Journal of Computational Mathematics
Vol.28, No.5, 2010, 645–675.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2009.10-m1016

ADAPTIVE FINITE ELEMENT APPROXIMATION FOR A CLASS OF PARAMETER ESTIMATION PROBLEMS*

Karl Kunisch
Department of Mathematics, University of Graz, Austria
Email: karl.kunisch@uni-graz.at
Wenbin Liu
KBS, University of Kent, Canterbury, CT2 7NF, England
Email: w.b.liu@kent.ac.uk
Yanzhen Chang and Ningning Yan
LSEC, Academy of Mathematics and Systems Science,
Chinese Academy of Sciences, Beijing, China
Email: ynn@amss.ac.cn
Ruo Li
Department of Computational Mathematics, Peking University, China
Email: rli@math.pku.edu.cn

Abstract

In this paper, we study adaptive finite element discretisation schemes for a class of parameter estimation problem. We propose to use adaptive multi-meshes in developing efficient algorithms for the estimation problem. We derive equivalent a posteriori error estimators for both the state and the control approximation, which particularly suit an adaptive multi-mesh finite element scheme. The error estimators are then implemented and tested with promising numerical results.

Mathematics subject classification: 49J20, 65N30.
Key words: Parameter estimation, Finite element approximation, Adaptive finite element methods, A posteriori error estimate.

1. Introduction

Adaptive finite element approximation is very important in improving accuracy and efficiency of the finite element discretisation because it ensures a higher density of nodes in certain area of the computational domain, where the solution is more difficult to approximate. By now the theory and application of adaptive finite element methods for the numerical solutions of partial differential equations (PDEs)have reached some state of maturity as documented by a series of monographs. There has been so extensive research on developing adaptive finite element algorithms for PDEs in the scientific literature that it is simply impossible to give even a very brief review here.

Recently, there has been intensive research in adaptive finite element method for optimal control problems, see, e.g., [2–4,16,19–22]. The main existing approaches are the goal-orientated a posteriori error estimators, see, e.g., [3,4], and the residual based a posteriori error estimators, see, e.g., [16, 19, 20], where a posteriori error estimates equivalent to the energy norm of the

* Received March 9, 2009 / Revised version received xxx, 2009 / Accepted June 6, 2009 /
Published online May 1, 2010 /

approximation error were derived for several types of optimal control problems. The readers can refer to the recent monograph [21] for more details.

Furthermore, it has been found that for constrained control problems, different adaptive meshes are often needed for the control and the states, see [15, 16]. Using different adaptive meshes for the control and the state allows very coarse meshes to be used in solving the state and co-state equations. Thus much computational work can be saved since one of the major computational loads in computing optimal control is to solve the state and co-state equations repeatedly. This will be also seen from our numerical experiments in Section 6.

In this paper, we are interested in the least-square formulation of the following parameter estimation problem in R^n $(n \leq 3)$:

$$\min_{u \in K}\{g(y) + j(u)\}, \tag{1.1}$$

subject to

$$-\mathrm{div}(A\nabla y(u)) + uy(u) = f \quad \text{in } \Omega. \tag{1.2}$$

where u is defined on Ω, and Ω is a bounded and simply connected open sets in R^n $(n \leq 3)$ with Lipschitz continuous boundaries $\partial\Omega$. Here $j(u) = \int_\Omega h(u)$ is convex functional, $f \in L^2(\Omega)$, and K is a closed convex set. We also assume that g and j are convex functionals which are continuously differentiable, and j is further strictly convex with $j(u) \to +\infty$ as $\|u\|_U \to \infty$, $g(\cdot)$ is bounded below. For the matrix A we assume that $A(\cdot) = (a_{i,j}(\cdot))_{n\times n} \in (W^{1,\infty}(\Omega))^{n\times n}$, such that there is a constant $c > 0$ satisfying that for any vector $X \in R^n$, it gives $X^tAX \geq c\|X\|^2_{R^n}$. The above problem is of course a class of optimal control problem. In comparison with the standard optimal control problems, there were relatively fewer known results in developing adaptive finite element approximation for parameter estimation problems due to the lower regularity of the parameter that often is discontinuous. In [5], goal-orientated a posteriori error estimators were developed for a class of parameter identification problem, and computational tests were presented. In [7, 8, 13], a posteriori error estimators of residual type were developed for the same problem but with stronger assumptions on the estimated parameter as required by the techniques used. In particular, these assumptions eliminate any jumps in the estimated parameter. Very recently a priori error estimates and super-convergence were presented in [26] for the above estimation problem, although much more work in convexity of the functional, regularity of the parameter, and a posteriori error estimation techniques was still needed before a posteriori error estimators of residual type can be rigorously derived.

The purpose of this work is to develop residual a posteriori error estimators for the adaptive finite element approximation of the above problem. In our work, the estimated parameter is assumed just in L^2 so that jumps in value are allowed for the estimated parameter. The plan of the paper is as follows. In Section 2, we introduce some notations and preliminaries. In Section 3, we will construct the finite element approximation for the parameter estimation problem. In Sections 4 and 5, sharp a posteriori error estimators are derived for the parameter identification problem. Finally numerical test results are presented in Section 6.

2. Notations and Preliminaries

2.1. Some notations

We adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{W^{m,q}(\Omega)}$ and seminorm $|\cdot|_{W^{m,q}(\Omega)}$ (or $\|\cdot\|_{m,q,\Omega}$, $|\cdot|_{m,q,\Omega}$ for simplification). Further we set $W^{1,q}_0(\Omega) \equiv$

$\{w \in W^{1,q}(\Omega) : w|_{\partial\Omega} = 0\}$ and abbreviate $W^{m,2}(\Omega)$ by $H^m(\Omega)$ ($W_0^{1,2}(\Omega)$ by $H_0^1(\Omega)$). In addition c and C denote generic positive constant independent of h. We will take the state space $V = H_0^1(\Omega)$, the parameter space $U = L^2(\Omega)$ and $H = L^2(\Omega)$ with the inner product $(\cdot,\cdot)$.

Let

$$\begin{aligned} a(y,v) &= \int_\Omega (A\nabla y)\cdot\nabla v \quad \forall y,v \in V,\\ (f_1,f_2) &= \int_\Omega f_1 f_2 \qquad\qquad \forall (f_1,f_2) \in H\times H. \end{aligned}$$

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

$$a(y,y) \geq c\|y\|_V^2, \quad |a(y,w)| \leq C\|y\|_V\|w\|_V. \tag{2.1}$$

Then the standard weak formula for the state equation (1.2) reads: find $y(u) \in V$ such that

$$a(y(u),v) + (uy(u),v) = (f,v) \qquad \forall v \in V. \tag{2.2}$$

Then, problem (1.1) and (1.2) can be rewritten as: (OBC) find $(y(u),u) \in V\times U$ such that

$$\min_{u\in K}\{g(y) + j(u)\}, \tag{2.3}$$

subject to

$$a(y(u),v) + (uy(u),v) = (f,v) \qquad \forall v \in V.$$

2.2. Well-posedness for state equation

In this paper we will consider the case where u may take negative values. Thus the state equation (2.2) may not be even well-posed. We need the following assumption on K for well-posedness of the state equation.

Assumption (H): For any $u \in K$ and $f \in H^{-1}(\Omega)$, the equation (2.2) admits a unique solution $y = y(u,f) \in V$. Furthermore there exist a neighborhood Q of u in K and a constant $C(Q) > 0$ independent of f, such that

$$\|y(u,f)\|_{H^1} \leq C(Q)\|f\|_{H^{-1}} \qquad \forall\, u \in Q.$$

Some examples of K, which satisfy Assumption (H), will be given in Appendix. Also in the Appendix, we will show:

Proposition 2.1. *If $\partial\Omega$ is $C^{1,1}$ regular or that Ω is a parallelepiped. Suppose that Assumption (H) holds, then for all $u \in K$, the solution $y = y(u,f)$ of (2.2) is in $H^2(\Omega)$. Furthermore for any $v \in K$, there exist a neighborhood $O(v)$ of v and a constant $C(v) > 0$ such that*

$$\|y(u,f)\|_{H^2} \leq C(v)\,\|f\|_{L^2} \qquad \forall\, u \in O(v)\cap K.$$

2.3. Convexity of control problem

We now in the position of examining the convexity of the parameter estimation problem. Let

$$J(u) = g(y(u)) + j(u), \tag{2.4}$$

where $y(u)$ is the solution of (2.2). Then it follows that J is continuous and weakly lower semi-continuous on $L^2(\Omega)$. The control problem (2.3) can be reformulated as the following minimization problem: (M)

$$\min_{u\in K} J(u).$$

Let u be a solution of (2.3). We shall further assume that there is a neighborhood of u such that J is a uniformly convex functional in the neighborhood. This is a strong assumption. Nevertheless, we notice that there are many cases where this assumption does hold. The argument is that in many applications, J is regularized. Thus if it is not strictly convex locally, then the regularization should be further improved.

One of the most frequently met cases is

$$J(u) = g(y(u)) + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}.$$

If α is large enough, then J is uniformly convex globally. For some important problems we can further show that J is convex locally for any $\alpha > 0$ (thus uniformly convex).

For example, for any $\alpha > 0$, consider the following regularized parameter (potential) estimation problem (P_α):

$$\min_{u\in K}\left\{\frac{1}{2}\|y-z\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}\right\}, \tag{2.5}$$

subject to

$$-\mathrm{div}(A\nabla y) + uy = f \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \partial\Omega,$$

where K satisfies Assumption (H).

Let us assume that z is identifiable in the sense that there exists $u_z \in K$ such that $y(u_z) = z$, where $y(u_z)$ is the solution of the equation:

$$-\mathrm{div}(A\nabla y) + u_z y = f \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \partial\Omega. \tag{2.6}$$

Under this condition, we will show the convexity of the control problem in Appendix.

From the local convexity any solution of (2.3) is locally unique, and it can be shown that if $(y(u), u)$ is a local solution of (2.3) then there is a $p \in V$ such that $(y, p, u) \in V \times V \times K$ satisfying

$$(A\nabla y, \nabla v) + (uy, v) = (f, v) \qquad \forall v \in V, \tag{2.7}$$

$$(\nabla q, A^*\nabla p) + (up, q) = (g'(y), q) \quad \forall q \in V, \tag{2.8}$$

$$(h'(u), u - w) - (py, u - w) \le 0 \quad \forall w \in K \subset U, \tag{2.9}$$

where A^* is the adjoint matrix of A.

3. Finite Element Approximation

3.1. The finite element space

In this section we consider the finite element approximation of the estimation problem. Here we only consider the n-simplex elements. Also we only consider the conforming finite elements.

For the problem (OBC) we assume $\Omega \subset R^n$ $(n \le 3)$. Let Ω^h be a polygonal approximation to Ω with boundary $\partial\Omega^h$. Let T^h be a partitioning of Ω^h into disjoint regular n-simplices 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 $\tau' \in T^h$. We further require that $P_i \in \partial\Omega^h \Rightarrow P_i \in \partial\Omega$ where $\{P_i\}(i = 1....J)$ is the vertex set associated with the triangulation T^h. For simplicity, we assume that Ω is a convex polygon so that $\Omega = \Omega^h$.

Associated with T^h is a finite dimensional subspace W^h of $C(\bar{\Omega}^h)$, such that $\chi|_\tau$ are polynomials of m-order ($m \geq 1$) for all $\chi \in W^h$ and $\tau \in T^h$. Let $V^h = W^h \cap V \subset V$.

Similarly, let T_U^h be a partitioning of Ω^h into disjoint regular n-simplices τ_U, so that $\bar{\Omega}^h = \bigcup_{\tau_U \in T_U^h} \bar{\tau}_U$. $\bar{\tau}_U$ and $\bar{\tau}'_U$ have either only one common vertex or a whole edge or face if τ_U and $\tau'_U \in T_U^h$.

Associated with T_U^h is another finite dimensional subspace W_U^h of $L^2(\Omega^h)$, such that $\chi|_{\tau_U}$ are polynomials of m-order ($m \geq 0$) for all $\chi \in W_U^h$ and $\tau_U \in T_U^h$. Here there is no requirement for the continuity. Let $U^h = W_U^h$. In this work, we only consider the piecewise constants or discontinuous linear elements. It is easy to see that $U^h \subset U$.

Let h_τ (h_{τ_U}) denote the maximum diameter of the element τ (τ_U) in T^h (T_U^h), let ρ_τ (ρ_{τ_U}) denote the diameter of the largest ball contained in τ (τ_U). Assume that there is a regularity constant R such that $1 \leq \max_{\tau \in T^h}(h_\tau/\rho_\tau) \leq R$ ($1 \leq \max_{\tau_U \in T_U^h}(h_{\tau_U}/\rho_{\tau_U}) \leq R$). Let $h = \max_{\tau \in T^h} h_\tau$ ($h_U = \max_{\tau_U \in T_U^h} h_{\tau_U}$).

3.2. The discrete scheme

The finite element approximation of the state equation reads:

$$a(y_h, w_h) + (u_h y_h, v_h) = (f, v_h) \qquad \forall v_h \in V^h \subset V. \tag{3.1}$$

The next lemma follows immediately.

