J. Nonlinear Funct. Anal. 2025 (2025) 25 https://doi.org/10.23952/jnfa.2025.25

MATHRES
Journal of Nonlinear Functional Analysis
Available online at http://jnfa.mathres.org

ERROR ANALYSIS OF A BDF2 SCHEME COMBINED WITH FINITE ELEMENTS FOR PARABOLIC OPTIMAL CONTROL PROBLEMS

JUN PAN[1], YUELONG TANG[2], YUCHUN HUA[2,*]

[1]School of Foundation Studies, Zhejiang Pharmaceutical University, Ningbo 315500, China
[2]College of Science, Hunan University of Science and Engineering, Yongzhou 425100, China

Abstract. This paper proposes a second order convergent discretization scheme in both space and time for parabolic optimal control problems (OCPs). For the state and co-state, the BDF2 scheme and finite elements (FEs) are used for the temporal and spatial discretization, respectively. The control is obtained by variational discretization. The second-order convergence results of all variables in the L^2-norm are rigorously derived. Superconvergence between the projections of the state and co-state and their numerical solutions is established. Two numerical examples are provided to confirm the theoretical analysis results.
Keywords. Error analysis; Finite elements; Parabolic optimal control problems.
2020 MSC. 49J20, 65M30.

1. INTRODUCTION

Constrained parabolic OCPs plays an increasingly important role in economics, biologic, social sciences, engineering physics, etc. Numerical solutions of parabolic OCPs were studied extensively. A systematic introduction can be found in [1, 2, 3, 4]. Many numerical approaches were successfully applied to solve parabolic OCPs, such as finite difference [5], multigrid method [6, 7], FE [8, 9], mixed FE [10, 11, 12], immersed FE [13], least-squares FE [14], finite volume element [15, 16], spectral method [17, 18], and so on. However, due to the low regularity of the control variables, most of the above methods only obtain first-order convergence results $\mathscr{O}(k+h)$ by using backward difference and piecewise constant functions discrete time and space variables, respectively. To improve the accuracy of the FE for solving elliptic OCPs, Meyer and Rösch [19] obtained a superconvergence result $\mathscr{O}(h^2)$ by post-processing

*Corresponding author.
E-mail address: yuchunhua@huse.edu.cn (Y. Hua).
Received February 14, 2025; Accepted August 31, 2025.

techniques, and Hinze [20] obtained a convergence result $\mathscr{O}(h^2)$ by using the variational discretization (VD). Recently, in order to improve the time error to $\mathscr{O}(k^2)$, the Crank-Nicolson scheme [21, 22, 23] and the BDF2 scheme [24] have been used to solve parabolic OCPs.

The primary contributions of this paper are threefold: (1) Design of a combined BDF2 scheme and VD approximation for parabolic OCPs with control constraints; (2) Rigorous second-order convergence results in L^2-norm for control, state and co-state are achieved in both time and space; (3) Derivation of superconvergence for the state and co-state in the spatial H^1-norm.

We focus on the following parabolic OCPs:

$$J(u,y) = \min_{u\in U_{ad}} \frac{1}{2}\int_0^T \left(\|u\|^2 + \|y-y_d\|^2\right) dt \tag{1.1}$$

subject to

$$\begin{cases} y_t - \operatorname{div}(\boldsymbol{A}(\boldsymbol{x})\nabla y) + c(\boldsymbol{x})y = u + f, & \forall t\in I, \boldsymbol{x}\in\Omega, \\ y(t,\boldsymbol{x}) = 0, & \forall t\in I, \boldsymbol{x}\in\partial\Omega, \\ y(0,\boldsymbol{x}) = y_0(\boldsymbol{x}), & \forall \boldsymbol{x}\in\Omega, \end{cases} \tag{1.2}$$

where $I=(0,T]$ with $T>0$ and $\Omega\subset\mathbb{R}^n\,(n=2,3)$ is a bounded convex polygon or polyhedron with boundary $\partial\Omega$. Let $c^*\ge c(\boldsymbol{x})\ge c_*>0$, $y_d, f\in\widetilde{U}$ with $\widetilde{U}=L^2(I;U)$ and $U=L^2(\Omega)$, $\boldsymbol{A}(\boldsymbol{x})=(a_{ij}(\boldsymbol{x}))_{n\times n}\in W^{1,\infty}(\Omega)^{n\times n}$ be a symmetric and positive definite function matrix satisfying $c^*\|\boldsymbol{X}\|^2\ge \boldsymbol{A}^T AX\ge c_*\|\boldsymbol{X}\|^2$, $\forall \boldsymbol{X}\in\mathbb{R}^n$. The control set is defined by

