Tang and Hua *Journal of Inequalities and Applications* (2020) 2020:240
https://doi.org/10.1186/s13660-020-02506-6

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Adaptive variational discretization approximation method for parabolic optimal control problems

Yuelong Tang[1*] and Yuchun Hua[2]

*Correspondence:
tangyuelonga@163.com
[1]Institute of Computational Mathematics, College of Science, Hunan University of Science and Engineering, 425199 Yongzhou, Hunan, China
Full list of author information is available at the end of the article

Abstract

In this paper, we study variational discretization method for parabolic optimization problems. Firstly, we obtain some convergence and superconvergence analysis results of the approximation scheme. Secondly, we derive a posteriori error estimates of the approximation solutions. Finally, we present variational discretization approximation algorithm and adaptive variational discretization approximation algorithm for parabolic optimization problems and do some numerical experiments to confirm our theoretical results.

MSC: 49M25; 65M60

Keywords: Adaptive; Variational discretization approximation; Parabolic optimization problems

1 Introduction

Optimal control problems (OCPs) play an important role in scientific and engineering numerical simulation, and nowadays are strongly utilized in biology, economics, and finance. Finite element approximation is one of the widely used numerical methods in computing OCPs. A systematic introduction of finite element method (FEM) for partial differential equations or OPCs can be found in [1, 2, 8, 20, 30, 33].

There have been abundant researches on FEM approximation for elliptic optimal control problems (EOCPs). The pioneering work of the late 1970s in the area of finite element approximation for EOCPs includes [10, 11], where a priori error estimates were established. Then a lot of superconvergence and a posteriori error estimate results of FEM solving different kinds of EOCPs were obtained. For instance, superconvergence of FEM for linear, bilinear, and semilinear EOCPs were derived in [27, 32], and [4], respectively, while residual- and recovery-based a posteriori error estimates of FEM for distributed OCPs were constructed in [22] and [18], respectively. Moreover, adaptive FEM for EOCPs were presented in [3, 13, 17]. It is worth mentioning that some similar results of mixed FEM for EOCPs can be found in [5, 7, 16].

In the last decade, numerical solution of parabolic optimal control problems (POCPs) became a hot research topic. A priori error estimates of FEM, space-time FEM, and

Petrov–Galerkin Crank–Nicolson approximation for POCPs were given in [12, 25], and [26], respectively. In [24], the authors constructed an adaptive space-time FEM for POCPs. There are also some results on residual-type a posteriori error estimates of FEM or mixed FEM for POCPs which can be found in [6, 23, 31], where the authors do not give any adaptive FEM approximation for POCPs.

Hinze presented a variational discretization (VD) concept for control constrained optimization problems in [14]. It can not only save some computation cost but also improve the error of the control variable. Recently, VD approximation of convection dominated diffusion EOCP with control constraints and POCP with pointwise state constraints were investigated in [15] and [9], respectively. We have investigated VD approximation for a linear POCP in [28].

The purpose of this work is to investigate VD approximation for a POCP with control constraints. We first analyze the convergence and superconvergence of the VD approximation scheme and derive a posteriori error estimates, then construct an adaptive VD approximation algorithm for the POCP. Finally, two numerical examples are provided to verify theoretical results.

We are interested in the following POCP:

$$\begin{cases} \min\limits_{u(x,t)\in K} \frac{1}{2}\int_0^T(\|y(x,t)-y_d(x,t)\|^2+\nu\|u(x,t)\|^2)\,dt, \\ y_t(x,t)-\operatorname{div}(A(x)\nabla y(x,t))=f(x,t)+u(x,t), \quad x\in\Omega, t\in J, \\ y(x,t)=0, \quad x\in\partial\Omega, t\in J, \\ y(x,0)=y_0(x), \quad x\in\Omega, \end{cases} \tag{1}$$

where $\nu>0$ represents the weight of the cost of the control, Ω is a bounded domain in $\mathbb{R}^2$ with a Lipschitz boundary $\partial\Omega$, $0<T<+\infty$ and $J=[0,T]$. The coefficient $A(x)=(a_{ij}(x))_{2\times 2}\in(W^{1,\infty}(\bar{\Omega}))^{2\times 2}$ is such that for any $\xi\in\mathbb{R}^2$, $(A(x)\xi)\cdot\xi\geq c|\xi|^2$ with $c>0$. We assume that K is a nonempty closed convex set in $L^2(J;L^2(\Omega))$, defined by

$$K=\bigl\{v(x,t)\in L^\infty\bigl(J;L^2(\Omega)\bigr):a\leq v(x,t)\leq b, \text{a.e. } (x,t)\in\Omega\times J\bigr\},$$

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 denote $W^{m,2}(\Omega)$ by $H^m(\Omega)$ and set $H_0^1(\Omega)\equiv\{v\in H^1(\Omega):v|_{\partial\Omega}=0\}$. 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 $\|v\|_{L^s(J;W^{m,q}(\Omega))}=(\int_0^T\|v\|_{W^{m,q}(\Omega)}^s\,dt)^{\frac{1}{s}}$ for $s\in[1,\infty)$ and the standard modification for $s=\infty$. As in [21], we can define the space $H^l(J;W^{m,q}(\Omega))$. In addition, c or C denotes a generic positive constant independent of h and k.

The outline of the paper is as follows. In Sect. 2, we give a VD approximation for the model problem. In Sect. 3, we derive some convergence and superconvergence analysis results for the control, state, and costate variables. In Sect. 4, we establish a posteriori error estimates for the approximation scheme. We present a VD approximation algorithm and adaptive VD approximation algorithm for the POCP and do some numerical experiments to confirm our theoretical results in the last section.

