RESEARCH　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Error analysis of variational discretization solving temperature control problems

Yuelong Tang[*]

[*]Correspondence:
tangyuelonga@163.com
Department of Mathematics and
Computational Science, Hunan
University of Science and
Engineering, Yongzhou, Hunan
425100, China

Abstract

In this paper, we consider variational discretization solving temperature control problems with pointwise control constraints, where the state and the adjoint state are approximated by piecewise linear finite element functions, while the control is not directly discretized. We derive *a priori* error estimates of second-order for the control, the state and the adjoint state. Moreover, we obtain *a posteriori* error estimates. Finally, we present some numerical algorithms for the control problem and do some numerical experiments to illustrate our theoretical results.

Keywords: variational discretization; finite element; optimal control problems; *a priori* error estimates; *a posteriori* error estimates

1 Introduction

We are interested in a material plate defined in a two-dimensional convex domain Ω with a Lipschitz boundary $\partial\Omega$. For the state y of the material, we choose the temperature distribution which is maintained equal to zero along the boundary. We denote thermal radiation or positive temperature feedback due to chemical reactions by the term $\phi(y)$ (see, *e.g.*, [1]) and assume that there exists a source $f \in L^2(\Omega)$. This system is governed by the following equation:

$$\begin{cases} -\operatorname{div}(A(x)\nabla y(x)) + \phi(y(x)) = f(x), & x \in \Omega, \\ y(x) = 0, & x \in \partial\Omega. \end{cases}$$

The setting above suggests that we may control the temperature distribution y to come close to a given target by acting with an additional distributed source term u, namely the control function. The corresponding optimal control problem is formulated as follows:

$$\begin{cases} \min_{u \in K}\{\int_{\Omega}(g(y) + h(u))\,dx\}, \\ -\operatorname{div}(A(x)\nabla y(x)) + \phi(y(x)) = f(x) + Bu(x), & x \in \Omega, \\ y(x) = 0, & x \in \partial\Omega, \end{cases} \tag{1.1}$$

where $g(\cdot)$ and $h(\cdot)$ are strictly convex continuous differentiable functions, $h(u) \to +\infty$ as $\|u\|_{L^2(\Omega)} \to \infty$, $A(x) = (a_{ij}(x))_{2\times 2} \in (W^{1,\infty}(\bar{\Omega}))^{2\times 2}$ such that $(A(x)\xi)\cdot\xi \geq c\,|\,\xi\,|^2$, $\forall\xi \in \mathbb{R}^2$, $f(x) \in L^2(\Omega)$, B is a linear continuous operator, and K is defined by

$$K = \{v(x) \in L^2(\Omega) : a \leq v(x) \leq b, \text{ a.e. } x \in \Omega\},$$

where a and b are two constants.

Optimal control problems have been extensively used in many aspects of the modern life such as social, economic, scientific and engineering numerical simulation [2]. Finite element approximation seems to be the most widely used method in computing optimal control problems. A systematic introduction of finite element method for PDEs and optimal control problems can be found in [1–8]. Concerning elliptic optimal control problems, *a priori* error estimates were investigated in [9, 10], *a posteriori* error estimates based on recovery techniques have been obtained in [11–13], *a posteriori* error estimates of residual type have been derived in [14–22], some error estimates and superconvergence results have been established in [23–27], and some adaptive finite element methods can be found in [28–31]. For parabolic optimal control problems, *a priori* error estimates are established in [32–34], *a posteriori* error estimates of residual type are investigated in [35, 36]. Recently, error estimates of spectral method for optimal control problems have been derived in [37, 38], and numerical methods for constrained elliptic control problems with rapidly oscillating coefficients are studied in [39].

For a constrained optimal control problem, the control has lower regularity than the state and the adjoint state. So most researchers considered using piecewise linear finite element functions to approximate the state and the adjoint state and using piecewise constant functions to approximate the control. They constructed a projection gradient algorithm where the *a priori* error estimates of the control is first-order in [11, 12]. Recently, Borzì considered a second-order discretization and multigrid solution of elliptic nonlinear constrained control problems in [40], Hinze introduced a variational discretization concept for optimal control problems and derived *a priori* error estimates for the control which is second-order in [41, 42]. The purpose of this paper is to consider variational discretization for convex temperature control problems governed by nonlinear elliptic equations with pointwise control constraints.

In this paper, we adopt the standard notation $W^{m,q}(\Omega)$ for Sobolev spaces on Ω with the norm $\| \cdot \|_{W^{m,q}(\Omega)}$ and seminorm $| \cdot |_{W^{m,q}(\Omega)}$. We set $H_0^1(\Omega) \equiv \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$ and denote $W^{m,2}(\Omega)$ by $H^m(\Omega)$. In addition, c or C denotes a generic positive constant.

The paper is organized as follows: In Section 2, we introduce a variational discretization approximation scheme for the model problem. In Section 3, we derive *a priori* error estimates. In Section 4, we derive sharp *a posteriori* error estimates of residual type. We present some numerical algorithms and do some numerical experiments to verify our theoretical results in the last section.

2 Variational discretization approximation for the model problem

We now consider a variational discretization approximation for the model problem (1.1). For ease of exposition, we set $V = H_0^1(\Omega)$, $U = L^2(\Omega)$, $\| \cdot \| = \| \cdot \|_{L^2(\Omega)}$, $\| \cdot \|_{1,\Omega} = \| \cdot \|_{H^1(\Omega)}$, $\| \cdot \|_{2,\Omega} = \| \cdot \|_{H^2(\Omega)}$ and

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

$$(u, w) = \int_\Omega u \cdot w, \quad \forall u, w \in U.$$

It follows from the assumptions on $A(x)$ that

$$a(y, y) \geq c\|y\|_{1,\Omega}^2, \qquad \big|a(y, w)\big| \leq C\|y\|_{1,\Omega}\|w\|_{1,\Omega}, \quad \forall y, w \in V. \tag{2.1}$$

Then the standard weak formula for the state equation is

$$a(y, w) + (\phi(y), w) = (f + Bu, w), \quad \forall w \in V, \tag{2.2}$$

where we assume that the function $\phi(\cdot) \in W^{1,\infty}(-R, R)$ for any $R > 0$, $\phi'(\cdot) \geq 0$ and $\phi'(y) \in L^2(\Omega)$ for any $y \in H^1(\Omega)$. Thus, the equation above has a unique solution.

