Recovery Type a Posteriori Error Estimates of Fully Discrete Finite Element Methods for General Convex Parabolic Optimal Control Problems

Yanping Chen\textsuperscript{1,\ast} and Yueling Tang\textsuperscript{2}

\textsuperscript{1}School of Mathematical Sciences, South China Normal University, Guangzhou 510631, Guangdong, China.

\textsuperscript{2}Hunan Key Laboratory for Computation and Simulation in Science and Engineering, School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, Hunan, China.

Received 28 June 2011; Accepted (in revised version) 5 February 2012

Abstract. In this paper, recovery type a posteriori error estimates of fully discrete finite element approximation for general convex parabolic optimal control problems with pointwise control constraints are investigated. The time discretization is based on the backward Euler method. The state and the adjoint state are approximated by piecewise linear functions and the control is approximated by piecewise constant functions. We derive the superconvergence properties of finite element solutions. By using the superconvergence results, we obtain recovery type a posteriori error estimates. Some numerical examples are presented to verify our theoretical results.

AMS subject classifications: 35B37, 49J20, 65N30

Key words: General convex optimal control problems, Fully discrete finite element approximation, A posteriori error estimates, Superconvergence, Recovery operator.

1. Introduction

It is well known that finite element methods are undoubtedly the most widely used numerical method in computing optimal control problems. A systematic introduction of finite element methods for PDEs and optimal control problems can be found in [7, 15, 17–20, 23, 28, 32, 33]. The literature on a posteriori error estimation of finite element method is huge. Some internationally known works can be found in [1–4, 6]. Concerning finite element methods of elliptic optimal control problems, a posteriori error estimates of residual type were investigated in [26], a posteriori error estimates of recovery type were derived in [21].

\footnote{Corresponding author. Email addresses: yanpingchen@scnu.edu.cn (Y. Chen), tangyuelong@163.com (Y. Tang)}
For parabolic optimal control problems, a priori error estimates of space-time finite element discretization were investigated in [29, 30], a priori error estimates of finite element methods were established in [24], and residual type a posteriori error estimates of finite element methods were established in [27, 34]. Recently, Fu and Rui considered a characteristic finite element approximation of control problems governed by transient advection-diffusion equations in [16].

Superconvergence properties of finite element methods for elliptic optimal control problems were studied in [10, 11, 31]. Yang and Chang showed the superconvergence properties for optimal control problem of bilinear type in [35]. The superconvergence of optimal control problems governed by Stokes equations were derived in [25]. Some superconvergence results of mixed finite element methods for elliptic optimal control problems can be found in [5, 8, 9, 12, 13, 36]. Recently, we discussed the superconvergence of finite element methods for quadratic parabolic optimal control problems in [14].

The purpose of this work is to study the superconvergence and recovery type a posteriori error estimates of the fully discrete finite element approximation for general convex parabolic optimal control problems with control constraints.

We are interested in the following parabolic optimal control problem:

\[
\begin{align*}
\min_{u(x,t) \in K} & \left\{ \int_0^T (g(y(x,t)) + h(u(x,t))) \, dt \right\}, \\
y_i(x,t) - \nabla(A(x) \nabla y(x,t)) & = f(x,t) + Bu(x,t), \quad x \in \Omega, \ t \in J, \\
y(x,t) & = 0, \quad x \in \partial \Omega, \ t \in J, \\
y(x,0) & = y_0(x), \quad x \in \Omega,
\end{align*}
\]

(1.1)

where \( \Omega \) be a bounded domain in \( \mathbb{R}^2 \) with a Lipschitz boundary \( \partial \Omega \), \( 0 < T < +\infty \) and \( J = [0, T] \). \( g(\cdot) \) and \( h(\cdot) \) are convex functionals on \( L^2(\Omega) \). The coefficient \( A(x) = (a_{ij}(x))_{2 \times 2} \in (W^{1,\infty}(\Omega))^{2 \times 2} \), such that for any \( \xi \in \mathbb{R}^2 \), \( (A(x)\xi) \cdot \xi \geq c | \xi |^2 \) with \( c > 0 \). Let \( B \) be a continuous linear operator from \( L^2(\Omega) \) to \( L^2(\Omega) \) and \( f(x,t) \in C(J; L^2(\Omega)) \). Moreover, we assume that \( g(\cdot) \) is bounded below, \( h(u) \to +\infty \) as \( ||u||_{L^q(\Omega)} \to \infty \) and \( K \) is a nonempty closed convex set in \( L^2(J; L^2(\Omega)) \), defined by

\[
K = \left\{ v(x,t) \in L^2(J; L^2(\Omega)) : a \leq v(x,t) \leq b, \ a.e. \ (x,t) \in \Omega \times J \right\},
\]

where \( a \) and \( b \) are constants.

In this paper, we adopt the standard notation \( W^{m,q}(\Omega) \) for Sobolev spaces on \( \Omega \) with norm \( || \cdot ||_{W^{m,q}(\Omega)} \) and seminorm \( | \cdot |_{W^{m,q}(\Omega)} \). We set \( H^1_0(\Omega) \equiv \left\{ v \in H^1(\Omega) : v|_{\partial \Omega} = 0 \right\} \) and denote \( W^{m,2}(\Omega) \) by \( H^m(\Omega) \). We denote by \( L^s(J; W^{m,q}(\Omega)) \) the Banach space of \( L^s \) integrable functions from \( J \) into \( W^{m,q}(\Omega) \) with norm \( ||v||_{L^s(J; W^{m,q}(\Omega))} = \left( \int_0^T ||v||_{W^{m,q}(\Omega)}^s \, dt \right)^{\frac{1}{s}} \) for \( s \in [1, \infty) \) and the standard modification for \( s = \infty \). We can define the space \( H^s(J; W^{m,q}(\Omega)) \). The details can be found in [24]. In addition, \( c \) or \( C \) denotes a generic positive constant independent of \( h \) and \( \Delta t \).

The plan of the paper is as follows. In Section 2, we formulate the fully discrete finite element approximation for general convex parabolic optimal control problems.
3, we give some useful intermediate error estimates. In Section 4, we derive the superconvergence properties for the control, the state and the adjoint state. In Section 5, we obtain a posteriori error estimates of recovery type for the fully discrete approximation scheme. We do some numerical experiments to verify our theoretical results in the last section.


For ease of exposition, we set \( V = L^2(J; W) \) with \( W = H^1_0(\Omega) \) and \( X = L^2(J; U) \) with \( U = L^2(\Omega) \). Moreover, we denote \( \| \cdot \|_{H^m(\Omega)} \) and \( \| \cdot \|_{L^2(\Omega)} \) by \( \| \cdot \|_m \) and \( \| \cdot \| \), respectively.