Lemma 3.1. *If K satisfies Assumption (H), then* (3.1) *has a unique solution for h sufficiently small.*

Let K^h be 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. The finite element approximation of control problem (2.3) reads (M^h):

$$\min_{u_h \in K^h} J_h(u_h),$$

where $J_h(u_h) = g(y_h(u_h)) + j(u_h)$. Or equivalently ($OBC^h$)

$$\min_{u_h \in K^h} \{g(y_h) + j(u_h)\}, \tag{3.2}$$

subject to $y_h \in V^h$, and

$$a(y_h, w_h) + (u_h y_h, v_h) = (f, v_h) \qquad \forall v_h \in V^h \subset V.$$

It is can be shown that the problem (3.2) has at least one solution (y_h, u_h), and that if a pair (y_h, u_h) is the solution of (3.2), 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:

$$(A\nabla y_h, \nabla v_h) + (u_h y_h, v_h) = (f, v_h) \qquad \forall v_h \in V^h \subset V, \tag{3.3}$$

$$(\nabla q_h, A^*\nabla p_h) + (u_h p_h, q_h) = (g'(y_h), q_h) \quad \forall q_h \in V^h \subset V, \tag{3.4}$$

$$(h'(u_h) - y_h p_h, w_h - u_h) \geq 0 \qquad \forall w_h \in K^h \subset U^h \subset L^2(\Omega). \tag{3.5}$$

3.3. Some useful lemmas

For the residual type a posteriori error estimates of the control problem, we need the following well-known interpolation error estimates.

Lemma 3.2. ([6]) *Let π_h be the standard Lagrange interpolation operator. For $m = 0$ or 1, $q > n/2$ and $v \in W^{2,q}(\Omega^h)$,*

$$|v - \pi_h v|_{W^{m,q}(\Omega^h)} \leq Ch^{2-m}|v|_{W^{2,q}(\Omega^h)}. \tag{3.6}$$

Lemma 3.3. ([23]) *Let $\hat{\pi}_h$ be the average interpolation operator defined in* ([23]). *For $m = 0$ or 1, $1 \leq q \leq \infty$ and $v \in W^{1,q}(\Omega^h)$,*

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

Lemma 3.4. ([12]) *For $v \in W^{1,q}(\Omega^h)$, $1 \leq q < \infty$,*

$$\|v\|_{W^{0,q}(\partial\tau)} \leq C\left(h_\tau^{-\frac{1}{q}}\|v\|_{W^{0,q}(\tau)} + h_\tau^{1-\frac{1}{q}}|v|_{W^{1,q}(\tau)}\right).$$

4. Equivalent a Posteriori Error Estimators

In this section, we will present a posteriori error estimate of the problem (2.7)-(2.9) and its approximation (3.3)-(3.5), before which, let us give some denotations.

4.1. Denotations

Firstly, we adopt the control set which satisfies the Assumption (H)

$$K = \{v \in L^2(\Omega) : \ v \geq \beta\},$$

where β is a constant. Secondly, divide $\bar{\Omega}$ into three subsets which are not intersected, i.e. $\bar{\Omega} = \bar{\Omega}_0 \cup \bar{\Omega}_0^+ \cup \bar{\Omega}_0^-$, where

$$\begin{aligned}
\Omega_0^- &= \{x \in \Omega : \ y_h(x)p_h(x) \geq h'(\beta)\},\\
\Omega_0^+ &= \{x \in \Omega : \ y_h(x)p_h(x) < h'(\beta), \ u_h(x) > \beta\},\\
\Omega_0 &= \{x \in \Omega : \ y_h(x)p_h(x) < h'(\beta), \ u_h(x) = \beta\}.
\end{aligned}$$

Thirdly, define $J(\cdot)$ and $J_h(\cdot)$ as before. Then we have that

$$\begin{aligned}
&(J'(u), w) = (h'(u), w) - (py, w),\\
&(J_h'(u_h), w_h) = (h'(u_h), w_h) - (p_h y_h, w_h),\\
&(J'(u_h), w) = (h'(u_h), w) - (p(u_h)y(u_h), w),
\end{aligned}$$

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

$$(A\nabla y(u_h), \nabla v) + (u_h y(u_h), v) = (f, v) \qquad \forall v \in V, \tag{4.1}$$

$$(\nabla q, A^*\nabla p(u_h)) + (u_h p(u_h), q) = (g'(y(u_h)), q) \quad \forall q \in V. \tag{4.2}$$

Fourthly, let us introduce

$$e^2 = \int_{\Omega_*} \Big((h'(u) - yp) - P_h(h'(u) - yp) \Big)^2,$$

where P_h is the L^2-project operator from $L^2(\Omega)$ to U^h, and U^h is the piecewise constant finite element space. When U^h is piecewise discontinuous linear finite element space, let us introduce

$$e^2 = \int_{\Omega_*} (h'(u_h) - y_h p_h)^2,$$

where

$$\Omega_* = \Big\{ x \in \Omega_0^+ : \; u(x) = \beta, \; u_h(x) > \beta \Big\}.$$

4.2. A upper bound

We shall only present the details for the case $V = H_0^1(\Omega)$, and give the results for Neumann boundary condition in Section 5. Moreover, we will assume that V^h is a conforming piecewise linear finite element space, and U^h is piecewise constant finite element space. For the case of piecewise discontinuous linear element space U^h, we can obtain the same results, and the proof is similar so we do not include the details here.

Lemma 4.1. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of* (2.7)-(2.9) *and* (3.3)-(3.5), *respectively. Assume J is locally uniform convex in the sense that there exists $c > 0$ such that $\forall w \in U$ in a neighborhood of u:*

$$(J'(w) - J'(u), w - u) \geq c\|w - u\|_{0,\Omega}^2. \tag{4.3}$$

Moreover, assume that $h'(\cdot)$ is locally Lipschitz continuous in a neighborhood of u. Then we have

$$e^2 + \|u - u_h\|_{0,\Omega}^2 \leq C\Big(\eta_1^2 + \|p(u_h)y(u_h) - p_h y_h\|_{0,\Omega}^2 \Big), \tag{4.4}$$

where $(y(u_h), p(u_h))$ is the solution of the equations (4.1)-(4.2), *e is defined in the last subsection, and*

$$\eta_1^2 = \int_{\Omega_0^- \cup \Omega_0^+} |h'(u_h) - y_h p_h|^2.$$

Proof. It follows from the assumption (4.3) that

$$\begin{aligned} c\|u - u_h\|_{0,\Omega}^2 \leq & J'(u), u - u_h)_U - (J'(u_h), u - u_h) \\ \leq & -(J'(u_h), u - u_h) \\ = & (J_h'(u_h), u_h - u) + (J_h'(u_h) - J'(u_h), u - u_h). \end{aligned} \tag{4.5}$$

Note that

$$\begin{aligned} & (J_h'(u_h), u_h - u) \\ = & \int_{\Omega_0^- \cup \Omega_0^+} (h'(u_h) - y_h p_h)(u_h - u) + \int_{\Omega_0} (h'(\beta) - y_h p_h)(\beta - u). \end{aligned} \tag{4.6}$$

It follows from the Schwarz's inequality and the inequality $2ab \leq a^2/\delta + \delta b^2$ that

$$\begin{aligned}
&\int_{\Omega_0^- \cup \Omega_0^+} (h'(u_h) - y_h p_h)(u_h - u) \\
\leq &\frac{1}{2\delta} \int_{\Omega_0^- \cup \Omega_0^+} (h'(u_h) - y_h p_h)^2 + \frac{\delta}{2} \|u_h - u\|_{0,\Omega}^2 \\
= &\frac{1}{2\delta} \eta_1^2 + \frac{\delta}{2} \|u_h - u\|_{0,\Omega}^2,
\end{aligned} \tag{4.7}$$

where $\delta > 0$ is a suitable positive constant. Moreover, note that the definition of Ω_0 implies that $(h'(\beta) - y_h p_h) > 0$ on Ω_0. Because that $\beta - u \leq 0$, we have that

$$\int_{\Omega_0} (h'(\beta) - y_h p_h)(\beta - u) \leq 0. \tag{4.8}$$

Then (4.6)-(4.8) imply that

$$(J_h'(u_h), u_h - u) \leq \frac{1}{2\delta} \eta_1^2 + \frac{\delta}{2} \|u_h - u\|_{0,\Omega}^2. \tag{4.9}$$

By using the formulas of J' and J_h', we have that

$$\begin{aligned}
&(J_h'(u_h) - J'(u_h), u - u_h) \\
= &(h'(u_h) - y_h p_h, u - u_h) - (h'(u_h) - y(u_h) p(u_h), u - u_h) \\
= &(y_h p_h - y(u_h) p(u_h), u_h - u) \\
\leq &\frac{1}{2\delta} \|y_h p_h - y(u_h) p(u_h)\|_{0,\Omega}^2 + \frac{\delta}{2} \|u_h - u\|_{0,\Omega}^2.
\end{aligned} \tag{4.10}$$

Therefore, it follows from (4.5) and (4.9)-(4.10) by setting δ to be small enough that

$$\|u - u_h\|_{0,\Omega}^2 \leq C(\eta_1^2 + \|p(u_h) y(u_h) - p_h y_h\|_{0,\Omega}^2). \tag{4.11}$$

Next let us consider the estimation of e. For all $x^* \in \Omega_*$, we have that $u_h(x^*) > \beta$ according to the definition of Ω_*. Therefore, there exists an element $\tau_U^* \in T_U^h$ such that $x^* \subset \tau_U^*$ and $u_h|_{\tau_U^*} = u_h(x^*) > \beta$. Then there exists an $\epsilon > 0$ such that $v_h = u_h \pm \epsilon \phi_{\tau_U^*} > \beta$, where $\phi_{\tau_U^*}$ is the basis function of U^h on the element τ_U. Thus we have $v^h \in K^h$, and it follows from (3.5) that

$$\int_{\tau_U^*} (h'(u_h) - y_h p_h) \phi_{\tau_U^*} = 0.$$

This implies that

$$P_h(h'(u_h) - y_h p_h) = 0 \text{ on } \tau_U^*.$$

Then we have that

$$P_h(h'(u_h) - y_h p_h)(x^*) = 0,$$

because that $P_h v \in U^h$ is a constant on τ_U^* for all $v \in L^2(\Omega)$. Noting that above proof is valid for all $x^* \in \Omega_*$, we have that $P_h(h'(u_h) - y_h p_h) = 0$ on Ω_*. Therefore,

$$\begin{aligned}
e^2 =& \int_{\Omega_*} \Big((h'(u) - yp) - P_h(h'(u) - yp) \Big)^2 \\
\leq& C \int_{\Omega_*} \Big((h'(u) - yp) - (h'(u_h) - y_h p_h) \Big)^2 + C \int_{\Omega_*} \Big(h'(u_h) - y_h p_h \Big)^2 \\
&+ C \int_{\Omega_*} \Big(P_h(h'(u_h) - y_h p_h) \Big)^2 + C \int_{\Omega_*} \Big(P_h(h'(u_h) - y_h p_h) - P_h(h'(u) - yp) \Big)^2 \\
\leq& C \Big(\|yp - y_h p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \Big) + C \int_{\Omega_*} \Big(h'(u_h) - y_h p_h \Big)^2 + 0 \\
&+ C \Big(\|yp - y_h p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \Big) \\
\leq& C \int_{\Omega_0^+} \Big(h'(u_h) - y_h p_h \Big)^2 + C \Big(\|yp - y_h p_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \Big),
\end{aligned} \tag{4.12}$$

where in the last step, we use the fact that $\Omega_* \subset \Omega_0^+$ according to the definition of Ω_*. Note that

$$\|yp - y_h p_h\|_{0,\Omega} \leq \|yp - y(u_h)p(u_h)\|_{0,\Omega} + \|y(u_h)p(u_h) - y_h p_h\|_{0,\Omega}. \tag{4.13}$$

Moreover, it follows from the well known embed Theorem and the equations (2.7)-(2.9) and (4.1)-(4.2), we obtain that

$$\begin{aligned}
&\|yp - y(u_h)p(u_h)\|_{0,\Omega} \\
\leq& C \Big(\|p\|_{0,4,\Omega} \|y - y(u_h)\|_{0,4,\Omega} + \|y(u_h)\|_{0,4,\Omega} \|p - p(u_h)\|_{0,4,\Omega} \Big) \\
\leq& C \Big(\|p\|_{1,\Omega} \|y - y(u_h)\|_{1,\Omega} + \|y(u_h)\|_{1,\Omega} \|p - p(u_h)\|_{1,\Omega} \Big) \\
\leq& C \|u - u_h\|_{0,\Omega}.
\end{aligned} \tag{4.14}$$

Therefore, it can be deduced from (4.11)-(4.14) that

$$e^2 \leq C(\eta_1^2 + \|p(u_h)y(u_h) - p_h y_h\|_{0,\Omega}^2). \tag{4.15}$$

Summing up, (4.4) is the direct result of (4.11) and (4.15). □

In order to have the a posteriori error estimates, we only need to estimate the term $\|p(u_h)y(u_h) - p_h y_h\|_{0,\Omega}$.