Throughout the paper, we impose the following assumptions:

(A1) $g'(\cdot)$ and $h'(\cdot)$ are Lipschitz continuous, namely,

$$\left| g'(y_1) - g'(y_2) \right| \leq C|y_1 - y_2|, \quad \forall y_1, y_2 \in L^2(\Omega),$$

$$\left| h'(u(x_1)) - h'(u(x_2)) \right| \leq C|x_1 - x_2|, \quad \forall u \in K, x_1, x_2 \in \bar{\Omega}.$$

(A2) There exists a positive constant m such that

$$h''(u) \geq m, \quad \forall u \in K.$$

Then the model problem (1.1) can be restated as

$$\begin{cases} \min_{u \in K} \{ \int_\Omega (g(y) + h(u)) \, dx \}, \\ a(y, w) + (\phi(y), w) = (f + Bu, w), \quad \forall w \in V. \end{cases} \tag{2.3}$$

It is well known (see, *e.g.*, [2]) that the control problem (2.3) has a solution $(y, u) \in V \times K$, and that if the pair $(y, u) \in V \times K$ is the solution of (2.3), then there is an adjoint state $p \in V$ such that the triplet $(y, p, u) \in V \times V \times K$ satisfies the following optimality conditions:

$$a(y, w) + (\phi(y), w) = (f + Bu, w), \quad \forall w \in V, \tag{2.4}$$

$$a(q, p) + (\phi'(y)p, q) = (g'(y), q), \quad \forall q \in V, \tag{2.5}$$

$$(h'(u) + B^*p, v - u) \geq 0, \quad \forall v \in K, \tag{2.6}$$

where B^* is the adjoint operator of B.

Lemma 2.1 *Suppose that assumptions* (A1)-(A2) *are satisfied. Let* $p \in V$ *be the solution of* (2.4)-(2.6). *Then the following equation*:

$$h'(s(x)) + B^*p(x) = 0, \tag{2.7}$$

admits a unique solution $s(x)$ *and* $s(x) \in C^{0,1}(\bar{\Omega})$.

Proof It follows from $h''(u) \geq m > 0$ that (2.7) has a unique solution. Note that $g'(y) \in L^2(\Omega)$. From the regularity theory of patrial differential equations (see, *e.g.*, [7]), we have

$$p(x) \in H_0^1(\Omega) \cap W^{2,2}(\Omega).$$

Because Ω is a two-dimension convex domain, according to embedding theorem, we get

$$p(x) \in C^{0,1}(\bar{\Omega}).$$

From (A2) and (2.7), we get

$$m\big|s(x) - s(x_0)\big|$$

$$\leq \left| \int_0^1 h''\big(\theta s(x) + (1-\theta)s(x_0)\big)\, d\theta \cdot \big(s(x) - s(x_0)\big) \right|$$

$$= \big|h'\big(s(x)\big) - h'\big(s(x_0)\big)\big|$$

$$= \big|-B^*p(x) + B^*p(x_0)\big|$$

$$\leq C|x - x_0|.$$

Consequently, we complete the proof of equation (2.7). □

We introduce the following pointwise projection operator:

$$\Pi_{[a,b]}\big(g(x)\big) = \max\big(a, \min\big(b, g(x)\big)\big). \tag{2.8}$$

It is clear that $\Pi_{[a,b]}(\cdot)$ is Lipschitz continuous with constant 1. As in [9], it is easy to prove the following lemma.

Lemma 2.2 *Let* (y, p, u) *and* $s(x)$ *be the solutions of* (2.4)-(2.6) *and* (2.7), *respectively. Assume that assumptions* (A1)-(A2) *are satisfied. Then*

$$u(x) = \Pi_{[a,b]}\big(s(x)\big). \tag{2.9}$$

Remark 2.1 We should point out that (2.6) and (2.9) are equivalent. This theory can be used to more complex situation, for example, K is characterized by a bound on the integral on u over Ω, namely, $\int_\Omega u(x)\, dx \geq 0$, we have similar results.

Let $\mathcal{T}^h$ be a regular triangulation of Ω, such that $\bar{\Omega} = \bigcup_{\tau \in \mathcal{T}^h} \bar{\tau}$. Let $h = \max_{\tau \in \mathcal{T}^h}\{h_\tau\}$, where h_τ denotes the diameter of the element τ. Associated with $\mathcal{T}^h$ is a finite dimensional subspace S^h of $C(\bar{\Omega})$, such that $\chi|_\tau$ are polynomials of m-order ($m \geq 1$) for all $\chi \in S^h$ and $\tau \in \mathcal{T}^h$. Let $V^h = \{v_h \in S^h : v_h|_{\partial\Omega} = 0\}$. It is easy to see that $V^h \subset V$.

Then a possible variational discretization approximation scheme of (2.1) is as follows:

$$\begin{cases} \min_{u_h \in K}\{\int_\Omega (g(y_h) + h(u_h))\, dx\}, \\ a(y_h, w_h) + (\phi(y_h), w_h) = (f + Bu_h, w_h), \quad \forall w_h \in V^h. \end{cases} \tag{2.10}$$

It is well known (see, *e.g.,* [35]) that control problem (2.10) has a solution $(y_h, u_h) \in V^h \times K$, and that if the pair $(y_h, u_h) \in V^h \times K$ is the solution of (2.10), then there is an adjoint state $p_h \in V^h$ such that the triplet $(y_h, p_h, u_h) \in V^h \times V^h \times K$ satisfies the following optimality conditions:

$$a(y_h, w_h) + \big(\phi(y_h), w_h\big) = (f + Bu_h, w_h), \quad \forall w_h \in V^h, \tag{2.11}$$

$$a(q_h, p_h) + \big(\phi'(y_h)p_h, q_h\big) = \big(g'(y_h), q_h\big), \quad \forall q_h \in V^h, \tag{2.12}$$

$$\big(h'(u_h) + B^*p_h, v - u_h\big) \geq 0, \quad \forall v \in K. \tag{2.13}$$

Similar to Lemma 2.2, it is easy to show the following lemma.

Lemma 2.3 *Suppose that assumptions* (A1)-(A2) *are satisfied. Let* (y_h, p_h, u_h) *be the solution of* (2.11)-(2.13), *and* $s_h(x)$ *is the solution of the following equation*:

$$h'\big(s_h(x)\big) + B^* p_h(x) = 0. \tag{2.14}$$

Then we have

$$u_h(x) = \Pi_{[a,b]}\big(s_h(x)\big). \tag{2.15}$$

Remark 2.2 In many applications, the objective functional is uniform convex near the solution u, which is assumed in many studies on numerical methods of the problem, see, for example, [20, 43]. In this paper, we assumed that $g(\cdot)$ and $h(\cdot)$ are strictly convex continuous differentiable functions, for instance, $h(u) = \frac{\alpha}{2}\int_\Omega u^2$, which is frequently met, then the exact solution of the variational inequality (2.13) is $u_h(x) = \max(a, \min(b, -\frac{1}{\alpha}B^* p_h(x)))$, and for numerically solving the problem, we can replace $u_h(x)$ by $\max(a, \min(b, -\frac{1}{\alpha}B^* p_h(x)))$ in our program.