Throughout the paper we impose the following assumptions:

(A1) \( g'(\cdot) \) is locally Lipschitz continuous and there exists a constant \( c > 0 \), such that
\[
(g'(y_1) - g'(y_2), y_1 - y_2) \geq c \| y_1 - y_2 \|^2, \quad \forall y_1, y_2 \in W.
\]

(A2) There exists a constant \( c > 0 \), such that
\[
(h'(u_1) - h'(u_2), u_1 - u_2) \geq c \| u_1 - u_2 \|^2, \quad \forall u_1, u_2 \in U.
\]

(A3) Let \( h(u) = \int_\Omega j(u) dx \) then \( (h'(u), v) = (j'(u), v) \), where \( j(\cdot) \) is a smooth and convex function such that \( j''(\cdot) \in W^{1,\infty}(\Omega) \) and \( j'''(\cdot) \in L^\infty(\mathbb{R}) \).

Let
\[
a(v, w) = \int_\Omega (A \nabla v) \cdot \nabla w \, dx, \quad \forall v, w \in W,
\]
\[
(f_1, f_2) = \int_\Omega f_1 \cdot f_2 \, dx, \quad \forall f_1, f_2 \in U.
\]

It follows from the assumptions on \( A \) that
\[
a(v, v) \geq c \| v \|^2, \quad |a(v, w)| \leq C \| v \|_1 \| w \|_1, \quad \forall v, w \in W.
\]

Thus a weak formula for the problem (1.1) reads: Find \( (y, u) \in \left( H^1(J; L^2(\Omega)) \cap V \right) \times K \), such that

\[
\begin{align*}
\min_{u \in K} \left\{ \int_0^T (g(y) + h(u)) \, dt \right\}, \\
(y_t, w) + a(y, w) = (f + Bu, w), \quad \forall w \in W, t \in J, \\
y(x, 0) = y_0(x), \quad x \in \Omega.
\end{align*}
\]

It is well known (see, e.g., [23, 27]) that the problem (2.3) has a unique solution \( (y, u) \), and the pair \( (y, u) \in \left( H^1(J; L^2(\Omega)) \cap V \right) \times K \) is the solution of (2.3) if and only if there is a adjoint state \( p \in H^1(J; L^2(\Omega)) \cap V \) such that the triplet \( (y, p, u) \) satisfies the following optimality conditions:

\[
\begin{align*}
(y_t, w) + a(y, w) = (f + Bu, w), \quad \forall w \in W, t \in J, \\
y(x, 0) = y_0(x), \quad x \in \Omega.
\end{align*}
\]
We define for \( p \), where \( p \) is the standard modification for \( Y \).

\[ |||\cdot||| \]

\( \Delta \) where \( \tau \) denotes the diameter of the element \( \tau \). Further more, we set

\[
U^h = \{ v_h \in L^2(\Omega) : v_h|_{\tau} = \text{constant}, \forall \tau \in \mathcal{T}^h \},
\]

\[ K^h = \{ v_h \in U^h : a \leq v_h \leq b \}, \]

\[ W^h = \{ v_h \in C(\Omega) : v_h|_{\tau} \in \mathbb{P}_1, \forall \tau \in \mathcal{T}^h, v_h|_{\partial \Omega} = 0 \}, \]

where \( \mathbb{P}_1 \) is the space of polynomials up to order 1.

We now consider the fully discrete finite element approximation of the problem (2.3). Let \( \Delta t > 0, N = T/\Delta t \in \mathbb{Z}^+, t_n = n\Delta t, n = 0, 1, \cdots, N \). Set \( \varphi^n = \varphi(x, t_n) \) and

\[
d_t \varphi^n = \frac{\varphi^n - \varphi^{n-1}}{\Delta t}, \quad n = 1, 2, \cdots, N.
\]

We define for \( 1 \leq p < \infty \) the discrete time-dependent norms

\[
|||\varphi|||_{l^p(J; W^{m,q}(\Omega))} := \left( \Delta t \sum_{n=k}^{N} |||\varphi^n|||_{W^{m,q}(\Omega)}^p \right)^{\frac{1}{p}},
\]

where \( k = 0 \) for the control \( u \) and the state \( y \) and \( k = 1 \) for the adjoint state \( p \), with the standard modification for \( p = \infty \). For convenience, we denote \( |||\cdot|||_{l^p(J; W^{m,q}(\Omega))} \) by \( |||\cdot|||_{l^p} \) and let

\[
l^p(J; W^{m,q}(\Omega)) := \{ \varphi : |||\varphi|||_{l^p(J; W^{m,q}(\Omega))} < \infty \}, \quad 1 \leq p \leq \infty.
\]

Then a possible fully discrete finite element approximation of (2.3) is as follows:

\[
\begin{cases}
\min_{u^n_h \in K^h} \left\{ \sum_{n=1}^{N} \Delta t \left( g(y^n_h) + h(u^n_h) \right) \right\}, \\
(d_t y^n_h, w_h) + a(y^n_h, w_h) = (f^n + Bu^n_h, w_h), \quad \forall w_h \in W^h, n = 1, 2, \cdots, N, \\
y^0_h(x) = y^h_0(x), \quad x \in \Omega,
\end{cases}
\]

where \( y^0_h(x) \in W^h \) is an approximation of \( y_0(x) \).

It follows (see e.g. [27]) that the control problem (2.7) has a unique solution \( \left( y^n_h, u^n_h \right) \), \( n = 1, 2, \cdots, N \), and \( \left( y^n_h, u^n_h \right) \in W^h \times K^h, n = 1, 2, \cdots, N \), is the solution of (2.7) if and only
Generally speaking, we select \( y_h^n(x) = P_h(y_0(x)) \) and \( P_h \) is an elliptic projection operator which will be specified later.

### 3. Error estimates of intermediate variables

We define some intermediate variables. For any \( v \in K \), let \((y(v), p(v)) \in V \times V\) be the solution of the following equations:

\[
(y_t(v), w) + a(y(v), w) = (f + Bv, w), \quad \forall w \in W, \quad t \in J,  \\
y(v)(x, 0) = y_0(x), \quad x \in \Omega,  \\
-p_t(v) + a(q, p(v)) = (g'(y(v)), q), \quad \forall q \in W, \quad t \in J,  \\
p(v)(x, T) = 0, \quad x \in \Omega.  
\]

For any \( v \in K \), a pair \((y_h^n(v), p_h^{n-1}(v)) \in W^h \times W^h\), satisfies the following system:

\[
\begin{align*}
\left( d_t y_h^n(v), w_h \right) &+ a\left( y_h^n(v), w_h \right) = (f^n + Bv^n, w_h), \quad \forall w_h \in W^h, \quad n = 1, 2, \ldots,  N, \quad \forall v \in K,  \\
y_h^n(v)(x) &= y_0(x), \quad x \in \Omega,  \\
-p_h^n(v) + a\left( q_h, p_h^{n-1}(v) \right) &= (g'(y_h^n(v)), q_h), \quad \forall q_h \in W^h, \quad n = 1, 2, \ldots,  N.  
\end{align*}
\]

Let \( u \) and \( u_h \) be the solutions of (2.4)-(2.6) and (2.8)-(2.10), respectively. It is clear that \((y, p) = (y(u), p(u))\) and \((y_h, p_h) = (y(u_h), p_h(u_h))\).