2 VD approximation of POCP

In this section, we present a VD approximation for the model problem (1). For ease of exposition, we denote $L^p(J; W^{m,q}(\Omega))$ by $L^p(W^{m,q})$. Let $W = H_0^1(\Omega)$ and $U = L^2(\Omega)$. Moreover, we denote $\|\cdot\|_{H^m(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$ by $\|\cdot\|_m$ and $\|\cdot\|$, respectively.

Let

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

$$(f_1,f_2) = \int_\Omega f_1 \cdot f_2, \quad \forall f_1,f_2 \in U.$$

It follows from the assumptions on A that

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

Thus a possible weak formula for the model problem (1) reads:

$$\begin{cases} \min\limits_{u\in K} \frac{1}{2}\int_0^T (\|y-y_d\|^2 + \nu\|u\|^2)\,dt, \\ (y_t,w) + a(y,w) = (f+u,w), \quad \forall w \in W, t \in J, \\ y(x,0) = y_0(x), \quad x \in \Omega. \end{cases} \tag{2}$$

It is well known (see, e.g., [20, 23]) that problem (2) has a unique solution (y,u), and the pair $(y,u) \in (H^1(L^2)\cap L^2(H^1)) \times K$ is the solution of (2) if and only if there is an adjoint state $p \in H^1(L^2)\cap L^2(H^1)$ such that the triplet (y,p,u) satisfies the following optimality conditions:

$$(y_t,w) + a(y,w) = (f+u,w), \quad \forall w \in W, t \in J, \tag{3}$$

$$y(x,0) = y_0(x), \quad x \in \Omega, \tag{4}$$

$$-(p_t,q) + a(q,p) = (y_d - y,q), \quad \forall q \in W, t \in J, \tag{5}$$

$$p(x,T) = 0, \quad x \in \Omega, \tag{6}$$

$$(\nu u - p, v-u) \geq 0, \quad \forall v \in K, t \in J. \tag{7}$$

We introduce the following pointwise projection operator:

$$\Pi_{[a,b]}\big(g(x,t)\big) = \min\left(b, \max\left(a, \frac{1}{\nu}g(x,t)\right)\right), \quad \forall (x,t) \in \Omega \times J.$$

Similar to [15], the variational inequality (7) can be rewritten as

$$u(x,t) = \Pi_{[a,b]}\big(p(x,t)\big), \quad \forall (x,t) \in \Omega \times J. \tag{8}$$

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

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

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

We now consider the time discretization for problem (2). Let $k > 0$, $N = T/k \in \mathbb{Z}^+$, $t_n = nk$, $n = 0, 1, \dots, N$. Set $\varphi^n = \varphi(x, t_n)$ and

$$d_t\varphi^n = \frac{\varphi^n - \varphi^{n-1}}{k}, \quad n = 1, 2, \dots, N.$$

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

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

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

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

Then a VD approximation of (2) is as follows:

$$\begin{cases} \min\limits_{u_h^n \in K} \frac{1}{2} \sum_{n=1}^N k(\|y_h^n - y_d^n\|^2 + \nu \|u_h^n\|^2), \\ (d_t y_h^n, w_h) + a(y_h^n, w_h) = (f^n + u_h^n, w_h), \quad \forall w_h \in W^h, n = 1, 2, \dots, N, \\ y_h^0(x) = P_h y_0(x), \quad x \in \Omega, \end{cases} \tag{9}$$

where P_h is an elliptic projection operator which will be specified later.

It follows (see, e.g., [15, 23]) that the control problem (9) has a unique solution (y_h^n, u_h^n), $n = 1, 2, \dots, N$, and $(y_h^n, u_h^n) \in W^h \times K$, $n = 1, 2, \dots, N$, is the solution of (9) if and only if there is an adjoint state $p_h^{n-1} \in W^h$, $n = 1, 2, \dots, N$, such that the triplet $(y_h^n, p_h^{n-1}, u_h^n) \in W^h \times W^h \times K$, $n = 1, 2, \dots, N$, satisfies the following optimality conditions:

$$\big(d_t y_h^n, w_h\big) + a\big(y_h^n, w_h\big) = \big(f^n + u_h^n, w_h\big), \quad \forall w_h \in W^h, \tag{10}$$

$$y_h^0(x) = P_h y_0(x), \quad x \in \Omega, \tag{11}$$

$$-\big(d_t p_h^n, q_h\big) + a\big(q_h, p_h^{n-1}\big) = \big(y_d^n - y_h^n, q_h\big), \quad \forall q_h \in W^h, \tag{12}$$

$$p_h^N(x) = 0, \quad x \in \Omega, \tag{13}$$

$$\big(\nu u_h^n - p_h^{n-1}, v - u_h^n\big) \ge 0, \quad \forall v \in K. \tag{14}$$