3 *A priori* error estimates

We now derive *a priori* error estimates of the variational discretization approximation scheme. Just for ease of exposition, let

$$J(u) = \int_\Omega \big(g(y) + h(u)\big)\, dx,$$

and $J'(u)$ is the Fréchet derivative of $J(u)$ at u. Similarly to (2.4)-(2.6), we can prove that

$$\big(J'(u), v\big) = \big(h'(u) + B^* p, v\big), \quad \forall v \in K,$$

$$\big(J'(u_h), v\big) = \big(h'(u_h) + B^* p(u_h), v\big), \quad \forall v \in K,$$

where $p(u_h)$ satisfies the following system:

$$a\big(y(u_h), w\big) + \big(\phi\big(y(u_h)\big), w\big) = (f + Bu_h, w), \quad \forall w \in V, \tag{3.1}$$

$$a\big(q, p(u_h)\big) + \big(\phi'\big(y(u_h)\big)p(u_h), q\big) = \big(g'\big(y(u_h)\big), q\big), \quad \forall q \in V. \tag{3.2}$$

Let $\pi_h : C^0(\bar{\Omega}) \to V^h$ be the standard Lagrange interpolation operator such that for any $v \in C^0(\bar{\Omega})$, $\pi_h v(P_i) = v(P_i)$ for all $P_i \in \mathbf{P}$, where $\mathbf{P}$ is the vertex set associated with the triangulation $\mathcal{T}^h$, and n is the dimension of the domain Ω, we have the following result:

Lemma 3.1 [3] *Let* π_h *be the standard Lagrange interpolation operator. For* $m = 0$ *or* 1, $q > \frac{n}{2}$ *and* $\forall v \in W^{2,q}(\Omega)$, *we have*

$$|v - \pi_h v|_{W^{m,q}(\Omega)} \le C h^{2-m} |v|_{W^{2,q}(\Omega)}.$$

Lemma 3.2 *Let* (y_h, p_h, u_h) *and* $(y(u_h), p(u_h))$ *be the solutions of* (2.11)-(2.13) *and* (3.1)-(3.2), *respectively. Assume that* $p(u_h), y(u_h) \in H^2(\Omega)$ *and* $\phi'(\cdot)$ *is locally Lipschitz continuous. Then there exists a constant C independent of h such that*

$$\big\|y(u_h) - y_h\big\|_{1,\Omega} + \big\|p(u_h) - p_h\big\|_{1,\Omega} \le Ch. \tag{3.3}$$

Proof From $\phi'(\cdot) \geq 0$, (2.12), (3.2) and embedding theorem $\|v\|_{L^4(\Omega)} \leq C\|v\|_{H^1(\Omega)}$, we have

$$
\begin{aligned}
c\big\|p(u_h) &- p_h\big\|_{1,\Omega}^2 \\
&\leq a\big(p(u_h) - p_h, p(u_h) - p_h\big) + \big(\phi'\big(y(u_h)\big)\big(p(u_h) - p_h\big), p(u_h) - p_h\big) \\
&= a\big(p(u_h) - \pi_h p(u_h), p(u_h) - p_h\big) + \big(\phi'\big(y(u_h)\big)\big(p(u_h) - p_h\big), p(u_h) - \pi_h p(u_h)\big) \\
&\quad + \big(g'\big(y(u_h)\big) - g'(y_h), \pi_h p(u_h) - p_h\big) + \big(\big(\phi'(y_h) - \phi'\big(y(u_h)\big)\big)p_h, \pi_h p(u_h) - p_h\big) \\
&\leq C\big\|p(u_h) - p_h\big\|_{1,\Omega}\big\|p(u_h) - \pi_h p(u_h)\big\|_{1,\Omega} + C\big\|y(u_h) - y_h\big\|\big\|\pi_h p(u_h) - p_h\big\| \\
&\quad + C\big\|\phi'\big(y(u_h)\big)\big\|\big\|p(u_h) - p_h\big\|_{L^4(\Omega)}\big\|p(u_h) - \pi_h p(u_h)\big\|_{L^4(\Omega)} \\
&\quad + C\big\|y(u_h) - y_h\big\|\big\|p_h\big\|_{L^4(\Omega)}\big\|\pi_h p(u_h) - p_h\big\|_{L^4(\Omega)} \\
&\leq C(\delta)\big\|p(u_h) - \pi_h p(u_h)\big\|_{1,\Omega}^2 + C(\delta)\big\|y(u_h) - y_h\big\|^2 \\
&\quad + C\delta\big(\big\|p(u_h) - p_h\big\|_{1,\Omega}^2 + \big\|\pi_h p(u_h) - p_h\big\|_{1,\Omega}^2\big) \\
&\leq C(\delta)\big\|p(u_h) - \pi_h p(u_h)\big\|_{1,\Omega}^2 + C(\delta)\big\|y(u_h) - y_h\big\|^2 + C\delta\big\|p(u_h) - p_h\big\|_{1,\Omega}^2.
\end{aligned} \tag{3.4}
$$

Note that $p(u_h) \in H^2(\Omega)$, by using Lemma 3.1, we obtain

$$
\begin{aligned}
\big\|p(u_h) - p_h\big\|_{1,\Omega} &\leq C\big\|p(u_h) - \pi_h\big(p(u_h)\big)\big\|_{1,\Omega} + C\big\|y(u_h) - y_h\big\| \\
&\leq Ch\big\|p(u_h)\big\|_{2,\Omega} + C\big\|y(u_h) - y_h\big\| \\
&\leq Ch + C\big\|y(u_h) - y_h\big\|.
\end{aligned} \tag{3.5}
$$

Similarly, we can prove that

$$
\big\|y(u_h) - y_h\big\|_{1,\Omega} \leq Ch\big\|y(u_h)\big\|_{2,\Omega} \leq Ch. \tag{3.6}
$$

