. Then, it can be shown that

$$\begin{aligned}
c\|e^p\|^2_{1,\Omega} &\le a(e^p, p(u_h)-p_h) = a(e^p - e_I^p, p(u_h)-p_h) + a(e_I^p, p(u_h)-p_h) \\
&= (g'(y(u_h)), e^p - e_I^p) - \sum_\tau \int_{\partial\tau} A^*\nabla p_h \cdot n(e^p - e_I^p) + (g'(y(u_h)) - g'(y_h), e_I^p) \\
&= (g'(y_h), e^p - e_I^p) - \sum_{l\cap\partial\Omega=\emptyset} \int_l [A^*\nabla p_h \cdot n](e^p - e_I^p) + (g'(y(u_h)) - g'(y_h), e^p) \\
&\le C\sum_{z\in\Lambda}\int_{\omega_z} h_z^2|g'(y_h) - \overline{g'(y_h)}_z|^2 + C\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A^*\nabla p_h\cdot n]^2 \\
&\quad + C\|g'(y(u_h)) - g'(y_h)\|^2_{0,\Omega} + C\delta\|e^p\|^2_{1,\Omega} \\
&\le C\epsilon_3^2 + C\sum_{l\cap\partial\Omega=\emptyset} h_l\int_l [A^*\nabla p_h\cdot n]^2 + C\|y(u_h) - y_h\|^2_{0,\Omega} + C\delta\|e^p\|^2_{1,\Omega}.
\end{aligned}$$

Hence,

$$\|p(u_h) - p_h\|^2_{1,\Omega} \le C\sum_{l\cap\partial\Omega=\emptyset} h_l\int_l [A^*\nabla p_h\cdot n]^2 + C\|y(u_h)-y_h\|^2_{0,\Omega} + C\epsilon_3^2. \tag{3.14}$$

Note that $G_h p_h$ is continuous on Ω. We have that

$$\begin{aligned}
\sum_{l\cap\partial\Omega=\emptyset} h_l\int_l [A^*\nabla p_h\cdot n]^2 &\le C\sum_{l\cap\partial\Omega=\emptyset} h_l\int_l [\nabla p_h]^2 = C\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [\nabla p_h - G_h p_h]^2 \\
&\le C\sum_{l\cap\partial\Omega=\emptyset}\sum_{l\cap\partial\tau\ne\emptyset}\int_\tau (\nabla p_h - G_h p_h)^2 \le C\|\nabla p_h - G_h p_h\|^2_{0,\Omega}.
\end{aligned} \tag{3.15}$$

It follows from (3.14) and (3.15) that

$$\|p(u_h) - p_h\|_{1,\Omega}^2 \le C\|\nabla p_h - G_h p_h\|_{0,\Omega}^2 + C\|y(u_h) - y_h\|_{0,\Omega}^2 + C\epsilon_3^2. \tag{3.16}$$

Similarly, let $e^y = y(u_h) - y_h$, $e_I^y = \pi e^y$. We have that

$$\begin{aligned}
c\|e^y\|_{1,\Omega}^2 &\le a(y(u_h) - y_h, e^y) = a(y(u_h) - y_h, e^y - e_I^y) \\
&= (f + Bu_h, e^y - e_I^y) - \sum_\tau \int_{\partial\tau} A\nabla y_h \cdot n(e^y - e_I^y) \\
&= (f + Bu_h, e^y - e_I^y) - \sum_{l\cap\partial\Omega=\emptyset} \int_l [A\nabla y_h \cdot n](e^y - e_I^y) \\
&\le C\sum_{z\in\Lambda} \int_{\omega_z} h_z^2 |f + Bu_h - (\overline{f + Bu_h})_z|^2 \\
&\quad + C\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A\nabla y_h \cdot n]^2 + C\delta\|e^y\|_{1,\Omega}^2 \\
&\le C\epsilon_2^2 + C\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A\nabla y_h \cdot n]^2 + C\delta\|e^y\|_{1,\Omega}^2.
\end{aligned}$$

Again,

$$\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A\nabla y_h \cdot n]^2 \le C\|\nabla y_h - G_h y_h\|_{0,\Omega}^2.$$

Thus,

$$\|y(u_h) - y_h\|_{1,\Omega}^2 \le C\|\nabla y_h - G_h y_h\|_{0,\Omega}^2 + C\epsilon_2^2. \tag{3.17}$$

Therefore, it follows from (3.13), (3.16) and (3.17) that

$$\begin{aligned}
&\|y(u_h) - y_h\|_{1,\Omega}^2 + \|p(u_h) - p_h\|_{1,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2 \\
&\quad\le C\|u_h - R_h u_h\|_{0,\Omega_U}^2 + C\|\nabla y_h - G_h y_h\|_{0,\Omega}^2 + C\|\nabla p_h - G_h p_h\|_{0,\Omega}^2 \\
&\qquad + C\epsilon_1^2 + C\epsilon_2^2 + C\epsilon_3^2 \\
&\quad= C\eta^2 + C\epsilon_1^2 + C\epsilon_2^2 + C\epsilon_3^2.
\end{aligned} \tag{3.18}$$

Note that

$$\begin{aligned}
\|y - y_h\|_{1,\Omega} &\le \|y - y(u_h)\|_{1,\Omega} + \|y(u_h) - y_h\|_{1,\Omega}, \\
\|p - p_h\|_{1,\Omega} &\le \|p - p(u_h)\|_{1,\Omega} + \|p(u_h) - p_h\|_{1,\Omega},
\end{aligned}$$

and

$$\begin{aligned}
\|y - y(u_h)\|_{1,\Omega} &\le C\|u - u_h\|_{0,\Omega_U}, \\
\|p - p(u_h)\|_{1,\Omega} &\le C\|y - y(u_h)\|_{0,\Omega} \le C\|u - u_h\|_{0,\Omega_U}.
\end{aligned}$$

We have that

$$\|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{1,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2 \le C\eta^2 + C\epsilon_1^2 + C\epsilon_2^2 + C\epsilon_3^2. \tag{3.19}$$

Furthermore, note that U^h is discontinuous finite element space. The L^2-project operator $\mathcal{P}_h$ from $L^2(\Omega_U)$ to U^h can be defined on every element, i.e.,

$$\int_{\tau_U}(\mathcal{P}_h w - w)\phi_{\tau_U} = 0 \quad \forall \tau_U \in T_U^h,$$

where ϕ_{τ_U} is the base function of U^h on the element τ_U. For all element $\tau_U^* \in T_U^h$ such that $u_h|_{\tau_U^*} > 0$, there exists an $\epsilon > 0$ such that $v_h = u_h \pm \epsilon\phi_{\tau_U^*} > 0$, and hence $v^h \in K^h$. Thus, it follows from (2.13) that

$$\int_{\tau_U^*}(\alpha u_h + B^* p_h)\phi_{\tau_U^*} = 0.$$

This implies that $\mathcal{P}_h(\alpha u_h + B^* p_h) = 0$ on τ_U^*. Note that $u_h > 0$ on Ω^*. Then we have $\mathcal{P}_h(\alpha u_h + B^* p_h) = 0$ on Ω^*. Therefore,

$$\begin{aligned} e^2 &= \int_{\Omega^*}((\alpha u + B^* p) - \mathcal{P}_h(\alpha u + B^* p))^2 \\ &\le C\int_{\Omega^*}((\alpha u + B^* p) - (\alpha u_h + B^* p_h))^2 \\ &\quad + C\int_{\Omega^*}(\alpha u_h + B^* p_h)^2 + C\int_{\Omega^*}\mathcal{P}_h(\alpha u_h + B^* p_h)^2 \\ &\quad + C\int_{\Omega^*}(\mathcal{P}_h(\alpha u_h + B^* p_h) - \mathcal{P}_h(\alpha u + B^* p))^2 \\ &\le C(\|p - p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2) + C\int_{\Omega^*}(\alpha u_h + B^* p_h)^2 \\ &\quad + 0 + C(\|p - p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2) \\ &\le C(\|p - p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2) + C\int_{\Omega_U^{++}}(\alpha u_h + B^* p_h)^2. \end{aligned} \tag{3.20}$$

It follows from (3.10) that

$$\begin{aligned} \int_{\Omega_U^{++}}(\alpha u_h + B^* p_h)^2 &\le \int_{\Omega_U^{-+}}(\alpha u_h + B^* p_h)^2 \\ &\le C\|R_h u_h - u_h\|_{0,\Omega_U}^2 + C\epsilon_1^2 + C\|u - u_h\|_{0,\Omega_U}^2 \\ &\le C\eta^2 + C\epsilon_1^2 + C\|u - u_h\|_{0,\Omega_U}^2. \end{aligned} \tag{3.21}$$