We introduce the standard \( L^2(\Omega)\)-orthogonal projection \( Q_h : U \rightarrow U^h \), which satisfies: for all \( \psi \in U \)

\[
(\psi - Q_h \psi, v_h) = 0, \quad \forall v_h \in U^h.  
\]
and the elliptic projection $P_h : W \to W^h$, which satisfies: for any $\phi \in W$
\[ a(\phi - P_h \phi, w_h) = 0, \quad \forall w_h \in W^h. \] (3.6)

We have the following approximation properties:
\[ \| \psi - Q_h \psi \|_{1,s} \leq Ch^{1+s} |\psi|_1, \quad \forall \psi \in H^1(\Omega), s = 0, 1, \] (3.7)
\[ \| \phi - P_h \phi \| \leq Ch^2 \| \phi \|_{H^2(\Omega)}, \quad \forall \phi \in H^2(\Omega). \] (3.8)

**Lemma 3.1.** Let $(y_h(Q_hu), p_h(Q_hu))$ and $(y_h(u), p_h(u))$ be the discrete solutions of (3.3)-(3.4) with $v = Q_hu$ and $v = u$, respectively. Suppose that $u \in L^2_J(H^1(\Omega))$ and the assumption (A1) is satisfied. Then
\[ \| y_h(Q_hu) - y_h(u) \|_{2(H^1)} + \| p_h(Q_hu) - p_h(u) \|_{2(H^1)} \leq Ch^2. \] (3.9)

**Proof.** Set $v = Q_hu$ and $v = u$ in (3.3), respectively. Then we obtain the following error equation:
\[ (d_t y^h_n(Q_hu) - d_t y^h_n(u), y^h_n(Q_hu) - y^h_n(u)) + a(y^h_n(Q_hu) - y^h_n(u), w_h) = (B(Q_hu^n - u^n), w_h), \quad \forall w_h \in W^h, n = 1, 2, \ldots, N. \] (3.10)

From Cauchy’s inequality, we have
\[ (d_t y^h_n(Q_hu) - d_t y^h_n(u), y^h_n(Q_hu) - y^h_n(u)) \geq \frac{1}{2\Delta t} \left( \| y^h_n(Q_hu) - y^h_n(u) \|^2 - \| y^{n-1}_h(Q_hu) - y^{n-1}_h(u) \|^2 \right), \] (3.11)
and
\[ (B(Q_hu^n - u^n), y^h_n(Q_hu) - y^h_n(u)) \leq C \| Q_hu^n - u^n \|_{-1} \| y^h_n(Q_hu) - y^h_n(u) \|_1 \leq C h^2 \| u^n \|_1 \| y^h_n(Q_hu) - y^h_n(u) \|_1 \leq C(\delta) h^4 \| u^n \|_1^2 + \delta \| y^h_n(Q_hu) - y^h_n(u) \|_1^2. \] (3.12)

By choosing $w_h = y^h_n(Q_hu) - y^h_n(u)$ in (3.10) and multiplying both sides of (3.10) by $2\Delta t$, then summing $n$ from 1 to $N$, we get
\[ \| y^h_n(Q_hu) - y^h_n(u) \|^2 + \sum_{n=1}^N \Delta t \| y^h_n(Q_hu) - y^h_n(u) \|_1^2 \leq C(\delta) h^4 \sum_{n=1}^N \Delta t \| u^n \|_1^2 + \delta \sum_{n=1}^N \Delta t \| y^h_n(Q_hu) - y^h_n(u) \|_1^2. \] (3.13)
Thus, we have
\[ \| y_h(Q_hu) - y_h(u) \|_{2(H^1)} \leq Ch^2 \| u \|_{2(H^1)}. \] (3.14)

Similarly, we obtain that
\[ \| p_h(Q_hu) - p_h(u) \|_{2(H^1)} \leq C \| y_h(Q_hu) - y_h(u) \|_{2(L^2)}. \] (3.15)

Then (3.9) follows from (3.14) and (3.15).
Lemma 3.2. Let \((y(v), p(v))\) and \((y_h(v), p_h(v))\) be the solutions of (3.1)-(3.2) and (3.3)-(3.4), respectively. Assume that \(y(v), p(v) \in L^2(J; H^2(\Omega)) \cap H^1(J; H^2(\Omega)) \cap H^2(J; L^2(\Omega))\) and the assumption (A1) is satisfied. Then, we have

\[
|||P_h y(v) - y_h(v)|||_{\mathcal{Q}(H^1)} + |||P_h p(v) - p_h(v)|||_{\mathcal{Q}(H)} \leq C \left( h^2 + \Delta t \right). \tag{3.16}
\]

Proof. From (3.1) and (3.3), we obtain

\[
\left( y^n_t(v) - d_t y^n_h(v), w_h \right) + a \left( y^n(v) - y^n_h(v), w_h \right) = 0,
\forall w_h \in W_h, \ n = 1, 2, \cdots, N. \tag{3.17}
\]

By using the definition of \(P_h\), we get

\[
\left( d_t P_h y^n(v) - d_t y^n_h(v), w_h \right) + a \left( P_h y^n(v) - y^n_h(v), w_h \right) = \left( d_t P_h y^n(v) - d_t y^n_h(v) + d_t y^n(v) - y^n_h(v), w_h \right). \tag{3.18}
\]

