Numer. Math. Theor. Meth. Appl. Vol. **5**, No. 4, pp. 573-591
doi: 10.4208/nmtma.2012.m1117 November 2012

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

Yuelong Tang[1] and Yanping Chen[2,*]

[1] *Hunan Key Laboratory for Computation and Simulation in Science and Engineering, School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, Hunan, China.*
[2] *School of Mathematical Sciences, South China Normal University, Guangzhou 510631, Guangdong, China.*

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

Available online 22 August 2012

Abstract. This paper is concerned with recovery type a posteriori error estimates of fully discrete finite element approximation for general convex parabolic optimal control problems with pointwise control constraints. 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 the 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].

*Corresponding author. *Email addresses:* `tangyuelonga@163.com` (Y. Tang), `yanpingchen@scnu.edu.cn` (Y. Chen)

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:

$$\begin{cases} \min\limits_{u(x,t)\in K} \left\{ \displaystyle\int_0^T \big(g(y(x,t)) + h(u(x,t))\big)\, dt \right\}, \\ y_t(x,t) - \operatorname{div}(A(x)\nabla y(x,t)) = f(x,t) + Bu(x,t), \quad x \in \Omega,\ t \in J, \\ y(x,t) = 0, \quad x \in \partial\Omega,\ t \in J, \\ y(x,0) = y_0(x), \qquad x \in \Omega, \end{cases} \tag{1.1}$$

where Ω 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 $\xi \in \mathbb{R}^2$, $(A(x)\xi)\cdot\xi \ge c \mid \xi \mid^2$ 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 $\|u\|_{L^2(\Omega)} \to \infty$ and K is a nonempty closed convex set in $L^2(J; L^2(\Omega))$, defined by

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

where a and b are constants.

In this paper, we adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with norm $\|\cdot\|_{W^{m,q}(\Omega)}$ and seminorm $|\cdot|_{W^{m,q}(\Omega)}$. We set $H_0^1(\Omega) \equiv \left\{ v \in H^1(\Omega) : v|_{\partial\Omega} = 0 \right\}$ 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 $\|v\|_{L^s(J;W^{m,q}(\Omega))} = (\int_0^T \|v\|_{W^{m,q}(\Omega)}^s dt)^{1/s}$ for $s \in [1, \infty)$ and the standard modification for $s = \infty$. We can define the space $H^l(J; W^{m,q}(\Omega))$. The details can be found in [24]. In addition, c or C denotes a generic positive constant independent of h and Δt.

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. In Section

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.

2. A fully discrete finite element approximation of parabolic control problem

For ease of exposition, we set $V = L^2(J;W)$ with $W = H_0^1(\Omega)$ 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

$$\big(g'(y_1)-g'(y_2), y_1-y_2\big) \geq c\|y_1-y_2\|^2, \qquad \forall\, y_1,\, y_2 \in W. \tag{2.1}$$

(A2) There exists a constant $c>0$, such that

$$\big(h'(u_1)-h'(u_2), u_1-u_2\big) \geq c\|u_1-u_2\|^2, \qquad \forall\, u_1,\, u_2 \in U. \tag{2.2}$$

(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

$$\begin{aligned} a(v,w) &= \int_\Omega (A\nabla v)\cdot\nabla w dx, && \forall\, v,\, w\in W,\\ (f_1,f_2) &= \int_\Omega f_1\cdot f_2 dx, && \forall\, f_1,\, f_2\in U. \end{aligned}$$

It follows from the assumptions on A that

$$a(v,v)\geq c\|v\|_1^2, \qquad |a(v,w)|\leq C\|v\|_1\|w\|_1, \qquad \forall\, v,\, w\in W.$$

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

$$\begin{cases} \min\limits_{u\in K}\left\{\displaystyle\int_0^T \big(g(y)+h(u)\big)\,dt\right\},\\ (y_t,w)+a(y,w)=(f+Bu,w), \qquad \forall\, w\in W,\ t\in J,\\ y(x,0)=y_0(x), \qquad x\in\Omega. \end{cases} \tag{2.3}$$

It is well known (see, e.g., [23,27]) that the problem (2.3) has a unique solution (y,u), and the pair $(y,u)\in (H^1(J;L^2(\Omega))\cap V)\times K$ 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:

$$
\begin{aligned}
&(y_t, w) + a(y, w) = (f + Bu, w), && \forall\, w \in W,\ t \in J,\\
&y(x,0) = y_0(x), && x \in \Omega,
\end{aligned}
\tag{2.4}
$$

$$
\begin{aligned}
&-(p_t, q) + a(q, p) = (g'(y), q), && \forall\, q \in W,\ t \in J,\\
&p(x,T) = 0, && x \in \Omega,
\end{aligned}
\tag{2.5}
$$

$$
(h'(u) + B^* p, v - u) \ge 0, \qquad \forall\, v \in K, t \in J, \tag{2.6}
$$

where B^* is the adjoint operator of B.

Let $\mathscr{T}^h$ be regular triangulations of Ω and $\bar{\Omega} = \cup_{\tau\in\mathscr{T}^h}\bar{\tau}$. Let $h = \max_{\tau\in\mathscr{T}^h}\{h_\tau\}$, where h_τ denotes the diameter of the element τ. Further more, we set

$$
\begin{aligned}
U^h &= \left\{\ v_h \in L^2(\Omega)\ :\ v_h|_\tau = constant,\ \forall\, \tau \in \mathscr{T}^h\ \right\},\\
K^h &= \left\{\ v_h \in U^h\ :\ a \le v_h \le b\ \right\},\\
W^h &= \left\{\ v_h \in C(\bar{\Omega})\ :\ v_h|_\tau \in \mathbb{P}_1,\ \forall\, \tau \in \mathscr{T}^h, v_h|_{\partial\Omega} = 0\ \right\},
\end{aligned}
$$

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

$$
d_t\varphi^n = \frac{\varphi^n - \varphi^{n-1}}{\Delta t}, \qquad n = 1, \cdots, N.
$$

We define for $1 \le p < \infty$ the discrete time-dependent norms

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

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 $|||\cdot|||_{l^p(W^{m,q})}$ and let

$$
l^p_D(J; W^{m,q}(\Omega)) := \left\{\ \varphi : |||\varphi|||_{l^p(W^{m,q})} < \infty\ \right\}, \qquad 1 \le p \le \infty.
$$

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

$$
\begin{cases}
\min\limits_{u_h^n \in K^h} \left\{ \sum\limits_{n=1}^{N} \Delta t \left(g\left(y_h^n\right) + h\left(u_h^n\right)\right)\right\},\\
\left(d_t y_h^n, w_h\right) + a\left(y_h^n, w_h\right) = \left(f^n + Bu_h^n, w_h\right), \qquad \forall\, w_h \in W^h,\ n = 1, \cdots, N,\\
y_h^0(x) = y_0^h(x), \qquad x \in \Omega,
\end{cases}
\tag{2.7}
$$

where $y_0^h(x) \in W^h$ is an appropriate approximation of $y_0(x)$.