Theorem 4.1. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of* (2.7)-(2.9) *and* (3.3)-(3.5)*, respectively. Assume that all the conditions in Lemma* 4.1 *hold. Moreover, assume that $g'(\cdot)$ is locally Lipschitz continuous in a neighborhood of y. Then,*

$$e^2 + \|u - u_h\|_{0,\Omega}^2 + \|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{1,\Omega}^2 \leq C(\eta_1^2 + \eta_2^2 + \eta_3^2), \tag{4.16}$$

where e and η_1 are defined in Lemma 4.1*, and*

$$\eta_2^2 = \sum_{l \cap \partial\Omega = \emptyset} h_l \int_l [(A\nabla y_h) \cdot n]^2 + \sum_{\tau \in T^h} h_\tau^2 \int_\tau \Big(f + \mathrm{div}(A\nabla y_h) - u_h y_h \Big)^2,$$

$$\eta_3^2 = \sum_{l \cap \partial\Omega = \emptyset} h_l \int_l [(A^*\nabla p_h) \cdot n]^2 + \sum_{\tau \in T^h} h_\tau^2 \int_\tau \Big(g'(y_h) + \mathrm{div}(A^*\nabla p_h) - u_h p_h \Big)^2,$$

where l is a face of an element τ, $[(A^*\nabla p_h\cdot n)]$ and $[(A\nabla y_h\cdot n)]$ are the A-normal derivative jumps over the interior face l, defined by

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

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

Proof. By the Assumption (H) and noting that $\|p(u_h)\|_{1,\Omega}\le C$ and $\|y_h\|_{1,\Omega}\le C$ when h is small enough because that $\|u_h\|_{0,\Omega}$ is bounded, we have

$$\begin{aligned}
&\|p(u_h)y(u_h)-p_hy_h\|_{0,\Omega}\\
\le&\|p(u_h)(y(u_h)-y_h)\|_{0,\Omega}+\|y_h(p(u_h)-p_h)\|_{0,\Omega}\\
\le&\|p(u_h)\|_{0,4,\Omega}\|y(u_h)-y_h\|_{0,4,\Omega}+\|y_h\|_{0,4,\Omega}\|p(u_h)-p_h\|_{0,4,\Omega}\\
\le&C\|p(u_h)\|_{1,\Omega}\|y(u_h)-y_h\|_{1,\Omega}+C\|y_h\|_{1,\Omega}\|p(u_h)-p_h\|_{1,\Omega}\\
\le&C\Big(\|y(u_h)-y_h\|_{1,\Omega}+\|p(u_h)-p_h\|_{1,\Omega}\Big).
\end{aligned}\tag{4.17}$$

Let $e^p=p_h-p(u_h)$. It follows from [1] that there exists a function $\phi\in H_0^1(\Omega)$ such that

$$c\|e^p\|_{1,\Omega}\|\phi\|_{1,\Omega}\le(\nabla\phi,A^*\nabla e^p)+(u_he^p,\phi).\tag{4.18}$$

Let $\phi_I\in V^h$ be the interpolation of ϕ defined in Lemma 3.3. Using the equations (3.4), (4.2) and Lemma 3.3, we have that

$$\begin{aligned}
&c\|e^p\|_{1,\Omega}\|\phi\|_{1,\Omega}\le(\nabla\phi,A^*\nabla e^p)+(u_he^p,\phi)\\
=&(\nabla(\phi-\phi_I),A^*\nabla e^p)+(u_he^p,\phi-\phi_I)+(\nabla\phi_I,A^*\nabla e^p)+(u_he^p,\phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau(A^*\nabla(p_h-p(u_h))\nabla(\phi-\phi_I)+u_h(p_h-p(u_h))(\phi-\phi_I))\\
&+(g'(y_h)-g'(y(u_h)),\phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau(-\mathrm{div}(A^*\nabla p_h)+u_hp_h)(\phi-\phi_I)-(g'(y(u_h)),\phi-\phi_I)\\
&+(g'(y_h)-g'(y(u_h)),\phi_I)+\sum_{l\cap\partial\Omega=\emptyset}\int_l[(A^*\nabla p_h)\cdot n](\phi-\phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau(-\mathrm{div}(A^*\nabla p_h)+u_hp_h-g'(y_h))(\phi-\phi_I)+(g'(y_h)-g'(y(u_h)),\phi)\\
&+\sum_{l\cap\partial\Omega=\emptyset}\int_l[(A^*\nabla p_h)\cdot n](\phi-\phi_I)\\
\le&C\Big(\sum_{\tau\in T^h}h_\tau^2\int_\tau(\mathrm{div}(A^*\nabla p_h)-u_hp_h+g'(y_h))^2\Big)^{\frac12}\|\phi\|_{1,\Omega}\\
&+C\Big(\sum_{l\cap\partial\Omega=\emptyset}h_l\int_l[(A^*\nabla p_h)\cdot n]^2\Big)^{\frac12}\|\phi\|_{1,\Omega}+C\|y_h-y(u_h)\|_{0,\Omega}\|\phi\|_{0,\Omega}.
\end{aligned}$$

Then,

$$\|p(u_h) - p_h\|_{1,\Omega}^2 \le C\Big(\sum_{\tau \in T^h} h_\tau^2 \int_\tau (g'(y_h) + \mathrm{div}(A^* \nabla p_h) - u_h p_h)^2 + \sum_{l \cap \partial\Omega = \emptyset} h_l \int_l [(A^* \nabla p_h) \cdot n]^2 + \|y(u_h) - y_h\|_{0,\Omega}^2 \Big). \tag{4.19}$$

Similarly, it can be proved that

$$\|y(u_h) - y_h\|_{1,\Omega}^2 \le C \sum_{l \cap \partial\Omega = \emptyset} h_l \int_l [(A\nabla y_h) \cdot n]^2 + C \sum_{\tau \in T^h} h_\tau^2 \int_\tau \Big(f + \mathrm{div}(A\nabla y_h) - u_h y_h \Big)^2. \tag{4.20}$$

Hence, it follows from Lemma 4.1 and (4.17), (4.19)-(4.20) that

$$e^2 + \|u - u_h\|_{0,\Omega}^2 + \|y(u_h) - y_h\|_{1,\Omega}^2 + \|p(u_h) - p_h\|_{1,\Omega}^2 \le C(\eta_1^2 + \eta_2^2 + \eta_3^2). \tag{4.21}$$

Note that

$$\|y - y_h\|_{1,\Omega} \le \|y - y(u_h)\|_{1,\Omega} + \|y(u_h) - y_h\|_{1,\Omega}, \tag{4.22a}$$

$$\|p - p_h\|_{1,\Omega} \le \|p - p(u_h)\|_{1,\Omega} + \|p(u_h) - p_h\|_{1,\Omega}, \tag{4.22b}$$

and

$$\|y(u_h) - y\|_{1,\Omega} + \|p(u_h) - p\|_{1,\Omega} \le C\|u - u_h\|_{0,\Omega}. \tag{4.22c}$$

Then, (4.16) follows from (4.21)-(4.22c). □

4.3. A lower bound

In order to derive the a posteriori lower bound, let $\bar{v}$ and $\tilde{A}$ be the integral averages of v and A on the element τ and edge l, respectively, such that

$$\bar{v}|_\tau = \frac{\int_\tau v}{\int_\tau 1}, \quad \tilde{A}|_l = \frac{\int_l A}{\int_l 1}.$$

Lemma 4.2. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of (2.7)-(2.9) and (3.3)-(3.5), respectively. Then we have*

$$\eta_2^2 \le C\|y - y_h\|_{1,\Omega}^2 + C\|u - u_h\|_{0,\Omega_U}^2 + C\epsilon_2^2, \tag{4.23}$$

where

$$\epsilon_2^2 = \sum_{\tau \in T^h} \int_\tau h_\tau^2 \Big((f + \mathrm{div}(A\nabla y_h) - u_h y_h) - (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h}) \Big)^2 + \sum_{l \cap \partial\Omega = \emptyset} h_l \int_l [(A - \tilde{A})\nabla y_h \cdot n]^2.$$

Proof. Using the standard bubble function technique (see, [25], for example), it can be proved that there are bubble functions $w_\tau \in H_0^1(\tau) \cap P_3$ and $w_l \in H_0^1(\tau_l^1 \cup \tau_l^2) \cap P_2$, where P_i denotes the polynomials of order i, such that

$$h_\tau^2 \int_\tau (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})^2 = \int_\tau (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h}) w_\tau, \tag{4.24}$$

$$h_l \int_l [(\tilde{A}\nabla y_h)\cdot n]^2 = \int_l [(\tilde{A}\nabla y_h)\cdot n] w_l, \tag{4.25}$$

$$\|w_\tau\|_{1,\tau}^2 \le C h_\tau^2 \int_\tau (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})^2, \tag{4.26}$$

$$\|w_\tau\|_{0,\infty,\tau}^2 \le C h_\tau^{-2} \|w_\tau\|_{0,\tau}^2 \le C h_\tau^2 \int_\tau (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})^2, \tag{4.27}$$

$$\|w_l\|_{1,\tau_l^1\cup\tau_l^2}^2 \le C h_l \int_l [(\tilde{A}\nabla y_h)\cdot n]^2, \tag{4.28}$$

$$h_l^{-1} \|w_l\|_{0,l}^2 \le C h_l \int_l [(\tilde{A}\nabla y_h)\cdot n]^2, \tag{4.29}$$

$$\|w_l\|_{0,\infty,\tau_l^1\cup\tau_l^2}^2 \le C h_l^{-2} \|w_l\|_{0,\tau_l^1\cup\tau_l^2}^2 \le C h_l \int_l [(\tilde{A}\nabla y_h)\cdot n]^2. \tag{4.30}$$

It follows from (4.24), (4.26) and (4.27) that for any $\tau \in T^h$,

$$\begin{aligned}
& h_\tau^2 \int_\tau |f + \mathrm{div}(A\nabla y_h) - u_h y_h|^2 \\
\le & C h_\tau^2 \int_\tau (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})^2 \\
& + C h_\tau^2 \int_\tau |(f + \mathrm{div}(A\nabla y_h) - u_h y_h) - (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})|^2 \\
= & C \int_\tau (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h}) w_\tau \\
& + C h_\tau^2 \int_\tau |(f + \mathrm{div}(A\nabla y_h) - u_h y_h) - (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})|^2 \\
= & C \int_\tau (\mathrm{div}(A\nabla(y_h - y)) + uy - u_h y_h) w_\tau \\
& + C \int_\tau ((\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h}) - (f + \mathrm{div}(A\nabla y_h) - u_h y_h)) w_\tau \\
& + C h_\tau^2 \int_\tau |(f + \mathrm{div}(A\nabla y_h) - u_h y_h) - (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})|^2 \\
\le & - C \int_\tau A\nabla(y_h - y)\nabla w_\tau + C \int_\tau (uy - u_h y_h) w_\tau + C\delta h_\tau^{-2} \|w_\tau\|_{0,\tau}^2 \\
& + C(\delta) h_\tau^2 \int_\tau |(f + \mathrm{div}(A\nabla y_h) - u_h y_h) - (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})|^2 \\
\le & C(\delta) \|y - y_h\|_{1,\tau}^2 + C(\delta) \|uy - u_h y_h\|_{0,1,\tau}^2 \\
& + C(\delta) h_\tau^2 \int_\tau |(f + \mathrm{div}(A\nabla y_h) - u_h y_h) - (\overline{f + \mathrm{div}(A\nabla y_h) - u_h y_h})|^2 \\
& + C\delta (h_\tau^{-2} \|w_\tau\|_{0,\tau}^2 + \|w_\tau\|_{0,\infty,\tau}^2 + \|w_\tau\|_{1,\tau}^2)
\end{aligned}$$

$$
\begin{aligned}
&\le C(\delta)\|y-y_h\|_{1,\tau}^2 + C(\delta)\|u\|_{0,\tau}^2\|y-y_h\|_{0,\tau}^2 + C(\delta)\|y_h\|_{0,\tau}^2\|u-u_h\|_{0,\tau}^2 \\
&\quad + C(\delta)h_\tau^2\int_\tau |(f+\mathrm{div}(A\nabla y_h)-u_h y_h) - (\overline{f+\mathrm{div}(A\nabla y_h)-u_h y_h})|^2 \\
&\quad + C\delta h_\tau^2\int_\tau (\overline{f+\mathrm{div}(A\nabla y_h)-u_h y_h})^2 \\
&\le C(\delta)\|y-y_h\|_{1,\tau}^2 + C(\delta)\|u-u_h\|_{0,\tau}^2 \\
&\quad + C(\delta)h_\tau^2\int_\tau |(f+\mathrm{div}(A\nabla y_h)-u_h y_h) - (\overline{f+\mathrm{div}(A\nabla y_h)-u_h y_h})|^2 \\
&\quad + C\delta h_\tau^2\int_\tau |f+\mathrm{div}(A\nabla y_h)-u_h y_h|^2,
\end{aligned}
$$

where δ is an arbitrary small positive number. Hence,

$$
\begin{aligned}
&\sum_{\tau\in T^h} h_\tau^2\int_\tau |f+\mathrm{div}(A\nabla y_h)-u_h y_h|^2 \\
&\le C(\|y-y_h\|_{1,\Omega}^2 + \|u-u_h\|_{0,\Omega}^2) \\
&\quad + C\sum_{\tau\in T^h} h_\tau^2\int_\tau |(f+\mathrm{div}(A\nabla y_h)-u_h y_h) - (\overline{f+\mathrm{div}(A\nabla y_h)-u_h y_h})|^2 \\
&\le C(\|y-y_h\|_{1,\Omega}^2 + \|u-u_h\|_{0,\Omega}^2) + C\epsilon_2^2.
\end{aligned} \tag{4.31}
$$