Note that

\[
\left( d_t P_h y^n(v) - d_t y^n_h(v), P_h y^n(v) - y^n_h(v) \right) \\
\leq \left\| d_t P_h y^n(v) - d_t y^n(v) \right\| \left\| P_h y^n(v) - y^n_h(v) \right\| \\
\leq C h^2 \left\| P_h y^n(v) - y^n_h(v) \right\| \\
\leq C h^2 (\Delta t)^{-1} \sum_{t_n=1}^{t_n} \left\| y^n(v) \right\|_2 \left\| P_h y^n(v) - y^n_h(v) \right\| \\
\leq C h^2 (\Delta t)^{-1} \left\| y^n(v) \right\|_{L^2(t_{n-1}; t_n; H^2(\Omega))} \left\| P_h y^n(v) - y^n_h(v) \right\|, \tag{3.19}
\]

and

\[
\left( d_t y^n(v) - y^n_t(v), P_h y^n(v) - y^n_h(v) \right) \\
\leq (\Delta t)^{-1} \left\| y^n(v) - y^{n-1}(v) - \Delta t y^n(v) \right\| \left\| P_h y^n(v) - y^n_h(v) \right\| \\
= (\Delta t)^{-1} \sum_{t_n=1}^{t_n} \left\| y^n(v) \right\|_2 \left\| P_h y^n(v) - y^n_h(v) \right\| \\
\leq C (\Delta t)^{\frac{1}{2}} \left\| y^n(v) \right\|_{L^2(t_{n-1}; t_n; L^2(\Omega))} \left\| P_h y^n(v) - y^n_h(v) \right\|. \tag{3.20}
\]

Similar to Lemma 3.1, from (3.18)-(3.20) and Young’s inequality, we have

\[
\left\| P_h y^n(v) - y^n_h(v) \right\|^2 + c \sum_{n=1}^{N} \Delta t \left\| P_h y^n(v) - y^n_h(v) \right\|^2 \\
\leq C(\delta) \left( h^4 \left\| y^n(v) \right\|^2_{L^2(J; H^2(\Omega))} + (\Delta t)^2 \left\| y^n(v) \right\|^2_{L^2(J; L^2(\Omega))} \right) + \delta \sum_{n=1}^{N} \Delta t \left\| P_h y^n(v) - y^n_h(v) \right\|^2. \tag{3.21}
\]
Thus, we get
\[ \|P_h y(v) - y_h(v)\|_{L^2(H^1)} \leq C \left( h^2 + \Delta t \right). \] (3.22)

Similarly, we derive that
\[ \|P_h p(v) - p_h(v)\|_{L^2(H^1)} \leq C \left( h^2 + \Delta t \right). \] (3.23)

From (3.22) and (3.23), we obtain (3.17).

### 4. Superconvergence analysis

Let \( u \) be the solutions of (2.4)-(2.6). For a fixed \( t^* (0 \leq t^* \leq T) \), we divide \( \Omega \) into the following subsets:
\[
\begin{align*}
\Omega^+ &= \{ \tau : \tau \subset \Omega, a < u(\cdot, t^*) < b \}, \\
\Omega^0 &= \{ \tau \subset \Omega, u(\cdot, t^*)|_{\tau} = a \text{ or } u(\cdot, t^*)|_{\tau} = b \}, \\
\Omega^- &= \Omega \setminus (\Omega^+ \cup \Omega^0).
\end{align*}
\]

It is easy to see that the above three subsets are not intersected with each other and \( \Omega = \Omega^+ \cup \Omega^0 \cup \Omega^- \). We assume that \( u \) and \( \mathcal{T}_h \) are regular such that \( \text{meas}(\Omega^-) \leq Ch, \) (see e.g., [30]).

**Theorem 4.1.** Let \( u \) and \( u_h \) be the solutions of (2.4)-(2.6) and (2.8)-(2.10), respectively. Assume that all the conditions in lemmas 3.1-3.2 are valid and the assumptions (A1)-(A3) are satisfied. Moreover, we suppose that the exact control and adjoint state solution satisfy
\[
\begin{align*}
u, h'(u) + B^* p &\in L^2_P(J; W^{1,\infty}(\Omega)), \\
\end{align*}
\]

Then, we have
\[ \|Q_h u - u_h\|_{L^2(\Omega^2)} \leq C \left( h^{\frac{5}{4}} + \Delta t \right). \] (4.1)

**Proof.** Let \( v = u_h \) in (2.6) and \( v = Q_h u^n \) in (2.10), we derive
\[
\begin{align*}
0 \leq & \left( h'(u^n) + B^* p^n, u^n_h - u^n \right) + \left( h' \left( u^n_h \right) + B^* p^{n-1}_h, Q_h u^n - u^n_h \right) \\
= & \left( h' \left( u^n_h \right) - h'(u^n) + B^* p^{n-1}_h - B^* p^n, Q_h u^n - u^n_h \right) + \left( h'(u^n) + B^* p^n, Q_h u^n - u^n_h \right). \\
\end{align*}
\] (4.2)

According to the assumption (A2) and (4.2), we obtain
\[
\begin{align*}
c \|Q_h u^n - u^n_h\|^2 &\leq \left( h' \left( Q_h u^n \right) - h' \left( u^n_h \right), Q_h u^n - u^n_h \right) \\
= & \left( h' \left( u^n \right) - h' \left( u^n_h \right), Q_h u^n - u^n_h \right) + \left( h' \left( Q_h u^n \right) - h'(u^n), Q_h u^n - u^n_h \right) \\
\leq & \left( h' \left( u^n \right) + B^* p^n, Q_h u^n - u^n \right) + \left( h' \left( Q_h u^n \right) - h'(u^n) + B^* p^{n-1}_h - B^* p^n, Q_h u^n - u^n_h \right). \\
\end{align*}
\] (4.3)
From the assumption (A3), we have that there exists a constant $0 \leq \theta \leq 1$ such that
\[\left(h'(Q_h u^n) - h'(u^n), Q_h u^n - u^n_{h} \right)\]
= \left( j''(u^n) (Q_h u^n - u^n) + \frac{1}{2} j'''(u^n + \theta (Q_h u^n - u^n)) (Q_h u^n - u^n)^2, Q_h u^n - u^n_{h} \right)\]
= \left( (j''(u^n) - \pi (j''(u^n))) (Q_h u^n - u^n), Q_h u^n - u^n_{h} \right) + \frac{1}{2} (j'''(u^n + \theta (Q_h u^n - u^n)) (Q_h u^n - u^n)^2, Q_h u^n - u^n_{h} \right)\]
\leq C \|j''(\cdot)\|_{1,\infty} \|Q_h u^n - u^n\| \|Q_h u^n - u^n_{h}\|
+ \frac{C}{2} \|j'''(\cdot)\|_{0,\infty} \|Q_h u^n - u^n\|^2 \|Q_h u^n - u^n_{h}\|
\leq C h^2 \|u^n\|_1 \|Q_h u^n - u^n_{h}\|.

It is clear that
\[\left( B^* p^{n-1}_h - B^* p^n, Q_h u^n - u^n_{h} \right)\]
= \left( B^* p^{n-1}_h(u_h) - B^* p^{n-1}_h(Q_h u^n) + B^* p^{n-1}_h(Q_h u^n) - B^* p^{n-1}_h(u_h), Q_h u^n - u^n_{h} \right)\]
+ \left( B^* p^{n-1}_h(u_h) - B^* p^{n-1}_h(u_h) + B^* p^{n-1}_h(u_h) - B^* p^n(u_h), Q_h u^n - u^n_{h} \right)\).