It follows (see e.g. [27]) that the control problem (2.7) has a unique solution (y_h^n, u_h^n), $n = 1, \cdots, N$, and $(y_h^n, u_h^n) \in W^h \times K^h$, $n = 1, \cdots, N$, is the solution of (2.7) if and only if there is a adjoint state $p_h^{n-1} \in W^h$, $n = 1, \cdots, N$, such that the triplet $(y_h^n, p_h^{n-1}, u_h^n) \in W^h \times W^h \times K^h$, $n = 1, \cdots, N$, satisfies the following optimality conditions:

$$\begin{aligned} &\left(d_t y_h^n, w_h\right) + a\left(y_h^n, w_h\right) = \left(f^n + Bu_h^n, w_h\right), && \forall\, w_h \in W^h,\ n = 1, \cdots, N, \\ &y_h^0(x) = y_0^h(x), && x \in \Omega, \end{aligned} \tag{2.8}$$

$$\begin{aligned} &-\left(d_t p_h^n, q_h\right) + a\left(q_h, p_h^{n-1}\right) = \left(g'\left(y_h^n\right), q_h\right), && \forall\, q_h \in W^h,\ n = N, \cdots, 1, \\ &p_h^N(x) = 0, && x \in \Omega, \end{aligned} \tag{2.9}$$

$$\left(h'\left(u_h^n\right) + B^* p_h^{n-1}, v - u_h^n\right) \geq 0, \qquad \forall\, v \in K^h,\ n = 1, \cdots, N. \tag{2.10}$$

Generally speaking, we select $y_0^h(x) = P_h\left(y_0(x)\right)$ 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:

$$\begin{aligned} &\left(y_t(v), w\right) + a\left(y(v), w\right) = \left(f + Bv, w\right), && \forall\, w \in W,\ t \in J, \\ &y(v)(x, 0) = y_0(x), && x \in \Omega, \end{aligned} \tag{3.1}$$

$$\begin{aligned} &-\left(p_t(v), q\right) + a\left(q, p(v)\right) = \left(g'(y(v)), q\right), && \forall\, q \in W,\ t \in J, \\ &p(v)(x, T) = 0, && x \in \Omega. \end{aligned} \tag{3.2}$$

For any $v \in K$, a pair $(y_h^n(v), p_h^{n-1}(v)) \in W^h \times W^h$, $n = 1, 2, \cdots, N$, satisfies the following system:

$$\begin{aligned} &\left(d_t y_h^n(v), w_h\right) + a\left(y_h^n(v), w_h\right) = \left(f^n + Bv^n, w_h\right), && \forall\, w_h \in W^h,\ n = 1, \cdots, N, \\ &y_h^0(v)(x) = y_0^h(x), && x \in \Omega, \end{aligned} \tag{3.3}$$

$$\begin{aligned} &-\left(d_t p_h^n(v), q_h\right) + a\left(q_h, p_h^{n-1}(v)\right) = \left(g'\left(y_h^n(v)\right), q_h\right), && \forall\, q_h \in W^h,\ n = N, \cdots, 1, \\ &p_h^N(v)(x) = 0, && x \in \Omega. \end{aligned} \tag{3.4}$$

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 $(y_h, p_h) = (y_h(u_h), p_h(u_h))$.

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

$$(\psi - Q_h\psi, v_h) = 0, \qquad \forall\, v_h \in U^h, \tag{3.5}$$

and the elliptic projection $P_h: W \to W^h$, which satisfies: for any $\phi \in W$

$$a(\phi - P_h\phi, w_h) = 0, \qquad \forall\, w_h \in W^h. \tag{3.6}$$

We have the following approximation properties:

$$\|\psi - Q_h\psi\|_{-s} \le Ch^{1+s}|\psi|_1, \qquad \forall\, \psi \in H^1(\Omega),\ s = 0, 1, \tag{3.7}$$

$$\|\phi - P_h\phi\| \le Ch^2\|\phi\|_{H^2(\Omega)}, \qquad \forall\, \phi \in H^2(\Omega). \tag{3.8}$$

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 $u \in l^2_D(J; H^1(\Omega))$ and the assumption (A1) is satisfied. Then*

$$|||y_h(Q_hu) - y_h(u)|||_{l^2(H^1)} + |||p_h(Q_hu) - p_h(u)|||_{l^2(H^1)} \le Ch^2. \tag{3.9}$$

Proof. Set $v = Q_hu$ and $v = u$ in (3.3), respectively. Then we obtain the following error equation:

$$\Big(d_t y_h^n(Q_hu) - d_t y_h^n(u), w_h\Big) + a\left(y_h^n(Q_hu) - y_h^n(u), w_h\right) = \left(B\left(Q_hu^n - u^n\right), w_h\right),$$
$$\forall\, w_h \in W^h,\ n = 1, \cdots, N. \tag{3.10}$$

From Cauchy's inequality, we have

$$\begin{aligned}
&\Big(d_t y_h^n(Q_hu) - d_t y_h^n(u), y_h^n(Q_hu) - y_h^n(u)\Big)\\
\ge&\frac{1}{2\Delta t}\left(\left\|y_h^n(Q_hu) - y_h^n(u)\right\|^2 - \left\|y_h^{n-1}(Q_hu) - y_h^{n-1}(u)\right\|^2\right),
\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}
&\Big(B\left(Q_hu^n - u^n\right), y_h^n(Q_hu) - y_h^n(u)\Big)\\
\le& C\left\|Q_hu^n - u^n\right\|_{-1}\left\|y_h^n(Q_hu) - y_h^n(u)\right\|_1 \le Ch^2\|u^n\|_1\left\|y_h^n(Q_hu) - y_h^n(u)\right\|_1\\
\le& C(\delta)h^4\|u^n\|_1^2 + \delta\left\|y_h^n(Q_hu) - y_h^n(u)\right\|_1^2.
\end{aligned} \tag{3.12}$$