Similar to [15], the optimality condition (14) can be equivalently expressed as

$$u_h^n = \Pi_{[a,b]}\big(p_h^{n-1}\big), \quad n = 1, 2, \dots, N. \tag{15}$$

3 Convergence and superconvergence analysis

For the approximation scheme (10)–(14), some convergence and superconvergence results will be derived in this section. We define the following intermediate variables. Let $(y_h^n(u), p_h^{n-1}(u)) \in W^h \times W^h$, $n = 1, 2, \dots, N$, satisfy the following system:

$$\big(d_t y_h^n(u), w_h\big) + a\big(y_h^n(u), w_h\big) = \big(f^n + u^n, w_h\big), \quad \forall w_h \in W^h, \tag{16}$$

$$y_h^0(u)(x) = P_h y_0(x), \quad x \in \Omega, \tag{17}$$

$$-\big(d_t p_h^n(u), q_h\big) + a\big(q_h, p_h^{n-1}(u)\big) = \big(y_d^n - y_h^n(u), q_h\big), \quad \forall q_h \in W^h, \tag{18}$$

$$p_h^N(u)(x) = 0, \quad x \in \Omega, \tag{19}$$

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

$$a(\phi - P_h\phi, w_h) = 0, \quad \forall w_h \in W^h. \tag{20}$$

It has the following approximation properties (see, e.g., [4]):

$$\|\phi - P_h\phi\| \le Ch^2 \|\phi\|_2, \quad \forall \phi \in H^2(\Omega). \tag{21}$$

Lemma 3.1 *Let (y, p, u) and $(y_h(u), p_h(u))$ be the solutions of (3)–(7) and (16)–(18), respectively. Suppose that $y, p \in l_D^2(H^2) \cap H^1(H^2) \cap H^2(L^2)$. Then*

$$\big\|y_h(u) - y\big\|_{l^2(L^2)} + \big\|p_h(u) - p\big\|_{l^2(L^2)} \le C\big(h^2 + k\big). \tag{22}$$

Proof From the definition of P_h, (3) and (16), for any $w_h \in W^h$, $n = 1, 2, \ldots, N$, we have

$$\begin{aligned}
&\big(d_t y_h^n(u) - d_t P_h y^n, w_h\big) + a\big(y_h^n(u) - P_h y^n, w_h\big) \\
&\quad = -\big(d_t P_h y^n, w_h\big) - a\big(y^n, w_h\big) + \big(f^n + u^n, w_h\big) \\
&\quad = -\big(d_t P_h y^n - d_t y^n, w_h\big) - \big(d_t y^n - y_t^n, w_h\big).
\end{aligned} \tag{23}$$

By selecting $w_h = y_h^n(u) - P_h y^n$, $n = 1, 2, \ldots, N$, we obtain

$$\begin{aligned}
&\big(d_t y_h^n(u) - d_t y^n, y_h^n(u) - P_h y^n\big) + a\big(y_h^n(u) - P_h y^n, y_h^n(u) - P_h y^n\big) \\
&\quad = -\big(d_t P_h y^n - d_t y^n, y_h^n(u) - P_h y^n\big) - \big(d_t y^n - y_t^n, y_h^n(u) - P_h y^n\big).
\end{aligned} \tag{24}$$

Note that $a(y_h^n(u) - P_h y^n, y_h^n(u) - P_h y^n) \ge 0$, and

$$\begin{aligned}
&\big(d_t y_h^n(u) - d_t P_h y^n, y_h^n(u) - P_h y^n\big) \\
&\quad \ge \frac{1}{k}\big(\big\|y_h^n(u) - P_h y^n\big\|^2 - \big\|y_h^n(u) - P_h y^n\big\|\,\big\|y_h^{n-1}(u) - P_h y^{n-1}\big\|\big).
\end{aligned} \tag{25}$$

It follows from (24)–(25) and Hölder's inequality that

$$\begin{aligned}
&\big\|y_h^n(u) - P_h y^n\big\| \\
&\quad \le \big\|y_h^{n-1}(u) - P_h y^{n-1}\big\| + \big\|(P_h - I)\big(y^n - y^{n-1}\big)\big\| + \big\|y^n - y^{n-1} - k y_t^n\big\|.
\end{aligned} \tag{26}$$

Summing over n from 1 to N^* ($1 \le N^* \le N$), we get

$$\begin{aligned}
\|y_h^{N^*}(u) - P_h y^{N^*}\| &\le \sum_{n=1}^{N^*} \|(P_h - I)(y^n - y^{n-1})\| + \sum_{n=1}^{N^*} \|y^n - y^{n-1} - k y_t^n\| \\
&\le \sum_{n=1}^{N^*} Ch^2 \|y^n - y^{n-1}\|_2 + \sum_{n=1}^{N^*} \int_{t_{n-1}}^{t_n} \|(t_{n-1} - t) y_{tt}\|\, dt \\
&\le Ch^2 \sum_{n=1}^{N^*} \int_{t_{n-1}}^{t_n} \|y_t\|_2\, dt + k \sum_{n=1}^{N^*} \int_{t_{n-1}}^{t_n} \|y_{tt}\|\, dt \\
&\le Ch^2 \int_0^{t_{N^*}} \|y_t\|_2\, dt + k \int_0^{t_{N^*}} \|y_{tt}\|\, dt \\
&\le C\big(h^2 \|y_t\|_{L^2(H^2)} + k \|y_{tt}\|_{L^2(L^2)}\big).
\end{aligned} \tag{27}$$

Thus, we have

$$\big\|y_h(u) - P_h y\big\|_{l^\infty(L^2)} \le C\big(h^2 + k\big). \tag{28}$$

From (21), we derive

$$|||P_h y - y|||_{l^2(L^2)}^2 = \sum_{n=1}^{N} k \|P_h y^n - y^n\|^2 \le Ch^4 \sum_{n=1}^{N} k \|y^n\|_2^2 = Ch^4 |||y|||_{l^2(H^2)}^2. \tag{29}$$

According to embedding theorem and (28)–(29), we have

$$\big\|y_h(u) - y\big\|_{l^2(L^2)} \le C\big(h^2 + k\big). \tag{30}$$

Similarly, we can prove that

$$\big\|p_h(u) - p\big\|_{l^2(L^2)} \le C\big(h^2 + k\big). \tag{31}$$

Then (22) follows from (30) and (31). □

For the control variable, we derive the following convergence result.