By using (A1) and (3.3)-(3.4), we have
\[\left( B^* p^{n-1}_h(u_h) - B^* p^{n-1}_h(Q_h u^n), Q_h u^n - u^n_{h} \right)\]
= - \left( g' \left(y^n_h(u_h)\right) - g' \left(y^n_h(Q_h u^n)\right), y^n_h(u_h) - y^n_h(Q_h u^n) \right)\)
\leq - c \|y^n_h(u_h) - y^n_h(Q_h u^n)\|^2 \leq 0.

From (4.3)-(4.6), we obtain
\[\|Q_h u^n - u^n_{h}\|_{L^2(\Omega)}^2 = \sum_{n=1}^{N} \Delta t \left( Q_h u^n - u^n_{h}, Q_h u^n - u^n_{h} \right)\]
\leq \sum_{n=1}^{N} \Delta t \left( h'(u^n) + B^* p^n, Q_h u^n - u^n \right)\]
+ \sum_{n=1}^{N} \Delta t \left( B^* p^{n-1}_h(Q_h u^n) - B^* p^{n-1}_h(u_h), Q_h u^n - u^n_{h} \right)\]
+ \sum_{n=1}^{N} \Delta t \left( B^* p^{n-1}_h(u_h) - B^* p^{n-1}_h(u_h), Q_h u^n - u^n_{h} \right)\)
+ \sum_{n=1}^{N} \Delta t \left( B^* p^{n-1}_h(u_h) - B^* p^n(u_h), Q_h u^n - u^n_{h} \right)\)
+ C h^2 \Delta t \sum_{n=1}^{N} \|u^n\|_1 \|Q_h u^n - u^n_{h}\|
:= I_1 + I_2 + I_3 + I_4 + I_5.
For the first term, we have
\[
(h'(u^n) + B^* p^n, Q_h u^n - u^n) = \int_{\Omega^+} + \int_{\Omega^-} + \int_{\Omega^0} (h'(u^n) + B^* p^n) (Q_h u^n - u^n) \, dx,
\]
and \((Q_h u^n - u^n)|_{\Omega^0} = 0\). From (2.6), we have \(h'(u^n) + B^* p^n = 0\) on \(\Omega^+\). Let \(\pi^c\) be the element average operator defined in [21]. Then
\[
I_1 = \Delta t \sum_{n=1}^{N} \int_{\Omega^-} (h'(u^n) + B^* p^n) (Q_h u^n - u^n) \, dx
\]
\[
= \Delta t \sum_{n=1}^{N} \int_{\Omega^-} (h'(u^n) + B^* p^n - \pi^c(h'(u^n) + B^* p^n)) (Q_h u^n - u^n) \, dx
\]
\[
\leq C \Delta t \sum_{n=1}^{N} \|h'(u^n) + B^* p^n\|_{1,\Omega^-} \|u^n\|_{1,\Omega^-}
\]
\[
\leq C \Delta t \sum_{n=1}^{N} \|h'(u^n) + B^* p^n\|_{W^{1,\infty}(\Omega^-)} \|u^n\|_{W^{1,\infty}(\Omega^-)} \text{meas}(\Omega^-)
\]
\[
\leq C \Delta t \left( \|h'(u^n) + B^* p^n\|_{L^2(\Omega^-)}^2 + \|u^n\|_{L^2(\Omega^-)}^2 \right).
\]

From Young’s inequality and Lemma 3.1, we derive
\[
I_2 = \sum_{n=1}^{N} \Delta t \left( B^* \left( p_{h,n}^{n-1} (Q_h u) - p_h^{n-1}(u) \right), Q_h u^n - u_h^n \right)
\]
\[
\leq C(\delta) \sum_{n=1}^{N} \Delta t \left( \|p_{h,n}^{n-1} (Q_h u) - p_h^{n-1}(u)\|_{L^2}^2 + \|p_{h,n}^{n-1} (Q_h u) - p_h^{n-1}(u)\|_{L^2}^2 \right)
\]
\[
= C(\delta) \|p_h (Q_h u) - p_h(u)\|_{L^2}^2 + C(\delta) \|Q_h u - u_h\|_{L^2}^2
\]
\[
\leq C(\delta) \|Q_h u - u_h\|_{L^2}^2.
\]

By using (3.8), Young’s inequality and Lemma 3.2, we get
\[
I_3 = \sum_{n=1}^{N} \Delta t \left( B^* \left( p_{h,n}^{n-1}(u) - P_h p_{h,n}^{n-1}(u) + P_h p_{h,n}^{n-1}(u) - p_{h,n}^{n-1}(u) \right), Q_h u^n - u_h^n \right)
\]
\[
\leq C(\delta) \sum_{n=1}^{N} \Delta t \|p_{h,n}^{n-1}(u) - P_h p_{h,n}^{n-1}(u)\|_{L^2}^2 + C(\delta) \|Q_h u - u_h\|_{L^2}^2
\]
\[
\leq C(\delta) \|p_h(u) - P_h p(u)\|_{L^2}^2 + C(\delta) \|Q_h u - u_h\|_{L^2}^2
\]
\[
\leq C(\delta) \left( h^4 + (\Delta t)^2 + \|Q_h u - u_h\|_{L^2}^2 \right).
\]
Note that
\[ I_4 = \sum_{n=1}^{N} \Delta t \left( B^* \left( p_{n-1}^n(u) - p_n^u(u) \right), Q_h u^n - u_h^n \right) \]
\[ \leq C \sum_{n=1}^{N} \Delta t \int_{t_{n-1}}^{t_n} \| P_t(u) \| dt \| Q_h u^n - u_h^n \| \]
\[ \leq C(\delta)(\Delta t)^2 \| P_t(u) \|_{L^2(J; L^2(\Omega))}^2 + \delta \| Q_h u - u_h \|_{L^2(\Omega)}^2, \] (4.12)
and
\[ I_5 = C h^2 \Delta t \sum_{n=1}^{N} \| u^n \|_1 \| Q_h u^n - u_h^n \| \]
\[ \leq C(\delta) h^3 \sum_{n=1}^{N} \Delta t \| u^n \|_1^2 + \delta \sum_{n=1}^{N} \Delta t \| Q_h u^n - u_h^n \|_1^2 \]
\[ \leq C(\delta) h^4 \| u^n \|_{L^2(\Omega)}^2 + \delta \| Q_h u - u_h \|_{L^2(\Omega)}^2. \] (4.13)

Let \( \delta \) be small enough, then (4.1) follows from (4.7)-(4.13).