Then (3.3) follows from (3.5)-(3.6). □

In order to derive sharp *a priori* estimates, we introduce the following auxiliary problems:

$$
-\operatorname{div}\big(A^*\nabla\xi\big) + \Phi\xi = F_1, \quad \text{in } \Omega, \xi|_{\partial\Omega} = 0, \tag{3.7}
$$

$$
-\operatorname{div}(A\nabla\zeta) + \phi'\big(y(u_h)\big)\zeta = F_2, \quad \text{in } \Omega, \zeta|_{\partial\Omega} = 0, \tag{3.8}
$$

where

$$
\Phi = \begin{cases} \frac{\phi(y(u_h)) - \phi(y_h)}{y(u_h) - y_h}, & y(u_h) \neq y_h, \\ \phi'(y_h), & y(u_h) = y_h. \end{cases}
$$

From the regularity estimates (see, *e.g.*, [6]), we obtain

$$
\|\xi\|_{2,\Omega} \leq C\|F_1\|, \qquad \|\zeta\|_{2,\Omega} \leq C\|F_2\|.
$$

Lemma 3.3 *Let (y_h, p_h, u_h) be the solution of* (2.11)-(2.13). *Suppose that $y(u_h), p(u_h) \in H^2(\Omega)$ and $\phi'(\cdot)$ is locally Lipschitz continuous. Then there exists a constant C independent of h such that*

$$\|y(u_h) - y_h\| + \|p(u_h) - p_h\| \le Ch^2. \tag{3.9}$$

Proof Let $F_1 = y(u_h) - y_h$ and $\xi_h = \pi_h \xi$. We have

$$\begin{aligned}
\|y(u_h) - y_h\|^2 &= a(y(u_h) - y_h, \xi) + (\Phi\xi, y(u_h) - y_h) \\
&= a(y(u_h) - y_h, \xi - \xi_h) + (\Phi\xi, y(u_h) - y_h) - (\phi(y(u_h)) - \phi(y_h), \xi_h) \\
&= a(y(u_h) - y_h, \xi - \xi_h) + (\phi(y(u_h)) - \phi(y_h), \xi - \xi_h) \\
&\le C\|y(u_h) - y_h\|_{1,\Omega}\|\xi - \xi_h\|_{1,\Omega} + C\|y(u_h) - y_h\|\|\xi - \xi_h\| \\
&\le C\|y(u_h) - y_h\|_{1,\Omega}\|\xi - \xi_h\|_{1,\Omega}.
\end{aligned} \tag{3.10}$$

Note that

$$\|\xi - \xi_h\|_{1,\Omega} \le Ch\|\xi\|_{2,\Omega} \le Ch\|y(u_h) - y_h\|. \tag{3.11}$$

Thus,

$$\|y(u_h) - y_h\| \le Ch\|y(u_h) - y_h\|_{1,\Omega} \le Ch^2. \tag{3.12}$$

Similarly, let $F_2 = p(u_h) - p_h$ and $\zeta_h = \pi_h \zeta$, we obtain

$$\|p(u_h) - p_h\| \le Ch^2. \tag{3.13}$$

From (3.12) and (3.13), we get (3.9). $\qquad \square$

Lemma 3.4 *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of* (2.4)-(2.6) *and* (2.11)-(2.13), *respectively. Assume that all the conditions in Lemma 3.3 are valid. Then there exists a constant C independent of h such that*

$$\|u - u_h\| \le Ch^2. \tag{3.14}$$

Proof It is clear that

$$(J'(v) - J'(u), v - u) \ge c\|v - u\|^2, \quad \forall v, u \in K. \tag{3.15}$$

By using (2.6) and (2.13), we have

$$\begin{aligned}
c\|u - u_h\|^2 &\le (J'(u) - J'(u_h), u - u_h) \\
&= (h'(u) + B^*p, u - u_h) - (h'(u_h) + B^*p(u_h), u - u_h) \\
&\le -(h'(u_h) + B^*p_h, u - u_h) + (B^*p_h - B^*p(u_h), u - u_h)
\end{aligned}$$

$$\leq \left(B^* p_h - B^* p(u_h), u - u_h\right)$$

$$\leq C \left\| p_h - p(u_h) \right\| \| u - u_h \|. \tag{3.16}$$

From (3.9) and (3.16), we derive (3.14). □

Now we combine Lemmas 3.2-3.4 to come up with the following main result.

Theorem 3.1 *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of (2.4)-(2.6) and (2.11)-(2.13), respectively. Assume that all the conditions in Lemmas 3.2-3.4 are valid. Then we have*

$$\| u - u_h \| + \| y - y_h \| + \| p - p_h \| \leq C h^2. \tag{3.17}$$

Proof Note that

$$\| p - p_h \| \leq \left\| p - p(u_h) \right\| + \left\| p(u_h) - p_h \right\|, \tag{3.18}$$

$$\| y - y_h \| \leq \left\| y - y(u_h) \right\| + \left\| y(u_h) - y_h \right\|. \tag{3.19}$$

From (2.4)-(2.5), (3.1)-(3.2) and the regularity estimates, we have

$$\left\| p - p(u_h) \right\| \leq \left\| p - p(u_h) \right\|_{2,\Omega} \leq C \left\| y - y(u_h) \right\|, \tag{3.20}$$

$$\left\| y - y(u_h) \right\| \leq \left\| y - y(u_h) \right\|_{2,\Omega} \leq C \| u - u_h \|. \tag{3.21}$$

Then, (3.17) follows from (3.9), (3.14) and (3.18)-(3.21). □

4 *A posteriori* error estimates

We now derive *a posteriori* error estimates for the variational discretization approximation scheme. The following lemmas are very important in deriving *a posteriori* error estimates of residual type.

Lemma 4.1 [5] $\forall v \in W^{1,q}(\Omega), 1 \leq q < \infty$,

$$\| v \|_{W^{0,q}(\partial \tau)} \leq C \left(h_\tau^{-\frac{1}{q}} \| v \|_{W^{0,q}(\tau)} + h_\tau^{1-\frac{1}{q}} |v|_{W^{1,q}(\tau)} \right). \tag{4.1}$$

Lemma 4.2 *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of (2.4)-(2.6) and (2.11)-(2.13), respectively. Then we have*

$$\| u - u_h \| \leq C \left\| p_h - p(u_h) \right\|, \tag{4.2}$$

where $p(u_h)$ is defined in (3.2).