Similarly, it follows from (4.25), (4.28)-(4.30) that for any edge l such that $l\cap\partial\Omega=\emptyset$, we have that

$$
\begin{aligned}
&h_l\int_l [(A\nabla y_h)\cdot n]^2 \le Ch_l\int_l [(\tilde{A}\nabla y_h)\cdot n]^2 + Ch_l\int_l [(A-\tilde{A})\nabla y_h\cdot n]^2 \\
&= C\int_l [(\tilde{A}\nabla y_h)\cdot n]w_l + Ch_l\int_l [(A-\tilde{A})\nabla y_h\cdot n]^2 \\
&= C\int_l [(A\nabla y_h)\cdot n]w_l + C\int_l [(\tilde{A}-A)\nabla y_h\cdot n]w_l + Ch_l\int_l [(A-\tilde{A})\nabla y_h\cdot n]^2 \\
&\le C\int_l [A\nabla(y_h-y)\cdot n]w_l + C(\delta)h_l\int_l [(A-\tilde{A})\nabla y_h\cdot n]^2 + C\delta h_l^{-1}\int_l w_l^2 \\
&= C\int_{\tau_l^1\cup\tau_l^2} A\nabla(y_h-y)\nabla w_l + C\int_{\tau_l^1\cup\tau_l^2}\mathrm{div}(A\nabla(y_h-y))w_l \\
&\quad + C(\delta)h_l\int_l [(A-\tilde{A})\nabla y_h\cdot n]^2 + C\delta h_l^{-1}\int_l w_l^2 \\
&= C\int_{\tau_l^1\cup\tau_l^2} A\nabla(y_h-y)\nabla w_l + C\int_{\tau_l^1\cup\tau_l^2}(\mathrm{div}(A\nabla y_h)+f-yu)w_l \\
&\quad + C(\delta)h_l\int_l [(A-\tilde{A})\nabla y_h\cdot n]^2 + C\delta h_l^{-1}\int_l w_l^2 \\
&= C\int_{\tau_l^1\cup\tau_l^2} A\nabla(y_h-y)\nabla w_l + C\int_{\tau_l^1\cup\tau_l^2}(\mathrm{div}(A\nabla y_h)+f-y_h u_h)w_l \\
&\quad + C\int_{\tau_l^1\cup\tau_l^2}(y_h u_h - yu)w_l + C(\delta)h_l\int_l [(A-\tilde{A})\nabla y_h\cdot n]^2 + C\delta h_l^{-1}\int_l w_l^2 \\
&\le C(\delta)\|y-y_h\|_{1,\tau_l^1\cup\tau_l^2}^2 + C(\delta)h_l^2\int_{\tau_l^1\cup\tau_l^2}|f+\mathrm{div}(A\nabla y_h)-u_h y_h|^2
\end{aligned}
$$

$$
\begin{aligned}
&+C(\delta)\|uy-u_hy_h\|^2_{0,1,\tau_l^1\cup\tau_l^2}+C(\delta)h_l\int_l[(A-\tilde{A})\nabla y_h\cdot n]^2\\
&+C\delta\bigg(\|w_l\|^2_{1,\tau_l^1\cup\tau_l^2}+h_l^{-2}\|w_l\|^2_{0,\tau_l^1\cup\tau_l^2}+\|w_l\|^2_{0,\infty,\tau_l^1\cup\tau_l^2}+h_l^{-1}\|w_l\|^2_{0,l}\bigg)\\
\le&C(\delta)\|y-y_h\|^2_{1,\tau_l^1\cup\tau_l^2}+C(\delta)h_l^2\int_{\tau_l^1\cup\tau_l^2}|f+\mathrm{div}(A\nabla y_h)-u_hy_h|^2\\
&+C(\delta)\|u-u_h\|^2_{0,\tau_l^1\cup\tau_l^2}+C(\delta)h_l\int_l[(A-\tilde{A})\nabla y_h\cdot n]^2+C\delta h_l\int_l[(\tilde{A}\nabla y_h)\cdot n]^2.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\sum_{l\cap\partial\Omega=\emptyset}h_l\int_l[(A\nabla y_h)\cdot n]^2\\
\le&C(\|y-y_h\|^2_{1,\Omega}+\|u-u_h\|^2_{0,\Omega})+C\epsilon_2^2\\
&+C\sum_{\tau\in T^h}h_\tau^2\int_\tau|f+\mathrm{div}(A\nabla y_h)-u_hy_h|^2.
\end{aligned}
\tag{4.32}
$$

Therefore, it follows from (4.31) and (4.32) that

$$
\eta_2^2\le C\bigg(\|y-y_h\|^2_{1,\Omega}+\|u-u_h\|^2_{0,\Omega}\bigg)+C\epsilon_2^2. \tag{4.33}
$$

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

Lemma 4.3. *Let (y,p,u) and (y_h,p_h,u_h) be the solutions of* (2.7)-(2.9) *and* (3.3)-(3.5)*, respectively. Moreover, assume that $g'(\cdot)$ is locally Lipschitz continuous in a neighborhood of y. Then we have*

$$
\eta_3^2\le C\|p-p_h\|^2_{1,\Omega}+C\|y-y_h\|^2_{1,\Omega}+C\|u-u_h\|^2_{0,\Omega}+C\epsilon_3^2, \tag{4.34}
$$

where

$$
\begin{aligned}
\epsilon_3^2=&\sum_{\tau\in T^h}\int_\tau h_\tau^2\Big((\mathrm{div}(A^*\nabla p_h)-u_hp_h+g'(y_h))-(\overline{\mathrm{div}(A^*\nabla p_h)-u_hp_h+g'(y_h)})\Big)^2\\
&+\sum_{l\cap\partial\Omega=\emptyset}h_l\int_l[(A^*-\tilde{A}^*)\nabla p_h\cdot n]^2.
\end{aligned}
$$

Using the lemmas above, we can prove the following a posteriori lower bound.

Theorem 4.2. *Let (y,p,u) and (y_h,p_h,u_h) be the solutions of* (2.7)-(2.9) *and* (3.3)-(3.5)*, respectively. Assume all the conditions in Lemmas* 4.2 *and* 4.3 *are valid. Then,*

$$
\sum_{i=1}^3\eta_i^2\le C\bigg(e^2+\|u-u_h\|^2_{0,\Omega}+\|y-y_h\|^2_{1,\Omega}+\|p-p_h\|^2_{1,\Omega}\bigg)+C\epsilon^2, \tag{4.35}
$$

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

Proof. It can be concluded from (2.9) that $h'(u)-yp=0$ when $u>\beta$, and $h'(u)-yp=h'(\beta)-yp\ge0$ when $u=\beta$. Moreover, we have that $(h'(\beta)-y_h(x)p_h(x))\le0$ when $x\in\Omega_0^-$ from the definition of Ω_0^-. Let

$$
\Omega_{00}=\{x\in\Omega_0^-:\ u(x)=\beta\},\quad \Omega_0^*=\Omega_0^-\setminus\Omega_{00}.
$$

Then we have that $(h'(\beta)-y_hp_h)^2 \le (h'(\beta)-y_hp_h-(h'(\beta)-yp))^2$ on Ω_{00}. It follows that

$$
\begin{aligned}
&\int_{\Omega_0^-}(h'(u_h)-y_hp_h)^2\\
=&\int_{\Omega_0^*}\Big((h'(u_h)-y_hp_h)-(h'(u)-yp)\Big)^2+\int_{\Omega_{00}}\Big(h'(u_h)-y_hp_h-h'(u)+h'(\beta)\Big)^2\\
\le& C\|u-u_h\|_{0,\Omega}^2+C\|yp-y_hp_h\|_{0,\Omega}^2+C\int_{\Omega_{00}}(h'(\beta)-y_hp_h)^2\\
\le& C\|u-u_h\|_{0,\Omega}^2+C\|yp-y_hp_h\|_{0,\Omega}^2+C\int_{\Omega_{00}}(h'(\beta)-y_hp_h-(h'(\beta)-yp))^2\\
\le& C\|u-u_h\|_{0,\Omega}^2+C\|yp-y_hp_h\|_{0,\Omega}^2\\
\le& C\Big(\|u-u_h\|_{0,\Omega}^2+\|p\|_{0,4,\Omega}^2\|y-y_h\|_{0,4,\Omega}^2+\|y_h\|_{0,4,\Omega}^2\|p-p_h\|_{0,4,\Omega}^2\Big)\\
\le& C\Big(\|u-u_h\|_{0,\Omega}^2+\|y-y_h\|_{1,\Omega}^2+\|p-p_h\|_{1,\Omega}^2\Big),
\end{aligned} \tag{4.36}
$$

where we used the properties that $\|v\|_{0,4,\Omega}\le C\|v\|_{1,\Omega}$, $\|p\|_{1,\Omega}\le C$ and $\|y_h\|_{1,\Omega}\le C$.

Moreover, we note that $u>\beta$ and hence $h'(u)-yp=0$ on $\Omega_0^+\backslash\Omega_*$. Furthermore, similarly to the proof of Lemma 4.1 we have that $u_h>\beta$ and hence $P_h(h'(u_h)-y_hp_h)=0$ on Ω_0^+. Then it can be deduced that

$$
\begin{aligned}
&\int_{\Omega_0^+}(h'(u_h)-y_hp_h)^2\\
=&\int_{\Omega_*}(h'(u_h)-y_hp_h)^2+\int_{\Omega_0^+\backslash\Omega_*}(h'(u_h)-y_hp_h)^2\\
=&\int_{\Omega_*}\Big((h'(u_h)-y_hp_h)-P_h(h'(u_h)-y_hp_h)\Big)^2\\
&+\int_{\Omega_0^+\backslash\Omega_*}\Big((h'(u_h)-y_hp_h)-(h'(u)-yp)\Big)^2\\
\le& C\int_{\Omega_*}\Big((h'(u)-yp)-P_h(h'(u)-yp)\Big)^2\\
&+C\int_{\Omega_*}\Big(P_h(h'(u)-yp)-P_h(h'(u_h)-y_hp_h)\Big)^2\\
&+C\int_{\Omega_0^+}\Big((h'(u_h)-y_hp_h)-(h'(u)-yp)\Big)^2\\
\le& C\Big(e^2+\|u-u_h\|_{0,\Omega}^2+\|y-y_h\|_{1,\Omega}^2+\|p-p_h\|_{1,\Omega}^2\Big).
\end{aligned} \tag{4.37}
$$

Thus it follows from (4.36) and (4.37) that

$$
\eta_1^2\le C\Big(e^2+\|u-u_h\|_{0,\Omega}^2+\|y-y_h\|_{1,\Omega}^2+\|p-p_h\|_{1,\Omega}^2\Big). \tag{4.38}
$$

Then the lower bound estimation (4.35) is proved from (4.38) and Lemmas 4.2-4.3. □

Remark 4.1. It can be shown that e^2 and ϵ^2 are of higher order in many cases, see [21] for the details.

4.4. Errors estimators in L^2-norm

In this section, we bound the errors in the L^2-norm to derive sharper estimators.

Lemma 4.4. *Suppose Ω is convex, and $g'(\cdot)$ is locally Lipschitz continuous in a neighborhood of y. Then,*

$$\|y(u_h) - y_h\|_{0,\Omega}^2 + \|p(u_h) - p_h\|_{0,\Omega}^2 \leq C(\hat{\eta}_2^2 + \hat{\eta}_3^2), \tag{4.39}$$

where

$$\hat{\eta}_2^2 = \sum_{l\cap\partial\Omega=\emptyset} h_l^3 \int_l [(A\nabla y_h)\cdot n]^2 + \sum_{\tau\in T^h} h_\tau^4 \int_\tau (f + \mathrm{div}(A\nabla y_h) - u_h y_h)^2, \tag{4.40}$$

$$\hat{\eta}_3^2 = \sum_{l\cap\partial\Omega=\emptyset} h_l^3 \int_l [(A^*\nabla p_h)\cdot n]^2 + \sum_{\tau\in T^h} h_\tau^4 \int_\tau (g'(y_h) + \mathrm{div}(A^*\nabla p_h) - u_h p_h)^2. \tag{4.41}$$

Proof. Let $e^p = p_h - p(u_h)$. Let ϕ be the solution of the equation:

$$(A\nabla\phi, \nabla w) + (u_h\phi, w) = (e^p, w), \quad \forall w \in H_0^1(\Omega). \tag{4.42}$$