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

$$\begin{aligned}
&\left\|y_h^N(Q_hu) - y_h^N(u)\right\|^2 + \sum_{n=1}^{N}\Delta t\left\|y_h^n(Q_hu) - y_h^n(u)\right\|_1^2\\
\le& C(\delta)h^4\sum_{n=1}^{N}\Delta t\|u^n\|_1^2 + \delta\sum_{n=1}^{N}\Delta t\left\|y_h^n(Q_hu) - y_h^n(u)\right\|_1^2.
\end{aligned} \tag{3.13}$$

Thus, we have

$$|||y_h(Q_hu) - y_h(u)|||_{l^2(H^1)} \le Ch^2|||u|||_{l^2(H^1)}. \tag{3.14}$$

Similarly, we obtain that

$$|||p_h(Q_hu) - p_h(u)|||_{l^2(H^1)} \le C|||y_h(Q_hu) - y_h(u)|||_{l^2(L^2)}. \tag{3.15}$$

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* $y(v),p(v)\in l_D^2(J;H^2(\Omega))\cap H^1(J;H^2(\Omega))\cap H^2(J;L^2(\Omega))$ *and the assumption (A1) is satisfied. Then we have*

$$|||P_hy(v)-y_h(v)|||_{l^2(H^1)}+|||P_hp(v)-p_h(v)|||_{l^2(H^1)}\le C\left(h^2+\Delta t\right). \tag{3.16}$$

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

$$\begin{aligned}&\left(y_t^n(v)-d_ty_h^n(v),w_h\right)+a\left(y^n(v)-y_h^n(v),w_h\right)=0,\\&\forall\, w_h\in W^h,\ n=1,\cdots,N.\end{aligned} \tag{3.17}$$

By using the definition of P_h, we get

$$\begin{aligned}&\left(d_tP_hy^n(v)-d_ty_h^n(v),w_h\right)+a\left(P_hy^n(v)-y_h^n(v),w_h\right)\\=&\left(d_tP_hy^n(v)-d_ty^n(v)+d_ty^n(v)-y_t^n(v),w_h\right).\end{aligned} \tag{3.18}$$

Note that

$$\begin{aligned}&\left(d_tP_hy^n(v)-d_ty^n(v),P_hy^n(v)-y_h^n(v)\right)\\\le&\left\|d_tP_hy^n(v)-d_ty^n(v)\right\|\left\|P_hy^n(v)-y_h^n(v)\right\|\\\le&Ch^2\left\|d_ty^n(v)\right\|_2\left\|P_hy^n(v)-y_h^n(v)\right\|\\\le&Ch^2(\Delta t)^{-1}\int_{t_{n-1}}^{t_n}\left\|y_t(v)\right\|_2dt\left\|P_hy^n(v)-y_h^n(v)\right\|\\\le&Ch^2(\Delta t)^{-\frac{1}{2}}||y_t(v)||_{L^2(t_{n-1},t_n;H^2(\Omega))}\left\|P_hy^n(v)-y_h^n(v)\right\|,\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}&\left(d_ty^n(v)-y_t^n(v),P_hy^n(v)-y_h^n(v)\right)\\\le&(\Delta t)^{-1}\left\|y^n(v)-y^{n-1}(v)-\Delta ty_t^n(v)\right\|\left\|P_hy^n(v)-y_h^n(v)\right\|\\=&(\Delta t)^{-1}\left\|\int_{t_{n-1}}^{t_n}(t_{n-1}-t)(y_{tt}(v))(t)dt\right\|\left\|P_hy^n(v)-y_h^n(v)\right\|\\\le&C(\Delta t)^{\frac{1}{2}}||y_{tt}(v)||_{L^2(t_{n-1},t_n;L^2(\Omega))}\left\|P_hy^n(v)-y_h^n(v)\right\|.\end{aligned} \tag{3.20}$$

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

$$\begin{aligned}&\left\|P_hy^N(v)-y_h^N(v)\right\|^2+c\sum_{n=1}^N\Delta t\left\|P_hy^n(v)-y_h^n(v)\right\|_1^2\\\le&C(\delta)\left(h^4||y_t(v)||^2_{L^2(J;H^2(\Omega))}+(\Delta t)^2||y_{tt}(v)||^2_{L^2(J;L^2(\Omega))}\right)\\&+\delta\sum_{n=1}^N\Delta t\left\|P_hy^n(v)-y_h^n(v)\right\|^2.\end{aligned} \tag{3.21}$$

Thus, we get

$$|||P_h y(v) - y_h(v)|||_{l^2(H^1)} \le C\left(h^2 + \Delta t\right). \tag{3.22}$$

Similarly, we derive that

$$|||P_h p(v) - p_h(v)|||_{l^2(H^1)} \le C\left(h^2 + \Delta t\right). \tag{3.23}$$

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^* $(0 \le t^* \le T)$, we divide Ω into the following subsets:

$$\begin{aligned}
&\Omega^+ = \{\ \cup\tau\ :\ \tau \subset \Omega,\ a < u(\cdot,t^*) < b\ \},\\
&\Omega^0 = \{\ \cup\tau\ :\ \tau \subset \Omega,\ u(\cdot,t^*)|_\tau = a \quad \text{or} \quad u(\cdot,t^*)|_\tau = b\ \},\\
&\Omega^- = \Omega \setminus \left(\Omega^+ \cup \Omega^0\right).
\end{aligned}$$

It is easy to see that the above three subsets are not intersected with each other and $\Omega = \bar{\Omega}^+ \cup \bar{\Omega}^0 \cup \bar{\Omega}^-$. We assume that u and $\mathscr{T}_h$ are regular such that $\operatorname{meas}\left(\Omega^-\right) \le Ch$, (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*

$$u,\ h'(u) + B^*p \in l_D^2(J; W^{1,\infty}(\Omega)).$$

Then, we have

$$|||Q_h u - u_h|||_{l^2(L^2)} \le C\left(h^{\frac{3}{2}} + \Delta t\right). \tag{4.1}$$

Proof. Letting $v = u_h$ in (2.6) and $v = Q_h u^n$ in (2.10), we derive

$$\begin{aligned}
0 \le& \left(h'(u^n) + B^*p^n, u_h^n - u^n\right) + \left(h'\left(u_h^n\right) + B^*p_h^{n-1}, Q_h u^n - u_h^n\right)\\
=& \left(h'\left(u_h^n\right) - h'(u^n) + B^*p_h^{n-1} - B^*p^n, Q_h u^n - u_h^n\right) + \left(h'(u^n) + B^*p^n, Q_h u^n - u^n\right).
\end{aligned} \tag{4.2}$$