Theorem 3.1 *Let (y, p, u) and (y_h, p_h, u_h) be the solutions of (3)–(7) and (10)–(14), respectively. Assume that all the conditions in Lemma 3.1 are satisfied. Then, we have*

$$|||u - u_h|||_{l^2(L^2)} \le C\big(h^2 + k\big). \tag{32}$$

Proof From (7) and (14), we obtain

$$\begin{aligned} &\nu|||u-u_h|||^2_{l^2(L^2)} \\ &\quad= \sum_{n=1}^{N} k\big(\nu\big(u^n-u_h^n\big), u^n-u_h^n\big) \\ &\quad\le \sum_{n=1}^{N} k\big(\nu u_h^n - p_h^{n-1}(u), u_h^n-u^n\big) + \sum_{n=1}^{N} k\big(p^n-p_h^{n-1}(u), u^n-u_h^n\big) \\ &\quad\le \sum_{n=1}^{N} k\big(p_h^{n-1} - p_h^{n-1}(u), u_h^n-u^n\big) + \sum_{n=1}^{N} k\big(p^n-p_h^{n-1}(u), u^n-u_h^n\big). \end{aligned} \tag{33}$$

It follows from (10)–(12) and (16)–(18) that

$$\sum_{n=1}^{N} k\big(p_h^{n-1} - p_h^{n-1}(u), u_h^n-u^n\big) = -|||y_h-y_h(u)|||^2_{l^2(L^2)} \le 0. \tag{34}$$

By using Hölder and Young inequalities, we get

$$\begin{aligned} &\sum_{n=1}^{N} k\big(p^n-p_h^{n-1}(u), u^n-u_h^n\big) \\ &\quad= \sum_{n=1}^{N} k\big(p^{n-1}-p_h^{n-1}(u), u^n-u_h^n\big) + \sum_{n=1}^{N} k\big(p^n-p^{n-1}, u^n-u_h^n\big) \\ &\quad\le C(\delta)\sum_{n=1}^{N} k\big\|p^{n-1}-p_h^{n-1}(u)\big\|^2 + C(\delta)\sum_{n=1}^{N} k\big\|p^n-p^{n-1}\big\|^2 + \delta\sum_{n=1}^{N} k\big\|u^n-u_h^n\big\|^2 \\ &\quad\le C(\delta)\big(|||p-p_h(u)|||^2_{l^2(L^2)} + (k)^2\|p_t\|^2_{L^2(J;L^2(\Omega))}\big) + \delta|||u-u_h|||^2_{l^2(L^2)}. \end{aligned} \tag{35}$$

From (22) and (33)–(35), we obtain (32). □

For the control, state and costate variables, we have the following results.

Theorem 3.2 *Let* (y,p,u) *and* (y_h,p_h,u_h) *be the solutions* (3)–(7) *and* (10)–(14), *respectively. Assume that all the conditions in Theorem* 3.1 *are valid. Then*

$$|||P_hy-y_h|||_{l^2(H^1)} + |||P_hp-p_h|||_{l^2(H^1)} \le C\big(h^2+k\big). \tag{36}$$

Proof From (3) and (10), for any $w_h \in W^h$, $n = 1,2,\ldots,N$, we obtain the following error equation:

$$\big(y_t^n - d_ty_h^n, w_h\big) + a\big(y^n-y_h^n, w_h\big) = \big(u^n-u_h^n, w_h\big). \tag{37}$$

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

$$\begin{aligned} &\big(d_tP_hy^n - d_ty_h^n, P_hy^n-y_h^n\big) + a\big(P_hy^n-y_h^n, P_hy^n-y_h^n\big) \\ &\quad= \big(d_tP_hy^n - d_ty^n + d_ty^n - y_t^n + u^n - u_h^n, P_hy^n-y_h^n\big). \end{aligned} \tag{38}$$

Note that

$$\bigl(d_tP_hy^n - d_ty_h^n, P_hy^n - y_h^n\bigr) \geq \frac{1}{2k}\bigl(\|P_hy^n - y_h^n\|^2 - \|P_hy^{n-1} - y_h^{n-1}\|^2\bigr) \tag{39}$$

and

$$\begin{aligned}\bigl(d_tP_hy^n - d_ty^n, P_hy^n - y_h^n\bigr) &\leq \|d_tP_hy^n - d_ty^n\|\|P_hy^n - y_h^n\| \\ &\leq Ch^2\|d_ty^n\|_2\|P_hy^n - y_h^n\| \\ &\leq Ch^2k^{-1}\int_{t_{n-1}}^{t_n}\|y_t\|_2\,dt\|P_hy^n - y_h^n\| \\ &\leq Ch^2k^{-\frac{1}{2}}\|y_t\|_{L^2(t_{n-1},t_n;H^2(\Omega))}\|P_hy^n - y_h^n\|.\end{aligned} \tag{40}$$