Then, (3.19–3.21) imply that

$$e^2 \le C(\|p - p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2) + C\eta^2 + C\epsilon_1^2 \le C\eta^2 + C\epsilon_1^2 + C\epsilon_2^2 + C\epsilon_3^2. \tag{3.22}$$

Summing up, (3.3) follows from (3.19) and (3.22). □

Theorem 3.2 *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of the systems (2.6–2.8) and (2.11–2.13), respectively. Assume that g' is Lipschitz continuous, and A is a constant positive definite matrix. Then,*

$$\eta^2 \le C(|y_h - y|_{1,\Omega}^2 + |p_h - p|_{1,\Omega}^2 + \|u_h - u\|_{0,\Omega_U}^2 + e^2) + C(\hat\epsilon_1^2 + \hat\epsilon_2^2 + \hat\epsilon_3^2), \tag{3.23}$$

where η *is defined by* (3.1), e *is defined by* (3.2), *and*

$$\hat{\epsilon}_1^2 = \int_{\Omega_U^d \setminus \Omega_U^-} (\alpha u_h + B^* p_h)^2,$$

$$\hat{\epsilon}_2^2 = \sum_\tau \int_\tau h_\tau^2 |f + Bu_h - (\overline{f + Bu_h})_\tau|^2,$$

$$\hat{\epsilon}_3^2 = \sum_\tau \int_\tau h_\tau^2 |g'(y_h) - \overline{g'(y_h)}_\tau|^2,$$

where

$$\Omega_U^d = \left\{ \bigcup \tau_U : u_h|_{\tau_U} = 0, \text{ and there exists } \tau_U' \in T_U^h \text{ such that } \bar{\tau}_U \cap \bar{\tau}_U' \neq \emptyset, u_h|_{\tau_U'} > 0 \right\},$$

and

$$\bar{f}_\tau = \frac{\int_\tau f}{\int_\tau 1}.$$

Proof Note that for any node z,

$$G_h y_h(z) = \sum_{j=1}^{J_z} \alpha_z^j (\nabla y_h)_{\tau_z^j},$$

with

$$\sum_{j=1}^{J_z} \alpha_z^j = 1.$$

Then, on any element τ,

$$\begin{aligned} |\nabla y_h - G_h y_h|^2 &= \left| (\nabla y_h)_\tau - \sum_{z \cap \bar{\tau} \neq \emptyset} \phi_z \left(\sum_{j=1}^{J_z} \alpha_z^j (\nabla y_h)_{\tau_z^j} \right) \right|^2 \\ &= \left| \sum_{z \cap \bar{\tau} \neq \emptyset} \phi_z \left(\sum_{j=1}^{J_z} \alpha_z^j ((\nabla y_h)_\tau - (\nabla y_h)_{\tau_z^j}) \right) \right|^2 \\ &\leq C \sum_{\tau' \subset S_\tau} |(\nabla y_h)_\tau - (\nabla y_h)_{\tau'}|^2, \end{aligned}$$

where $S_\tau = \bigcup_{\tau' \in T_{S_\tau}} \tau'$, $T_{S_\tau} = \{\tau' \in T^h : \bar{\tau}' \cap \bar{\tau} \neq \emptyset\}$. Given an edge l of an element in T^h, let $[w]_l$ be the jump across l. Note that $\forall \tau, \tau' \subset S_\tau$, there exist a finite positive integer m_τ, which is independent of h, and elements $\tau_i \subset S_\tau$, $i = 0, 1, \ldots, m_\tau$, such that $\bar{\tau}_{i-1} \cap \bar{\tau}_i = l_i$, $i = 1, \ldots, m_\tau$, where $l_i \subset S_\tau$ are edges of the elements, $l_i \cap \partial S_\tau = \emptyset$, and $\tau = \tau_0$, $\tau' = \tau_{m_\tau}$. Hence,

$$|(\nabla y_h)_\tau - (\nabla y_h)_{\tau'}| = \left| \sum_{i=1}^{m_\tau} [\nabla y_h]_{l_i} \right| \leq \sum_{i=1}^{m_\tau} |[\nabla y_h]_{l_i}| \leq \sum_{l \subset S_\tau \setminus \partial S_\tau} |[\nabla y_h]_l|.$$

Since y_h is continuous on Ω, $[\frac{\partial y_h}{\partial t}]_l = 0$, if $l \cap \partial\Omega = \emptyset$, where $\frac{\partial y_h}{\partial t}$ is the tangent derivative of y_h. Therefore, $|[\nabla y_h]_l| = |[\frac{\partial y_h}{\partial n}]_l|$, if $l \cap \partial\Omega = \emptyset$. Then, we have that

$$\|Gy_h - \nabla y_h\|_{0,\tau}^2 \le Ch_\tau^2 \sum_{l\subset(S_\tau\backslash\partial S_\tau)} \left|\left[\frac{\partial y_h}{\partial n}\right]_l\right|^2 \le C \sum_{l\subset(S_\tau\backslash\partial S_\tau)} h_l \int_l [\nabla y_h \cdot n]_l^2,$$

where h_l is the size of the edge l. Hence, noting that A is positive definite, and $n = k[\nabla y_h]_l$, with $k = \pm(1/|[\nabla y_h]_l|)$, because $|[\nabla y_h]_l \cdot n| = |[\nabla y_h]_l|$. We have that

$$\begin{aligned}[\nabla y_h \cdot n]_l^2 &= k^2|[\nabla y_h]_l|^4 \le Ck^2((A[\nabla y_h]_l)\cdot[\nabla y_h]_l)^2 = C((A[\nabla y_h]_l)\cdot k[\nabla y_h]_l)^2 \\ &= C((A[\nabla y_h]_l)\cdot n)^2 = C[(A\nabla y_h)\cdot n]_l^2.\end{aligned}$$

Then,

$$\sum_\tau \|Gy_h - \nabla y_h\|_{0,\tau}^2 \le C \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [\nabla y_h \cdot n]_l^2 \le C \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A\nabla y_h \cdot n]^2. \tag{3.24}$$

Using the standard bubble technique (see, e.g. [1, 22]), it can be shown that

$$\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A\nabla y_h\cdot n]^2 \le C\|y-y_h\|_{1,\Omega}^2 + C\sum_\tau h_\tau^2 \int_\tau (f+Bu_h)^2 + C\|u-u_h\|_{0,\Omega_U}^2,$$

and

$$\begin{aligned}\sum_\tau h_\tau^2 \int_\tau (f+Bu_h)^2 &\le C\|y-y_h\|_{1,\Omega}^2 + C\|u-u_h\|_{0,\Omega_U}^2 \\ &\quad + C\sum_\tau h_\tau^2 \int_\tau (f+Bu_h-(\overline{f+Bu})_\tau)^2.\end{aligned}$$

Hence,

$$\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A\nabla y_h\cdot n]^2 \le C\|y-y_h\|_{1,\Omega}^2 + C\|u-u_h\|_{0,\Omega_U}^2 + C\hat{\epsilon}_2^2. \tag{3.25}$$

Then, it follows from (3.24) and (3.25) that

$$\|Gy_h - \nabla y_h\|_{0,\Omega}^2 \le C\|y-y_h\|_{1,\Omega}^2 + C\|u-u_h\|_{0,\Omega_U}^2 + C\hat{\epsilon}_2^2. \tag{3.26}$$

Similarly, it can be shown that

$$\|Gp_h - \nabla p_h\|_{0,\Omega}^2 \le C \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A^*\nabla p_h\cdot n]^2,$$

and

$$\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [A^*\nabla p_h\cdot n]^2 \le C\|p-p_h\|_{1,\Omega}^2 + C\|g'(y)-g'(y_h)\|_{0,\Omega}^2 + C\sum_\tau h_\tau^2 \int_\tau (g'(y_h))^2,$$

$$\sum_\tau h_\tau^2 \int_\tau (g'(y_h))^2 \le C\|p - p_h\|_{1,\Omega}^2 + C\|g'(y) - g'(y_h)\|_{0,\Omega}^2 + C\sum_\tau h_\tau^2 \int_\tau (g'(y_h) - \overline{g'(y_h)}_\tau)^2.$$

Hence, we have that

$$\|Gp_h - \nabla p_h\|_{0,\Omega}^2 \le C\|p - p_h\|_{1,\Omega}^2 + C\|y - y_h\|_{1,\Omega}^2 + C\hat{\epsilon}_3^2. \tag{3.27}$$