According to the assumption (A2) and (4.2), we obtain

$$\begin{aligned}
& c\|Q_h u^n - u_h^n\|^2\\
\le& \left(h'\left(Q_h u^n\right) - h'\left(u_h^n\right), Q_h u^n - u_h^n\right)\\
=& \left(h'(u^n) - h'\left(u_h^n\right), Q_h u^n - u_h^n\right) + \left(h'\left(Q_h u^n\right) - h'(u^n), Q_h u^n - u_h^n\right)\\
\le& \left(h'(u^n) + B^*p^n, Q_h u^n - u^n\right) + \left(h'\left(Q_h u^n\right) - h'(u^n) + B^*p_h^{n-1} - B^*p^n, Q_h u^n - u_h^n\right).
\end{aligned} \tag{4.3}$$

From the assumption (A3), we have that there exists a constant $0 \leq \theta \leq 1$ such that

$$\begin{aligned}
&\left(h'\left(Q_h u^n\right) - h'\left(u^n\right), Q_h u^n - u_h^n\right)\\
=&\left(j''\left(u^n\right)\left(Q_h u^n - u^n\right) + \frac{1}{2} j'''\left(u^n + \theta\left(Q_h u^n - u^n\right)\right)\left(Q_h u^n - u^n\right)^2, Q_h u^n - u_h^n\right)\\
=&\left(\left(j''\left(u^n\right) - \pi^c\left(j''\left(u^n\right)\right)\right)\left(Q_h u^n - u^n\right), Q_h u^n - u_h^n\right)\\
&+\frac{1}{2}\left(j'''\left(u^n + \theta\left(Q_h u^n - u^n\right)\right)\left(Q_h u^n - u^n\right)^2, Q_h u^n - u_h^n\right)\\
\leq& Ch\left\|j''(\cdot)\right\|_{1,\infty}\left\|Q_h u^n - u^n\right\|\left\|Q_h u^n - u_h^n\right\|\\
&+\frac{C}{2}\left\|j'''(\cdot)\right\|_{0,\infty}\left\|Q_h u^n - u^n\right\|^2_{W^{0,4}(\Omega)}\left\|Q_h u^n - u_h^n\right\|\\
\leq& Ch^2\left\|u^n\right\|_1\left\|Q_h u^n - u_h^n\right\|.
\end{aligned} \tag{4.4}$$

It is clear that

$$\begin{aligned}
&\left(B^* p_h^{n-1} - B^* p^n, Q_h u^n - u_h^n\right)\\
=&\left(B^* p_h^{n-1}(u_h) - B^* p_h^{n-1}(Q_h u) + B^* p_h^{n-1}(Q_h u) - B^* p_h^{n-1}(u), Q_h u^n - u_h^n\right)\\
&+\left(B^* p_h^{n-1}(u) - B^* p^{n-1}(u) + B^* p^{n-1}(u) - B^* p^n(u), Q_h u^n - u_h^n\right).
\end{aligned} \tag{4.5}$$

By using (A1) and (3.3)-(3.4), we have

$$\begin{aligned}
&\left(B^* p_h^{n-1}(u_h) - B^* p_h^{n-1}(Q_h u), Q_h u^n - u_h^n\right)\\
=&-\left(g'\left(y_h^n(u_h)\right) - g'\left(y_h^n(Q_h u)\right), y_h^n(u_h) - y_h^n(Q_h u)\right)\\
\leq&-c\left\|y_h^n(u_h) - y_h^n(Q_h u)\right\|^2 \leq 0.
\end{aligned} \tag{4.6}$$

From (4.3)-(4.6), we obtain

$$\begin{aligned}
&|||Q_h u - u_h|||^2_{l^2(L^2)} = \sum_{n=1}^{N} \Delta t\left(Q_h u^n - u_h^n, Q_h u^n - u_h^n\right)\\
\leq&\sum_{n=1}^{N}\Delta t\left(h'\left(u^n\right) + B^* p^n, Q_h u^n - u^n\right) + \sum_{n=1}^{N}\Delta t\left(B^* p_h^{n-1}(Q_h u) - B^* p_h^{n-1}(u), Q_h u^n - u_h^n\right)\\
&+\sum_{n=1}^{N}\Delta t\left(B^* p_h^{n-1}(u) - B^* p^{n-1}(u), Q_h u^n - u_h^n\right) + \sum_{n=1}^{N}\Delta t\left(B^* p^{n-1}(u) - B^* p^n(u), Q_h u^n - u_h^n\right)\\
&+Ch^2\Delta t\sum_{n=1}^{N}\left\|u^n\right\|_1\left\|Q_h u^n - u_h^n\right\| := I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned} \tag{4.7}$$

For the first term, we have

$$\begin{aligned}
&\left(h'\left(u^n\right) + B^* p^n, Q_h u^n - u^n\right)\\
=&\int_{\Omega^+} + \int_{\Omega^0} + \int_{\Omega^-}\left(h'\left(u^n\right) + B^* p^n\right)\left(Q_h u^n - u^n\right) dx,
\end{aligned} \tag{4.8}$$

and $(Q_hu^n - u^n)|_{\Omega^0} = 0$. From (2.6), we have $h'(u^n) + B^*p^n = 0$ on Ω^+. Let π^c be the element average operator defined in [21]. Then

$$
\begin{aligned}
I_1 =&\Delta t \sum_{n=1}^{N} \int_{\Omega^-} \left(h'(u^n) + B^*p^n\right)\left(Q_hu^n - u^n\right) dx \\
=&\Delta t \sum_{n=1}^{N} \int_{\Omega^-} \left(h'(u^n) + B^*p^n - \pi^c(h'(u^n) + B^*p^n)\right)\left(Q_hu^n - u^n\right) dx \\
\le& Ch^2 \Delta t \sum_{n=1}^{N} \left\|h'(u^n) + B^*p^n\right\|_{1,\Omega^-} \|u^n\|_{1,\Omega^-} \\
\le& Ch^2 \Delta t \sum_{n=1}^{N} \left\|h'(u^n) + B^*p^n\right\|_{W^{1,\infty}(\Omega^-)} \|u^n\|_{W^{1,\infty}(\Omega^-)} \operatorname{meas}\left(\Omega^-\right) \\
\le& Ch^3 \left(|||h'(u^n) + B^*p^n|||^2_{l^2(w^{1,\infty})} + |||u|||^2_{l^2(w^{1,\infty})} \right). \qquad (4.9)
\end{aligned}
$$