Additionally,

$$\begin{aligned}\bigl(d_ty^n - y_t^n, P_hy^n - y_h^n\bigr) &= k^{-1}\bigl(y^n - y^{n-1} - ky_t^n, P_hy^n - y_h^n\bigr) \\ &\leq k^{-1}\|y^n - y^{n-1}ky_t^n\|\|P_hy^n - y_h^n\| \\ &= k^{-1}\left\|\int_{t_{n-1}}^{t_n}(t_{n-1} - t)y_{tt}\,dt\right\|\|P_hy^n - y_h^n\| \\ &\leq Ck^{\frac{1}{2}}\|y_{tt}\|_{L^2(t_{n-1},t_n;L^2(\Omega))}\|P_hy^n - y_h^n\|.\end{aligned} \tag{41}$$

Multiplying both sides of (38) by $2k$, summing over n from 1 to N, and by using Hölder and Young inequalities, we have

$$\begin{aligned}&\|P_hy^N - y_h^N\|^2 + c\sum_{n=1}^{N}k\|P_hy^n - y_h^n\|_1^2 \\ &\quad\leq C(\delta)\bigl(h^4\|y_t\|_{L^2(H^2)}^2 + k^2\|y_{tt}\|_{L^2(L^2)}^2 + |||u - u_h|||_{l^2(L^2)}^2\bigr) + \delta\sum_{n=1}^{N}k\|P_hy^n - y_h^n\|^2.\end{aligned} \tag{42}$$

From (32) and (42), we get

$$|||P_hy - y_h|||_{l^2(H^1)} \leq C\bigl(h^2 + k\bigr). \tag{43}$$

Similarly, we can prove that

$$|||P_hp - p_h|||_{l^2(H^1)} \leq C\bigl(h^2 + k\bigr). \tag{44}$$

Then (36) follows from (43)–(44). □

Theorem 3.3 *Assume that $u \in l^2(H^1)$ and all the conditions in Theorem 3.2 are valid. Then*

$$|||P_hu - u_h|||_{l^2(H^1)} \leq C\bigl(h^2 + k\bigr). \tag{45}$$

Proof Notice that $\Pi_{[a,b]}$ is Lipschitz continuous with constant 1. From (8) and (15), we have

$$\begin{aligned} |||P_h u - u_h|||_{l^2(H^1)} &\le C \big\| \Pi_{[a,b]}(P_h p - p_h) \big\|_{l^2(H^1)} \\ &\le C |||P_h p - p_h|||_{l^2(H^1)}. \end{aligned} \tag{46}$$

From (36) and (46), we obtain (45). □

4 A posteriori error estimates

In this section, a posteriori error estimates of recovery type will be established by utilizing the superconvergence results derived in Sect. 3. Similar to the Z–Z patch recovery in [34, 35], we introduce recovery operators R_h and G_h for the state and the adjoint state. Let $R_h v$ be a continuous piecewise linear function (without zero boundary constraint) and let the values of $R_h v$ on the nodes be defined by a least-squares argument on element patches surrounding the nodes. The gradient recovery operator $G_h v = (R_h v_{x_1}, R_h v_{x_2})$. The details also can be found in [18].

Theorem 4.1 *Let (y,p,u) and (y_h,p_h,u_h) be the solutions of (3)–(7) and (10)–(14), respectively. Suppose that all the conditions in Theorem 3.2 are valid and $y,p \in L^2_D(J;H^3(\Omega))$. Then*

$$|||G_h y_h - \nabla y|||_{l^2(L^2)} + |||G_h p_h - \nabla p|||_{l^2(L^2)} \le C\big(h^2 + k\big). \tag{47}$$

Proof Let y_I be the piecewise linear Lagrange interpolation of y. From Theorem 2.1.1 in [19], we have

$$\|G_h y - y_I\|_1 \le Ch^2 \|y\|_3. \tag{48}$$

According to the standard interpolation error estimate technique (see, e.g., [8]), we have

$$\|G_h y_I - \nabla y\| \le Ch^2 |y|_3. \tag{49}$$

By using (47)–(48), we get

$$\begin{aligned} \big\|G_h y_h^n - \nabla y^n\big\| &= \big\|G_h y_h^n - G_h P_h y^n\big\| + \big\|G_h P_h y^n - G_h y_I^n\big\| + \big\|G_h y_I^n - \nabla y^n\big\| \\ &\le C\big\|y_h^n - P_h y^n\big\|_1 + C\big\|P_h y^n - y_I^n\big\|_1 + \big\|G_h y_I^n - \nabla y^n\big\| \\ &\le C\big\|y_h^n - P_h y^n\big\|_1 + Ch^2\big\|y^n\big\|_3. \end{aligned} \tag{50}$$

Therefore,

$$\sum_{n=1}^N k\big\|G_h y_h^n - \nabla y^n\big\|^2 \le C\sum_{n=1}^N k\big\|y_h^n - P_h y^n\big\|_1^2 + Ch^4 \sum_{n=1}^N k\big\|y^n\big\|_3^2. \tag{51}$$

From Theorem 3.2 and (51), we derive

$$|||G_h y_h - \nabla y|||_{l^2(L^2)} \le C\big(h^2 + k\big). \tag{52}$$

Similarly, we can prove that

$$|||G_h p_h - \nabla p|||_{l^2(L^2)} \leq C\big(h^2 + k\big). \tag{53}$$

Then (47) follows from (52)–(53). □

By using the above superconvergence properties, it is easy to prove the following a posteriori error estimate results.