$$U_{ad} = \left\{u\in\widetilde{U} : u^*\ge u(t,\boldsymbol{x})\ge u_*, \quad \text{a.e. in } I\times\Omega, u^*, u_*\in\mathbf{R}\right\}. \tag{1.3}$$

From now on, we denote standard Sobolev spaces on Ω by $W^{m,p}(\Omega)$ with a semi-norm $|\cdot|_{m,p}$ and a norm $\|\cdot\|_{m,p}$. For $p=2$, we set $H^m(\Omega)=W^{m,2}(\Omega)$, $H_0^1(\Omega)=\{\varphi\in H^1(\Omega):\varphi|_{\partial\Omega}=0\}$, $\|\cdot\|_m=\|\cdot\|_{m,2}$, and $\|\cdot\|=\|\cdot\|_{0,2}$. Let $L^s(I;W^{m,p}(\Omega))$ be the Banach space of all L^s integrable functions from I into $W^{m,p}(\Omega)$ with norm

$$\|\varphi\|_{L^s(I;W^{m,p}(\Omega))} = \left(\int_0^T \|\varphi\|^s_{W^{m,p}(\Omega)}dt\right)^{\frac{1}{s}} \text{ for } s\in[1,\infty).$$

In addition, $C>0$ denotes a generic constant independent of the mesh-size k or h.

This article is structured as follows. In Section 2, we construct a combined BDF2 scheme and FEs discretization for problem (1.1)-(1.3). We derive second-order convergence results of all variables by introducing some projection operators and auxiliary variables in Section 3. We study superconvergence between numerical solutions and projections of the state and co-state in Section 4. To support our theoretical results, two examples are provided in Section 5, the last section.

2. BDF2 with FEs Approximation of Parabolic OCPs

We construct a combined BDF2 scheme and FEs discretization of (1.1)-(1.3) in this section. To fix the idea, we set $\widetilde{V}=L^2(I;V)$ with $V=H_0^1(\Omega)$ and

$$K = \{u\in U : u^*\ge u(\boldsymbol{x})\ge u_*, \quad \text{a.e. in } \Omega\}.$$

Moreover,

$$a(v,w) = \int_\Omega [v\cdot v\cdot w + (\boldsymbol{A}\nabla v)\cdot\nabla w], \quad \forall v,w\in V,$$
$$(v,w) = \int_\Omega v\cdot w, \quad \forall v,w\in U.$$

According to the assumptions on $\boldsymbol{A}(\boldsymbol{x})$ and $c(\boldsymbol{x})$, we have

$$|a(v,w)| \le C\|v\|_1\|w\|_1, \quad \forall v,w\in V$$

and $c_*\|w\|_1^2 \le a(w,w)$ for all $w\in V$. Then a weak form of the parabolic OCPs (1.1)-(1.3) reads:

$$\begin{cases} J(u,y) = \min\limits_{u\in U_{ad}} \dfrac{1}{2}\displaystyle\int_0^T \left(\|u\|^2 + \|y-y_d\|^2\right) dt, \\ (y_t,v) + a(y,v) = (u+f,v), \quad \forall t\in I, v\in V, \\ y(0,\boldsymbol{x}) = y_0(\boldsymbol{x}), \quad \forall \boldsymbol{x}\in\Omega. \end{cases} \tag{2.1}$$

It is known [1] that (2.1) has an unique solution (u,y), and $(u,y)\in U_{ad}\times(\widetilde{V}\cap H^1(I;V))$ fulfills (2.1) if and only if there exists a co-state $z\in(\widetilde{V}\cap H^1(I;V))$ such that (u,y,z) satisfies:

$$(y_t,v) + a(y,v) = (u+f,v), \quad \forall t\in I, v\in V, \tag{2.2}$$
$$y(0,\boldsymbol{x}) = y_0(\boldsymbol{x}), \quad \boldsymbol{x}\in\Omega, \tag{2.3}$$
$$-(z_t,w) + a(w,z) = (y-y_d,w), \quad \forall t\in I, w\in V, \tag{2.4}$$
$$z(T,\boldsymbol{x}) = 0, \quad \boldsymbol{x}\in\Omega, \tag{2.5}$$
$$(z+u, v-u) \ge 0, \quad \forall t\in I, v\in K. \tag{2.6}$$