**Theorem 4.2.** Let \((y, p, u)\) and \((y_h, p_h, u_h)\) be the solutions (2.4)-(2.6) and (2.8)-(2.10), respectively. Assume that all the conditions in Theorem 4.1 are valid. Then
\[ \| P_h y - y_h \|_{L^2(\Omega)} + \| P_h p - p_h \|_{L^2(\Omega)} \leq C \left( h^2 + \Delta t \right). \] (4.14)

**Proof.** From (2.4) and (2.8), we have the following error equation:
\[ \left( y^n_t - d_t y_h^n, w_h \right) + a \left( y^n - y_h^n, w_h \right) = \left( B \left( u^n - u_h^n \right), w_h \right), \]
\[ \forall w_h \in W^h, n = 1, 2, \cdots, N. \] (4.15)

By choosing \( w_h = P_h y^n - y_h^n \) and using the definition of \( P_h \), we get
\[ \left( d_t P_h y^n - d_t y_h^n, P_h y^n - y_h^n \right) + a \left( P_h y^n - y_h^n, P_h y^n - y_h^n \right) \]
\[ = \left( d_t P_h y^n - d_t y_h^n, P_h y^n - y_h^n \right) + B \left( u^n - Q_h u^n \right), P_h y^n - y_h^n \]. (4.16)

Note that
\[ \left( B \left( u^n - Q_h u^n \right), P_h y^n - y_h^n \right) \leq C \| u^n - Q_h u^n \|_{L^2(\Omega)} \| P_h y^n - y_h^n \|_1 \]
\[ \leq C h^2 \| u^n \|_{L^2(\Omega)} \| P_h y^n - y_h^n \|_1 \]
\[ \leq C(\delta) h^4 \| u^n \|_{L^2(\Omega)}^2 + \delta \| P_h y^n - y_h^n \|_1^2. \] (4.17)

Similar to Lemma 3.2, by using (4.1) and (4.16)-(4.17), we derive
\[ \| P_h y - y_h \|_{L^2(\Omega)} \leq C \left( h^2 + \Delta t \right). \] (4.18)

Similarly, we get that
\[ \| P_h p - p_h \|_{L^2(\Omega)} \leq C \left( h^2 + \Delta t \right). \] (4.19)

From (4.18) and (4.19), we derive (4.14).
5. A posteriori error estimates

We introduce recovery operators $R_h$ and $G_h$ for the control and the state and the adjoint state, respectively. Let $R_h v$ be a continuous piecewise linear function (without zero boundary constraint). Similar to the Z-Z patch recovery in [37, 38], the value of $R_h v$ on the nodes are defined by least-squares argument on an element patches surrounding the nodes. The gradient recovery operator $G_h v = \left( R_h v_x, R_h v_x \right)$, where $R_h$ is the recovery operator defined above for the recovery of the control. The details can be found in [21].

**Theorem 5.1.** Let $u$ and $u_h$ be the solutions of (2.4)-(2.6) and (2.8)-(2.10), respectively. Assume that all the conditions in Theorem 4.1 are valid. Moreover, we suppose that $u \in L^\infty(J; W^{1,\infty}(\Omega))$ and $\Omega$ is convex. Then

$$|||R_h u_h - u|||_{L^2(L^2)} \leq C \left( h^3 + \Delta t \right).$$

**Proof.** It follows from Lemma 4.2 in [21] that

$$||R_h u^n_h - u^n|| \leq ||R_h u^n_h - R_h Q_h u^n|| + ||R_h Q_h u^n - R_h u^n|| + ||R_h u^n - u^n||$$

$$\leq ||R_h u^n_h - R_h Q_h u^n|| + ||R_h Q_h u^n - R_h u^n|| + C h^2.$$  

By using the definition of $R_h$, we have

$$R_h u^n = R_h Q_h u^n,$$  

and

$$||R_h u^n_h - R_h Q_h u^n|| \leq ||u^n_h - Q_h u^n||.$$  

From Theorem 4.1 and (5.2)-(5.4), we obtain

$$||R_h u_h - u||_{L^2(L^2)}^2 \leq C h^3 + C ||Q_h u - u_h||_{L^2(L^2)}^2 \leq C \left( h^3 + (\Delta t)^2 \right).$$

Then (5.1) follows from (5.5).

**Theorem 5.2.** Let $(y, p, u)$ and $(y_h, p_h, u_h)$ be the solutions of (2.4)-(2.6) and (2.8)-(2.10), respectively. Suppose that all the conditions in Theorem 4.2 are valid and $y, p \in L^2(J; H^1(\Omega))$. Then, we have

$$||G_h y - \nabla y||_{L^2(L^2)} + ||G_h p - \nabla p||_{L^2(L^2)} \leq C \left( h^3 + \Delta t \right).$$

**Proof.** Let $y_I$ be the piecewise linear Lagrange interpolation of $y$. According to Theorem 2.1.1 in [22], we have

$$||P_h y - y_I||_1 \leq C h^2 ||y||_3.$$  

(5.7)
Recovery a posteriori error estimates

From the standard interpolation error estimate technique (see, e.g., [15]) that
\[ \| G_h y - \nabla y \| \leq C h^2 | y |_3. \] (5.8)

By using (5.7)-(5.8), we get
\[
\begin{align*}
\| G_h y_h^n - \nabla y_h^n \| &= \| G_h y_h^n - G_h P_h y_h^n \| + \| G_h P_h y_h^n - G_h y_h^n \| \\
&\leq C \| y_h^n - P_h y_h^n \|_1 + C \| P_h y_h^n - y_h^n \|_1 + \| G_h y_h^n - \nabla y_h^n \| \\
&\leq C \| y_h^n - P_h y_h^n \|_1 + C h^2 | y_h^n |_3.
\end{align*}
\] (5.9)

Therefore,
\[
\sum_{n=1}^{N} \Delta t \| G_h y_h^n - \nabla y_h^n \|^2 \leq C \sum_{n=1}^{N} \Delta t \| y_h^n - P_h y_h^n \|^2_1 + C h^4 \sum_{n=1}^{N} \Delta t \| y_h^n \|^2_3.
\] (5.10)

From Lemma 4.2 and (5.10), we derive
\[ ||| G_h y_h - \nabla y_h |||_{\ell^2(L^2)} \leq C \left( h^2 + \Delta t \right). \] (5.11)

Similarly, we can prove that
\[ ||| G_h p_h - \nabla p_h |||_{\ell^2(L^2)} \leq C \left( h^2 + \Delta t \right). \] (5.12)

Then (5.6) follows from (5.11)-(5.12).