Repeating the proof in the beginning, it can be shown that

$$\begin{aligned}
\|R_h u_h - u_h\|_{0,\Omega_U}^2 &\le C \sum_{l\subset(\Omega_U^d\cup\Omega_h^+),l\cap\partial(\Omega_U^d\cup\Omega_h^+)=\emptyset} h_l^2[u_h]^2 \\
&= \frac{C}{\alpha^2} \sum_{l\subset(\Omega_U^d\cup\Omega_h^+),l\cap\partial(\Omega_U^d\cup\Omega_h^+)=\emptyset} h_l \int_l [\alpha u_h + B^* p_h]^2 \\
&\le C\int_{\Omega_U^d\cup\Omega_h^+} (\alpha u_h + B^* p_h)^2 \\
&= C\int_{\Omega_U^{-+}} (\alpha u_h + B^* p_h)^2 + C\int_{(\Omega_U^d\cup\Omega_h^+)\backslash\Omega_U^{-+}} (\alpha u_h + B^* p_h)^2 \\
&= C\int_{\Omega_U^{-+}} (\alpha u_h + B^* p_h)^2 + C\int_{\Omega_U^d\backslash\Omega_U^-} (\alpha u_h + B^* p_h)^2 \\
&= C\int_{\Omega_U^{-+}} (\alpha u_h + B^* p_h)^2 + C\hat{\epsilon}_1^2.
\end{aligned} \tag{3.28}$$

It follows from the definition of Ω_U^{-+} that

$$\int_{\Omega_U^{-+}} (\alpha u_h + B^* p_h)^2 = \int_{\Omega_U^-} (\alpha u_h + B^* p_h)^2 + \int_{\Omega_U^{++}} (\alpha u_h + B^* p_h)^2. \tag{3.29}$$

Note that $(B^* p + \alpha u) = 0$ when $u > 0$ and $B^* p \ge 0$ when $u = 0$. Let

$$\Omega_0^0 = \{x \in \Omega_U^- : u(x) = 0\}.$$

We have that

$$\begin{aligned}
\int_{\Omega_U^-} (\alpha u_h + B^* p_h)^2 &= \int_{\Omega_0^0} (\alpha u_h + B^* p_h - \alpha u)^2 + \int_{\Omega_U^-\backslash\Omega_0^0} (\alpha u_h + B^* p_h - \alpha u - B^* p)^2 \\
&\le C\left(\|u - u_h\|_{0,\Omega_U}^2 + \|p - p_h\|_{0,\Omega}^2 + \int_{\Omega_0^0} (B^* p_h)^2\right) \\
&\le C\left(\|u - u_h\|_{0,\Omega_U}^2 + \|p - p_h\|_{0,\Omega}^2 + \int_{\Omega_0^0} (B^* p_h - B^* p)^2\right) \\
&\le C(\|u - u_h\|_{0,\Omega_U}^2 + \|p - p_h\|_{0,\Omega}^2),
\end{aligned} \tag{3.30}$$

where we used the facts that $B^*p_h \leq 0 \leq B^*p$ on Ω_0^0. Moreover, note that $u > 0$ and hence $B^*p + \alpha u = 0$ on $\Omega_U^{++} \setminus \Omega^*$, and $\mathcal{P}_h(B^*p_h + \alpha u_h) = 0$ because that $u_h > 0$ on Ω^*. It can be deduced that

$$\begin{aligned}
\int_{\Omega_U^{++}} (B^*p_h + \alpha u_h)^2 &= \int_{\Omega^*} (B^*p_h + \alpha u_h)^2 + \int_{\Omega_U^{++}\setminus\Omega^*} (B^*p_h + \alpha u_h)^2 \\
&= \int_{\Omega^*} ((B^*p_h + \alpha u_h) - \mathcal{P}_h(B^*p_h + \alpha u_h))^2 \\
&\quad + \int_{\Omega_U^{++}\setminus\Omega^*} (B^*p_h + \alpha u_h - (B^*p + \alpha u))^2 \\
&\leq C\int_{\Omega^*} ((B^*p + \alpha u) - \mathcal{P}_h(B^*p + \alpha u))^2 \\
&\quad + C\int_{\Omega^*} ((B^*p_h + \alpha u_h) - (B^*p + \alpha u))^2 \\
&\quad + C\int_{\Omega^*} (\mathcal{P}_h(B^*p + \alpha u) - \mathcal{P}_h(B^*p_h + \alpha u_h))^2 \\
&\quad + C(\|p - p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2) \\
&\leq Ce^2 + C(\|p - p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega_U}^2). \qquad (3.31)
\end{aligned}$$

Therefore, it follows from (3.29–3.31) that

$$\int_{\Omega_U^{-+}} (\alpha u_h + B^*p_h)^2 \leq C(e^2 + \|u - u_h\|_{0,\Omega_U}^2 + \|p - p_h\|_{0,\Omega}^2). \qquad (3.32)$$

Hence, it follows from (3.28) and (3.32) that

$$\|R_h u_h - u_h\|_{0,\Omega_U}^2 \leq C(e^2 + \|u - u_h\|_{0,\Omega_U}^2 + \|p - p_h\|_{0,\Omega}^2) + C\hat{\epsilon}_1^2. \qquad (3.33)$$

Then, (3.23) follows from (3.26), (3.27) and (3.33). □

Remark 3.2 It is shown from Theorem 3.1 and 3.2 that for the standard regular mesh and general solution, the a posteriori error estimator η^2 provides the equivalent error bound for the quantity:

$$E^2 = |y_h - y|_{1,\Omega}^2 + |p_h - p|_{1,\Omega}^2 + \|u_h - u\|_{0,\Omega_U}^2 + e^2,$$

if the extra terms ϵ_i and $\hat{\epsilon}_i$, $i = 1, 2, 3$ (they are all computable), are of higher order.

Note that in the quantity E^2, the terms $|y_h - y|_{1,\Omega}^2$, $|p_h - p|_{1,\Omega}^2$ and $\|u_h - u\|_{0,\Omega_U}^2$ present the errors of the finite element approximation for the state y, the costate p and the control u. Then we should consider the left term e^2 in E^2. It is clear that meas(Ω^*) indicates the approximation errors of the numerical coincident sets. Furthermore, note that

$$\begin{aligned}
e^2 &= \int_{\Omega^*} ((B^*p + \alpha u) - \mathcal{P}_h(B^*p + \alpha u))^2 \\
&\leq Ch_U^2(\|p\|_{1,\Omega^{**}}^2 + \|u\|_{1,\Omega^{**}}^2) \\
&\leq Ch_U^2(\|p\|_{1,\infty,\Omega^{**}}^2 + \|u\|_{1,\infty,\Omega^{**}}^2)\text{meas}(\Omega^{**}),
\end{aligned}$$

where

$$\Omega^{**}=\Big\{\bigcup\tau_U:\tau_U\cap\Omega^*\neq\emptyset\Big\}.$$

Thus e^2 will be of higher order (if U^h is the piecewise constant finite element space) as long as $\text{meas}(\Omega^{**})=o(1)$. Therefore only in some rare cases where numerical coincident sets may not converge, the term e^2 may take effects.

Next, let us consider the possible higher order terms ϵ_i and $\hat\epsilon_i$, $i=1,2,3$. It is easy to see that if f, u and g' are smooth, ϵ_i^2 and $\hat\epsilon_i^2$, $i=2,3$, are all higher order terms.

Consider $\hat\epsilon_1$. If $u_h|_{\tau_U}>0$, then $u_h|_{\tau_U}=-(\int_{\tau_U}B^*p_h/\int_{\tau_U}\alpha)$. Hence,

$$|\alpha u_h+B^*p_h|_{\tau_U}=\left|B^*p_h-\left(\int_{\tau_U}B^*p_h\Big/\int_{\tau_U}1\right)\right|_{\tau_U}\le Ch_U.$$

Consider the case that $u_h|_{\tau_U'}=0$. Note that $B^*p_h>0$ on $\Omega_U^d\setminus\Omega_U^-$. Moreover, for any $\tau_U'\subset\Omega_U^d$, there exists an element τ_U such that $\bar\tau_U\cap\bar\tau_U'\neq\emptyset$, and

$$u_h|_{\tau_U}=-\left(\int_{\tau_U}B^*p_h\Big/\int_{\tau_U}\alpha\right)=-(B^*p_h(x_{\tau_U})/\alpha)>0,$$

where x_{τ_U} is the center of gravity for the element τ_U. Then,

$$|\alpha u_h+B^*p_h|_{\tau_U'}=|B^*p_h|_{\tau_U'}\le|(B^*p_h)_{\tau_U'}-B^*p_h(x_{\tau_U})|\le Ch_U.$$

It follows that $|\alpha u_h+B^*p_h|\le Ch_U$ in $\Omega_U^d\setminus\Omega_U^-$. Therefore,

$$\hat\epsilon_1^2=\int_{\Omega_U^d\setminus\Omega_U^-}|\alpha u_h+B^*p_h|^2\le Ch_U^2\text{meas}(\Omega_U^d\setminus\Omega_U^-)\le Ch_U^2\text{meas}(\Omega_U^d).$$

Then $\hat\epsilon_1^2$ should be a higher term because $\text{meas}(\Omega_U^d)$ should be $o(1)$ if the measure of the free boundary ($u=0$) is zero.