As in [20], variational inequality (2.6) equals to $u = \max\{u_*, \min\{u^*, -z\}\}$. Partition I into $0 = t_0 < t_1 < \cdots < t_N = T$ with $k = T/N$ and $t_n = nk$, $n = 0,1,\cdots,N$. For any function $\psi(t,\boldsymbol{x})$, we set $\psi^n = \psi(t_n,\boldsymbol{x})$,

$$d_l\psi^n = \psi^n - \psi^{n-l}, l = 1,2,$$
$$D_t^+\psi^n = \frac{3\psi^n - 4\psi^{n-1} + \psi^{n-2}}{2k},$$
$$D_t^-\psi^n = -\frac{3\psi^n - 4\psi^{n+1} + \psi^{n+2}}{2k}$$

and time-dependent discrete norms

$$\|\psi\|_{l^s(I;W^{m,p}(\Omega))} = \left(\sum_{n=1-l}^{N-l} k\|\psi^n\|_{m,p}^s\right)^{1/s}, \quad 1\le s<\infty$$

with standard modification for $s=\infty$, where $l=0$ or 1. Furthermore, we set $\|\cdot\|_{\boldsymbol{\Theta}} = \|\cdot\|_{l^2(I;L^2(\Omega))}$, $\|\cdot\|_{\boldsymbol{\Xi}} = \|\cdot\|_{l^2(I;H^1(\Omega))}$, and $\|\cdot\|_{\boldsymbol{\Gamma}} = \|\cdot\|_{l^\infty(I;L^2(\Omega))}$.

Let $\mathscr{T}_h$ be regular triangular subdivision of Ω and $h = \max\limits_{E\in\mathscr{T}_h}\{h_E\}$ with $h_E = \mathrm{diam}(E)$. The piecewise linear function space [25] associated with $\mathscr{T}_h$ can be defined as

$$V_h := \left\{v_h\in C(\bar{\Omega}) : v_h|_E\in P_1(E), v_h|_{\partial\Omega} = 0, \forall E\in\mathscr{T}_h\right\}.$$

Then a combined BDF2 scheme and FEs discretization of (2.1) is as follows:

$$\begin{cases} J(u_h,y_h)=\min\limits_{u_h^n\in K}\frac{k}{2}\sum\limits_{n=0}^{N-1}\left(\|u_h^n\|^2+\|y_h^n-y_d^n\|^2\right), \\ \left(D_t^+y_h^n,v_h\right)+a\left(y_h^n,v_h\right)=\left(u_h+f^n,v_h\right), \quad \forall v_h\in V_h, n=1,2,\cdots,N, \\ y_h^0=P_hy_0, \end{cases} \tag{2.7}$$

where y_h^{-1} and P_h will be specific later. It follows from [1, 20, 24] that (2.7) has an unique solution $\left(u_h^n,y_h^n\right), n=0,1,\cdots,N$, and $\left(u_h^n,y_h^n\right)\in K\times V_h, n=0,1,\cdots,N$ fulfills (2.7) if and only if there exists a co-state $z_h^n\in V_h, n=N,\cdots,1,0$ such that $\left(u_h^n,y_h^n,z_h^n\right)$ satisfies:

$$\left(D_t^+y_h^n,v_h\right)+a\left(y_h^n,v_h\right)=\left(u_h^n+f^n,v_h\right), \quad \forall v_h\in V_h, n=1,2,\cdots,N, \tag{2.8}$$

$$y_h^0=P_hy_0, \tag{2.9}$$

$$-\left(D_t^-z_h^n,w_h\right)+a\left(w_h,z_h^n\right)=\left(y_h^n-y_d^n,w_h\right), \quad \forall w_h\in V_h, n=N-1,\cdots,1,0, \tag{2.10}$$

$$z_h^N=0, \tag{2.11}$$

$$\left(z_h^n+u_h^n,v-u_h^n\right)\geq 0, \quad \forall v\in K, n=0,1,\cdots,N. \tag{2.12}$$