Theorem 4.2 *Assume that all the conditions in Theorem 4.1 are valid. Then*

$$\eta_1 := |||G_h y_h - \nabla y_h|||_{l^2(L^2)} = \big|\big|\big|\nabla(y - y_h)\big|\big|\big|_{l^2(L^2)} + \mathcal{O}\big(h^2 + k\big), \tag{54}$$

$$\eta_2 := |||G_h p_h - \nabla p_h|||_{l^2(L^2)} = \big|\big|\big|\nabla(p - p_h)\big|\big|\big|_{l^2(L^2)} + \mathcal{O}\big(h^2 + k\big). \tag{55}$$

5 Numerical experiments

We do some numerical experiments to demonstrate our theoretical results. For an acceptable error *Tol*, we present a VD approximation algorithm (see Algorithm 5.1) for the constrained POCP (1). For ease of exposition, we have omitted the subscript h.

Similar to [18], by selecting the same meshes for the state and the adjoint state and using η_1 and η_2 as mesh refinement indicators for the state and the adjoint state, for an acceptable error Tol', we construct adaptive VD approximation algorithm (see Algorithm 5.2).

The following numerical examples were solved with AFEPack which is freely available. Just for simplicity, we let $\Omega = [0,1] \times [0,1]$, $T = 1$, $A(x) = E$ be the 2×2 identity matrix and denote $|||\cdot|||_{l^2(L^2)}$ and $|||\cdot|||_{l^2(H^1)}$ by $|||\cdot|||$ and $|||\cdot|||_1$, respectively. The convergence order is

Algorithm 5.1 VD approximation algorithm

Step 1. Initialize u_0.

Step 2. Solve the following equations:

$$\begin{cases} (\frac{y_n^i - y_n^{i-1}}{k}, w) + a(y_n^i, w) = (f^i + u_n^i, w), & y_n^i, y_n^{i-1} \in W^h, \forall w \in W^h, \\ (\frac{p_n^{i-1} - p_n^i}{k}, q) + a(q, p_n^{i-1}) = (y_d^i - y_n^i, q), & p_n^i, p_n^{i-1} \in W^h, \forall q \in W^h, \\ u_{n+1}^i = \Pi_{[a,b]}(p_n^{i-1}), \quad i = 1, 2, \ldots, N; \end{cases}$$

Step 3. Calculate the iterative error, $E_{n+1} = |||u_{n+1} - u_n|||_{l^2(L^2)}$;

Step 4. If $E_{n+1} > Tol$, go to Step 1; else stop.

Algorithm 5.2 Adaptive VD approximation algorithm

Step 1. Solve the discretized optimization problem (9) with the Algorithm 5.1 on the current meshes to obtain a numerical solution u_n' and calculate the error estimators η_1 and η_2;

Step 2. Adjust the meshes by using the estimators η_1 and η_2, then 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 $E_{n+1}' > Tol'$, go to Step 1; else stop.

computed by the following formula: $Rate = \frac{\log(e_{i+1})-\log(e_i)}{\log(h_{i+1})-\log(h_i)}$, where $e_i(e_{i+1})$ denotes the error when the spatial partition size is $h_i(h_{i+1})$. We solve the following POCP:

$$\begin{cases} \min_{u\in K} \frac{1}{2}\int_0^T (\|y(x,t)-y_d(x,t)\|^2 + \nu\|u(x,t)\|^2)\,dt, \\ y_t(x,t) - \operatorname{div}(A(x)\nabla y(x,t)) = f(x,t) + u(x,t), \quad x\in\Omega, t\in J, \\ y(x,t) = 0, \quad x\in\partial\Omega, t\in J, \\ y(x,0) = y_0(x), \quad x\in\Omega. \end{cases}$$

Moreover, we assume that

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

Example 1 The data are as follows:

$$\begin{aligned} &\nu = 1, \qquad a = -0.25, \qquad b = 0.25, \\ &p(x,t) = \sin(2\pi x_1)\sin(2\pi x_2)(1-t), \\ &y(x,t) = \sin(2\pi x_1)\sin(2\pi x_2)t, \\ &u(x,t) = \max\bigl(-0.25, \min\bigl(0.25, -p(x,t)\bigr)\bigr), \\ &f(x,t) = y_t(x,t) - \operatorname{div}\bigl(A(x)\nabla y(x,t)\bigr) - u(x,t), \\ &y_d(x,t) = y(x,t) + p_t(x,t) + \operatorname{div}\bigl(A^*(x)\nabla p(x,t)\bigr). \end{aligned}$$

We use the Algorithm 5.1 to solve the first example. In Fig. 1, we plot the profile of the numerical solution u_h at $t = 0.5$ when $h = \frac{1}{80}$ and $k = \frac{1}{640}$. In Table 1, the errors $|||u-u_h|||$, $|||P_hy-y_h|||_1$, and $|||P_hp-p_h|||_1$ on a sequence of uniformly refined meshes are shown. It is easy to see $|||u-u_h||| = \mathcal{O}(h^2+k)$, $|||P_hy-y_h|||_1 = \mathcal{O}(h^2+k)$, and $|||P_hp-p_h|||_1 = \mathcal{O}(h^2+k)$, which confirm our theoretical results.