By using the above superconvergence results, we obtain the following a posteriori error estimates of fully discrete finite element approximation for parabolic optimal control problems.

**Theorem 5.3.** Assume that all the conditions in Theorem 5.1 and Theorem 5.2 are valid. Then
\[
\begin{align*}
\eta_1 &:= ||| R_h u - u_h \|||_{\ell^2(L^2)} = \| u - u_h \|_{\ell^2(L^2)} + \mathcal{O} \left( h^2 + \Delta t \right), \quad (5.13) \\
\eta_2 &:= ||| G_h y - \nabla y_h \|||_{\ell^2(L^2)} = \| \nabla (y - y_h) \|_{\ell^2(L^2)} + \mathcal{O} \left( h^2 + \Delta t \right), \quad (5.14) \\
\eta_3 &:= ||| G_h p - \nabla p_h \|||_{\ell^2(L^2)} = \| \nabla (p - p_h) \|_{\ell^2(L^2)} + \mathcal{O} \left( h^2 + \Delta t \right). \quad (5.15)
\end{align*}
\]

**Proof.** From Theorem 5.1 and Theorem 5.2, it is easy to obtain the above results.

6. Numerical experiments

For a constrained parabolic optimal control problem:
\[
\min_{u \in K} J(u),
\]
where \( J(u) \) is a convex functional on \( X \) and \( K \) is a close convex subset of \( X \), the iterative scheme reads \((n = 0, 1, 2, \ldots)\):

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

(6.1)

where \( b(\cdot, \cdot) = \int_0^T (\cdot, \cdot) \) is a symmetric and positive definite bilinear form, \( \rho_n \) is a step size of iteration and the projection operator \( P_K^b \) can be computed in the similar way as [21].

The bilinear form \( b(\cdot, \cdot) \) provides a suitable precondition for the projection algorithm. For an acceptable error \( Tol \), by applying (6.1) and to the discretized parabolic optimal control problem (2.7), we present the following projection gradient algorithm:

**Projection gradient algorithm**

Step 1. Solve the following equations:

\[
\begin{aligned}
&\begin{cases}
 b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n \int_0^T (h'(u_n) + B^*p_n, v), \quad u_{n+\frac{1}{2}}, u_n \in U^h, \forall v \in U^h, \\
\frac{\eta_n - \eta_{n-1}}{\Delta t}, w + a \left( y_n^i, w \right) = \left( f^i + Bu_n, w \right), \quad y_n^i, y_{n-1}^i \in W^h, \forall w \in W^h, \\
\frac{p_n - p_{n-1}}{\Delta t}, q + a(q, p_{n-1}^i) = g' \left( y_n^i \right), \quad p_n, p_{n-1}^i \in W^h, \forall q \in W^h, u_{n+1} = P_K^b(u_{n+\frac{1}{2}}),
\end{cases}
\end{aligned}
\]

where we have omitted the subscript \( h \);

Step 2. Calculate the iterative error: \( E_{n+1} = ||u_{n+1} - u_n||_{L^2(U)} \);

Step 3. If \( E_{n+1} \leq Tol \), stop, else go to Step 1.

Similar to [21], by selecting different meshes for the control and the state and the adjoint state and using \( \eta_1 \) and \( \eta_2 + \eta_3 \) as meshes refinement indicators for the control and the state and the adjoint state, respectively. For an acceptable error \( Tol' \), we construct the following fully discrete adaptive finite element algorithm:

**Adaptive algorithm**

Step 1. Solve the discretized optimization problem with the Projection gradient algorithm on the current meshes get numerical solution \( u'_n \) and calculate the error estimators \( \eta_n \);

Step 2. Adjust the meshes by using the estimators \( \eta_n \) and update the numerical solution \( u'_n \) and obtain \( u'_{n+1} \) on new meshes;

Step 3. Calculate the iterative error: \( E'_{n+1} = ||u'_{n+1} - u'_n||_{L^2(U)} \);

Step 4. If \( E'_{n+1} \leq Tol' \), stop, else go to Step 1.

The following numerical examples were solved with codes developed based on AFEPack. The details can be found at http://www.acm.caltech.edu/~rlj/AFEPack/. Just for simplicity, we let \( I \) be the \( 2 \times 2 \) identity matrix and denote \( ||| \cdot |||_{L^2(U)} \) by \( || \cdot || \). The discretization
was described in Section 2: the state and the adjoint state are approximated by piecewise linear functions and the control is approximated by piecewise constant functions. Let \( \Omega = [0, 1] \times [0, 1], T = 1 \) and \( B \) be the identity operator. We solve the following type of parabolic optimal control problems:

\[
\min_{u \in K} \left\{ \frac{1}{2} \int_0^T \left( \| y(x, t) - y_d(x, t) \|^2 + \| u(x, t) - u_d(x, t) \|^2 \right) dt \right\},
\]

\[
y_t(x, t) - \text{div}(A(x) \nabla y(x, t)) = f(x, t) + Bu(x, t), \quad x \in \Omega, \; t \in J,
\]

\[
y(x, t) = 0, \quad x \in \partial \Omega, \; t \in J,
\]

\[
y(x, 0) = y_0(x), \quad x \in \Omega.
\]

Moreover, we assume that

\[
K = \left\{ v(x, t) \in L^2(J; L^2(\Omega)) : a \leq v(x, t) \leq b, \; (x, t) \in \Omega \times J \right\}.
\]

**Example 1.** The data are as follows:

\[
a = -0.4, \; b = 0.4,
\]

\[
A(x) = \begin{pmatrix} \sin(\frac{\pi x_1}{2}) & 0 \\ 0 & \sin(\frac{\pi x_2}{2}) \end{pmatrix}
\]

\[
p(x, t) = \sin(2\pi x_1) \sin(2\pi x_2)(1 - t),
\]

\[
y(x, t) = \sin(2\pi x_1) \sin(2\pi x_2) t,
\]

\[
u_d(x, t) = 2\sin(2\pi x_1) \sin(2\pi x_2) t,
\]

\[
u(x, t) = \max(-0.4, \min(0.4, u_d(x, t) - p(x, t))),
\]

\[
f(x, t) = y_t(x, t) - \text{div}(A(x) \nabla y(x, t)) - Bu(x, t),
\]

\[
y_d(x, t) = y(x, t) + p_t(x, t) + \text{div}(A'(x) \nabla p(x, t)).
\]