It follows from Assumption (H) and Proposition 2.1 that

$$\|\phi\|_{H^2(\Omega)} \leq C\|e^p\|_{L^2(\Omega)}. \tag{4.43}$$

Let $\phi_I \in V^h$ be the standard Lagrange interpolation of ϕ. Using the equations (3.4), (4.2), (4.42), and Lemmas 3.2, 3.4, we have that

$$\begin{aligned}
&\|e^p\|_{0,\Omega}^2 = (e^p, e^p) = (\nabla\phi, A^*\nabla e^p) + (u_h\phi, e^p)\\
=&(\nabla(\phi-\phi_I), A^*\nabla e^p) + (u_h e^p, \phi-\phi_I) + (\nabla\phi_I, A^*\nabla e^p) + (u_h e^p, \phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau (A^*\nabla(p_h - p(u_h))\nabla(\phi-\phi_I) + u_h(p_h - p(u_h))(\phi-\phi_I))\\
&\quad + (g'(y_h) - g'(y(u_h)), \phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau (-\mathrm{div}(A^*\nabla p_h) + u_h p_h)(\phi-\phi_I) - (g'(y(u_h)), \phi-\phi_I)\\
&\quad + (g'(y_h) - g'(y(u_h)), \phi_I) + \sum_{l\cap\partial\Omega=\emptyset}\int_l [(A^*\nabla p_h)\cdot n](\phi-\phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau (-\mathrm{div}(A^*\nabla p_h) + u_h p_h - g'(y_h))(\phi-\phi_I) + (g'(y_h) - g'(y(u_h)), \phi)\\
&\quad + \sum_{l\cap\partial\Omega=\emptyset}\int_l [(A^*\nabla p_h)\cdot n](\phi-\phi_I)\\
\leq& C\Big(\sum_{\tau\in T^h} h_\tau^4 \int_\tau (\mathrm{div}(A^*\nabla p_h) - u_h p_h + g'(y_h))^2\Big)^{\frac{1}{2}} \|\phi\|_{2,\Omega}\\
&\quad + C\Big(\sum_{l\cap\partial\Omega=\emptyset} h_l^3 \int_l [(A^*\nabla p_h)\cdot n]^2\Big)^{\frac{1}{2}} \|\phi\|_{2,\Omega} + C\|y_h - y(u_h)\|_{0,\Omega}\|\phi\|_{0,\Omega}.
\end{aligned}$$

Then, it follows from (4.43) that

$$\begin{aligned}
&\|p(u_h)-p_h\|_{0,\Omega}^2\\
\leq &C(\delta)\bigg(\sum_{l\cap\partial\Omega=\emptyset} h_l^3\int_l[(A^*\nabla p_h)\cdot n]^2+\sum_{\tau\in T^h} h_\tau^4\int_\tau(g'(y_h)+\mathrm{div}(A^*p_h)-u_hp_h)^2\bigg)\\
&+C(\delta)\|y(u_h)-y_h\|_{0,\Omega}^2+C\delta\|p(u_h)-p_h\|_{0,\Omega}^2.
\end{aligned}$$

The estimate about $\|y(u_h)-y_h\|_{0,\Omega}^2$ can be proved similarly, so we omit the details.

Theorem 4.3. *Let* (y,p,u) *and* (y_h,p_h,u_h) *be the solutions of* (2.7)-(2.9) *and* (3.3)-(3.5), *respectively. Assume that all the conditions in Lemma* 4.1 *and* 4.4 *hold. Then,*

$$e^2+\|u-u_h\|_{0,\Omega}^2+\|y-y_h\|_{0,\Omega}^2+\|p-p_h\|_{0,\Omega}^2\leq C\left(\eta_1^2+\hat{\eta}_2^2+\hat{\eta}_3^2\right). \tag{4.44}$$

Proof. It follows from Proposition 2.1 that $\|p\|_{0,\infty,\Omega}\leq C$ and $\|y\|_{0,\infty,\Omega}\leq C$ (see [14], for example). We have that $\|p(u_h)\|_{0,\infty,\Omega}\leq C$ and $\|y_h\|_{0,\infty,\Omega}\leq C$, when h is small enough. Hence,

$$\begin{aligned}
&\|p(u_h)y(u_h)-p_hy_h\|_{0,\Omega}\\
\leq&\|p(u_h)(y(u_h)-y_h)\|_{0,\Omega}+\|y_h(p(u_h)-p_h)\|_{0,\Omega}\\
\leq&\|p(u_h)\|_{0,\infty,\Omega}\|y(u_h)-y_h\|_{0,\Omega}+\|y_h\|_{0,\infty,\Omega}\|p(u_h)-p_h\|_{0,\Omega}\\
\leq&C\bigg(\|y(u_h)-y_h\|_{0,\Omega}+\|p(u_h)-p_h\|_{0,\Omega}\bigg).
\end{aligned} \tag{4.45}$$

Hence, it follows from Lemmas 4.1 and 4.4 that

$$e^2+\|u-u_h\|_{0,\Omega}^2+\|y(u_h)-y_h\|_{0,\Omega}^2+\|p(u_h)-p_h\|_{0,\Omega}^2\leq C\left(\eta_1^2+\hat{\eta}_2^2+\hat{\eta}_3^2\right). \tag{4.46}$$

Note that

$$\|y-y_h\|_{0,\Omega}\leq\|y-y(u_h)\|_{0,\Omega}+\|y(u_h)-y_h\|_{0,\Omega}, \tag{4.47}$$

$$\|p-p_h\|_{0,\Omega}\leq\|p-p(u_h)\|_{0,\Omega}+\|p(u_h)-p_h\|_{0,\Omega}, \tag{4.48}$$

$$\|y(u_h)-y\|_{0,\Omega}+\|p(u_h)-p\|_{0,\Omega}\leq C\|u-u_h\|_{0,\Omega}. \tag{4.49}$$

Then, (4.44) follows from (4.46)-(4.49). □

5. Neumann Boundary Condition

We now consider the parameter problem (1.1) governed by the elliptic problem with the Neumann boundary condition:

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

subject to $y\in H^1(\Omega)$ and

$$a(y(u),v)+(uy(u),v)=(f,v)+\int_{\partial\Omega}rv,\qquad\forall v\in V=H^1(\Omega).$$

It can be shown that the problem (5.1) has the locally unique solution $(y(u),u)$ and that $(y(u),u)$ is the solution of (5.1) only if there is $p\in H^1(\Omega)$ such that $(y,p,u)\in H^1(\Omega)\times H^1(\Omega)\times K$

satisfying

$$(A\nabla y, \nabla v) + (uy, v) = (f, v) + \int_{\partial\Omega} rv \quad \forall v \in H^1(\Omega), \tag{5.2}$$

$$(\nabla q, A^*\nabla p) + (up, q) = (g'(y), q) \qquad \forall q \in H^1(\Omega), \tag{5.3}$$

$$(h'(u), u - w) - (py, u - w) \le 0 \qquad \forall w \in K \subset U. \tag{5.4}$$

Construct the finite element space as in Section 3, excepted that here we let $V^h = W^h$ (not $V^h = W^h \cap H_0^1(\Omega)$). Then, the finite element approximation of (5.2)-(5.4) is: to find $(y_h, p_h, u_h) \in V^h \times V^h \times K^h$ such that

$$a(y_h, v_h) + (u_h y_h, v_h) = (f, v_h) + \int_{\partial\Omega} rv_h \quad \forall v_h \in V^h \subset H^1(\Omega), \tag{5.5}$$

$$(\nabla q_h, A^*\nabla p_h) + (u_h p_h, q_h) = (g'(y_h), q_h) \quad \forall q_h \in V^h \subset H^1(\Omega), \tag{5.6}$$

$$(h'(u_h) - y_h p_h, w_h - u_h) \ge 0 \qquad \forall w_h \in K^h \subset U^h \subset L^2(\Omega). \tag{5.7}$$

Using the similar techniques, we can extend the results of Theorems 4.1, 4.2 and 4.3 as follows.

Theorem 5.1. *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of* (5.2)-(5.4) *and* (5.5)-(5.7)*, respectively. Assume that all the conditions in Theorem* 4.1 *hold. Then,*

$$e^2 + \|u - u_h\|_{0,\Omega}^2 + \|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{1,\Omega}^2 \le C(\eta_1^2 + \bar{\eta}^2) \tag{5.8}$$

where e, η_1 are defined similarly in the last section, and

$$\begin{aligned}\bar{\eta}^2 = & \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [(A\nabla y_h)\cdot n]^2 + \sum_{\tau\in T^h} h_\tau^2 \int_\tau (f + \mathrm{div}(A\nabla y_h) - u_h y_h)^2 \\ & + \sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [(A^*\nabla p_h)\cdot n]^2 + \sum_{\tau\in T^h} h_\tau^2 \int_\tau (g'(y_h) + \mathrm{div}(A^*\nabla p_h) - u_h p_h)^2 \\ & + \sum_{l\subset\partial\Omega} h_l \int_l ((A\nabla y_h)\cdot n - r)^2 + \sum_{l\subset\partial\Omega} h_l \int_l ((A^*\nabla p_h)\cdot n)^2,\end{aligned}$$

where l is a face of an element τ, $[(A^*\nabla p_h \cdot n)]$ and $[(A\nabla y_h \cdot n)]$ are the A-normal derivative jumps over the interior face l, defined by

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

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

Proof. Let $e^p = p_h - p(u_h)$, where $p(u_h)$ is the solution of the equations:

$$(A\nabla y(u_h), \nabla v) + (u_h y(u_h), v) = (f, v) + \int_{\partial\Omega} rv \quad \forall v \in H^1(\Omega), \tag{5.9}$$

$$(\nabla q, A^*\nabla p(u_h)) + (u_h p(u_h), q) = (g'(y(u_h)), q) \quad \forall q \in H^1(\Omega). \tag{5.10}$$

It follows from [1] that there exists a function $\phi \in H^1(\Omega)$ such that

$$c\|e^p\|_{1,\Omega}\|\phi\|_{1,\Omega} \le (\nabla\phi, A^*\nabla e^p) + (u_h e^p, \phi). \tag{5.11}$$

Let $\phi_I \in V^h$ be the interpolation of ϕ defined in Lemma 3.3. Using the equations (5.6), (5.10) and Lemma 3.3, we have that

$$
\begin{aligned}
&c\|e^p\|_{1,\Omega}\|\phi\|_{1,\Omega} \le (\nabla\phi, A^*\nabla e^p) + (u_h e^p, \phi)\\
=&(\nabla(\phi-\phi_I), A^*\nabla e^p) + (u_h e^p, \phi-\phi_I) + (\nabla\phi_I, A^*\nabla e^p) + (u_h e^p, \phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau (A^*\nabla(p_h - p(u_h))\nabla(\phi-\phi_I) + u_h(p_h - p(u_h))(\phi-\phi_I))\\
&\quad + (g'(y_h) - g'(y(u_h)), \phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau (-\mathrm{div}A^*(\nabla p_h) + u_h p_h)(\phi-\phi_I) - (g'(y(u_h)), \phi-\phi_I) + (g'(y_h) - g'(y(u_h)), \phi_I)\\
&\quad + \sum_{l\cap\partial\Omega=\emptyset}\int_l [(A^*\nabla p_h)\cdot n](\phi-\phi_I) + \sum_{l\subset\partial\Omega}\int_l ((A^*\nabla p_h)\cdot n)(\phi-\phi_I)\\
=&\sum_{\tau\in T^h}\int_\tau (-\mathrm{div}(A^*\nabla p_h) + u_h p_h - g'(y_h))(\phi-\phi_I) + (g'(y_h) - g'(y(u_h)), \phi)\\
&\quad + \sum_{l\cap\partial\Omega=\emptyset}\int_l [(A^*\nabla p_h)\cdot n](\phi-\phi_I) + \sum_{l\subset\partial\Omega}\int_l ((A^*\nabla p_h)\cdot n)(\phi-\phi_I)\\
\le& C\Big(\sum_{\tau\in T^h} h_\tau^2 \int_\tau (\mathrm{div}(A^*\nabla p_h) - u_h p_h + g'(y_h))^2\Big)^{\frac{1}{2}}\|\phi\|_{1,\Omega}\\
&\quad + C\Big(\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [(A^*\nabla p_h)\cdot n]^2 + \sum_{l\subset\partial\Omega} h_l \int_l ((A^*\nabla p_h)\cdot n)^2\Big)^{\frac{1}{2}}\|\phi\|_{1,\Omega}\\
&\quad + C\|y_h - y(u_h)\|_{0,\Omega}\|\phi\|_{0,\Omega}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
\|p(u_h) - p_h\|_{1,\Omega}^2 \le& C\Big(\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [(A^*\nabla p_h)\cdot n]^2 + \sum_{l\subset\partial\Omega} h_l \int_l ((A^*\nabla p_h)\cdot n)^2\\
&\quad + \sum_{\tau\in T^h} h_\tau^2 \int_\tau (g'(y_h) + \mathrm{div}(A^* p_h) - u_h p_h)^2 + \|y(u_h) - y_h\|_{0,\Omega}^2\Big).
\end{aligned}
$$

Similarly, it can be proved that

$$
\begin{aligned}
\|y(u_h) - y_h\|_{1,\Omega}^2 \le& C\Big(\sum_{l\cap\partial\Omega=\emptyset} h_l \int_l [(A\nabla y_h)\cdot n]^2 + \sum_{l\subset\partial\Omega} h_l \int_l ((A\nabla y_h)\cdot n - r)^2\\
&\quad + \sum_{\tau\in T^h} h_\tau^2 \int_\tau (f + \mathrm{div}(A\nabla y_h) - u_h y_h)^2\Big).
\end{aligned}
$$