Proof It follows from (2.6) and (2.13) that

$$
\begin{aligned}
c\|u - u_h\|^2 &\le \big(J'(u), u - u_h\big) - \big(J'(u_h), u - u_h\big) \\
&\le -\big(J'(u_h), u - u_h\big) \\
&= -\big(h'(u_h) + B^* p_h, u - u_h\big) \\
&\quad + \big(B^* p_h - B^* p(u_h), u - u_h\big) \\
&\le C(\delta)\big\| p_h - p(u_h)\big\|^2 + \delta \|u - u_h\|^2.
\end{aligned}
\tag{4.3}
$$

Let δ be small enough, then (4.2) follows from (4.3). □

Lemma 4.3 *Let (y_h, p_h, u_h) and $(y(u_h), p(u_h))$ be the solutions of* (2.11)-(2.13) *and* (3.1)-(3.2), *respectively. Assume that $\phi'(\cdot)$ is locally Lipschitz continuous. Then there exists a positive constant C independent of h such that*

$$
\big\|y_h - y(u_h)\big\|_{1,\Omega}^2 + \big\|p_h - p(u_h)\big\|_{1,\Omega}^2 \le C\big(\eta_1^2 + \eta_2^2\big),
\tag{4.4}
$$

where

$$
\eta_1^2 = \sum_{\tau \in \mathcal{T}^h} h_\tau^2 \int_\tau \big(g'(y_h) + \operatorname{div}\big(A^* \nabla p_h\big) - \phi'(y_h) p_h\big)^2 dx + \sum_{l \cap \partial\Omega \neq \emptyset} h_l \int_l \big[A^* \nabla p_h \cdot \mathbf{n}\big]^2 ds,
$$

$$
\eta_2^2 = \sum_{\tau \in \mathcal{T}^h} h_\tau^2 \int_\tau \big(\operatorname{div}(A \nabla y_h) - \phi(y_h) + f + B u_h\big)^2 dx + \sum_{l \cap \partial\Omega \neq \emptyset} h_l \int_l [A \nabla y_h \cdot \mathbf{n}]^2 ds,
$$

where h_l is the size of the face $l = \bar{\tau}_l^1 \cap \bar{\tau}_l^2$, and τ_l^1, τ_l^2 are two neighboring elements in $\mathcal{T}^h$, $[A \nabla y_h \cdot \mathbf{n}]_l$, and $[A^* \nabla p_h \cdot \mathbf{n}]_l$ are the A-normal and A^*-normal derivative jumps over the interior face l, respectively, defined by

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

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

where $\mathbf{n}$ is the normal vector on $l = \tau_l^1 \cap \tau_l^2$ outwards τ_l^1. For later convenience, we defined $[A \nabla y_h \cdot \mathbf{n}]_l = 0$ and $[A^* \nabla p_h \cdot \mathbf{n}]_l = 0$ when $l \subset \partial\Omega$.

Proof Let $e^p = p_h - p(u_h)$ and $e_I^p = \pi_h e^p$, it follows from the Green formula, embedding theorem $\|v\|_{L^4(\Omega)} \le C\|v\|_{H^1(\Omega)}$, Lemma 4.1, (2.12) and (3.2) that

$$
\begin{aligned}
c\big\| p_h &- p(u_h)\big\|_{1,\Omega}^2 \\
&\le a\big(e^p, p_h - p(u_h)\big) + \big(\phi'\big(y(u_h)\big)\big(p_h - p(u_h)\big), e^p\big) \\
&= a\big(e^p - e_I^p, p_h - p(u_h)\big) + \big(\phi'(y_h) p_h - \phi'\big(y(u_h)\big) p(u_h), e^p - e_I^p\big) + a\big(e_I^p, p_h - p(u_h)\big) \\
&\quad + \big(\phi'(y_h) p_h - \phi'\big(y(u_h)\big) p(u_h), e_I^p\big) + \big(\phi'\big(y(u_h)\big) p_h - \phi'(y_h) p_h, e^p\big) \\
&= \sum_{\tau \in \mathcal{T}^h} \int_\tau \big(g'(y_h) + \operatorname{div}\big(A^* \nabla p_h\big) - \phi'(y_h) p_h\big)\big(e_I^p - e^p\big) dx + \big(g'(y_h) - g'\big(y(u_h)\big), e^p\big)
\end{aligned}
$$

$$+ \sum_{\tau \in \mathcal{T}^h} \int_{\partial \tau} \left(A^* \nabla p_h \cdot \mathbf{n}\right)\left(e^p - e_I^p\right) ds + \left(\phi'(y(u_h))p_h - \phi'(y_h)p_h, e^p\right)$$

$$\leq C\eta_1^2 + C(\delta)\left\|y_h - y(u_h)\right\|_{1,\Omega}^2 + \delta\left\|p_h - p(u_h)\right\|_{1,\Omega}^2. \tag{4.5}$$

Similarly, we obtain

$$c\left\|y_h - y(u_h)\right\|_{1,\Omega}^2$$

$$\leq a\left(y_h - y(u_h), e^y\right) + \left(\phi(y_h) - \phi(y(u_h)), e^y\right)$$

$$= \left(A\nabla\left(y_h - y(u_h)\right), \nabla\left(e^y - e_I^y\right)\right) + \left(\phi(y_h) - \phi(y(u_h)), e^y - e_I^y\right)$$

$$= \sum_{\tau \in \mathcal{T}^h} \int_{\tau} \left(\mathrm{div}(A\nabla y_h) - \phi(y_h) + f + Bu_h\right)\left(e_I^y - e^y\right) dx + \sum_{\tau \in \mathcal{T}^h} \int_{\partial\tau} \left(A\nabla y_h \cdot \mathbf{n}\right)\left(e^y - e_I^y\right) ds$$

$$\leq C(\delta)\eta_2^2 + \delta\left\|y_h - y(u_h)\right\|_{1,\Omega}^2. \tag{4.6}$$

From (4.5) and (4.6), we derive (4.4). □

Theorem 4.1 *Let* (y, p, u) *and* (y_h, p_h, u_h) *be the solutions of* (2.4)-(2.6) *and* (2.11)-(2.13), *respectively. Assume that all the conditions in Lemmas* 4.2-4.3 *are valid. Then there exists a constant C independent of h such that*

$$\|u - u_h\|^2 + \|y - y_h\|_{1,\Omega}^2 + \|p - p_h\|_{1,\Omega}^2 \leq C\left(\eta_1^2 + \eta_2^2\right), \tag{4.7}$$

where η_1 *and* η_2 *are defined in Lemma* 4.3.