This process is not self-starting. Like [24, 26], we can add initial conditions $y_h^{-1}=P_h\left(-kf^0\right)$ and $z_h^{N+1}=P_h\left(-k\left(y_h^N-y_d^N\right)\right)$. Similar to (2.6), inequality (2.12) is equivalent to

$$u_h^n=\max\left\{u_*,\min\left\{u^*,-z_h^n\right\}\right\}, \quad n=0,1,\cdots,N.$$

3. Convergence

We derive convergence of (2.8)-(2.12) in this section. For the need of error analysis later, we introduce the projection operator [25] $P_h: V\to V_h$ with

$$\begin{aligned} &a(P_hv-v,v_h)=0, \quad \forall v_h\in V_h, v\in V, \\ &\|P_hv-v\|+h\|\nabla(P_hv-v)\|\leq Ch^2\|v\|_2, \quad \forall v\in H^2(\Omega) \end{aligned} \tag{3.1}$$

and auxiliary variable $\left(y_h^n(u),z^n(u)\right)\in V_h\times V_h$, $n=0,1,\cdots,N$ fulfills

$$\left(D_t^+y_h^n(u),v_h\right)+a\left(y_h^n(u),v_h\right)=\left(u^n+f^n,v_h\right), \quad \forall v_h\in V_h, n=1,2,\cdots,N, \tag{3.2}$$

$$y_h^0(u)=y_h^0, y_h^{-1}(u)=y_h^{-1}, \tag{3.3}$$

$$-\left(D_t^-z_h^n(u),w_h\right)+a\left(w_h,z_h^n(u)\right)=\left(y_h^n(u)-y_d^n,w_h\right), \quad \forall w_h\in V_h, n=N-1,\cdots,1,0, \tag{3.4}$$

$$z_h^N(u)=0, z_h^{N+1}(u)=z_h^{N+1}. \tag{3.5}$$

Lemma 3.1. *Let* $\left(y_h^n,z_h^n\right)$ *and* $\left(y_h^n(u),z_h^n(u)\right)$ *with* $n=0,1,\cdots,N$ *be the solutions of (2.8)-(2.12) and (3.2)-(3.5), respectively. Then*

$$\|y_h-y_h(u)\|_{\mathbf{\Gamma}}+\|y_h-y_h(u)\|_{\mathbf{\Xi}}\leq C\|u_h-u\|_{\mathbf{\Theta}}. \tag{3.6}$$

$$\|z_h-z_h(u)\|_{\mathbf{\Gamma}}+\|z_h-z_h(u)\|_{\mathbf{\Xi}}\leq C\|u_h-u\|_{\mathbf{\Theta}}. \tag{3.7}$$

Proof. For simplicity, we set

$$\begin{aligned} \alpha_y^n &= y_h^n-y_h^n(u), n=-1,0,\cdots,N, \\ \beta_z^n &= z_h^n-z_h^n(u), n=0,1,\cdots,N+1. \end{aligned}$$

From (2.8)-(2.12) and (3.2)-(3.5), we have

$$\begin{aligned}\left(D_t^+\alpha_y^n, v_h\right) + a\left(\alpha_y^n, v_h\right) = (u_h^n - u^n, v_h), \quad \forall v_h \in V_h, n = 1,2,\cdots,N,\\ \alpha_y^0 = 0,\ \alpha_y^{-1} = 0,\end{aligned} \tag{3.8}$$

$$\begin{aligned}-\left(D_t^-\beta_z^n, w_h\right) + a\left(w_h, \beta_z^n\right) = \left(\alpha_y^n, w_h\right), \quad \forall w_h \in V_h, n = N-1,\cdots,1,0,\\ \beta_z^N = 0,\ \beta_z^{N+1} = 0.\end{aligned} \tag{3.9}$$