Then, by following the proof in Theorem 4.1, we have (5.8). □

Theorem 5.2. *Let* (y, p, u) *and* (y_h, p_h, u_h) *be the solutions of* (5.2)-(5.4) *and* (5.5)-(5.7), *respectively. Assume that all the conditions in Lemmas* 4.2 *and* 4.3 *hold. Then,*

$$
\eta_1^2 + \bar{\eta}^2 \le C\left(e^2 + \|u - u_h\|_{0,\Omega}^2 + \|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{1,\Omega}^2\right) + C\tilde{\epsilon}^2,
$$

where

$$
\begin{aligned}
\tilde{\epsilon}^2 = & \sum_{\tau\in T^h}\int_\tau h_\tau^2((f+\operatorname{div}(A\nabla y_h)-u_hy_h)-(\overline{f+\operatorname{div}(A\nabla y_h)-u_hy_h}))^2 \\
& + \sum_{\tau\in T^h}\int_\tau h_\tau^2((\operatorname{div}(A^*\nabla p_h)-u_hp_h+g'(y_h))-(\overline{\operatorname{div}(A^*\nabla p_h)-u_hp_h+g'(y_h)}))^2 \\
& + \sum_{l\cap\partial\Omega=\emptyset} h_l\int_l[(A-\tilde{A})\nabla y_h\cdot n]^2+\sum_{l\cap\partial\Omega=\emptyset} h_l\int_l[(A^*-\tilde{A}^*)\nabla p_h\cdot n]^2 \\
& + \sum_{l\subset\partial\Omega} h_l\int_l(((A-\tilde{A})\nabla y_h)\cdot n-r+\tilde{r})^2+\sum_{l\subset\partial\Omega} h_l\int_l(((A^*-\tilde{A}^*)\nabla p_h)\cdot n)^2.
\end{aligned}
$$

Theorem 5.3. *Let* (y,p,u) *and* (y_h,p_h,u_h) *be the solutions of* (5.2)-(5.4) *and* (5.5)-(5.7)*, respectively. Assume that all the conditions in Lemmas* 4.1 *and* 4.4 *hold. Then,*

$$
e^2+\|u-u_h\|_{0,\Omega}^2+\|y-y_h\|_{0,\Omega}^2+\|p-p_h\|_{0,\Omega}^2\le C(\eta_1^2+\tilde{\eta}^2),
$$

where

$$
\begin{aligned}
\tilde{\eta}^2 = & \sum_{l\cap\partial\Omega=\emptyset} h_l^3\int_l[(A\nabla y_h)\cdot n]^2+\sum_{\tau\in T^h} h_\tau^4\int_\tau(f+\operatorname{div}(A\nabla y_h)-u_hy_h)^2 \\
& + \sum_{l\cap\partial\Omega=\emptyset} h_l^3\int_l[(A^*\nabla p_h)\cdot n]^2+\sum_{\tau\in T^h} h_\tau^4\int_\tau\Big(g'(y_h)+\operatorname{div}(A^*\nabla p_h)-u_hp_h\Big)^2 \\
& + \sum_{l\subset\partial\Omega} h_l^3\int_l\Big((A\nabla y_h)\cdot n-r\Big)^2+\sum_{l\subset\partial\Omega} h_l^3\int_l\Big((A^*\nabla p_h)\cdot n\Big)^2.
\end{aligned}
$$

6. Numerical Experiments

In this section, we carry out some numerical experiments to demonstrate the error estimators developed in Section 4. Our numerical examples are the following type of parameter estimation problem:

$$
\min\left\{\frac{1}{2}\int_\Omega(y-y_0)^2+\frac{1}{2}\int_\Omega(u-u_0)^2\right\}
$$
$$
s.t.\quad \left\{\begin{array}{rcl} -\Delta y+uy & = & f, \\ y|_{\partial\Omega} & = & y_0|_{\partial\Omega}=0, \\ u & \ge & d. \end{array}\right.
$$

In our examples, $\Omega=(0,1)^2$, and d is a constant. Let Ω be partitioned into T^h and T_U^h as described in Section 3. We may use different meshes for the approximation of the state and the control. In all our experiments, we shall use η_1 as the control mesh refinement indicator, and $\eta_2+\eta_3$ as the state's and co-state's.

In solving our discretised optimal control problems, we use the preconditioned projection gradient method. We now briefly describe the solution algorithm to be used for solving the numerical examples in this section:

Algorithm

(i) solve the discretised 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.

Again, the readers can refer to [21] for the details of such algorithms.

Example 6.1. The first example is to solve the following problem on $\Omega = (0,1)^2$:

$$\begin{aligned} &\min \quad \frac{1}{2}\int_\Omega (y-y_0)^2 dx + \frac{\alpha}{2}\int_\Omega (u-u_0)^2 dx \\ &\text{s.t.} -\Delta y + uy = f, \quad u \geq 0, \end{aligned} \tag{6.1}$$

where $\alpha = 10^{-2}$ and

$$\begin{aligned} &u = \begin{cases} 1.0, & x_1 + x_2 > 1.0, \\ 0.0, & x_1 + x_2 \leq 1.0, \end{cases} \\ &y = \sin \pi x_1 \sin \pi x_2, \\ &u_0 = u, \quad y_0 = y, \\ &f = 2\pi^2 y + uy, \quad p = 0. \end{aligned} \tag{6.2}$$

We firstly compute Example 6.1 on a uniform mesh. In Figure 6.1, the exact solution u is plotted. The state and co-state are approximated by continuous piecewise linear elements, while discontinuous piecewise linear elements are used to approximate the control. In Table 6.1, the mesh information is displayed with L^2 approximation errors for the control and the states.

The adaptive multi-meshes presented in Figures 6.1 and 6.2 suggest that the u-mesh adapts very well to the neighborhood of discontinuities, and a higher density of nodes are indeed distributed along them. Furthermore the optimal meshes for the parameter and the states are very different as seen in Figures 6.1 and 6.2. It can be clearly seen from Table 6.1 that on the adaptive meshes one may use 10 times fewer degree of freedoms (DOFs) in all the variables to produce a given L^2 control error reduction.

Example 6.2. The second example is the following one also on $\Omega = (0,1)^2$:

$$\begin{aligned} &\min \quad \frac{1}{2}\int_\Omega (y-y_0)^2 dx + \frac{\alpha}{2}\int_\Omega (u-u_0)^2 dx \\ &\text{s.t.} -\Delta y + uy = f, \quad u \geq 0, \end{aligned} \tag{6.3}$$

Table 6.1: Computational results of Example 6.1.

		On uniform mesh			On adaptive mesh		
		u	y	p	u	y	p
Mesh info	# nodes	7777	7777	7777	602	501	501
	# sides	23008	23008	23008	1556	1416	1416
	# elements	15232	15232	15232	955	916	916
	# DOFs	45696	7777	7777	2865	501	501
L^2 error		1.85e-02	1.17e-04	5.64e-06	1.67e-02	1.52e-03	7.11e-05

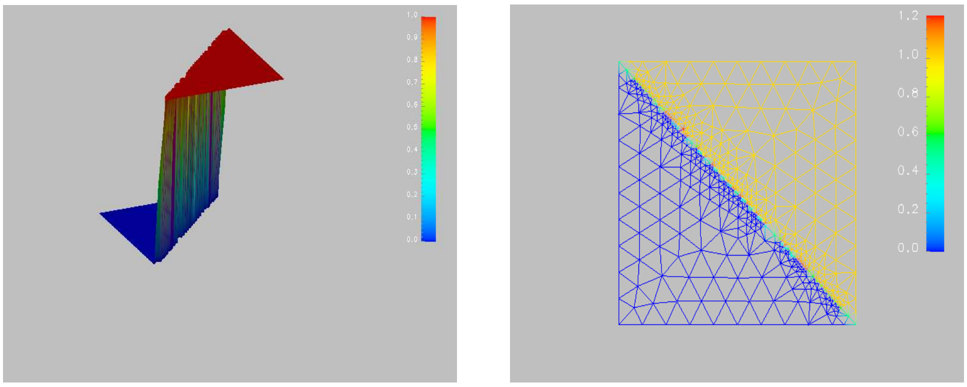

Fig. 6.1. Profile of u and mesh for u_h in Example 6.1.

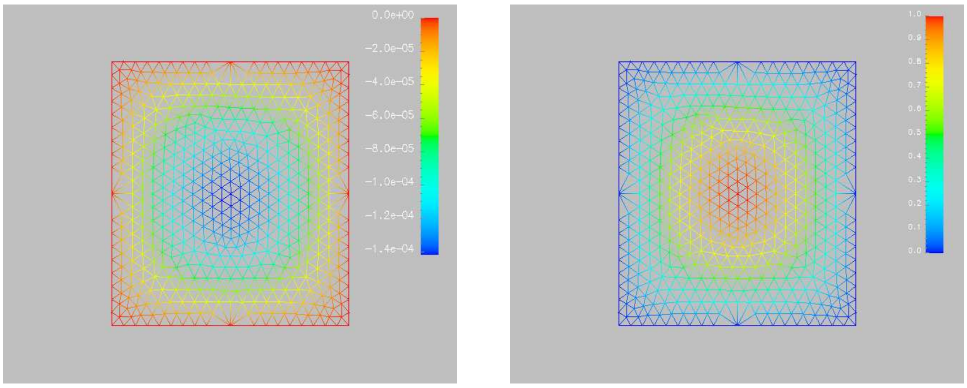

Fig. 6.2. Meshes for p_h (left) and for y_h (right) in Example 6.1.

where

$$
\begin{aligned}
y &= (x_1^2 + x_2^2 - 1)/4, \\
u_0 &= \begin{cases} 0.1/\alpha, & \text{if } x_1 + x_2 < 1, \\ 0, & \text{otherwise} \end{cases} \\
& p = \sin(2\pi x_1)\sin(2\pi x_2), \\
& u = \max(u_0 + yp/\alpha, 0), \\
& y_0 = y - 8\pi^2 p - up, \\
& f = uy - 1, \quad \alpha = 10^{-2}.
\end{aligned}
\tag{6.4}
$$

In Figure 6.3, the exact solution of u is plotted. The state and co-state are approximated by continuous piecewise linear elements, while discontinuous piecewise linear elements are used to approximate the control. The mesh information and numerical results are presented in Table 6.2. It was found that these adaptive meshes can further reduce also about 10 times of the DOFs in all the state variables to produce a given L^2 control error reduction.

Example 6.3. In this example on $\Omega = (0,1)^2$ we solve

$$
\begin{aligned}
\min \quad & \frac{1}{2}\int_\Omega (y - y_0)^2 dx + \frac{\alpha}{2}\int_\Omega (u - u_0)^2 dx \\
& \text{s.t.} - \Delta y + uy = f, \quad u \geq -1.
\end{aligned}
\tag{6.5}
$$

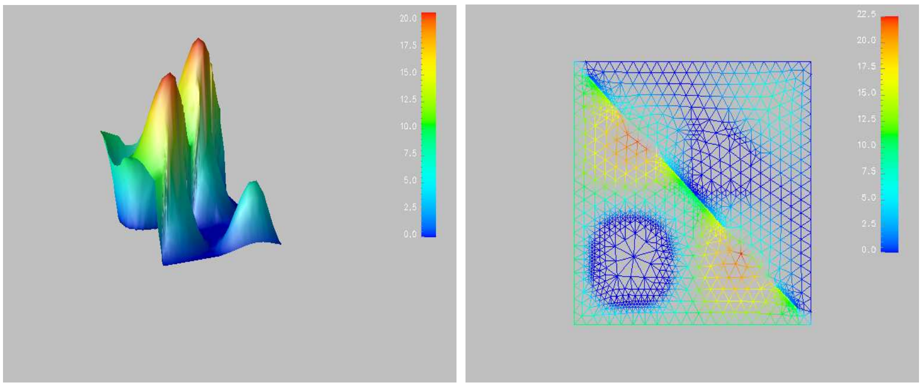

Fig. 6.3. Profile of u and mesh for u_h in Example 6.2.

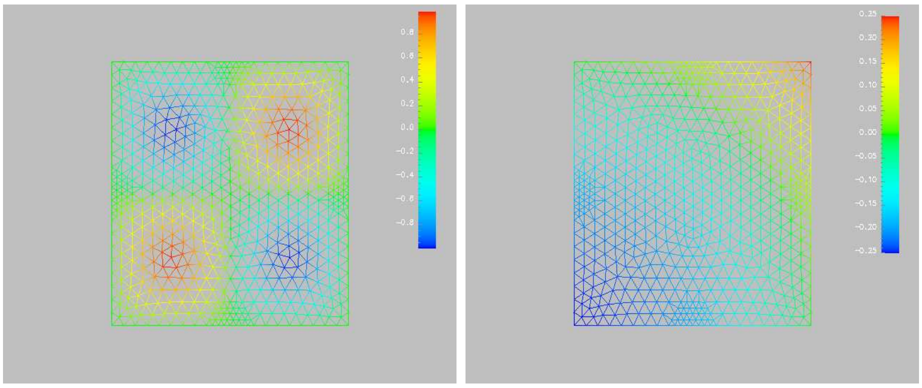

Fig. 6.4. Meshes for p_h (left) and for y_h (right) in Example 6.2.