Proof Note that

$$\|p - p_h\|_{1,\Omega} \leq \left\|p - p(u_h)\right\|_{1,\Omega} + \left\|p_h - p(u_h)\right\|_{1,\Omega}, \tag{4.8}$$

$$\|y - y_h\|_{1,\Omega} \leq \left\|y - y(u_h)\right\|_{1,\Omega} + \left\|y_h - y(u_h)\right\|_{1,\Omega}, \tag{4.9}$$

and

$$\left\|p - p(u_h)\right\|_{1,\Omega} \leq \left\|p - p(u_h)\right\|_{2,\Omega} \leq C\left\|y - y(u_h)\right\|, \tag{4.10}$$

$$\left\|y - y(u_h)\right\|_{1,\Omega} \leq \left\|y - y(u_h)\right\|_{2,\Omega} \leq C\|u - u_h\|. \tag{4.11}$$

Then (4.7) follows from (4.2), (4.4) and (4.8)-(4.11). □

5 Numerical experiments

For a constrained optimization problem

$$\min_{u \in K \subset U} J(u),$$

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

$$\begin{cases} b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n(J'(u_n), v), & \forall v \in U, \\ u_{n+1} = P_K^b(u_{n+\frac{1}{2}}), \end{cases} \tag{5.1}$$

where $b(\cdot, \cdot)$ is a symmetric and positive definite bilinear form, and similarly to [11], the projection operator $P_K^b : U \to K$ is defined: For given $w \in U$ find $P_K^b w \in K$ such that

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

The bilinear form $b(\cdot, \cdot)$ provides a suitable precondition for the projection gradient algorithm. Let $U^h = \{v_h \in L^2(\Omega), a \le v_h \le b : v_h|_\tau = constant, \forall \tau \in \mathcal{T}^h\}$. For an acceptable error *tol* and a fixed step size ρ_n, by applying (5.1) to the discretized nonlinear elliptic optimal control problem, we introduce the following projection gradient algorithm (see, *e.g.*, [11, 12]), for ease of exposition, we have omitted the subscript h.

Algorithm 5.1 (Projection gradient algorithm)

Step 1. Initialize u_0;

Step 2. Solve the following equations:

$$\begin{cases} b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n(h'(u_n) + B^* p_n, v), & u_{n+\frac{1}{2}}, u_n \in U^h, \forall v \in U^h, \\ a(y_n, w) + (\phi(y_n), w) = (f + Bu_n, w), & y_n \in V^h, \forall w \in V^h, \\ a(q, p_n) + (\phi'(y_n)p_n, q) = (y_n - y_d, q), & p_n \in V^h, \forall q \in V^h, \\ u_{n+1} = P_K^b(u_{n+\frac{1}{2}}); \end{cases} \tag{5.2}$$

Step 3. Calculate the iterative error: $e_{n+1} = \|u_{n+1} - u_n\|$;

Step 4. If $e_{n+1} \le tol$, stop; else, go to Step 2.

According to the preceding analysis, we construct the following variational discretization algorithm.

Algorithm 5.2 (Variational discretization algorithm)

Step 1. Initialize u_0;

Step 2. Solve the following equations:

$$\begin{cases} b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n(h'(u_n) + B^* p_n, v), & u_{n+\frac{1}{2}}, u_n \in U, \forall v \in U, \\ a(y_n, w) + (\phi(y_n), w) = (f + Bu_n, w), & y_n \in V^h, \forall w \in V^h, \\ a(q, p_n) + (\phi'(y_n)p_n, q) = (y_n - y_d, q), & p_n \in V^h, \forall q \in V^h, \\ u_{n+1} = \Pi_{[a,b]}(u_{n+\frac{1}{2}}). \end{cases} \tag{5.3}$$

Step 3. Calculate the iterative error: $e_{n+1} = \|u_{n+1} - u_n\|$;

Step 4. If $e_{n+1} \le tol$, stop; else, go to Step 2.

It is well known that there are four major types of adaptive finite element methods, namely, the h-methods (mesh refinement), the p-methods (order enrichment), the r-methods (mesh redistribution) and the hp-methods (the combination of h-method and p-method). For an acceptable error *Tol*, by using *a posteriori* error estimator η_1^2 and η_2^2 as the mesh refinement indicator and the Algorithm 5.2, we present the following adaptive variational discretization algorithm.

Algorithm 5.3 (Adaptive variational discretization algorithm)

Step 1. Solve the discretized optimization problem with Algorithm 5.2 on the current mesh obtain the numerical solution u'_n and calculate the error estimators η_1 and η_2;

Step 2. Adjust the mesh by using estimators η_1 and η_2, then update the numerical solution u'_n and obtain the new numerical solution u'_{n+1} on new mesh;

Step 3. If $\|u'_{n+1} - u'_n\| \leq Tol$, stop; else, go to Step 1.

All of the following numerical examples were solved numerically with codes developed based on AFEPack which provided a general tool of finite element approximation for PDEs. The package is freely available and the details can be found in [30].

We consider the following optimal control problems:

$$\begin{cases} \min_{u \in K} \{ \int_\Omega (g(y(x)) + h(u(x))) \, dx \}, \\ -\operatorname{div}(A(x)\nabla y(x)) + \phi(y(x)) = f(x) + Bu(x), \quad x \in \Omega, \\ y(x) = 0, \quad x \in \partial\Omega, \end{cases}$$

where

$$g(y(x)) = \frac{1}{2}[y(x) - y_0(x)]^2,$$

$$h(u(x)) = \frac{1}{2}[u(x) - u_0(x)]^2,$$

and

$$K = \{ v(x) \in L^2(\Omega) : a \leq v(x) \leq b, x \in \Omega \},$$

the domain Ω is the unit square $[0,1] \times [0,1]$ and $B = I$.

Example 1 In the first example, we compare the convergence order of $\|u - u_h\|$ in Algorithm 5.1 with that in Algorithm 5.2. The data are as follows:

$$A(x) = E, \qquad \phi(y) = y^3, \qquad a = 0, \qquad b = 1.5,$$

$$p(x) = 4x_1 x_2 \sin(2\pi x_1) \sin(2\pi x_2),$$

$$y(x) = p(x),$$

$$u_0(x) = 1 + \sin(2\pi x_1) \sin(2\pi x_2),$$

$$u(x) = \min(1.5, \max(0, u_0(x) - p(x))),$$

$$f(x) = -\operatorname{div}(A(x)\nabla y(x)) + \phi(y(x)) - u(x),$$

$$y_0(x) = y(x) + \operatorname{div}(A^*(x)\nabla p(x)) - \phi'(y(x))p(x).$$

The numerical results are listed in Table 1 and Table 2.

In Figure 1, we see clearly that in the projection gradient algorithm, $\|u - u_h\| = O(h)$, while in the variational discretization algorithm, $\|u - u_h\| = O(h^2)$. In Figure 2, we show the profiles of the exact solution u alongside the solution error.