The first example is solved by the Projection gradient algorithm. The errors \( ||u - u_h||, ||Q_h u - u_h|| \) and \( ||u - R_h u_h|| \) on a sequence of uniformly refined meshes are shown in Table 1. It is easy to see \( ||u - u_h|| = \mathcal{O}(h + \Delta t) \), \( ||Q_h u - u_h|| = \mathcal{O} \left( h^\frac{3}{2} + \Delta t \right) \) and \( ||u - R_h u_h|| = \mathcal{O} \left( h^\frac{3}{2} + \Delta t \right) \) which confirm our theoretical results. In Figure 1, we plot the profile of the numerical solution \( u_h \) at \( t = 0.5 \) when \( h = 1.25E - 2 \) and \( \Delta t = 1/270 \).

| Table 1: The error of the control variable, Example 1. |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| **h**       | **\Delta t** | **||u - u_h||** | **||Q_h u - u_h||** | **||u - R_h u_h||** |
| 1.0E-1     | 1/10         | 7.22045E-1      | 3.61315E-2          | 5.48766E-2          |
| 5.0E-2     | 1/30         | 3.67380E-2      | 8.82138E-3          | 1.78529E-2          |
| 2.5E-2     | 1/90         | 1.80089E-2      | 3.01585E-3          | 5.74956E-3          |
| 1.25E-2    | 1/270        | 9.01929E-3      | 9.00128E-4          | 1.93514E-3          |
Example 2. The data are as follows:

\[
A(x) = I, \ a = 0.5, \ b = 1.5, \\
p(x, t) = \sin(\pi x_1)\sin(\pi x_2)\sin(\pi t), \\
y(x, t) = \sin(\pi x_1)\sin(\pi x_2)\sin(\pi t), \\
u_d(x, t) = \begin{cases} 
2\sin(\pi x_1)\sin(\pi x_2)\sin(\pi t), & x_1 + x_2 \leq 1, \\
2\sin(\pi x_1)\sin(\pi x_2)\sin(\pi t) + 1, & x_1 + x_2 > 1, 
\end{cases} \\
u(x, t) = \max \left(0.5, \min \left(1.5, u_d(x, t) - p(x, t)\right)\right), \\
f(x, t) = y_t(x, t) - \text{div}(A(x)\nabla y(x, t)) - Bu(x, t), \\
y_d(x, t) = y(x, t) + p_t(x, t) + \text{div}(A^*(x)\nabla p(x, t)).
\]

We take a small time size \(\Delta t = 10^{-2}\) and solve this example by using the Projection gradient algorithm and the Adaptive algorithm.

### Table 2: Numerical results for Example 2 on uniform meshes.

<table>
<thead>
<tr>
<th></th>
<th>nodes ((u, y, p))</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(|u - u_h|)</td>
<td>1.12343E-1</td>
<td>7.09256E-2</td>
<td>6.39910E-2</td>
<td>4.99335E-2</td>
<td>3.51377E-2</td>
<td></td>
</tr>
<tr>
<td>(|\nabla y - \nabla y_h|)</td>
<td>2.45326E-1</td>
<td>1.23238E-1</td>
<td>6.17604E-2</td>
<td>3.10496E-2</td>
<td>1.58586E-2</td>
<td></td>
</tr>
<tr>
<td>(|\nabla p - \nabla p_h|)</td>
<td>2.48348E-1</td>
<td>1.30627E-1</td>
<td>7.60830E-2</td>
<td>5.44273E-2</td>
<td>4.74900E-2</td>
<td></td>
</tr>
<tr>
<td>(|R_{uh} - u_h|)</td>
<td>1.47331E-1</td>
<td>1.09594E-1</td>
<td>6.80918E-2</td>
<td>4.46566E-2</td>
<td>3.18298E-2</td>
<td></td>
</tr>
<tr>
<td>(|G_{yh} - \nabla y_h|)</td>
<td>2.49205E-1</td>
<td>1.24082E-1</td>
<td>6.17985E-2</td>
<td>3.08335E-2</td>
<td>1.54005E-2</td>
<td></td>
</tr>
<tr>
<td>(|G_{ph} - \nabla p_h|)</td>
<td>2.48576E-1</td>
<td>1.24027E-1</td>
<td>6.17976E-2</td>
<td>3.08386E-2</td>
<td>1.54043E-2</td>
<td></td>
</tr>
</tbody>
</table>

Numerical results based on a sequence of uniformly refined meshes and adaptive meshes are listed in Table 2 and Table 3, respectively. It is clear that the adaptive meshes generated via the error estimators \(\eta_i\) are able to save substantial computational work, in comparison with the uniform meshes. In Figure 2, it is easy to see that the mesh of \(u\) adapts...
Table 3: Numerical results for Example 2 on adaptive meshes.

<table>
<thead>
<tr>
<th></th>
<th>adaptive</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{nodes (u)}$</td>
<td>139</td>
<td>422</td>
<td>725</td>
<td>1179</td>
<td>1915</td>
<td></td>
</tr>
<tr>
<td>$\text{nodes (y,p)}$</td>
<td>139</td>
<td>513</td>
<td>1943</td>
<td>3283</td>
<td>4439</td>
<td></td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>\nabla y - \nabla y_h</td>
<td></td>
<td>$</td>
<td>1.74066E-1</td>
<td>8.75146E-2</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>\nabla p - \nabla p_h</td>
<td></td>
<td>$</td>
<td>1.79000E-1</td>
<td>9.79710E-2</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>R_h u_h - u_h</td>
<td></td>
<td>$</td>
<td>1.00369E-1</td>
<td>6.82909E-2</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>G_h y_h - \nabla y_h</td>
<td></td>
<td>$</td>
<td>1.82864E-1</td>
<td>8.87847E-2</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>G_h p_h - \nabla p_h</td>
<td></td>
<td>$</td>
<td>1.82770E-1</td>
<td>8.87683E-2</td>
</tr>
</tbody>
</table>

very well to the neighborhood of the discontinuous line $x_1 + x_2 = 1$, and a higher density of node points are indeed distributed along the line. In Figure 3, we plot the profile of the numerical solution $u_h$ at $t = 0.5$ on adaptive mesh when $\text{nodes} = 1915$.

![Figure 2: The adaptive mesh of $u$ when $\text{nodes} = 1915$.](image1)

![Figure 3: The numerical solution $u_h$ at $t = 0.5$ on adaptive mesh ($\text{nodes} = 1915$).](image2)
Acknowledgements This work is supported by Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2008), National Science Foundation of China (10971074), Specialized Research Fund for the Doctoral Program of Higher Education (20114407110009), and Hunan Provincial Innovation Foundation for Postgraduate (lx2009 B120). The authors would like to thank the referees for the helpful suggestions, which lead to improvements of the presentation.

References


32 D. Tiba, Lectures on The Optimal Control of Elliptic Equations, University of Jyvaskyla Press, Finland, 1995.