Table 6.2: Computational results of Example 6.2

		On uniform mesh			On adaptive mesh		
		u	y	p	u	y	p
Mesh info	# nodes	7777	7777	7777	2357	603	603
	# sides	23008	23008	23008	6131	1689	1689
	# elements	15232	15232	15232	3775	1087	1087
	# DOFs	45696	7777	7777	11325	603	603
L^2 error		7.63e-02	1.76e-05	3.63e-04	5.07e-02	1.53e-04	5.21e-03

But with $u^* = u + 1$, the problem is changed into:

$$\begin{aligned}
&-\Delta y + (u^* - 1)y = f, \\
&-\Delta p + (u^* - 1)p = y - y_0, \\
&(\alpha(u^* - 1) - py, v - u^*) \geq 0, \ \forall\, v \geq 0.
\end{aligned} \tag{6.6}$$

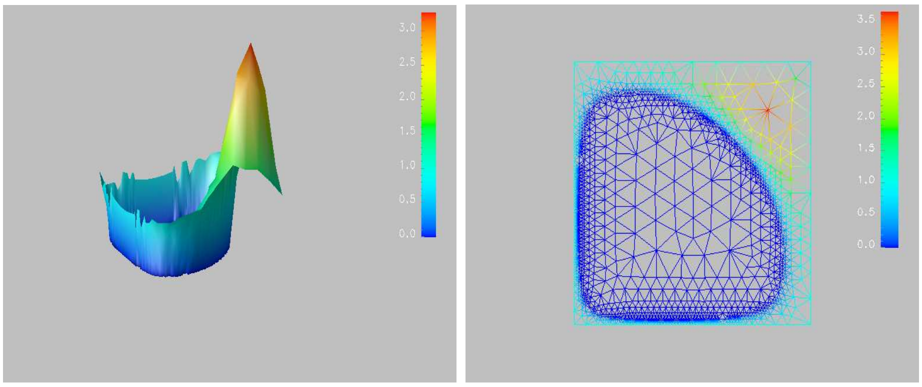
Fig. 6.5. Profile of u^* and mesh for u_h in Example 6.3

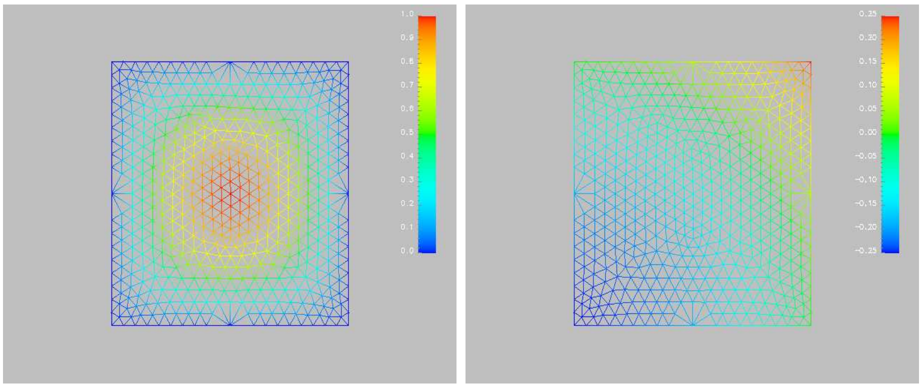
Fig. 6.6. Meshes for p_h (left) and y_h (right) in Example 6.3.

So, we only need to solve the new system. Here, we choose

$$\begin{aligned}
&y = (x_1^2 + x_2^2 - 1)/4, \quad u_0 = 0, \\
&u^* = \max(1 + py/\alpha, 0), \\
&p = \sin(\pi x_1)\sin(\pi x_2), \\
&y_0 = y - 2\pi^2 p - (u^* - 1)p, \\
&f = (u^* - 1)y - 1, \quad \alpha = 10^{-2}.
\end{aligned} \tag{6.7}$$

Table 6.3: Computational results of Example 6.3

		On uniform mesh			On adaptive mesh		
		u^*	y	p	u^*	y	p
Mesh info	# nodes	7777	7777	7777	1639	501	501
	# sides	23008	23008	23008	4258	1416	1416
	# elements	15232	15232	15232	2620	916	916
	# DOFs	45696	7777	7777	7860	501	501
L^2 error		5.44e-03	1.01e-05	1.03e-04	5.97e-03	2.64e-04	1.68e-03

The numerical results are summarized in the Table 6.3. It can be seen that for a given control error, such adaptive meshes can reduce the numbers of DOFs up to ten times. It is clear that the adaptive multi-mesh can save computational work substantially.

7. Appendix

In this appendix, we shall first show Proposition 2.1. Then we show the convexity of the reduced objective functional for the parameter estimation problem. Let us begin with two examples of K, which satisfy Assumption (H). Just for simplicity, we use $\|\cdot\|_{L^p}$ to present $\|\cdot\|_{L^p(\Omega)}$ in this section.

Example 7.1. Note that

$$\|\nabla y\|_{L^2}^2 + (u\,y, y) \geq \|\nabla y\|_{L^2}^2 + \inf_{x\in\Omega} u(x)\|y\|_{L^2}^2,$$

for all $y \in H_0^1(\Omega)$, there exist by Poincare's inequality constants $\underline{c} > 0$ and $C > 0$ such that

$$\|\nabla\, y\|_{L^2}^2 + (u\,y, y) \geq C\|\nabla y\|_{L^2}^2 \quad \forall y \in H_0^1(\Omega),$$

and all $u \in K_{\underline{c}} = \{u \in L^2(\Omega)\colon\ u(x) \geq -\underline{c}\}$. Let $c \geq \underline{c}$ and $K_c = \{u \in L^2(\Omega)\colon\ u(x) \geq -c\}$. By the Lax-Milgram lemma for any $u \in K_c$, exists a unique solution $y(u)$ to (2.2) satisfying $\|y(u)\|_{H^1} \leq \frac{1}{C}\|f\|_{H^{-1}}$. Hence K_c satisfies Assumption (H).

Example 7.2. Denote by T the operator in $H^{-1}(\Omega)$ with $\mathrm{dom}T = H_0^1(\Omega)$, and

$$T(u)\,\varphi = -div(A\nabla\varphi) + u\,\varphi,$$

where $u \in L^2(\Omega)$. Note that since $u\,\varphi \in L^{4/3}(\Omega)$, for $\varphi \in H_0^1(\Omega)(\subset\ L^4(\Omega))$ and $u \in L^2(\Omega)$ it follows that $u\,\varphi \in H^{-1}(\Omega)$.

Assume that $\{\inf\ \tilde{c}(x)\colon x \in \Omega\} > -\infty$. Then the resolvent $(T(\tilde{c}) + \lambda\, I)^{-1}$ exists for all λ sufficiently large. Moreover it is a compact operator on $H^{-1}(\Omega)$. As a consequence, $T(\tilde{c})$ has only point spectrum. In the follows we show that if 0 is not in the point spectrum of $T(\tilde{c})$, then for $\rho > 0$ sufficiently small

$$K_\rho = B(\tilde{c}, \rho) = \{u \in L^2(\Omega)\colon\ \|u - \tilde{c}\| \leq \rho\}$$

satisfies Assumption (H).
To this end, note that under our assumptions, $T(\tilde{c})\colon H_0^1(\Omega) \to H^1(\Omega)$ is an isomorphism. In particular, there exists $\kappa > 0$ such that

$$\|\varphi\|_{H_0^1} \leq \kappa\|T\,\varphi\|_{H^{-1}} \quad \forall\varphi \in H_0^1(\Omega). \tag{7.1}$$

Let $B(0, \rho) = \{u \in L^2(\Omega)\colon \|u\|_{L^2} \leq \rho\}$. Then for arbitrary $y \in H_0^1(\Omega)$ and $u \in B(0, \rho)$ we have

$$\begin{aligned}\|u\,y\|_{H^{-1}} &= \sup\,\{(u\,y, \varphi)\colon \|\varphi\|_{H_0^1} = 1\}\\ &\leq \sup\,\{\|u\|_{L^2}\|y\|_{L^4}\|\varphi\|_{L^4}\colon \|\varphi\|_{H_0^1} = 1\}.\end{aligned}$$

Let C denote the embedding constant of $H^1(\Omega)$ into $L^4(\Omega)$. Then by (7.1)

$$\|u\,y\|_{H^{-1}} \leq C^2\rho\,\|y\|_{H_0^1} \leq C^2\kappa\,\rho\|T(\tilde{c})y\|_{H^{-1}}.$$

Hence the operator $u\,I$, with $u \in B(0, \rho)$ is $T(\tilde{c})-$bounded. Let ρ be such that $C^2\kappa\,\rho < 1$. Then by perturbation analysis, $T(\tilde{c} + u)^{-1} \in \mathcal{L}(H^{-1}(\Omega))$ for every $u \in B(0, \rho)$, and

$$||T(\tilde{c} + u)^{-1}||_{\mathcal{L}(H^{-1})} \leq C, \tag{7.2}$$

for a constant C independent of $c \in B(0,\rho)$. We now show that $\{\|y(\tilde{c}+u)\|_{H_0^1} : c \in B(0,\rho)\}$ is bounded, where

$$T(\tilde{c}+u)y(\tilde{c}+u) = f. \tag{7.3}$$

Taking the inner product in $L^2(\Omega)$ with $y = y(\tilde{c}+u)$, we find

$$(\nabla y, \nabla y) = (f - (\tilde{c}+u)y, y).$$

This implies that

$$\begin{aligned}
&\|\nabla y\|_{L^2}^2 \\
\leq &\|f\|_{H^{-1}}\|\nabla y\|_{L^2} + \|\tilde{c}+u\|_{L^2}\|y\|_{L^4}^2 \\
\leq &\|f\|_{H^{-1}}^2 + \frac{1}{4}\|\nabla y\|_{L^2}^2 + (\|\tilde{c}\|_{L^2} + \rho)\|y\|_{L^4}^2.
\end{aligned}$$

We recall the well-known estimate: For every $\epsilon > 0$ there exists a constant k_ϵ such that

$$\|\varphi\|_{L^4}^2 \leq \epsilon\|\nabla\varphi\|_{L^2}^2 + k_\epsilon\|\varphi\|_{L^2}^2 \qquad \forall \varphi \in H_0^1(\Omega), \tag{7.4}$$

where we use that $n \leq 3$, see [14]. With $\epsilon = \frac{1}{4}(\|\tilde{c}\|_{L^2} + \rho)^{-1}$ we find

$$\|\nabla\, y(\tilde{c}+u)\|_{L^2}^2 \leq 2\|f\|_{H^{-1}}^2 + 2\, k_\epsilon \|y(\tilde{c}+u)\|_{L^2}^2.$$

By the interpolation inequality [18] there exists a constant $c > 0$ such that

$$\|\varphi\|_{L^2}^2 \leq c\|\varphi\|_{H_0^1}\|\varphi\|_{H^{-1}} \text{ for all } \varphi \in H_0^1(\Omega).$$

Consequently

$$\|\nabla y(\tilde{c}+u)\|_{L^2}^2 \leq 2\|f\|_{H^{-1}}^2 + 2\, c\, k_\epsilon \|\nabla y(\tilde{c}+u)\| \|y(\tilde{c}+u)\|_{H^{-1}},$$

and by (7.2)

$$\|\nabla y(\tilde{c}+u)\|_{L^2} \leq C\|f\|_{H^{-1}},$$

for a constant C independent of $u \in B(0,\rho)$. Thus K_ρ satisfies Assumption (H).

Furthermore if for any $u \in K$, $\inf_{x\in\Omega} u(x) > -\infty$ and 0 is not in the point spectrum of $T(u)$, then K satisfies the assumption (H).

Proposition 7.1. *Assume that $\partial\Omega$ is $C^{1,1}$ regular or that Ω is a parallelepiped. Suppose that Assumption (H) holds. Then for all $u \in K$, the solution $y = y(u,f)$ of (2.2) is in $H^2(\Omega)$. Furthermore for any $v \in K$, there exist a neighborhood $O(v)$ of v and a constant $C(v) > 0$ such that*

$$\|y(u,f)\|_{H^2} \leq C(v)\, \|f\|_{L^2} \qquad \forall u \in O(v) \cap K.$$

Proof. For any $v \in K$, let Q be the neighborhood defined in Assumption (H). We first show that there exists C_1 such that

$$\|u\, y\|_{L^{3/2}} \leq C_1\|f\|_{H^{-1}}, \tag{7.5}$$

where $y = y(u,f)$, uniformly for $u \in Q$ and $f \in H^{-1}$. In fact

$$\left(\int_\Omega u^{3/2}y^{3/2}dx\right)^{2/3} \leq \left(\int u^2 dx\right)^{1/2} \cdot \left(\int y^6 dx\right)^{1/6},$$

and since $H^1(\Omega)$ embeds continuously into $L^6(\Omega)$, estimate (7.5) follows from Assumption (H).

Recall [24], that $W^{1,3/2}(\Omega)$ embeds continuously into $L^3(\Omega)$. Hence the set $\{u\,y(u,f)\colon u \in Q\}$ can be considered as a uniformly bounded family of elements in $(W^{1,3/2}(\Omega))^*$. From [24], for every $u\,y$ there exists $\{(q_0, q_1, \cdots, q_n)\} \in (L^3(\Omega))^{n+1}$ such that

$$\langle u\,y, w\rangle_{W^{1,3/2}(\Omega)^*, W^{1,3/2}(\Omega)} = \langle q_0, w\rangle + \sum_{i=1}^{n} \langle q_i, D_i w\rangle \quad \forall w \in W^{1,3/2}(\Omega), \tag{7.6}$$

and

$$\sum_{i=0}^{n} \|q_i\|_{L^3} \leq C_2 \|f\|_{H^{-1}} \quad \forall u \in Q.$$

By global regularity results, see [24], applied to

$$\begin{cases} -div(A\nabla\, y) = f - u\,y \;\text{ in }\; \Omega, \\ \\ y = 0 \;\text{ on }\; \partial\Omega, \end{cases} \tag{7.7}$$

with $f - u\,y$ as inhomogeneity and $u\,y$ as in (7.6), we find

$$\|y\|_{W^{1,3}} \leq C_4 \|f\|_{H^{-1}}. \tag{7.8}$$

Hence, in the case of dimension two it follows that

$$\|y\|_{L^\infty} \leq C_5 \|f\|_{H^{-1}}.$$

Consequently by H^2-regularity results [14], applied to (7.7) we have $y \in H^2(\Omega)$ and

$$\|y\|_{H^2} \;\leq\; C_6 \|f\|_{L^2} + C_5 \sup_{u \in Q} \|u\|_{L^2} \|f\|_{H^{-1}} \;\leq\; C_7 \|f\|_{L^2},$$

with C_i independent of $u \in Q$ and $f \in L^2(\Omega)$. This is the desired estimate for $n = 2$.