Table 1 Algorithm 5.1, Example 1

Mesh	$\|u - u_h\|$	$\|y - y_h\|$	$\|p - p_h\|$
16 × 16	5.98051E-02	4.92490E-02	4.76241E-02
32 × 32	2.84008E-02	1.25725E-02	1.21182E-02
64 × 64	1.39765E-02	3.15952E-03	3.04294E-03
128 × 128	6.96692E-03	7.90942E-04	7.61570E-04
256 × 256	3.48077E-03	1.97800E-04	1.90445E-04

Table 2 Algorithm 5.2, Example 1

Mesh	$\|u - u_h\|$	$\|y - y_h\|$	$\|p - p_h\|$
16 × 16	4.31006E-02	4.92551E-02	4.76397E-02
32 × 32	1.09812E-02	1.25730E-02	1.21193E-02
64 × 64	2.77425E-03	3.15960E-03	3.04301E-03
128 × 128	6.92439E-04	7.90991E-04	7.61575E-04
256 × 256	1.73211E-04	1.97847E-04	1.90445E-04

Figure 1 The convergence order of $\|u - u_h\|$.
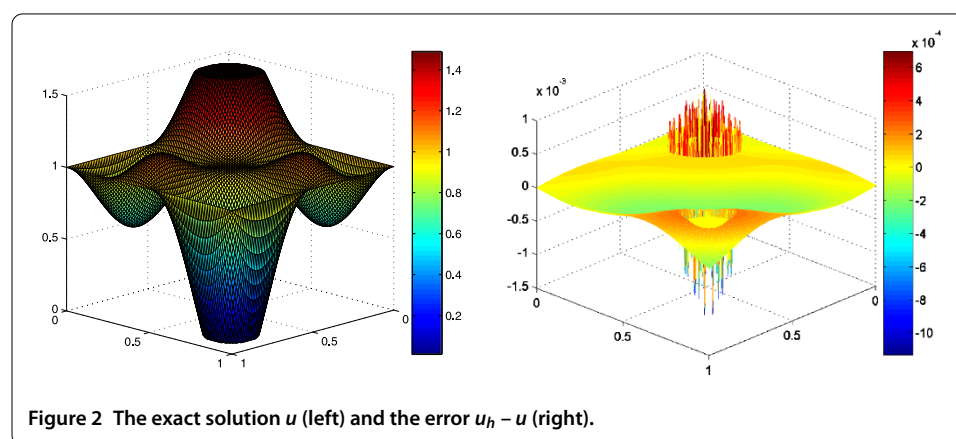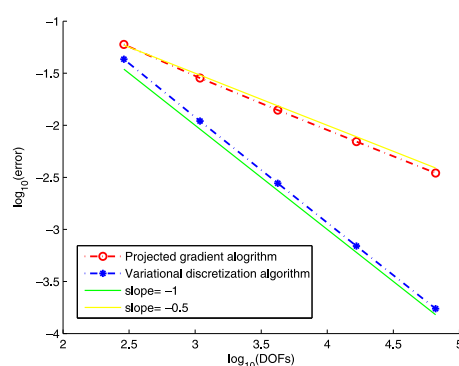

**Figure 2 The exact solution u (left) and the error $u_h - u$ (right).**

Example 2 In order to illustrate the reliability and efficiency of the *a posteriori* error estimates in Theorem 4.1, we use Algorithm 5.3 to solve this example. The data are as follows:

$$\phi(y) = y^3, \qquad a = -0.5, \qquad b = 0,$$

$$A(x) = \begin{cases} E, & x_1 + x_2 \geq 1, \\ 2 \cdot E, & x_1 + x_2 < 1, \end{cases}$$

Table 3 Numerical results, Example 2.

Mesh	Nodes	Sides	Elements	Dofs	$\|u - u_h\|$
Uniform mesh	513	1456	944	513	6.19419E-02
Adaptive mesh	145	392	248	145	6.04841E-02

Figure 3 The exact solution u (left) and the error $u_h - u$ (right).
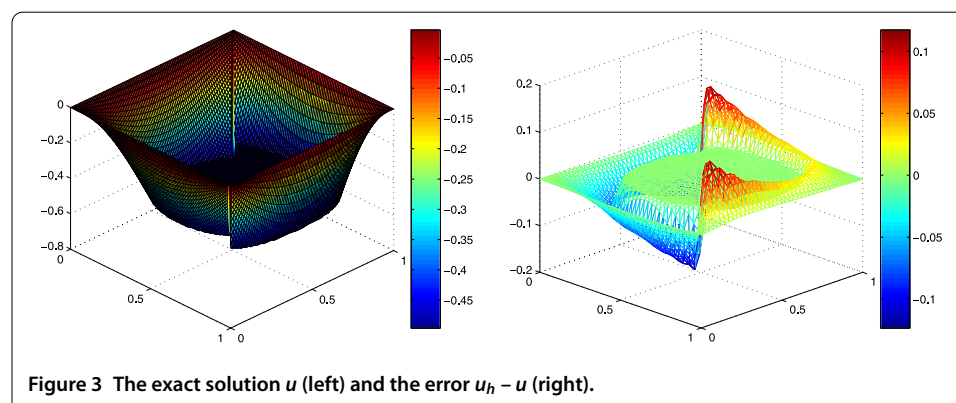
$$p(x) = \begin{cases} 2\sin(\pi x_1)\sin(\pi x_2), & x_1 + x_2 \geq 1, \\ \sin(\pi x_1)\sin(\pi x_2), & x_1 + x_2 < 1, \end{cases}$$

$$y(x) = p(x),$$

$$u_0(x) = 0,$$

$$u(x) = \min\big(0, \max\big(-0.5, u_0(x) - p(x)\big)\big),$$

$$f(x) = -\operatorname{div}\big(A(x)\nabla y(x)\big) + \phi\big(y(x)\big) - u(x),$$

$$y_0(x) = y(x) + \operatorname{div}\big(A^*(x)\nabla p(x)\big) - \phi'\big(y(x)\big)p(x).$$

The numerical results based on adaptive mesh and uniform mesh are presented in Table 3. In Figure 3, we show the profiles of the exact solution u alongside the solution error. From Table 3, it is clear that the adaptive mesh generated via the error indicators in Theorem 4.1 are able to save substantial computational work, in comparison with the uniform mesh. Our numerical results confirm our theoretical results.