According to the definitions of $d_l\psi^n$ and $D_t^+\psi^n$, we obtain

$$2\left(d_l\alpha_y^n, \alpha_y^n\right) = d_l\|\alpha_y^n\|^2 + \|d_l\alpha_y^n\|^2, \quad l = 1,2$$

and

$$\begin{aligned}k\left(D_t^+\alpha_y^n, \alpha_y^n\right) =& 2\left(d_1\alpha_y^n, \alpha_y^n\right) - \frac{1}{2}\left(d_2\alpha_y^n, \alpha_y^n\right)\\ =& d_1\|\alpha_y^n\|^2 - \frac{1}{4}d_2\|\alpha_y^n\|^2 + \|d_1\alpha_y^n\|^2 - \frac{1}{4}\|d_2\alpha_y^n\|^2.\end{aligned} \tag{3.10}$$

Summing n from 1 to $\tilde{N}\,(\tilde{N} \le N)$, we have

$$\sum_{n=1}^{\tilde{N}}\left(d_1\|\alpha_y^n\|^2 - \frac{1}{4}d_2\|\alpha_y^n\|^2\right) = \frac{3}{4}\|\alpha_y^{\tilde{N}}\|^2 - \frac{1}{4}\|\alpha_y^{\tilde{N}-1}\|^2 - \frac{3}{4}\|\alpha_y^0\|^2 + \frac{1}{4}\|\alpha_y^{-1}\|^2. \tag{3.11}$$

Since $d_2\alpha_y^n = d_1\alpha_y^n + d_1\alpha_y^{n-1}$, we derive

$$\begin{aligned}\sum_{n=1}^{\tilde{N}}\left(\|d_1\alpha_y^n\|^2 - \frac{1}{4}\|d_2\alpha_y^n\|^2\right) \ge& \sum_{n=1}^{\tilde{N}}\left(\|d_1\alpha_y^n\|^2 - \frac{1}{4}\left(\|d_1\alpha_y^n\|^2 + \|d_1\alpha_y^{n-1}\|^2\right)\right)\\ \ge& \frac{1}{2}\sum_{n=1}^{\tilde{N}}\left(\|d_1\alpha_y^n\|^2 - \|d_1\alpha_y^{n-1}\|^2\right)\\ \ge& -\|\alpha_y^0\|^2 - \|\alpha_y^{-1}\|^2.\end{aligned} \tag{3.12}$$

In view of $\alpha_y^0 = \alpha_y^{-1} = 0$, (3.10)-(3.12) yield

$$k\sum_{n=1}^{\tilde{N}}\left(D_t^+\alpha_y^n, \alpha_y^n\right) \ge \frac{3}{4}\|\alpha_y^{\tilde{N}}\|^2 - \frac{1}{4}\|\alpha_y^{\tilde{N}-1}\|^2. \tag{3.13}$$

Taking $v_h = \alpha_y^n$ in (3.8) and multiplying both sides of (3.8) with $\frac{4k}{3}$, then using (3.13) and ε-Cauchy inequality, we arrive at

$$\|\alpha_y^{\tilde{N}}\|^2 + \frac{4c}{3}\sum_{n=1}^{\tilde{N}} k\|\alpha_y^n\|_1^2 \le \frac{1}{3}\|\alpha_y^{\tilde{N}-1}\|^2 + \frac{4}{3}C(\varepsilon)\sum_{n=1}^{\tilde{N}} k\|u_h^n - u^n\|^2 + \frac{4\varepsilon}{3}\sum_{n=1}^{\tilde{N}} k\|\alpha_y^n\|^2. \tag{3.14}$$

Then (3.6) follows from (3.14) and the Poincaré inequality. Choosing $w_h = \alpha_z^n$ in (3.9), we can derive (3.7) from (3.6) and (3.9) analogously. □

Lemma 3.2. *Let (u,y,z) and $\left(y_h^n(u), z_h^n(u)\right)$ with $n = 0,1,\cdots,N$ be the solutions of (2.2)-(2.6) and (3.2)-(3.5), respectively. Suppose that $y,z,f \in l^2(I;H^2(\Omega))$, $y_t, z_t \in L^2(I;H^2(\Omega))$, $y_{ttt}, z_{ttt} \in$*

$l^2(I;L^2(\Omega))$ and $y_0 \in H^2(\Omega)$. Then

$$\|y-y_h(u)\|_{\Gamma}+\|z-z_h(u)\|_{\Gamma} \le C\left(k^2+h^2\right), \tag{3.15}$$

$$\|y-y_h(u)\|_{\Xi}+\|z-z_h(u)\|_{\Xi} \le C\left(k^2+h\right). \tag{3.16}$$

Proof. For convenience, we set

$$\begin{aligned}
&\eta_y^n = y^n - P_h y^n,\ \theta_y^n = P_h y^n - y_h^n(u), n=-1,0,\cdots,N,\\
&\eta_z^n = z^n - P_h z^n,\ \theta_z^n = P_h z^n - z_h^n(u), n=0,1,\cdots,N+1.
\end{aligned}$$