Turning to the case $n = 3$ we utilize (7.8) and the continuous embedding $W^{1,3}(\Omega) \hookrightarrow L^{12}(\Omega)$, see [24]. Using Hölder's inequality with $p = \frac{7}{6}$ and $p' = 7$, and (7.8)

$$\|u\,y\|_{L^{\frac{12}{7}}} \leq C_8 \|u\|_{L^2} \cdot \|y\|_{L^{12}} \leq C_9 \|y\|_{W^{1,3}} \leq C_4 C_9 \|f\|_{H^{-1}}. \tag{7.9}$$

Hence $u\,y$ can be considered as functional on $L^{\frac{12}{5}}(\Omega)$. Since $W^{1,4/3}(\Omega)$ embeds continuously into $L^{\frac{12}{5}}(\Omega)$ (for $n = 3$), $\{c\,y(u)\colon u \in Q\}$ can be considered as a uniformly bounded family of functionals on $W^{1,4/3}(\Omega)$. Hence for every $u\,y$ there exists $\{(q_0, \cdots, q_n)\} \in L^4(\Omega)^{n+1}$ such that

$$\langle u\,y, w\rangle_{W^{1,4/3}(\Omega)^*, W^{1,4/3}} = \langle q_0, w\rangle + \sum_{i=1}^{n} \langle q_i, D_i w\rangle, \quad \forall w \in W^{1,4/3}(\Omega),$$

and

$$\sum_{i=0}^{n} \|q_i\|_{L^4} \;\leq\; C_{10} \|f\|_{H^{-1}}, \;\text{ for all }\; u \in Q,$$

where $q_i = q_i(uy)$. Proceeding as in the case $n = 2$ we find

$$\|y\|_{W^{1,4}} \;\leq\; C_{11} \|f\|_{H^{-1}}. \tag{7.10}$$

Since for $W^{1,4}(\Omega) \hookrightarrow C(\Omega)$ for $n = 3$, the proof can be completed as for $n = 2$.

Remark 7.1. Assumption (H) can be strengthened in the following global formulation: For any bounded subset $Q \in L^2(\Omega)$, there exists a constant $C(Q) > 0$ such that

$$\|y(u,f)\|_{H^1} \leq C(Q)\|f\|_{H^{-1}} \quad \forall u \in Q \cap K.$$

It follows from the proof of Proposition 2.1 that there exists $C(Q) > 0$ such that

$$\|y(u,f)\|_{H^2} \leq C(Q)\|f\|_{L^2} \quad \forall u \in Q \cap K.$$

In the follows we shown that $J(u)$ is convex for any $\alpha > 0$. To this end, we first need a lemma.

Lemma 7.1. *Let (y_α, u_α) be the solution of* (2.5) *for $\alpha > 0$. Let $u_z \in K$ be a minimum norm solution of $y(u_z) = z$. Then*

$$\lim_{\alpha \to 0} \|u_\alpha\|_{L^2} = \|u_z\|_{L^2}. \tag{7.11}$$

Proof. Note that

$$\sup \|u_{\alpha_1}\|_{L^2} \leq \inf \|u_{\alpha_2}\|_{L^2} \leq \|u_z\|_{L^2} \quad \text{if } \alpha_1 \geq \alpha_2, \tag{7.12}$$

where inf (sup) is taken over all solutions to (P_{α_2}) (respectively (P_{α_1})). Note here that we do not know whether or not (P_α) has a unique solution. If (7.11) were not correct, then there would exist a sequence α_n with $\lim_{n\to\infty} \alpha_n = 0$ and associated solution u_{α_n} such that

$$\lim_{n\to\infty} \|u_{\alpha_n}\|_{L^2} < \|u_z\|_{L^2}.$$

But $\{u_{\alpha_n}\}_{n=1}^\infty$ is bounded in $L^2(\Omega)$, and hence there exists a subsequence denoted by the same symbol, and $\bar{u} \in L^2(\Omega)$ such that $\lim_{n\to\infty} u_{\alpha_n} = \bar{u}$ in $L^2(\Omega)$. We also have $y(u_{\alpha_n}) \to y(\bar{u})$ in $L^2(\Omega)$. Taking the limit in

$$\frac{1}{2}\|y(u_{\alpha_n}) - z\|_{L^2}^2 + \frac{\alpha_n}{2}\|u_{\alpha_n}\|_{L^2}^2 \leq \frac{1}{2}\|y(u) - z\|_{L^2}^2 + \frac{\alpha_n}{2}\|u\|_{L^2}^2 \quad \forall u \in K,$$

we find that

$$\frac{1}{2}\|y(\bar{u}) - z\|_{L^2}^2 \leq \frac{1}{2}\|y(u) - z\|_{L^2}^2 \quad \forall u \in K.$$

In particular, taking $u = u_z$, we have $y(\bar{u}) = z$. From (7.12) and weak lower semi-continuity of the norm, we have

$$\overline{\lim}_{n\to\infty}\|u_{\alpha_n}\|_{L^2} \leq \|u_z\|_{L^2} \leq \|\bar{u}\|_{L^2} \leq \underline{\lim}_{n\to\infty}\|u_{\alpha_n}\|_{L^2},$$

and thus u_{α_n} converges strongly to a minimum norm solution of $y(u) = z$. Thus it follows from (7.12) that for any $\epsilon > 0$, there exists $\alpha(\epsilon)$ such that

$$\|u_z\|_{L^2}^2 - \inf \|u_\alpha\|_{L^2}^2 < \epsilon^2 \quad \forall \alpha \in (0, \alpha(\epsilon)],$$

where inf is taken over all solutions of (P_α). □

Proposition 7.2. *Suppose z is identifiable. Let u be a solution of* (2.5). *Then for all $v \in L^2(\Omega)$*

$$J''(u)(v,v) \geq \frac{\alpha}{2}\|v\|_{L^2}^2, \tag{7.13}$$

where $J(u) = g(y(u)) + j(u)$ is defined in (2.4).

Proof. It is easy to see that

$$J''(u)(v,v) = (y',y') + (y-z, y''(v,v)) + \alpha(v,v), \tag{7.14}$$

and

$$\begin{aligned} -\mathrm{div}(A\nabla y'(v)) + uy'(v) &= -vy, \\ -\mathrm{div}(A\nabla y''(v,v)) + uy''(v,v) &= -2vy'(v). \end{aligned}$$

Let us denote

$$C(u)y := -\mathrm{div}(A\nabla y) + uy.$$

We have that

$$\begin{aligned} &(y',y') + (y-z, y''(v,v)) = \|C^{-1}(vy)\|_{L^2}^2 - 2(y-z, C^{-1}(vy'(v))) \\ =&\|C^{-1}(vy)\|_{L^2}^2 + 2(C^{-1}(y-z), vC^{-1}(vy)) \\ =&\|C^{-1}(vy)\|_{L^2}^2 + 2(C^{-1}(y-z), vC^{-1}(vy)) + \|vC^{-1}(y-z)\|_{L^2}^2 - \|vC^{-1}(y-z)\|_{L^2}^2 \\ =&\|C^{-1}(vy) + vC^{-1}(y-z)\|_{L^2}^2 - \|vC^{-1}(y-z)\|_{L^2}^2. \end{aligned}$$

Therefore, it follows from (7.14) that

$$J''(u)(v,v) \geq \alpha\|v\|_{L^2}^2 - \|vC^{-1}(y-z)\|_{L^2}^2 \geq \alpha\|v\|_{L^2}^2 \left(1 - \frac{1}{\alpha}\|C^{-1}(y-z)\|_{L^\infty}^2\right). \tag{7.15}$$

It follows from Lemma 7.1 that taken $\bar{\epsilon} = (1/2\|C^{-1}\|_{\mathcal{L}(L^2,L^\infty)}^2)^{\frac{1}{2}}$, we have a $\alpha(\bar{\epsilon})$ such that for all $\alpha \in (0, \alpha(\bar{\epsilon})]$,

$$\|u_z\|_{L^2}^2 - \|u_\alpha\|_{L^2}^2 \leq \bar{\epsilon}^2 = \frac{1}{2\|C^{-1}\|_{\mathcal{L}(L^2,L^\infty)}^2}.$$

Note that

$$\|y_\alpha - z\|_{L^2}^2 + \alpha\|u_\alpha\|_{L^2}^2 \leq \|y(u_z) - z\|_{L^2}^2 + \alpha\|u_z\|_{L^2}^2.$$

We have

$$\|y_\alpha - z\|_{L^2}^2 \leq \alpha\Big(\|u_z\|_{L^2}^2 - \|u_\alpha\|_{L^2}^2\Big) \leq \frac{\alpha}{2\|C^{-1}\|_{\mathcal{L}(L^2,L^\infty)}^2}.$$

That is

$$\|C^{-1}(y-z)\|_{L^\infty}^2 \leq \frac{\alpha}{2}.$$

Therefore, it follows from (7.15) that

$$J''(u)(v,v) \geq \frac{\alpha}{2}\|v\|_{L^2}^2.$$

This proves (7.13). $\square$

Acknowledgments. Supported by the National Basic Research Program under the Grant 2005CB321701, 2010CB731505 and the National Natural Science Foundation of China under the Grant 10771211.

References

[1] I. Babuška and A.K. Aziz, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Academic Press, 1972.

[2] M. Bergounioux, M. Haddou, M. Hintermuller and K. Kunisch, A comparison of a Moreau-Yosida based active set strategy and interior point methods for constrained optimal control problems, *SIAM J. Optimz.*, **11** (2000), 495-521.

[3] R. Becker and R. Rannacher, An optimal control approach to a posteriori error estimation, Acta Numerica, pp. 1-102, Cambridge University Press, 2001.

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

[5] R. Becker and B. Vexler, A posteriori error estimation for finite element discretization of parameter identication problems, *Numer. Math.*, **96**:3 (2004), 435-459.

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

[7] T. Feng, N. Yan and W. Liu, Adaptive finite element methods for the identification of distributed parameters in elliptic equation, *Adv. Comput. Math.*, **29**:1 (2008), 27-53.

[8] T. Feng, M. Gulliksson and W. Liu, Adaptive finite element methods for parameter estimation problems in linear elasticity, *International Journal of Numerical Analysis & Modeling*, **6** (2009), 17-32.

[9] M.B. Giles and N.A. Pierce, An introduction to the adjoint approach to design, *Flow. Turbul. Combust.*, **65** (2000), 393-415.

[10] M. Hintermüller, K. Ito and K. Kunisch, The primal-dual active set strategy as semi-smooth newton method, Report 214, Technische Universitat Graz, 2001.

[11] K. Ito and K. Kunisch, The augmented Lagrangian method for parameter estimation in elliptic systems, *SIAM J. Control. Optim.*, **28** (1990), 113-136.

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

[13] K. Kunisch, W.B. Liu and N. Yan, A posteriori error estimators for a model parameter estimation problem, in Numerical Mathematics and Advanced Applications (F.Brezzi, A. Buffa, S. Corsaro and A. Murli eds), pp. 723-730, Springer, Milan, 2001.

[14] O.A. Ladyzhenskaya and H.H. Urlatseva, Linear and Quasilinear Elliptic Equations, Academic Press, New York, 1968.

[15] R. Li, On multi-mesh h-adaptive methods, *J. Sci. Comput.*, **24**:3 (2005), 321-341.

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

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

[18] J.L. Lions and E. Magenes, Non-Homogeneous Boundary Value Problems and Applications, Springer Verlag, Belrin, 1972.

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

[20] W. Liu and N. Yan, A posteriori error analysis for convex distributed optimal control problems, *Adv. Comput. Math.*, **15**:1-4 (2001), 285-309.

[21] W.B. Liu and N. Yan, Adaptive Finite Element Methods for Optimal Control Governed by PDEs, Science Press, Beijing, 2008.

[22] W. Liu, Wei Gong and N. Yan A new finite element approximation of a state-constrained optimal control problem, *J. Comput. Math.*, **27** (2009), 97-114.

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

[24] G.M. Troianiello, Elliptic Differential Equations and Obstacle Problems, Plenum Press, New York, 1987.

[25] R. Verfürth, A Review of a Posteriori Error Estimation and Adaptive Mesh Refinement, Wiley-Teubner, London, UK, 1996.

[26] D. Yang, Y.Z. Chang and W.B. Liu, A priori error estimate and superconvergence analysis for an optimal control problem of bilinear type, *J. Comput. Math.*, **26** (2008), 471-487.