References
 1. Pao, C: Nonlinear Parabolic and Elliptic Equations. Plenum Press, New York (1992)
 2. Lions, J: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
 3. Ciarlet, P: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
 4. He, BS: Solving a class of linear projection equations. Numer. Math. **68**, 71-80 (1994)
 5. Kufner, A, John, O, Fucik, S: Function Spaces. Nordhoff, Leyden (1977)
 6. Ladyzhenskaya, O, Urlatseva, H: Linear and Quasilinear Elliptic Equations. Academic Press, New York (1968)
 7. Lions, J, Magenes, E: Non Homogeneous Boundary Value Problems and Applications. Springer, Berlin (1972)
 8. Neittaanmaki, P, Tiba, D: Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications. Dekker, New York (1994)

9. Arada, N, Casas, E, Tröltzsch, F: Error estimates for the numerical approximation of a semilinear elliptic control problem. Comput. Optim. Appl. **23**, 201-229 (2002)
10. Liu, W, Yan, N: Adaptive Finite Element Methods for Optimal Control Governed by PDEs. Science Press, Beijing (2008)
11. Li, R, Liu, W, Yan, N: *A posteriori* error estimates of recovery type for distributed convex optimal control problems. J. Sci. Comput. **33**, 155-182 (2007)
12. Liu, H, Yan, N: Recovery type superconvergence and *a posteriori* error estimates for control problems governed by Stokes equations. J. Comput. Appl. Math. **209**, 187-207 (2007)
13. Yan, N: *A posteriori* error estimates of gradient recovery type for FEM of optimal control problem. Adv. Comput. Math. **19**, 323-336 (2003)
14. Ainsworth, M, Oden, JT: A Posteriori Error Estimation in Finite Element Analysis. Wiley, New York (2000)
15. Babuška, I, Strouboulis, T, Upadhyay, CS, Gangaraj, SK: *A posteriori* estimation and adaptive control of the pollution error in the h-version of the finite element method. Int. J. Numer. Methods Eng. **38**(24), 4207-4235 (1995)
16. Carstensen, C, Verfürth, R: Edge residuals dominate *a posteriori* error estimates for low order finite element methods. SIAM J. Numer. Anal. **36**(5), 1571-1587 (1999)
17. Hintermüller, M, Hoppe, RHW, Iliash, Y, Lieweg, M: An *a posteriori* error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. ESAIM Control Optim. Calc. Var. **14**(3), 540-560 (2008)
18. Hintermüller, M, Hoppe, RHW: Goal-oriented adaptivity in control constrained optimal control of partial differential equations. SIAM J. Control Optim. **47**(3), 1721-1743 (2008)
19. Liu, W, Yan, N: *A posteriori* error estimates for distributed convex optimal control problems. Adv. Comput. Math. **15**, 285-309 (2001)
20. Liu, W, Yan, N: *A posteriori* error estimates for control problems governed by nonlinear elliptic equations. Appl. Numer. Math. **47**, 173-187 (2003)
21. Veeser, A: Efficient and reliable *a posteriori* error estimators for elliptic obstacle problems. SIAM J. Numer. Anal. **39**(1), 146-167 (2001)
22. Verfurth, R: A Review of Posteriori Error Estimation and Adaptive Mesh Refinement. Wiley, London (1996)
23. Brandts, J: Superconvergence and *a posteriori* error estimation for triangular mixed finite elements. Numer. Math. **68**, 311-324 (1994)
24. Chen, Y, Dai, Y: Superconvergence for optimal control problems governed by semi-linear elliptic equations. J. Sci. Comput. **39**, 206-221 (2009)
25. Chen, Y, Liu, B: Error estimates and superconvergence of mixed finite element for quadratic optimal control. Int. J. Numer. Anal. Model. **3**, 311-321 (2006)
26. Xing, X, Chen, Y: L^∞-error estimates for general optimal control problem by mixed finite element methods. Int. J. Numer. Anal. Model. **5**(3), 441-456 (2008)
27. Zhang, Z, Zhu, JZ: Analysis of the superconvergent patch recovery technique and *a posteriori* error estimator in the finite element method (I). Comput. Methods Appl. Math. **123**, 173-187 (1995)
28. Becker, R, Kapp, H, Rannacher, R: Adaptive finite element methods for optimal control of partial differential equations: basic concept. SIAM J. Control Optim. **39**(1), 113-132 (2000)
29. Hoppe, RHW, Kieweg, M: Adaptive finite element methods for mixed control-state constrained optimal control problems for elliptic boundary value problems. Comput. Optim. Appl. **46**, 511-533 (2010)
30. Li, R, Liu, W, Ma, H, Tang, T: Adaptive finite element approximation for distributed elliptic optimal control problems. SIAM J. Control Optim. **41**(5), 1321-1349 (2002)
31. Vexler, B, Wollner, W: Adaptive finite elements for elliptic optimization problems with control constraints. SIAM J. Control Optim. **47**(1), 509-534 (2008)
32. Alt, W, Mackenroth, U: Convergence of finite element approximation to state constrained convex parabolic boundary control problems. SIAM J. Control Optim. **27**, 718-736 (1989)
33. Chen, Y, Lu, Z: Error estimates for quadratic parabolic optimal control problems by fully discrete mixed finite element methods. Finite Elem. Anal. Des. **46**(11), 957-965 (2010)
34. Chen, Y, Yang, J: *A posteriori* error estimation for a fully discrete discontinuous Galerkin approximation to a kind of singularly perturbed problems. Finite Elem. Anal. Des. **43**(10), 757-770 (2007)
35. Liu, W, Yan, N: *A posteriori* error estimates for optimal control problems governed by parabolic equations. Numer. Math. **93**, 497-521 (2003)
36. Xiong, C, Li, Y: *A posteriori* error estimates for optimal distributed control governed by the evolution equations. Appl. Numer. Math. **61**, 181-200 (2011)
37. Chen, Y, Huang, Y, Yi, N: *A posteriori* error estimates of spectral method for optimal control problems governed by parabolic equations. Sci. China Ser. A **51**(8), 1376-1390 (2008)
38. Chen, Y, Yi, N, Liu, W: A legendre-Galerkin spectral method for optimal control problems governed by elliptic equations. SIAM J. Numer. Anal. **46**(5), 2254-2275 (2008)
39. Chen, Y, Tang, Y: Numerical methods for constrained elliptic optimal control problems with rapidly oscillating coefficients. East Asian J. Appl. Math. **1**, 235-247 (2011)
40. Borzì, A: High-order discretization and multigrid solution of elliptic nonlinear constrained control problems. J. Comput. Appl. Math. **200**, 67-85 (2005)
41. Hinze, M: A variational discretization concept in control constrained optimization: the linear-quadratic case. Comput. Optim. Appl. **30**, 45-63 (2005)
42. Hinze, M, Yan, N, Zhou, Z: Variational discretization for optimal control governed by convection dominated diffusion equations. J. Comput. Math. **27**, 237-253 (2009)
43. Liu, W, Tiba, D: Error estimates for the finite element approximation of a class of nonlinear optimal control problems. Numer. Funct. Anal. Optim. **22**, 953-972 (2001)