From Young's inequality and Lemma 3.1, we derive

$$
\begin{aligned}
I_2 =& \sum_{n=1}^{N} \Delta t \left(B^* \left(p_h^{n-1}(Q_hu) - p_h^{n-1}(u) \right), Q_hu^n - u_h^n \right) \\
\le& C(\delta) \sum_{n=1}^{N} \Delta t \left\| p_h^{n-1}(Q_hu) - p_h^{n-1}(u) \right\|^2 + \delta \sum_{n=1}^{N} \Delta t \left\| Q_hu^n - u_h^n \right\|^2 \\
=& C(\delta)|||p_h(Q_hu) - p_h(u)|||^2_{l^2(L^2)} + \delta|||Q_hu - u_h|||^2_{l^2(L^2)} \\
\le& C(\delta)h^4 + \delta|||Q_hu - u_h|||^2_{l^2(L^2)}. \qquad (4.10)
\end{aligned}
$$

By using (3.8), Young's inequality and Lemma 3.2, we get

$$
\begin{aligned}
I_3 =& \sum_{n=1}^{N} \Delta t \left(B^* \left(p_h^{n-1}(u) - P_hp^{n-1}(u) + P_hp^{n-1}(u) - p^{n-1}(u) \right), Q_hu^n - u_h^n \right) \\
\le& C(\delta) \sum_{n=1}^{N} \Delta t \left\| p_h^{n-1}(u) - P_hp^{n-1}(u) \right\|^2 + C(\delta)h^4 \sum_{n=1}^{N} \Delta t \|p^{n-1}(u)\|_2^2 \\
&+ \delta \sum_{n=1}^{N} \Delta t \left\| Q_hu^n - u_h^n \right\|^2 \\
\le& C(\delta)|||p_h(u) - P_hp(u)|||^2_{l^2(L^2)} + C(\delta)h^4|||p(u)|||^2_{l^2(H^2)} + \delta|||Q_hu - u_h|||^2_{l^2(L^2)} \\
\le& C(\delta)\left(h^4 + (\Delta t)^2\right) + \delta|||Q_hu - u_h|||^2_{l^2(L^2)}. \qquad (4.11)
\end{aligned}
$$

Note that

$$\begin{aligned}I_4 =&\sum_{n=1}^{N}\Delta t\left(B^*\left(p^{n-1}(u)-p^n(u)\right),Q_hu^n-u_h^n\right)\\ \leq& C\sum_{n=1}^{N}\Delta t\int_{t_{n-1}}^{t_n}\left\|p_t(u)\right\|dt\left\|Q_hu^n-u_h^n\right\|\\ \leq& C(\delta)(\Delta t)^2\|p_t(u)\|^2_{L^2(J;L^2(\Omega))}+\delta|||Q_hu-u_h|||^2_{l^2(L^2)},\end{aligned} \tag{4.12}$$

and

$$\begin{aligned}I_5 =&Ch^2\Delta t\sum_{n=1}^{N}\|u^n\|_1\left\|Q_hu^n-u_h^n\right\|\\ \leq& C(\delta)h^4\sum_{n=1}^{N}\Delta t\,\|u^n\|_1^2+\delta\sum_{n=1}^{N}\Delta t\left\|Q_hu^n-u_h^n\right\|^2\\ \leq& C(\delta)h^4|||u|||^2_{l^2(L^2)}+\delta|||Q_hu-u_h|||^2_{l^2(L^2)}.\end{aligned} \tag{4.13}$$

By letting δ be small enough, (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*

$$|||P_hy-y_h|||_{l^2(H^1)}+|||P_hp-p_h|||_{l^2(H^1)}\leq C\left(h^{\frac{3}{2}}+\Delta t\right). \tag{4.14}$$

Proof. From (2.4) and (2.8), we have the following error equation:

$$\begin{aligned}&\left(y_t^n-d_ty_h^n,w_h\right)+a\left(y^n-y_h^n,w_h\right)=\left(B\left(u^n-u_h^n\right),w_h\right),\\ &\forall\, w_h\in W^h,\ n=1,\cdots,N.\end{aligned} \tag{4.15}$$

By choosing $w_h=P_hy^n-y_h^n$ and using the definition of P_h, we get

$$\begin{aligned}&\left(d_tP_hy^n-d_ty_h^n,P_hy^n-y_h^n\right)+a\left(P_hy^n-y_h^n,P_hy^n-y_h^n\right)\\ =&\left(d_tP_hy^n-d_ty^n+d_ty^n-y_t^n+B\left(Q_hu^n-u_h^n\right)+B\left(u^n-Q_hu^n\right),P_hy^n-y_h^n\right).\end{aligned} \tag{4.16}$$

Note that

$$\begin{aligned}&\left(B\left(u^n-Q_hu^n\right),P_hy^n-y_h^n\right)\leq C\left\|u^n-Q_hu^n\right\|_{-1}\left\|P_hy^n-y_h^n\right\|_1\\ \leq& Ch^2\|u^n\|_1\left\|P_hy^n-y_h^n\right\|_1\leq C(\delta)h^4\|u^n\|_1^2+\delta\left\|P_hy^n-y_h^n\right\|_1^2.\end{aligned} \tag{4.17}$$

Similar to Lemma 3.2, by using (4.1) and (4.16)-(4.17), we derive

$$|||P_hy-y_h|||_{l^2(H^1)}\leq C\left(h^{\frac{3}{2}}+\Delta t\right). \tag{4.18}$$

Similarly, we get that

$$|||P_hp-p_h|||_{l^2(H^1)}\leq C\left(h^{\frac{3}{2}}+\Delta t\right). \tag{4.19}$$

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 $G_h v = \left(R_h v_{x_1}, R_h v_{x_2}\right)$, 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 Ω is convex. Then*

$$|||R_h u_h - u|||_{l^2(L^2)} \le C\left(h^{\frac{3}{2}} + \Delta t\right). \tag{5.1}$$

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

$$\begin{aligned} \left\|R_h u_h^n - u^n\right\| &\le \left\|R_h u_h^n - R_h Q_h u^n\right\| + \left\|R_h Q_h u^n - R_h u^n\right\| + \left\|R_h u^n - u^n\right\| \\ &\le \left\|R_h u_h^n - R_h Q_h u^n\right\| + \left\|R_h Q_h u^n - R_h u^n\right\| + C h^{\frac{3}{2}}. \end{aligned} \tag{5.2}$$