Choosing $t=t_n$ in (2.2), subtracting (3.2), then utilizing the definition of P_h, we obtain

$$\left(D_t^+\theta_y^n, v_h\right)+a\left(\theta_y^n, v_h\right)=\left(D_t^+y^n-y_t^n, v_h\right)-\left(D_t^+\eta_y^n, v_h\right), \quad \forall v_h\in V_h, n=1,2,\cdots,N. \tag{3.17}$$

Similar to (3.13), we can derive

$$k\sum_{n=1}^{\tilde N}\left(D_t^+\theta_y^n,\theta_y^n\right) \ge \frac{3}{4}\|\theta_y^{\tilde N}\|^2-\frac{1}{4}\|\theta_y^{\tilde N-1}\|^2-\frac{7}{4}\|\theta_y^0\|^2-\frac{3}{4}\|\theta_y^{-1}\|^2. \tag{3.18}$$

From Taylor expansions, we obtain

$$D_t^+y^n-y_t^n=\frac{3y^n-4y^{n-1}+y^{n-2}}{2k}-y_t^n=\frac{k^2}{3}y_{ttt}^{n-1}+\mathscr{O}(k^3). \tag{3.19}$$

Moreover,

$$\sum_{n=1}^{\tilde N}k\|D_t^+y^n-y_t^n\|^2 \le Ck^4\sum_{n=1}^{\tilde N}k\|y_{ttt}^{n-1}\|^2 \le Ck^4\|y_{ttt}\|_{\Theta}^2. \tag{3.20}$$

From the definition of $D_t^+\eta_y^n=\frac{3d_1\eta_y^n-d_1\eta_y^{n-1}}{2k}$, (3.1), and the Cauchy inequality, we arrive at

$$\begin{aligned}
\sum_{n=1}^{\tilde N}k\|D_t^+\eta_y^n\|^2 \le&\frac{3}{2}\sum_{n=1}^{\tilde N}k\left\|\frac{d_1\eta_y^n}{k}\right\|^2+\frac{1}{2}\sum_{n=1}^{\tilde N}k\left\|\frac{d_1\eta_y^{n-1}}{k}\right\|^2\\
\le&\frac{3k^2}{2}\sum_{n=1}^{\tilde N}\int_{t_{n-1}}^{t_n}\|(\eta_y)_t\|^2dt+\frac{k^2}{2}\sum_{n=1}^{\tilde N}\int_{t_{n-2}}^{t_{n-1}}\|(\eta_y)_t\|^2dt\\
\le&Ch^4\Big(\sum_{n=1}^{\tilde N}\int_{t_{n-1}}^{t_n}\|y_t\|_2^2dt+\sum_{n=1}^{\tilde N}\int_{t_{n-2}}^{t_{n-1}}\|y_t\|_2^2dt\Big)\\
\le&2Ch^4\|y_t\|_{L^2(I;H^2(\Omega))}^2.
\end{aligned} \tag{3.21}$$

Thus

$$
\begin{aligned}
\|\theta_y^{\tilde N}\|^2+\frac{4c_*}{3}\sum_{n=1}^{\tilde N}k\|\theta_y^n\|_1^2 \le& \frac{4}{3}C(\varepsilon)\sum_{n=1}^{\tilde N}k\|D_t^+y^n-y_t^n\|^2+\frac{4}{3}C(\varepsilon)\sum_{n=1}^{\tilde N}k\|D_t^+\eta_y^n\|^2+\frac{8\varepsilon}{3}\sum_{n=1}^{\tilde N}\|\eta_y^n\|^2\\
&+\frac{1}{3}\|\theta_y^{\tilde N-1}\|^2+\frac{7}{3}\|\theta_y^0\|^2+\|\theta_y^{-1}\|^2\\
\le& \frac{4}{3}C(\varepsilon)\left(k^4\|y_{ttt}\|_{\mathbf{\Theta}}^2+h^4\left(\|y_t\|^2_{L^2(I;H^2(\Omega))}+\|y_0\|_2^2+\|f^0\|_2^2\right)\right)\\
&+\frac{1}{3}\|\theta_y^{\tilde N-1}\|^2+\frac{8\varepsilon}{3}\sum_{n=1}^{\tilde N}\|\eta_y^n\|^2.
\end{aligned}
\tag{3.22}
$$