Example 2 The data are as follows:

$$\begin{aligned} &\nu = 1, \qquad a = 0, \qquad b = 10, \\ &p(x,t) = \frac{(t-1)\sin(\pi x_1)\sin(\pi x_2)}{(x_1-0.5)^2+(x_2-0.5)^2+0.05}, \end{aligned}$$

Figure 1 The numerical solution u_h at $t = 0.5$ when $h = \frac{1}{80}$ and $k = \frac{1}{640}$

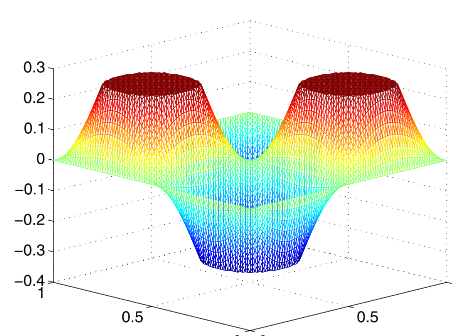

Table 1 Numerical results, Example 1

h	k	$\|\|\|u-u_h\|\|\|$	Rate	$\|\|\|P_hy-y_h\|\|\|_1$	Rate	$\|\|\|P_hp-p_h\|\|\|_1$	Rate
$\frac{1}{10}$	$\frac{1}{10}$	4.59597e–02	–	7.21635e–03	–	9.51033e–03	–
$\frac{1}{20}$	$\frac{1}{40}$	1.23011e–02	1.90	1.70370e–03	2.08	2.24615e–03	2.08
$\frac{1}{40}$	$\frac{1}{160}$	3.11887e–03	1.97	4.22468e–04	2.01	5.55263e–04	2.01
$\frac{1}{80}$	$\frac{1}{640}$	7.81133e–04	1.99	1.05482e–04	2.00	1.38467e–04	2.00

Table 2 Numerical results on uniform meshes, Example 2

Uniform meshes	1	2	3	4
Nodes	121	441	1681	6561
$\|\|\|u-u_h\|\|\|$	1.51170e–01	5.67167e–02	5.65961e–02	5.65652e–02
$\|\|\|\nabla y-\nabla y_h\|\|\|$	5.22324e+00	2.69670e+00	1.35903e+00	6.80921e–01
$\|\|\|\nabla p-\nabla p_h\|\|\|$	5.22335e+00	2.69667e+00	1.35902e+00	6.80944e–01
$\|\|\|G_hy_h-\nabla y_h\|\|\|$	5.50904e+00	2.78499e+00	1.37386e+00	6.82949e–01
$\|\|\|G_hp_h-\nabla p_h\|\|\|$	5.51247e+00	2.78475e+00	1.37345e+00	6.82704e–01

Table 3 Numerical results on adaptive meshes, Example 2

Adaptive meshes	1	2	3	4
Nodes	139	445	716	739
$\|\|\|u-u_h\|\|\|$	9.33794e–02	4.53875e–02	4.53241e–02	4.53216e–02
$\|\|\|\nabla y-\nabla y_h\|\|\|$	4.29891e+00	2.19612e+00	1.45671e+00	1.42490e+00
$\|\|\|\nabla p-\nabla p_h\|\|\|$	4.30707e+00	2.21794e+00	1.49066e+00	1.45962e+00
$\|\|\|G_hy_h-\nabla y_h\|\|\|$	4.78789e+00	2.29456e+00	1.46836e+00	1.43600e+00
$\|\|\|G_hp_h-\nabla p_h\|\|\|$	4.71907e+00	2.26074e+00	1.44664e+00	1.41476e+00

$$
\begin{aligned}
&y(x,t)=\frac{t\sin(\pi x_1)\sin(\pi x_2)}{(x_1-0.5)^2+(x_2-0.5)^2+0.05},\\
&u(x,t)=\min\bigl(10,\max\bigl(0,-p(x,t)\bigr)\bigr),\\
&f(x,t)=y_t(x,t)-\operatorname{div}\bigl(A(x)\nabla y(x,t)\bigr)-u(x,t),\\
&y_d(x,t)=y(x,t)+p_t(x,t)+\operatorname{div}\bigl(A^*(x)\nabla p(x,t)\bigr).
\end{aligned}
$$

We take a small time size $k=\frac{1}{100}$ and solve the second example by using the Algorithms 5.1 and 5.2, respectively. Numerical results based on a sequence of uniformly refined meshes and adaptive meshes are listed in Tables 2 and 3, respectively. It is clear that the adaptive meshes generated via the error estimators η_1 and η_2 are able to save substantial computational work, in comparison with the uniform meshes.

6 Conclusions

Although there has been extensive research on FEMs for various POCPs, they mostly focused on convergence and superconvergence (see, e.g., [12, 23–26, 28, 29]), and the results on convergence and superconvergence were $\mathcal{O}(h+k)$ and $\mathcal{O}(h^{\frac{3}{2}}+k)$, respectively. Recently, VD were used to deal with different OCPs in [9, 14, 15, 28], while there was little work on POCPs. Hence, our results on adaptive VD for POCPs are new.

Acknowledgements
The authors are very grateful to both referees for carefully reading of the paper and their comments and suggestions.

Funding
The first author is supported by the National Natural Science Foundation of China (11401201), the Natural Science Foundation of Hunan Province (2020JJ4323), the Scientific Research Project of Hunan Provincial Department of Education

(20A211), the construct program of applied characteristic discipline in Hunan University of Science and Engineering. The second author is supported by the Scientific Research Project of Hunan Provincial Department of Education (20C0854), the scientific research program in Hunan University of Science and Engineering (20XKY059).

Abbreviations
OCP, optimal control problem; FEM, finite element method; EOCP, elliptic optimal control problem; POCP, parabolic optimal control problem; VD, variational discretization.

Availability of data and materials
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
The authors have participated in the sequence alignment and drafted the manuscript. They have approved the final manuscript.