By using the definition of R_h, we have

$$R_h u^n = R_h Q_h u^n, \tag{5.3}$$

$$\left\|R_h u_h^n - R_h Q_h u^n\right\| \le C\left\|u_h^n - Q_h u^n\right\|. \tag{5.4}$$

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

$$|||R_h u_h - u|||^2_{l^2(L^2)} \le C h^3 + C|||Q_h u - u_h|||^2_{l^2(L^2)} \le C\left(h^3 + (\Delta t)^2\right). \tag{5.5}$$

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 $y, p \in l_D^2(J; H^3(\Omega))$. Then, we have*

$$|||G_h y_h - \nabla y|||_{l^2(L^2)} + |||G_h p_h - \nabla p|||_{l^2(L^2)} \le C\left(h^{\frac{3}{2}} + \Delta t\right). \tag{5.6}$$

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

$$\|P_h y - y_I\|_1 \le C h^2 \|y\|_3. \tag{5.7}$$

From the standard interpolation error estimate technique (see, e.g., [15]) that

$$\|G_h y_I - \nabla y\| \le C h^2 |y|_3. \tag{5.8}$$

By using (5.7)-(5.8), we get

$$\begin{aligned}\left\|G_h y_h^n-\nabla y^n\right\| =&\left\|G_h y_h^n-G_h P_h y^n\right\|+\left\|G_h P_h y^n-G_h y_I^n\right\|+\left\|G_h y_I^n-\nabla y^n\right\| \\ \leq& C\left\|y_h^n-P_h y^n\right\|_1+C\left\|P_h y^n-y_I^n\right\|_1+\left\|G_h y_I^n-\nabla y^n\right\| \\ \leq& C\left\|y_h^n-P_h y^n\right\|_1+C h^2\|y^n\|_3. \end{aligned} \tag{5.9}$$

Therefore,

$$\sum_{n=1}^{N}\Delta t\left\|G_h y_h^n-\nabla y^n\right\|^2 \leq C\sum_{n=1}^{N}\Delta t\left\|y_h^n-P_h y^n\right\|_1^2+Ch^4\sum_{n=1}^{N}\Delta t\left\|y^n\right\|_3^2. \tag{5.10}$$

From Lemma 4.2 and (5.10), we derive

$$|||G_h y_h-\nabla y|||_{l^2(L^2)} \leq C\left(h^{\frac{3}{2}}+\Delta t\right). \tag{5.11}$$

Similarly, we can prove that

$$|||G_h p_h-\nabla p|||_{l^2(L^2)} \leq C\left(h^{\frac{3}{2}}+\Delta t\right). \tag{5.12}$$

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 Theorems* 5.1 *and* 5.2 *are valid. Then*

$$\eta_1 := |||R_h u_h-u_h|||_{l^2(L^2)}=|||u-u_h|||_{l^2(L^2)}+\mathscr{O}\left(h^{\frac{3}{2}}+\Delta t\right), \tag{5.13}$$

$$\eta_2 := |||G_h y_h-\nabla y_h|||_{l^2(L^2)}=|||\nabla\left(y-y_h\right)|||_{l^2(L^2)}+\mathscr{O}\left(h^{\frac{3}{2}}+\Delta t\right), \tag{5.14}$$

$$\eta_3 := |||G_h p_h-\nabla p_h|||_{l^2(L^2)}=|||\nabla\left(p-p_h\right)|||_{l^2(L^2)}+\mathscr{O}\left(h^{\frac{3}{2}}+\Delta t\right). \tag{5.15}$$

Proof. From Theorems 5.1 and 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 $(n=0,1,2,\cdots)$:

$$\left\{\begin{array}{l} b(u_{n+\frac{1}{2}},v)=b(u_n,v)-\rho_n\left(J'(u_n),v\right), \quad \forall\, v\in U^h, \\ u_{n+1}=P_K^b(u_{n+\frac{1}{2}}), \end{array}\right. \tag{6.1}$$

where $b(\cdot,\cdot)=\int_0^T(\cdot,\cdot)$ is a symmetric and positive definite bilinear form, ρ_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:

Algorithm 6.1. Projection gradient algorithm

Step 1. Solve the following equations:

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

where we have omitted the subscript h;
Step 2. Calculate the iterative error: $E_{n+1}=|||u_{n+1}-u_n|||_{l^2(L^2)}$;
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 η_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:

Algorithm 6.2. 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 η_i;
Step 2. Adjust the meshes by using the estimators η_i and update the numerical solution u_n' and obtain u_{n+1}' on new meshes;
Step 3. Calculate the iterative error: $E_{n+1}'=|||u_{n+1}'-u_n'|||_{l^2(L^2)}$;
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/~rli/AFEPack/. Just for simplicity, we let I be the 2×2 identity matrix and denote $|||\cdot|||_{l^2(L^2)}$ 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

Table 1: The error of the control variable, Example 1.

h	Δt	$\|\|\|u-u_h\|\|\|$	$\|\|\|Q_hu-u_h\|\|\|$	$\|\|\|u-R_hu_h\|\|\|$
1.0E-1	1/10	7.22E-1	3.61E-2	5.48E-2
5.0E-2	1/30	3.67E-2	8.82E-3	1.78E-2
2.5E-2	1/90	1.80E-2	3.01E-3	5.74E-3
1.25E-2	1/270	9.01E-3	9.00E-4	1.93E-3

parabolic optimal control problems:

$$\begin{cases} \min\limits_{u\in K}\left\{\dfrac{1}{2}\displaystyle\int_0^T \left(\left\|y(x,t)-y_d(x,t)\right\|^2+\left\|u(x,t)-u_d(x,t)\right\|^2\right)dt\right\}, \\ y_t(x,t)-\mathrm{div}(A(x)\nabla y(x,t))=f(x,t)+Bu(x,t), \quad x\in\Omega, \ \ t\in J, \\ y(x,t)=0, \quad x\in\partial\Omega, \ \ t\in J, \\ y(x,0)=y_0(x), \qquad x\in\Omega. \end{cases}$$

Moreover, we assume that

$$K=\Big\{\ v(x,t)\in L^2(J;L^2(\Omega)) : a\le v(x,t)\le b, \quad (x,t)\in\Omega\times J\ \Big\}.$$