Last, let us consider ϵ_1. Let $\tilde p_h\in H^2(\Omega_U)$ be such that $\|\tilde p_h-B^*p_h\|_{0,\Omega_U}\le Ch_U^2$, and

$$\int_{\tau_U}\tilde p_h=\int_{\tau_U}B^*p_h\quad\forall\tau_U\in T_U^h.$$

Let

$$\Omega_{00}=\Big\{\bigcup\tau_U:u_h|_{\tau_U'}>0,\forall\tau_U'\in T_U^h,\bar\tau_U'\cap\bar\tau_U\neq\emptyset\Big\}.$$

Then $-\alpha R_hu_h=R_hB^*p_h=R_h\tilde p_h$ on Ω_{00}. Hence,

$$\|\alpha R_hu_h+\tilde p_h\|_{0,\Omega_{00}}=\|\tilde p_h-R_h\tilde p_h\|_{0,\Omega_{00}}\le Ch_U^2.$$

Hence,

$$\begin{aligned}\epsilon_1^2&=\int_{\Omega_U^{-+}}(B^*p_h+\alpha R_hu_h)^2=\int_{\Omega_{00}}(B^*p_h+\alpha R_hu_h)^2+\int_{\Omega_U^{-+}\setminus\Omega_{00}}(B^*p_h+\alpha R_hu_h)^2\\&\le Ch_U^4+Ch_U^2\text{meas}(\Omega_U^{-+}\setminus\Omega_{00}).\end{aligned}$$

Then ϵ_1^2 should be a higher term because $\text{meas}(\Omega_U^{-+}\setminus\Omega_{00})$ should be $o(1)$ if the measure of the free boundary is zero.

Remark 3.3 The matrix A in the Theorems 3.1 and 3.2 can be extended variable $A \in (W^{1,\infty}(\Omega))^{2\times 2}$. In this case, ϵ_i and $\hat{\epsilon}_i$, $i = 2, 3$, should be replaced by

$$\epsilon_2^2 = \sum_{z\in\Lambda} \int_{\omega_z} h_z^2 |f + Bu_h + \mathrm{div}(A\nabla y_h) - (\overline{f + Bu_h + \mathrm{div}(A\nabla y_h)})_z|^2,$$

$$\epsilon_3^2 = \sum_{z\in\Lambda} \int_{\omega_z} h_z^2 |g'(y_h) + \mathrm{div}(A^*\nabla p_h) - (\overline{g'(y_h) + \mathrm{div}(A^*\nabla p_h)})_z|^2,$$

$$\hat{\epsilon}_2^2 = \sum_{\tau} \int_{\tau} h_\tau^2 |f + Bu_h + \mathrm{div}(A\nabla y_h) - (\overline{f + Bu_h + \mathrm{div}(A\nabla y_h)})_\tau|^2$$
$$+ \sum_l h_l \int_l [(A\nabla y_h - \bar{A}_l \nabla y_h)\cdot n]^2,$$

$$\hat{\epsilon}_3^2 = \sum_{\tau} \int_{\tau} h_\tau^2 |g'(y_h) + \mathrm{div}(A^*\nabla p_h) - (\overline{g'(y_h) + \mathrm{div}(A^*\nabla p_h)})_\tau|^2$$
$$+ \sum_l h_l \int_l [(A^*\nabla p_h - \bar{A}_l^* \nabla p_h)\cdot n]^2,$$

where $\bar{A}_l$ ($\bar{A}_l^*$) is the integral average of A (A^*) on the edge l. Furthermore the Theorems 3.1 and 3.2 can be further extended to the case where the objective functional reads

$$J(u) = g(y) + \frac{\alpha}{2}\|u - u_0\|_{0,\Omega_U}^2,$$

where $u_0 \in U$. One only needs to replace B^*p (B^*p_h) by $B^*p - u_0$ ($B^*p_h - u_0$) in the proofs.

4 Superconvergence Analysis and Recovery

In this section, we will provide the superconvergence results. Firstly, let us consider the superconvergence analysis for the control u. Let

$$\Omega_U^+ = \left\{\bigcup \tau_U : \tau_U \subset \Omega_U, u|_{\tau_U} > 0\right\},$$
$$\Omega_U^0 = \left\{\bigcup \tau_U : \tau_U \subset \Omega_U, u|_{\tau_U} = 0\right\},$$
$$\Omega_U^b = \Omega_U \setminus (\Omega_U^+ \cup \Omega_U^0).$$

In this section, we assume that u and T_U^h are regular such that $\mathrm{meas}(\Omega_U^b) \le Ch_U$.

Lemma 4.1 *Let u and u_h be the solutions of (2.8) and (2.13), respectively. Let $u_I \in K^h$ be the L^2-project of u, such that*

$$u_I|_{\tau_U} = \frac{\int_{\tau_U} u}{\int_{\tau_U} 1}.$$

Assume that g' is Lipschitz continuous, and Ω is convex. Then,

$$\|u_h - u_I\|_{0,\Omega_U} \le C(h_U^{\frac{3}{2}} + h^2). \tag{4.1}$$

Proof Note that $u_h, u_I \in K^h \subset K$. It follows from (2.8) and (2.13) that

$$(\alpha u + B^* p, u - u_h) \leq 0,$$

and

$$(\alpha u_h + B^* p_h, u_h - u_I) \leq 0.$$

It follows that

$$\begin{aligned}
\alpha\|u_h - u_I\|^2_{0,\Omega_U} &= \alpha(u_h - u_I, u_h - u_I)_U \leq -(B^* p_h, u_h - u_I)_U - (\alpha u_I, u_h - u_I)_U \\
&= (B^* p, u - u_h)_U + (B^* p, u_I - u)_U + (\alpha u_I, u_I - u_h)_U + (B^*(p - p_h), u_h - u_I)_U \\
&\leq -\alpha(u, u - u_h)_U + (B^* p, u_I - u)_U + \alpha(u_I, u_I - u_h)_U + (B^*(p - p_h), u_h - u_I)_U \\
&= \alpha(u_I - u, u_I - u_h)_U + (\alpha u + B^* p, u_I - u)_U + (B^*(p - p(u_h)), u - u_I)_U \\
&\quad + (B^*(p - p(u_h)), u_h - u)_U + (B^*(p(u_h) - p_h), u_h - u_I)_U, \qquad (4.2)
\end{aligned}$$

where $p(u_h)$ is defined by (3.5–3.6). Note that Ω is convex. We have that $p \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$, and $u \in W^{1,\infty}(\Omega_U)$.

Let π^c be the integral average operator such that $\pi^c u = u_I$. It follows from the definition of u_I that

$$(u_I - u, u_I - u_h)_U = \sum_{\tau_U} (u_I - u_h) \int_{\tau_U} (\pi^c u - u) = 0, \qquad (4.3)$$

and

$$\begin{aligned}
(B^*(p - p(u_h)), u - u_I)_U &= \sum_{\tau_U} \int_{\tau_U} (B^*(p - p(u_h)) - \pi^c(B^*(p - p(u_h))))(u - \pi^c u) \\
&\leq C \sum_{\tau_U} h^2_{\tau_U} |B^*(p - p(u_h))|_{1,\tau_U} |u|_{1,\tau_U} \\
&\leq C h^2_U \|p - p(u_h)\|_{1,\Omega} \|u\|_{1,\Omega}. \qquad (4.4)
\end{aligned}$$

Using (2.6–2.7) and (3.5–3.6), we have that

$$\begin{aligned}
\|p - p(u_h)\|_{1,\Omega} &\leq C\|g'(y) - g'(y(u_h))\|_{0,\Omega} \leq C\|y - y(u_h)\|_{0,\Omega} \leq C\|B(u - u_h)\|_{0,\Omega} \\
&\leq C\|u - u_h\|_{0,\Omega_U} \leq C\|u - u_I\|_{0,\Omega_U} + C\|u_I - u_h\|_{0,\Omega_U} \\
&\leq C h_U |u|_{1,\Omega_U} + C\|u_I - u_h\|_{0,\Omega_U} \qquad (4.5)
\end{aligned}$$

and

$$\begin{aligned}
(B^*(p - p(u_h)), u_h - u)_U &= (p - p(u_h), B(u_h - u)) = a(y(u_h) - y, p - p(u_h)) \\
&= (g'(y) - g'(y(u_h)), y(u_h) - y) \leq 0. \qquad (4.6)
\end{aligned}$$