Let ε be small enough. From (3.22) and Poincaré inequality, we obtain that

$$\|\theta_y\|_{\mathbf{\Gamma}}+\|\theta_y\|_{\mathbf{\Xi}}\le C\left(k^2+h^2\right). \tag{3.23}$$

From (3.1), (3.23), and the triangle inequality, we find that

$$\|y-y_h(u)\|_{\mathbf{\Gamma}}\le\|\eta_y\|_{\mathbf{\Gamma}}+\|\theta_y\|_{\mathbf{\Gamma}}\le C(k^2+h^2) \tag{3.24}$$

and

$$\|y-y_h(u)\|_{\mathbf{\Xi}}\le\|\eta_y\|_{\mathbf{\Xi}}+\|\theta_y\|_{\mathbf{\Xi}}\le C(k^2+h). \tag{3.25}$$

Selecting $t=t_n$ in (2.4) and subtracting (3.4), we have

$$
\begin{aligned}
-\left(D_t^-\theta_z^n,w_h\right)+a\left(w_h,\theta_z^n\right)=&\left(D_t^-\eta_z^n,w_h\right)+\left(z_t^n-D_t^-z^n,w_h\right)\\
&+\left(\eta_y^n,w_h\right)+\left(\theta_y^n,w_h\right),\quad \forall w_h\in V_h, n=N-1,\cdots,1,0.
\end{aligned}
$$

Similar to the derivation process of (3.23), we can obtain

$$\|\theta_z\|_{\mathbf{\Gamma}}+\|\theta_z\|_{\mathbf{\Xi}}\le C\left(k^2+h^2\right). \tag{3.26}$$

From the triangle inequality, (3.1), and (3.26), we derive

$$\|z-z_h(u)\|_{\mathbf{\Gamma}}\le\|\eta_z\|_{\mathbf{\Gamma}}+\|\theta_z\|_{\mathbf{\Gamma}}\le C(k^2+h^2) \tag{3.27}$$

and

$$\|z-z_h(u)\|_{\mathbf{\Xi}}\le\|\eta_z\|_{\mathbf{\Xi}}+\|\theta_z\|_{\mathbf{\Xi}}\le C(k^2+h). \tag{3.28}$$

Then (3.15)-(3.16) follows from (3.24)-(3.25) and (3.27)-(3.28) immediately. □

Theorem 3.3. *Let (u,y,z) and (u_h,y_h,z_h) be the solutions of (2.2)-(2.6) and (2.8)-(2.12), respectively. Suppose that conditions in Lemma 3.2 are satisfied. Then*

$$\|u-u_h\|_{\mathbf{\Theta}}\le C(k^2+h^2), \tag{3.29}$$

$$\|y-y_h\|_{\mathbf{\Gamma}}+\|z-z_h\|_{\mathbf{\Gamma}}\le C\left(k^2+h^2\right), \tag{3.30}$$

$$\|y-y_h\|_{\mathbf{\Xi}}+\|z-z_h\|_{\mathbf{\Xi}}\le C\left(k^2+h\right). \tag{3.31}$$

Proof. Taking $v = u_h$ with $t = t_n$ in (2.6) and $v = u^n$ in (2.12), for $n = 0, 1, \cdots, N$, we have $\left(z^n + u^n, u_h^n - u^n\right) \geq 0$ and $\left(z_h^n + u_h^n, u^n - u_h^n\right) \geq 0$. Thus

$$\begin{aligned}\|u-u_h\|_{\Theta}^2 &= \sum_{n=1}^{N} k\left(u^n-u_h^n, u^n-u_h^n\right) \\ &\leq \sum_{n=1}^{N} k\left(z_h^n(u)-z^n, u^n-u_h^n\right) + \sum_{n=1}^{N} k\left(z_h^n - z_h^n(u), u^n-u_h^n\right).\end{aligned} \tag{3.32}$$