Example 1. The data are as follows:

$$\begin{aligned} &a=-0.4, \quad b=0.4,\\ &A(x)=\begin{pmatrix} \sin(\pi x_1/2) & 0\\ 0 & \sin(\pi x_2/2)\end{pmatrix}\\ &p(x,t)=\sin(2\pi x_1)\sin(2\pi x_2)(1-t),\\ &y(x,t)=\sin(2\pi x_1)\sin(2\pi x_2)t,\\ &u_d(x,t)=2\sin(2\pi x_1)\sin(2\pi x_2)t,\\ &u(x,t)=\max\Big(-0.4,\min\big(0.4,u_d(x,t)-p(x,t)\big)\Big),\\ &f(x,t)=y_t(x,t)-\mathrm{div}\big(A(x)\nabla y(x,t)\big)-Bu(x,t),\\ &y_d(x,t)=y(x,t)+p_t(x,t)+\mathrm{div}\big(A^*(x)\nabla p(x,t)\big). \end{aligned}$$

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

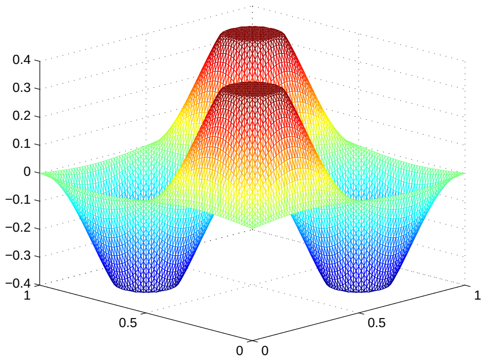

Figure 1: The numerical solution u_h at $t=0.5$ when $h=1.25E-2$ and $\Delta t=1/270$ for Example 1.

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

uniform	1	2	3	4	5
nodes (u,y,p)	121	441	1681	6561	25921
$\|\|\|u-u_h\|\|\|$	1.12E-1	7.09E-2	6.39E-2	4.99E-2	3.51E-2
$\|\|\|\nabla y-\nabla y_h\|\|\|$	2.45E-1	1.23E-1	6.17E-2	3.10E-2	1.58E-2
$\|\|\|\nabla p-\nabla p_h\|\|\|$	2.48E-1	1.30E-1	7.60E-2	5.44E-2	4.74E-2
$\|\|\|R_hu_h-u_h\|\|\|$	1.47E-1	1.09E-1	6.80E-2	4.46E-2	3.18E-2
$\|\|\|G_hy_h-\nabla y_h\|\|\|$	2.49E-1	1.24E-1	6.17E-2	3.08E-2	1.54E-2
$\|\|\|G_hp_h-\nabla p_h\|\|\|$	2.48E-1	1.24E-1	6.17E-2	3.08E-2	1.54E-2

Example 2. The data are as follows:

$$
\begin{aligned}
&A(x)=I, \quad a=0.5, \quad b=1.5,\\
&p(x,t)=\sin(\pi x_1)\sin(\pi x_2)\sin(\pi t),\\
&y(x,t)=\sin(\pi x_1)\sin(\pi x_2)\sin(\pi t),\\
&u_d(x,t)=\begin{cases} 2\sin(\pi x_1)\sin(\pi x_2)\sin(\pi t), & x_1+x_2\le 1,\\ 2\sin(\pi x_1)\sin(\pi x_2)\sin(\pi t)+1, & x_1+x_2>1,\end{cases}\\
&u(x,t)=\max\Big(0.5,\min\big(1.5,u_d(x,t)-p(x,t)\big)\Big),\\
&f(x,t)=y_t(x,t)-\mathrm{div}\big(A(x)\nabla y(x,t)\big)-Bu(x,t),\\
&y_d(x,t)=y(x,t)+p_t(x,t)+\mathrm{div}\big(A^*(x)\nabla p(x,t)\big).
\end{aligned}
$$

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

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

Table 3: Numerical results for Example 2 on adaptive meshes.

adaptive	1	2	3	4	5
nodes (u)	139	422	725	1179	1915
nodes (y,p)	139	513	1943	3283	4439
$\|\|\|u-u_h\|\|\|$	7.25E-2	5.79E-2	4.31E-2	3.02E-2	2.24E-2
$\|\|\|\nabla y-\nabla y_h\|\|\|$	1.74E-1	8.75E-2	4.42E-2	3.80E-2	3.67E-2
$\|\|\|\nabla p-\nabla p_h\|\|\|$	1.79E-1	9.79E-2	6.28E-2	5.86E-2	5.78E-2
$\|\|\|R_hu_h-u_h\|\|\|$	1.00E-1	6.82E-2	4.71E-2	3.50E-2	2.59E-2
$\|\|\|G_hy_h-\nabla y_h\|\|\|$	1.82E-1	8.87E-2	4.44E-2	3.79E-2	3.66E-2
$\|\|\|G_hp_h-\nabla p_h\|\|\|$	1.82E-1	8.87E-2	4.44E-2	3.80E-2	3.66E-2

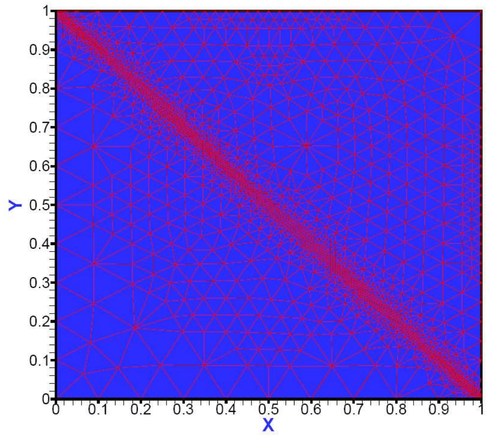

Figure 2: The adaptive mesh of u when $nodes=1915$ for Example 2.

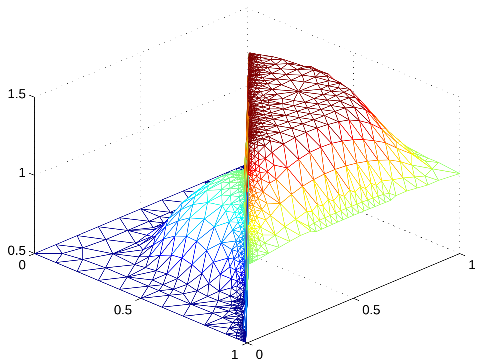

Figure 3: The numerical solution u_h at $t=0.5$ on adaptive mesh $(nodes=1915)$ for Example 2.

generated via the error estimators η_i are able to save substantial computational work, in comparison with the uniform meshes. In Fig. 2, it is easy to see that the mesh of u adapts 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 Fig. 3, we plot the profile of the numerical solution u_h at $t = 0.5$ on adaptive mesh when $nodes = 1915$.