It follows from Schwarz inequality that

$$\begin{aligned}
(B^*(p(u_h) - p_h), u_h - u_I)_U &\leq C\|B^*(p(u_h) - p_h)\|_{0,\Omega_U} \|u_h - u_I\|_{0,\Omega_U} \\
&\leq C\|p(u_h) - p_h\|^2_{0,\Omega} + C\delta\|u_h - u_I\|^2_{0,\Omega_U}, \qquad (4.7)
\end{aligned}$$

where δ is an arbitrary small positive constant. Then, it follows from (4.2–4.7) that

$$\begin{aligned}\alpha\|u_h-u_I\|^2_{0,\Omega_U}&\le(\alpha u+B^*p,u_I-u)_U+Ch_U^2(h_U+\|u_I-u_h\|_{0,\Omega_U})\\&\quad+C\|p(u_h)-p_h\|^2_{0,\Omega}+C\delta\|u_I-u_h\|^2_{0,\Omega_U}\\&\le(\alpha u+B^*p,u_I-u)_U+Ch_U^3+C\|p(u_h)-p_h\|^2_{0,\Omega}\\&\quad+C\delta\|u_I-u_h\|^2_{0,\Omega_U}.\end{aligned}$$

Hence,

$$\|u_h-u_I\|^2_{0,\Omega_U}\le C(\alpha u+B^*p,u_I-u)_U+Ch_U^3+C\|p(u_h)-p_h\|^2_{0,\Omega}. \tag{4.8}$$

Let $p(y_h)\in H_0^1(\Omega)$ be the solution of the equation:

$$a(q,p(y_h))=(g'(y_h),q)\quad\forall q\in H_0^1(\Omega).$$

Then,

$$\|p(y_h)-p(u_h)\|_{0,\Omega}\le C\|g'(y_h)-g'(y(u_h))\|_{0,\Omega}\le C\|y_h-y(u_h)\|_{0,\Omega}. \tag{4.9}$$

Note that y_h and p_h are the standard finite element approximations of $y(u_h)$ and $p(y_h)$, respectively. We have that (see, e.g., [5])

$$\|y_h-y(u_h)\|_{0,\Omega}\le Ch^2|y(u_h)|_{2,\Omega}\le Ch^2 \tag{4.10}$$

and

$$\|p_h-p(y_h)\|_{0,\Omega}\le Ch^2|p(y_h)|_{2,\Omega}\le Ch^2. \tag{4.11}$$

Therefore, it follows from (4.9–4.11) that

$$\|p(u_h)-p_h\|_{0,\Omega}\le\|p(u_h)-p(y_h)\|_{0,\Omega}+\|p(y_h)-p_h\|_{0,\Omega}\le Ch^2. \tag{4.12}$$

Note that

$$\begin{aligned}(\alpha u+B^*p,u_I-u)_U&=\int_{\Omega_U^+}(\alpha u+B^*p)(u_I-u)+\int_{\Omega_U^0}(\alpha u+B^*p)(u_I-u)\\&\quad+\int_{\Omega_U^b}(\alpha u+B^*p)(u_I-u),\end{aligned}$$

and

$$(\alpha u+B^*p)|_{\Omega_U^+}=0,\qquad(u_I-u)|_{\Omega_U^0}=0.$$

Then,

$$\begin{aligned}(\alpha u+B^*p,u_I-u)_U&=\int_{\Omega_U^b}(\alpha u+B^*p)(u_I-u)\\&=\sum_{\tau_U\subset\Omega_U^b}\int_{\tau_U}(\alpha u+B^*p-\pi^c(\alpha u+B^*p))(\pi^cu-u)\end{aligned}$$

$$\leq C \sum_{\tau_U \subset \Omega_U^b} h_{\tau_U}^2 |\alpha u + B^* p|_{1,\tau_U} |u|_{1,\tau_U}$$

$$\leq C h_U^2 (\|u\|_{1,\infty,\Omega_U}^2 + \|p\|_{1,\infty,\Omega}^2) \mathrm{meas}(\Omega_U^b) \leq C h_U^3. \qquad (4.13)$$

Therefore, (4.1) follows from (4.8), (4.12) and (4.13). □

Lemma 4.2 *Suppose that all the conditions of Lemma 4.1 are valid. Then,*

$$\|R_h u - u\|_{0,\Omega_U} \leq C h_U^{\frac{3}{2}}, \qquad (4.14)$$

where R_h is the recovery operator defined in the beginning of Sect. 3.

Proof Note that $u \in W^{1,\infty}(\Omega_U)$, and $u \in H^2(\Omega_U^+ \cup \Omega_U^0)$. Let

$$\Omega_h^{++} = \left\{ \bigcup \tau_U : \omega_z \subset \Omega_U^+, \forall z \in \bar{\tau}_U \right\},$$

and

$$\Omega_h^{00} = \left\{ \bigcup \tau_U : \omega_z \subset \Omega_U^0, \forall z \in \bar{\tau}_U \right\}.$$

Then,

$$R_h u(x) = u(x) = 0 \quad \forall x \in \Omega_U^{00}. \qquad (4.15)$$

It can be proved by the standard technique (see, e.g., [5]) that

$$\|R_h u - u\|_{0,\Omega_U^{++}} \leq C h_U^2 \|u\|_{2,\Omega_U^+}, \qquad (4.16)$$

and

$$\|R_h u - u\|_{0,\Omega_U \setminus (\Omega_U^{++} \cup \Omega_U^{00})}^2 \leq C h_U^2 \|u\|_{1,\Omega_U^{bb}}^2 \leq C h_U^2 \|u\|_{1,\infty,\Omega_U}^2 \mathrm{meas}(\Omega_U^{bb}),$$

where

$$\Omega_U^{bb} = \Omega_U \setminus (\Omega_U^{++} \cup \Omega_U^{00}).$$

Note that $\mathrm{meas}(\Omega_U^b) = O(h_U)$ and hence $\mathrm{meas}(\Omega_U^{bb}) = O(h_U)$. We have that

$$\|R_h u - u\|_{0,\Omega_U \setminus (\Omega_U^{++} \cup \Omega_U^{00})} \leq C h_U^{\frac{3}{2}}. \qquad (4.17)$$

Therefore, it follows from (4.15–4.17) that

$$\|R_h u - u\|_{0,\Omega_U}^2 = \|R_h u - u\|_{0,\Omega_U^{++}}^2 + \|R_h u - u\|_{0,\Omega_U^{00}}^2 + \|R_h u - u\|_{0,\Omega_U \setminus (\Omega_U^{++} \cup \Omega_U^{00})}^2$$
$$\leq C h_U^4 + 0 + C h_U^3 \leq C h_U^3.$$

This proves (4.14). □

Theorem 4.1 *Suppose that all the conditions of Lemma 4.1 are valid. Then,*

$$\|R_h u_h - u\|_{0,\Omega_U} \leq C (h_U^{\frac{3}{2}} + h^2). \qquad (4.18)$$

Proof Let u_I be defined in Lemma 4.1. Then,

$$\|R_h u_h - u\|_{0,\Omega_U} \le \|u - R_h u\|_{0,\Omega_U} + \|R_h u - R_h u_I\|_{0,\Omega_U} + \|R_h u_I - R_h u_h\|_{0,\Omega_U}. \quad (4.19)$$

It follows from Lemma 4.2 that

$$\|u - R_h u\|_{0,\Omega_U} \le C h_U^{\frac{3}{2}}. \quad (4.20)$$

Noting the definition of R_h, we have that

$$R_h u = R_h u_I, \quad (4.21)$$

and

$$\|R_h u_I - R_h u_h\|_{0,\Omega_U} \le C\|u_I - u_h\|_{0,\Omega_U}. \quad (4.22)$$

It has been proved in Lemma 4.1 that

$$\|u_I - u_h\|_{0,\Omega_U} \le C(h_U^{\frac{3}{2}} + h^2). \quad (4.23)$$

Therefore, (4.18) follows from (4.19–4.23). □

Corollary 4.1 *Let u and u_h be the solutions of* (2.8) *and* (2.13), *respectively. Assume that g' is Lipschitz continuous, and Ω is convex. Then,*

$$\|u - u_h\|_{-1,\Omega_U} \le C(h_U^{\frac{3}{2}} + h^2). \quad (4.24)$$

Proof For any function $\phi \in H^1(\Omega_U)$, let $\phi_I \in U^h$ be the L^2-project of ϕ, such that

$$\phi_I|_{\tau_U} = \frac{\int_{\tau_U} \phi}{\int_{\tau_U} 1}.$$

Then,

$$(u - u_h, \phi)_U = (u - u_h, \phi - \phi_I)_U + (u - u_h, \phi_I)_U. \quad (4.25)$$

Note that

$$(u - u_h, \phi - \phi_I)_U \le \|u - u_h\|_{0,\Omega_U}\|\phi - \phi_I\|_{0,\Omega_U} \le C(h_U + h^2)h_U\|\phi\|_{1,\Omega_U}, \quad (4.26)$$

and it follows from Lemma 4.1 that

$$(u - u_h, \phi_I)_U = (u_I - u_h, \phi_I)_U \le \|u_I - u_h\|_{0,\Omega_U}\|\phi_I\|_{0,\Omega_U} \le C(h_U^{\frac{3}{2}} + h^2)\|\phi\|_{1,\Omega_U}. \quad (4.27)$$

Therefore, it follows from (4.25–4.27) that

$$\|u - u_h\|_{-1,\Omega_U} = \sup_{\phi \in H^1(\Omega_U)} \frac{(u - u_h, \phi)_U}{\|\phi\|_{1,\Omega_U}} \le C(h_U^{\frac{3}{2}} + h^2).$$

This proves the corollary. □

Then, we consider the superconvergence for the state y and the co-state p.