Author details
[1]Institute of Computational Mathematics, College of Science, Hunan University of Science and Engineering, 425199 Yongzhou, Hunan, China. [2]College of Science, Hunan University of Science and Engineering, 425199 Yongzhou, Hunan, China.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 30 June 2020 Accepted: 29 October 2020 Published online: 11 November 2020

References
1. Ainsworth, M., Oden, J.: A Posteriori Error Estimation in Finite Element Analysis. Wiley-Interscience, New York (2000)
2. Babuvška, I., Rheinboldt, W.: Error estimates for adaptive finite element computations. SIAM J. Numer. Anal. **44**(4), 75–102 (1978)
3. Becker, R., Kapp, H., Rannacher, R.: Adaptive finite element methods for optimal control of partial differential equations: basic concept. SIAM J. Control Optim. **39**(1), 113–132 (2000)
4. Chen, Y., Dai, Y.: Superconvergence for optimal control problems governed by semi-linear elliptic equations. J. Sci. Comput. **39**, 206–221 (2009)
5. Chen, Y., Huang, Y., Liu, W., Yan, N.: Error estimates and superconvergence of mixed finite element methods for convex optimal control problems. J. Sci. Comput. **42**(3), 382–403 (2010)
6. Chen, Y., Lin, Z.: A posteriori error estimates of semidiscrete mixed finite element methods for parabolic optimal control problems. East Asian J. Appl. Math. **5**(1), 85–108 (2015)
7. Chen, Y., Liu, W.: A posteriori error estimates for mixed finite element solutions of convex optimal control problems. J. Comput. Appl. Math. **211**, 76–89 (2008)
8. Ciarlet, P.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
9. Deckelnick, K., Hinze, M.: Variational discretization of parabolic control problems in the presence of pointwise state constraints. J. Comput. Math. **29**(1), 1–16 (2011)
10. Falk, F.: Approximation of a class of optimal control problems with order of convergence estimates. J. Math. Anal. Appl. **44**, 28–47 (1973)
11. Geveci, T.: On the approximation of the solution of an optimal control problem governed by an elliptic equation. RAIRO. Anal. Numér. **13**, 313–328 (1979)
12. Gong, W., Hinze, M., Zhou, Z.: A priori error analysis for finite element approximation of parabolic optimal control problems with pointwise control. SIAM J. Control Optim. **52**(1), 97–119 (2014)
13. Gong, W., Yan, N.: Adaptive finite element method for elliptic optimal control problems: convergence and optimality. Numer. Math. **135**, 1121–1170 (2017)
14. Hinze, M.: A variational discretization concept in control constrained optimization: the linear-quadratic case. Comput. Optim. Appl. **30**, 45–61 (2005)
15. Hinze, M., Yan, N., Zhou, Z.: Variational discretization for optimal control governed by convection dominated diffusion equations. J. Comput. Math. **27**, 237–253 (2009)
16. Hoppe, R., 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)
17. 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 (2012)
18. 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)
19. Lin, Q., Zhu, Q.: The Preprocessing and Postprocessing for the Finite Element Method. Shanghai Scientific and Technical Publishers, Shanghai (1994)
20. Lions, J.: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
21. Lions, J., Magenes, E.: Non Homogeneous Boundary Value Problems and Applications. Springer, Berlin (1972)
22. Liu, W., Yan, N.: A posteriori error estimates for distributed convex optimal control problems. Adv. Comput. Math. **15**, 285–309 (2001)
23. Liu, W., Yan, N.: A posteriori error estimates for optimal control problems governed by parabolic equations. Numer. Math. **93**, 497–521 (2003)

24. Meidner, D., Vexler, B.: Adaptive space-time finite element methods for parabolic optimization problems. SIAM J. Control Optim. **46**(1), 116–142 (2007)
25. Meidner, D., Vexler, B.: 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), 1301–1329 (2008)
26. Meidner, D., Vexler, B.: A priori error analysis of the Petrov–Galerkin Crank–Nicolson scheme for parabolic optimal control problems. SIAM J. Control Optim. **49**(5), 2183–2211 (2011)
27. Meyer, C., Rösch, A.: Superconvergence properties of optimal control problems. SIAM J. Control Optim. **43**(3), 970–985 (2004)
28. Tang, Y., Chen, Y.: Variational discretization for parabolic optimal control problems with control constraints. J. Syst. Sci. Complex. **25**, 880–895 (2012)
29. Tang, Y., Hua, Y.: Convergence and superconvergence of variational discretization for parabolic bilinear optimization problems. J. Inequal. Appl. **2019**, 239 (2019)
30. Vallejos, M., Borzì, A.: Multigrid optimization methods for linear and bilinear elliptic optimal control problems. Computing **82**, 31–52 (2008)
31. Xiong, C., Li, Y.: A posteriori error estimates for optimal distributed control governed by the evolution equations. Appl. Numer. Math. **61**, 181–200 (2011)
32. Yang, D., Chang, Y., Liu, W.: A priori error estimate and superconvergence analysis for an optimal control problem of bilinear type. J. Comput. Math. **26**(4), 471–487 (2008)
33. Yang, J.: A unified a posteriori error estimate of local discontinuous Galerkin approximation for reactive transport problems. Indian J. Pure Appl. Math. **46**(6), 759–772 (2015)
34. Zienkiwicz, O.C., Zhu, J.Z.: The superconvergence patch recovery and a posteriori error estimates. Int. J. Numer. Methods Eng. **33**, 1331–1382 (1992)
35. Zienkiwicz, O.C., Zhu, J.Z.: The superconvergence patch recovery (SPR) and adaptive finite element refinement. Comput. Methods Appl. Math. **101**, 207–224 (1992)