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 Provinical 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

[1] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley Interscience, New York, 2000.

[2] I. Babuška, T. Strouboulis, C.S. Upadhyay and S.K. Gangaraj, *A posteriori estimation and adaptive control of the pollution error in the h-version of the finite element method*, Int. J. Numer. Meth. Eng., 38(24) (1995), pp. 4207–4235.

[3] 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(1) (2000), pp. 113–132.

[4] R. Becker, D. Meidner and B. Vexler, *Efficient numerical solution of parabolic optimization problems by finite element methods*, Optim. Meth. Soft., 22(5) (2007), pp. 813–833.

[5] J. Brandts, *Superconvergence and a posteriori error estimation for triangular mixed finite elements*, Numer. Math., 68 (1994), pp. 311–324.

[6] C. Carstensen and R. Verfürth, *Edge residuals donminate a posteriori error estimates for low order finite element methods*, SIAM J. Numer. Anal., 36(5) (1999), pp. 1571–1587.

[7] C. Chen and Y. Huang, *High Accuracy Theory of Finite Element Methods*, Hunan Science and Technology Press, Hunan, 1995.

[8] Y. Chen, *Superconvergence of mixed finite element methods for optimal control problems*, Math. Comput., 77 (2008), pp. 1269–1291.

[9] Y. Chen, *Superconvergence of quadratic optimal control problems by triangular mixed finite elements*, Int. J. Numer. Meth. Eng., 75(8) (2008), pp. 881–898.

[10] Y. Chen and Y. Dai, *Superconvergence for optimal control problems governed by semi-linear elliptic equations*, J. Sci. Comput., 39 (2009), pp. 206–221.

[11] Y. Chen, Y. Fu, H. Liu, Y. Dai and H. Wei, *Recovery a posteriori error estimates for general convex elliptic optimal control problems subject to pointwise control constraints*, J. Comput. Math., 27(4) (2009), pp. 543–560.

[12] Y. Chen, Y. Huang, W. Liu and N. Yan, *Error estimates and superconvergence of mixed finite element methods for convex optimal control problems*, J. Sci. Comput., 42 (2010), pp. 382–403.

[13] Y. Chen and W. Liu, *Error estimates and superconvergence of mixed finite element for quadratic optimal control*, Int. J. Numer. Anal. Model., 3 (2006), pp. 311–321.

[14] Y. Chen and Y. Tang, *Superconvergence analysis of fully discrete finite element methods for semilinear parabolic optimal control problems*, Front. Math. China, appear.

[15] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amstterdam, 1978.

[16] H. FU AND H. RUI, *A priori error estimates for optimal control problems governed by transient advection-diffusion equationg*, J. Sci. Comput., 38 (2009), pp. 290–315.

[17] M. HUANG, *Numerical Methods for Evolution Equations*, Science Press, Beijing, 2004.

[18] Y. HUANG AND J. XU, *Superconvergence of quadratic finite elements on mildly structured grids*, Math. Comp., 77 (2008), pp. 1253–1268.

[19] G. KNOWLES, *Finite element approxiamtion of parabolic time optimal control problems*, SIAM J. Control Optim., 20(3) (1982), pp. 414–427.

[20] A. KUFNER, O. JOHN AND S. FUCK, *Function Spaces*, Nordhoff, Leiden, The Netherlands, 1997.

[21] R. LI, W. LIU AND N. YAN, *A posteriori error estimates of recovery type for distributed convex optimal control problems,* J. Sci. Comput., 33 (2007), pp. 155–182.

[22] Q. LIN AND Q. ZHU, *The Preprocessing and Postprocessing for the Finite Element Method*, Shanghai Scientific and Technical Publishers, Shanghai, 1994.

[23] J. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.

[24] J. LIONS AND E. MAGENES, *Non Homogeneous Boundary Value Problems and Applications*, Springer-verlag, Berlin, 1972.

[25] H. LIU AND N. YAN, *Recovery type superconvergence and a posteriori error estimates for control problems governed by Stokes equations*, J. Comput. Appl. Math., 209 (2007), pp. 187–207.

[26] W. LIU AND N. YAN, *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math. 15 (2001) 285-309.

[27] W. LIU AND N. YAN, *A posteriori error estimates for optimal control problems governed by parabolic equations*, Numer. Math., 93 (2003), pp. 497–521.

[28] W. LIU AND N. YAN, *Adaptive Finite Element Methods for Optimal Control Governed by PDEs*, Science Press, Beijing, 2008.

[29] D. MEIDNER AND B. VEXLER, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems Part I: problems without control constraints*, SIAM J. Control Optim., 47(3) (2008), pp. 1150–1177.

[30] D. MEIDNER AND B. VEXLER, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems Part II: problems with control constraints*, SIAM J. Control Optim., 47(3) (2008), pp. 1301–1329.

[31] C. MEYER AND A. RÖSCH, *Superconvergence properties of optimal control problems*, SIAM J. Control Optim., 43(3) (2004), pp. 970–985.

[32] P. NEITTAANMAKI AND D. TIBA, *Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications*, Dekker, New York, 1994.

[33] D. TIBA, *Lectures on The Optimal Control of Elliptic Equations*, University of Jyvaskyla Press, Finland, 1995.

[34] C. XIONG AND Y. LI, *A posteriori error estimates for optimal distributed control governed by the evolution equations*, Appl. Numer. Math., 61 (2011), pp. 181–200.

[35] D. YANG, Y. CHANG AND W. LIU, *A priori error estimate and superconvergence annalysis for an opitmal control problem of bilinear type*, J. Comput. Math., 26(4) (2008), pp. 471–487.

[36] J. ZHOU, Y. CHEN AND Y. DAI, *Superconvergence of triangular mixed finite elements for optimal control problems with an integral constraint*, Appl. Math. Comput., 217 (2010), pp. 2057–2066.

[37] O. C. ZIENKIWICZ AND J.Z. ZHU, *The superconvergence patch recovery and a posteriori error estimates*, Int. J. Numer. Meth. Eng., 33 (1992), pp. 1331–1382.

[38] O. C. ZIENKIWICZ AND J.Z. ZHU, *The superconvergence patch recovery (SPR) and adaptive finite element refinement*, Comput. Meth. Appl. Math., 101 (1992), pp. 207–224.