Lemma 4.3 *Let y, p be the solutions of* (2.6) *and* (2.7), *and y_h, p_h be the solutions of* (2.11) *and* (2.12). *Let y_I and p_I be the piecewise linear Lagrange interpolations of y and p. Assume that all the conditions in Lemma* 4.1 *are valid. Moreover, assume that the mesh T^h is uniform, and $y, p \in H^3(\Omega)$. Then,*

$$|y_h - y_I|_{1,\Omega} + |p_h - p_I|_{1,\Omega} \le C(h^2 + h_U^{\frac{3}{2}}). \tag{4.28}$$

Proof Let $w_h = y_h - y_I$. It follows from (2.6) and (2.11) that

$$\begin{aligned} c|y_h - y_I|_{1,\Omega}^2 &\le a(y_h - y_I, w_h) = a(y - y_I, w_h) + a(y_h - y, w_h) \\ &= a(y - y_I, w_h) + (B(u_h - u), w_h). \end{aligned} \tag{4.29}$$

Note that the mesh T^h is uniform. It can be shown that (see, e.g., [10, 11, 19])

$$|a(y - y_I, w_h)| \le Ch^2 |y|_{3,\Omega} |w_h|_{1,\Omega}. \tag{4.30}$$

It follows from Corollary 4.1 that

$$|(B(u_h - u), w_h)| = |(u_h - u, B^* w_h)_U| \tag{4.31}$$

$$\le C\|u_h - u\|_{-1,\Omega_U} \|B^* w_h\|_{1,\Omega_U} \le C(h_U^{\frac{3}{2}} + h^2)|w_h|_{1,\Omega}. \tag{4.32}$$

It follows from (4.29–4.31) that

$$|y_h - y_I|_{1,\Omega} \le C(h_U^{\frac{3}{2}} + h^2). \tag{4.33}$$

Similarly, let $q_h = p_h - p_I$. Then, it follows from (2.7) and (2.12) that

$$\begin{aligned} c|p_h - p_I|_{1,\Omega}^2 &\le a(q_h, p_h - p_I) = a(q_h, p - p_I) + a(q_h, p_h - p) \\ &= a(q_h, p - p_I) + (g'(y_h) - g'(y), q_h). \end{aligned} \tag{4.34}$$

Again, it can be shown that

$$|a(q_h, p - p_I)| \le Ch^2 |p|_{3,\Omega} |q_h|_{1,\Omega}. \tag{4.35}$$

It follows from (4.33), Poincare inequality and the standard interpolation error estimate (see, e.g., [5]) that

$$\begin{aligned} |(g'(y_h) - g'(y), q_h)| &\le C\|y_h - y\|_{0,\Omega}\|q_h\|_{0,\Omega} \le C(|y_h - y_I|_{1,\Omega} + \|y_I - y\|_{0,\Omega})|q_h|_{1,\Omega} \\ &\le C(h_U^{\frac{3}{2}} + h^2 + h^2|y|_{2,\Omega})|q_h|_{1,\Omega}. \end{aligned} \tag{4.36}$$

Therefore, it follows from (4.34–4.36) that

$$|p_h - p_I|_{1,\Omega} \le C(h_U^{\frac{3}{2}} + h^2). \tag{4.37}$$

Then, (4.28) follows from (4.33) and (4.37). □

Theorem 4.2 *Suppose that all the conditions in Lemmas* 4.1 *and* 4.3 *are valid. Then,*

$$\|G_h y_h - \nabla y\|_{0,\Omega} + \|G_h p_h - \nabla p\|_{0,\Omega} \le C(h^2 + h_U^{\frac{3}{2}}). \tag{4.38}$$

Proof Note that

$$\|G_h y_h - \nabla y\|_{0,\Omega} \le \|G_h y_h - G_h y_I\|_{0,\Omega} + \|G_h y_I - \nabla y\|_{0,\Omega}. \tag{4.39}$$

It follows from Lemma 4.3 that

$$\|G_h y_h - G_h y_I\|_{0,\Omega} \le C\|\nabla(y_h - y_I)\|_{0,\Omega} \le C(h^2 + h_U^{\frac{3}{2}}). \tag{4.40}$$

It has been proved in [24] (Remark 3.2 and Theorem 3.2) that $G_h v_I = \nabla v$ on τ_U if v is a quadratic function on the neighborhood of τ_U $(\bigcup_{\bar{\tau}'_U \cap \bar{\tau}_U \neq \emptyset}\{\tau'_U\})$. Then, it follows from the standard interpolation error estimate technique (see, e.g., [5]) that

$$\|G_h y_I - \nabla y\|_{0,\Omega} \le Ch^2 |y|_{3,\Omega}. \tag{4.41}$$

Therefore, it follows from (4.39–4.41) that

$$\|G_h y_h - \nabla y\|_{0,\Omega} \le C(h^2 + h_U^{\frac{3}{2}}). \tag{4.42}$$

Similarly, it can be proved that

$$\|G_h p_h - \nabla p\|_{0,\Omega} \le C(h^2 + h_U^{\frac{3}{2}}). \tag{4.43}$$

Therefore, (4.38) follows from (4.42) and (4.43). □

Based on the superconvergence analysis, we have the following results for the recovery type a posteriori error estimator.

Theorem 4.3 *Suppose that all the conditions of Theorems 4.1 and 4.2 are valid. Then,*

$$\|R_h u_h - u_h\|_{0,\Omega_U} = \|u - u_h\|_{0,\Omega_U} + O(h_U^{\frac{3}{2}} + h^2), \tag{4.44}$$

$$\|G_h y_h - \nabla y_h\|_{0,\Omega} = \|\nabla(y - y_h)\|_{0,\Omega} + O(h_U^{\frac{3}{2}} + h^2), \tag{4.45}$$

$$\|G_h p_h - \nabla p_h\|_{0,\Omega} = \|\nabla(p - p_h)\|_{0,\Omega} + O(h_U^{\frac{3}{2}} + h^2). \tag{4.46}$$

That is

$$\eta^2 = \|u - u_h\|^2_{0,\Omega_U} + \|\nabla(y - y_h)\|^2_{0,\Omega} + \|\nabla(p - p_h)\|^2_{0,\Omega} + o(h_U^{\frac{3}{2}} + h^2), \tag{4.47}$$

where η is defined in (3.1).

Proof Note that

$$\big|\|R_h u_h - u_h\|_{0,\Omega_U} - \|u - u_h\|_{0,\Omega_U}\big| \le \|R_h u_h - u\|_{0,\Omega_U}.$$

It follows from Theorem 4.1 that

$$\|R_h u_h - u\|_{0,\Omega_U} \le C(h_U^{\frac{3}{2}} + h^2).$$

Hence,

$$|\|R_h u_h - u_h\|_{0,\Omega_U} - \|u - u_h\|_{0,\Omega_U}| \le C(h_U^{\frac{3}{2}} + h^2).$$

This proves (4.44). The results (4.45) and (4.46) can be proved from Theorem 4.2 similarly, and (4.47) is a direct result from (4.44–4.46). □

Remark 4.1 In Theorem 4.3, it is shown that under some strong conditions, the a posteriori error estimator η^2 is not only equivalent as shown in the last section, but also asymptotically exact.

5 Numerical Examples

In this section, we carry out some numerical experiments to demonstrate the error estimators developed in Sect. 3. In most control problems, the optimal control is often of prime interest. Thus it is important to develop mesh refinement schemes which are most efficient to reduce the error $\|u - u_h\|$. Here we use the h-method, which will be briefly explained below. The general idea is to refine the meshes such that the error estimators are *equally* distributed over the computational mesh. Assume that an *a posteriori* error estimator η has the form $\eta^2 = \sum_{e_i} \eta_{e_i}^2$, where e_i is a finite element. In mesh refinement method for instance, at each iteration, an average quantity of $\{\eta_{e_i}^2\}$ is calculated, and each $\eta_{e_i}^2$ is then compared with this quantity. The element e_i is to be refined or coarsed if $\eta_{e_i}^2$ is larger or smaller than a certain proportional of this quantity. As $\eta_{e_i}^2$ reflects the distribution of the total approximation error over e_i, this strategy guarantees that a higher density of nodes is distributed over the area where the error is larger, see [22]. Let us mention that the r-method is used in [9] for the purpose of clear comparisons.

Our numerical examples are the following type of optimal control problems (OCP):

$$\min \quad \frac{1}{2}\int_\Omega (y - y_0)^2 + \frac{1}{2}\int_{\Omega_U} (u - u_0)^2$$

$$\text{s.t.} \quad \begin{cases} -\Delta y + \phi(y) = Bu + f, \\ y|_{\partial\Omega} = y_0|_{\partial\Omega}, \\ u \ge 0 \quad \text{in } \Omega_U. \end{cases}$$

In our examples, $\Omega_U = \Omega = [0, 1] \times [0, 1]$ and $B = I$. Let Ω^h and Ω_U^h be partitioned into T^h and T_U^h as described Sect. 2. We may use different meshes for the approximation of the state and the control. We utilize a C++ library AFEPack to provide a general tool of mesh adaptation for multi-meshes. The package is freely available and the details can be found at http://www.ukc.ac.uk/cbs/staff/homepage/wbl/data/AFELAB-AFEPACK.htm. In all our experiments, we shall use $\|R_h u_h - u_h\|_{0,\Omega_U}$ as the control mesh refinement indicator, and $\|G_h y_h - \nabla y_h\|_{0,\Omega} + \|G_h p_h - \nabla p_h\|_{0,\Omega}$ as the state's and co-state's.

With the error estimators derived in the above section, one can generate adaptive meshes and discretise the control problems into some finite dimensional optimization problems, which can be solved via mathematical programming. With our discontinues discretisation on control, here we use a simple and yet fast preconditioned projection algorithm for the control problem. It works very well with the adaptive multi-mesh discretisation, and can be used to solve large scale control problems. For a constrained optimization problem:

$$\min_{u\in K} J(u), \tag{5.1}$$

where $J(u)$ is a convex functional on U and K is a convex subset of U, the iterative scheme reads ($n = 0, 1, 2, \dots$)

$$\begin{cases} b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n(J'(u_n), v) \quad \forall v \in U, \\ u_{n+1} = P_K^b(u_{n+\frac{1}{2}}), \end{cases} \tag{5.2}$$

where $b(\cdot,\cdot)$ is a symmetric and positive definite bilinear form such that there exist constant c_0 and c_1 satisfying

$$|b(u,v)| \le c_1 \|u\|_U \|v\|_U \quad \forall u, v \in U, \tag{5.3}$$

$$b(u,u) \ge c_0 \|u\|_U^2 \tag{5.4}$$

and the projection operator $P_K^b\ U \to K$ is defined: For given $w \in U$ find $P_K^b w \in K$ such that

$$b(P_K^b w - w, P_K^b w - w) = \min_{u \in K} b(u - w, u - w). \tag{5.5}$$

The bilinear form $b(\cdot,\cdot)$ provides suitable preconditioning for the projection algorithm. Otherwise its speed may be slow when h is very small. One can just use a fixed step size, or variable ones from a line search procedure. When the step sizes are small enough, its convergence can be shown with the standard techniques. Let $U = U^h$. An application of (5.2) to the discretised elliptic control problem yields the following algorithm (PPGA).

$$\begin{cases} b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n(u_n - u_0 + B^* p_n, v), \quad u_{n+\frac{1}{2}}, u_n \in U^h, \forall v \in U^h, \\ a(y_n, w) = (f + Bu_n, w), \quad y_n \in V^h, \forall w \in V^h, \\ a(q, p_n) = (y_n - y_0, q), \quad p_n \in V^h, \forall w \in V^h, \\ u_{n+1} = P_K^b(u_{n+\frac{1}{2}}), \end{cases} \tag{5.6}$$

where we have omitted the subscript h. The main computational effort is to solve the two state equations, and to compute the projection $P_K^b u_{n+\frac{1}{2}}$. In this paper we use a fast algebraic multigrid solver to solve the state equations. Then it is clear that the key to saving computing time is how to compute $P_K^b u_{n+\frac{1}{2}}$ efficiently. If one uses the C^0 finite elements to approximate the control, then one has to solve a global variational inequality, via, e.g., Semi-smooth Newton method. The computational load is not trivial. Our discontinuous discretization in control makes it possible to explicitly compute P_K^b. For the piecewise constant elements, $K^h = \{u_h : u_h \ge 0\}$ and $b(u,v) = (u,v)_U$, then

$$P_K^b p_h|_{\tau_U} = \max(0, \operatorname{avg}(p_h)|_{\tau_U}),$$

where $\operatorname{avg}(p_h)|_{\tau_U}$ is the average of p_h over τ_U.

In solving our discretised optimal problem, we use the preconditioned projection gradient method (5.6) with $b(u,v) = (u,v)_{U^h}$ and a fixed step size ρ. We now briefly describe the solution algorithm to be used for solving the numerical examples in this section:

Algorithm 0

(i) solve the discritized optimization problem with the projection gradient method on the current meshes and calculate the error estimators η_i;
(ii) adjust the meshes using the estimators and update the solution on new meshes, as described.

Table 1 Comparison between uniform mesh and adaptive mesh for Example 1, the errors are $\|u-u_h\|_{L^2}$, $|y-y_h|_{H^1}$ and $|p-p_h|_{H^1}$, respectively

Item		On uniform mesh			On adaptive mesh		
		u	y	p	u	y	p
Mesh	# nodes		8097		721	804	804
	# sides info		23968		1868	2215	2215
	# elements		15872		1148	1412	1412
	# dofs	47616	8097	8097	3444	804	804
Error		6.07904e-03	5.87182e-01	5.94960e-02	5.09988e-03	2.07212e+00	1.05092e-01

Table 2 Example 1 on sequential uniform meshes

	1	2	3	4	5
# nodes	537	2065	8097	32065	127617
$\|u-u_h\|_{L^2}$	1.26596e-02	8.68499e-03	6.07904e-03	4.30752e-03	3.07156e-03
$\|y-y_h\|_{H^1}$	2.34528e+00	1.17398e+00	5.87182e-01	2.93618e-01	1.46812e-01
$\|p-p_h\|_{H^1}$	1.18983e-01	5.94960e-02	2.97496e-02	1.48752e-02	7.43766e-03
$\|R_h u_h-u_h\|_{L^2}$	1.22897e-02	8.68152e-03	5.77539e-03	4.06993e-03	2.91860e-03
$\|G_h y_h-\nabla y_h\|_{L^2}$	2.38738e+00	1.18247e+00	5.89036e-01	2.94047e-01	1.46915e-01
$\|G_h p_h-\nabla p_h\|_{L^2}$	1.20586e-01	5.98600e-01	2.98353e-02	1.48959e-02	7.44275e-03

Example 1 The first example is to solve the following model problem as

$$\begin{aligned} \min \quad & \frac{1}{2}\int_\Omega (y-y_0)^2dx+\frac{1}{2}\int_\Omega (u-u_0)^2dx \\ \text{s.t.} \quad & -\Delta y = u+f, \quad u\ge 0, \end{aligned} \tag{5.7}$$

in which $\Omega=(0,1)\times(0,1)$, and

$$\begin{aligned} & z=\begin{cases} 0.5, & x_1+x_2>1.0, \\ 0.0, & x_1+x_2\le 1.0, \end{cases} \\ & p=\sin\pi x_1 \sin\pi x_2, \\ & u_0 = 1.0-\sin\frac{\pi x_1}{2}-\sin\frac{\pi x_2}{2}+z, \\ & u=\max(u_0-p,0), \\ & y_0=0, \\ & f=4\pi^4 p-u, \\ & y=2\pi^2 p. \end{aligned} \tag{5.8}$$

In the numerical simulation, we use piecewise linear finite element space for the approximation of y and p, and piecewise constant for u.

It is clear that the adaptive meshes generated via the error indicators are able to save substantial computational work, in comparison with the uniform meshes. We now examine exactness of the error indicators.

Table 3 Example 1 on sequential adaptive meshes

Adaptive step	1	2	3	4	5
# nodes(u)	145	204	398	519	721
# nodes(y, p)	145	537	804	804	804
$\|u-u_h\|_{L^2}$	2.08452e-02	1.32012e-02	9.07738e-03	6.65461e-03	5.09988e-03
$\|y-y_h\|_{H^1}$	4.67225e+00	2.34528e+00	2.07212e+00	2.07212e+00	2.07212e+00
$\|p-p_h\|_{H^1}$	2.38026e-01	1.18983e-01	1.05092e-01	1.05092e-01	1.05092e-01
$\|R_h u_h - u_h\|_{L^2}$	1.79538e-02	1.23715e-02	8.81438e-03	6.11814e-03	4.62809e-03
$\|G_h y_h - \nabla y_h\|_{L^2}$	4.89647e+00	2.38738e+00	2.09266e+00	2.09266e+00	2.09266e+00
$\|G_h p_h - \nabla p_h\|_{L^2}$	2.45173e-01	1.20586e-01	1.05759e-01	1.05759e-01	1.05759e-01

Table 4 Comparison between uniform mesh and adaptive mesh for Example 2, the errors are $\|u-u_h\|_{L^2}$, $|y-y_h|_{H^1}$ and $|p-p_h|_{H^1}$, respectively

Item		On uniform mesh			On adaptive mesh		
		u	y	p	u	y	p
Mesh	# nodes		8097		4290	174	174
info	# sides		23968		12007	463	463
	# elements		15872		7718	290	290
	# dofs	47616	8097	8097	23154	174	174
Error		3.55106e-02	1.24857e-01	4.11802e-05	2.27846e-02	9.12034e-01	2.21293e-03

The above two tables confirm that the error indicators are not only equivalent but also asymptotically exact.

Example 2 This is an example of nonlinear control problems,

$$\begin{aligned} \min \quad & \frac{1}{2}\int_\Omega (y-y_0)^2 dx + \frac{1}{2}\int_\Omega (u-u_0)^2 dx \\ \text{s.t.} \quad & -\Delta y + y^3 = u + f, \quad u \ge 0, \end{aligned} \tag{5.9}$$

in which $\Omega = (0,1)\times(0,1)$, and

$$\begin{aligned} y_0 &= \sin 2\pi x_1 + \sin 2\pi x_2, \\ y &= y_0, \\ u_0 &= \max(4\pi^2 y_0, 0), \\ u &= u_0, \\ f &= 4\pi^2 y_0 + y_0^3 - u, \\ p &= 0. \end{aligned} \tag{5.10}$$

The dual equation of the state equation is

$$-\Delta p + 3y^2 p = y - y_0. \tag{5.11}$$

Table 5 Example 2 on sequential uniform meshes

	1	2	3	4	5
# nodes	537	2065	8097	32065	127617
$\|u-u_h\|_{L^2}$	2.70932e-01	9.71364e-02	3.55106e-02	1.24375e-02	4.36575e-03
$\|y-y_h\|_{H^1}$	4.98455e-01	2.49610e-01	1.24857e-01	6.24357e-02	3.12188e-02
$\|p-p_h\|_{H^1}$	6.56826e-04	1.64519e-04	4.11802e-05	1.02861e-05	2.55767e-06
$\|R_hu_h-u_h\|_{L^2}$	2.24279e-01	8.20346e-02	3.02998e-02	1.07037e-02	3.74490e-03
$\|G_hy_h-\nabla y_h\|_{L^2}$	5.03510e-01	2.50141e-01	1.24894e-01	6.24328e-02	3.12165e-02
$\|G_hp_h-\nabla p_h\|_{L^2}$	6.55469e-05	8.16034e-06	1.01707e-06	1.26914e-07	1.58380e-08

Table 6 Example 2 on sequential adaptive meshes

Adaptive step	1	2	3	4	5
# nodes(u)	145	384	1288	2710	4290
# nodes(y, p)	145	174	174	174	174
$\|u-u_h\|_{L^2}$	8.13922e-01	2.70932e-01	9.71366e-02	4.01130e-02	2.27846e-02
$\|y-y_h\|_{H^1}$	9.91391e-01	9.12057e-01	9.12057e-01	9.12057e-01	9.12057e-01
$\|p-p_h\|_{H^1}$	2.58517e-03	2.22193e-03	2.22193e-03	2.22193e-03	2.22193e-03
$\|R_hu_h-u_h\|_{L^2}$	6.95352e-01	2.24280e-01	8.20353e-02	3.09396e-02	1.26466e-02
$\|G_hy_h-\nabla y_h\|_{L^2}$	1.03237e+00	9.49570e-01	9.49570e-01	9.49570e-01	9.49570e-01
$\|G_hp_h-\nabla p_h\|_{L^2}$	5.20497e-04	4.48978e-04	4.48978e-04	4.48978e-04	4.48978e-04

Although our error indicators were derived for a linear model problem, the analysis is applicable to this extension as well. Thus we present the example to demonstrate this. Again, we use piecewise linear finite element space for the approximation of y and p, and piecewise constant for u.

Let us mention that the main part of the computation is to solve the two state equations repeatedly. Thus it is clear that the adaptive scheme is able to solve huge computational work in this case.

Again we are able to confirm the equivalence and exactness of the estimators in this case.

Acknowledgement The first author was partially supported by EPSRC GR/R31980 and NSFC grant NSFC10301003. N. Yan's research was supported by the National Natural Science Foundation of China (No. 60474027).

References

1. Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis. Wiley Interscience, New York (2000)
2. Becker, R., Rannacher, R.: An optimal control approach to a-posteriori error estimation. In: Iserles, A. (ed.) Acta Numerica 2001, pp. 1–102. Cambridge University Press, Cambridge (2001)
3. Becker, R., Kapp, H., Rannacher, R.: Adaptive finite element methods for optimal control of partial differential equations: basic concept. SIAM J. Control Optim. **39**(1), 113–132 (2000)
4. Carstensen, C., Verfürth, R.: Edge residual dominate a posteriori error estimates for low order finite element methods. SIAM J. Numer. Anal. **36**, 1571–1587 (1999)

5. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
6. Falk, F.S.: Approximation of a class of optimal control problems with order of convergence estimates. J. Math. Anal. Appl. **44**, 28–47 (1973)
7. Haslinger, J., Neittaanmaki, P.: Finite Element Approximation for Optimal Shape Design. Wiley, Chichester (1989)
8. Huang, Y., Li, R., Liu, W., Yan, N.: Multi-adaptive mesh discretization for finite element approximation of constrained optimal control problems. SIAM. J. Control Optim. (to appear)
9. Li, R., Lin, W., Ma, H., Tang, T.: Adaptive finite element approximation for distributed elliptic optimal control problems. SIAM J. Control Optim. **41**(5), 1321–1349 (2002)
10. Krizek, M., Neitaanmaki, P., Stenberg, R.: Finite Element Methods: Superconvergence, Post-processing, and A Posteriori Estimates. Lectures Notes in Pure and Applied Mathematics, vol. 196. Dekker, New York (1998)
11. Lin, Q., Yan, N.: Construction and Analysis of High Efficient Finite Element. Hebei University Press (1996) (in Chinese)
12. Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
13. Liu, W., Yan, N.: A posteriori error analysis for convex distributed optimal control problems. Adv. Comput. Math. **15**, 285–309 (2001)
14. Liu, W., Yan, N.: A posteriori error estimates for convex boundary control problems. SIAM J. Numer. Anal. **39**, 73–99 (2001)
15. Liu, W., Yan, N.: A posteriori error estimates for control problems governed by Stokes equations. SIAM J. Numer. Anal. **40**(5), 1850–1869 (2002)
16. Liu, W., Yan, N.: A posteriori error estimates for optimal control problems governed by parabolic equations. Numer. Math. **93**(3), 497–521 (2003)
17. Malanowski, K.: Convergence of approximations vs. regularity of solutions for convex, control constrained, optimal control systems. Appl. Math. Optim. **8** (1982)
18. Neittaanmaki, P., Tiba, D.: Optimal Control of Nonlinear Parabolic Systems. Theory, Algorithms and Applications. Dekker, New York (1994)
19. Oganesyan, L.A., Rukhovetz, L.A.: Study of the rate of convergence of variational difference scheme for second order elliptic equation in two-dimensional field with a smooth boundary. V.S.S.R. Comput. Math. Math. Phys. **9**, 158–183 (1969)
20. Pironneau, O.: Optimal Shape Design for Elliptic Systems. Springer, Berlin (1984)
21. Tiba, D., Troltzsch, F.: Error estimates for the discretization of state constrained convex control problems. Numer. Funct. Anal. Optim. **17**, 1005–1028 (1996)
22. Verfurth, R.: A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement. Wiley-Teubner, London (1996)
23. Yan, N., Zhou, A.: Gradient recovery type a posteriori error estimates for finite filament approximations on irregular meshes. Comput. Methods Appl. Mech. Eng. **190**, 4289–4299 (2001)
24. Zhang, Z., Yan, N.: Recovery type a posteriori error estimates in finite element methods. Korean J. Comput. Appl. Math. **8**, 235–251 (2001)
25. Zienkiewicz, O.C., Zhu, J.Z.: The superconvergent patch recovery and a posteriori error estimates. Int. J. Numer. Methods Eng. **33**, 1331–1382 (1992)
26. Zienkiewicz, O.C., Zhu, J.Z.: The superconvergence patch recovery (SPR) and adaptive finite element refinement. Comput. Methods Appl. Mech. Eng. **101**, 207–224 (1992)