According to (2.8)-(2.11) and (3.2)-(3.5), we have

$$\sum_{n=1}^{N} k\left(z_h^n - z_h^n(u), u^n - u_h^n\right) = -\|y_h - y_h(u)\|^2 \leq 0. \tag{3.33}$$

From (3.15) and the ε-Cauchy inequality, we obtain

$$\begin{aligned}\sum_{n=1}^{N} k\left(z_h^n(u)-z^n, u^n-u_h^n\right) \leq & C(\varepsilon)\sum_{n=1}^{N} k\|z_h^n(u)-z^n\|^2 + \varepsilon\sum_{n=1}^{N} k\|u^n-u_h^n\|^2 \\ \leq & C(\varepsilon)\left(k^2+h^2\right)^2 + \varepsilon\sum_{n=1}^{N} k\|u^n-u_h^n\|^2.\end{aligned} \tag{3.34}$$

Then (3.29) follows from (3.32)-(3.34) immediately. From Lemmas 3.1-3.2, the triangle inequality, and (3.29), it is easy to obtain (3.30)-(3.31). □

4. SUPERCONVERGENCE

Superconvergence between the projections and numerical solutions of the state and co-state is analyzed in this section.

Theorem 4.1. *Let (u, y, z) and (u_h, y_h, z_h) be the solutions of (2.2)-(2.6) and (2.8)-(2.12), respectively. Suppose that conditions in Lemma 3.2 are satisfied. Then*

$$\|P_h y - y_h\|_{\Xi} + \|P_h z - z_h\|_{\Xi} \leq C\left(k^2 + h^2\right). \tag{4.1}$$

Proof. For the sake of simplicity, we set

$$\rho_y^n = P_h y^n - y_h^n,\ \rho_z^n = P_h z^n - z_h^n,\ n = 0, 1, \cdots, N.$$

From (2.2)-(2.5) and (2.8)-(2.11), we find by the definition of P_h that

$$\left(D_t^+\rho_y^n, v_h\right) + a\left(\rho_y^n, v_h\right) = \left(D_t^+ y^n - y_t^n, v_h\right) + \left(u^n - u_h^n, v_h\right) + \left(D_t^+\eta_y^n, v_h\right),$$
$$\forall v_h \in V_h, n = 1, 2, \cdots, N, \tag{4.2}$$

$$\rho_y^0 = 0,\ \rho_y^{-1} = 0, \tag{4.3}$$

$$-\left(D_t^-\rho_z^n, w_h\right) + a\left(w_h, \rho_z^n\right) = \left(z_t^n - D_t^- z^n, w_h\right) + \left(D_t^-\eta_z^n, w_h\right) + \left(\eta_y^n, w_h\right) + \left(\rho_y^n, w_h\right),$$
$$\forall w_h \in V_h, n = N-1, \cdots, 1, 0, \tag{4.4}$$

$$\rho_z^N = 0,\ \rho_z^{N+1} = 0. \tag{4.5}$$

Note that $\rho_y^0 = \rho_y^{-1} = 0$. As (3.13), we obtain

$$k\sum_{n=1}^{\tilde{N}}\left(D_t^+\rho_y^n, \rho_y^n\right) \geq \frac{3}{4}\|\rho_y^{\tilde{N}}\|^2 - \frac{1}{4}\|\rho_y^{N^*-1}\|^2. \tag{4.6}$$

Combing (3.20)-(3.21), (3.29), (4.2)-(4.3), (4.6), and the ε-Cauchy inequality, we obtain

$$\begin{aligned}\frac{3}{4}\|\rho_y^{\tilde{N}}\|^2+c_*\sum_{n=1}^{\tilde{N}}k\|\rho_y^n\|_1^2 \leq & C(\varepsilon)\left(k^4\|y_{ttt}\|_{\boldsymbol{\Theta}}^2+h^4\|y_t\|^2_{L^2(I;H^2(\Omega))}+\left(k^2+h^2\right)^2\right)\\ &+\frac{1}{4}\|\rho_y^{\tilde{N}-1}\|^2+3\varepsilon\|\rho_y\|_{\boldsymbol{\Theta}}^2.\end{aligned} \tag{4.7}$$

Let ε be small enough. It follows from the Poincaré inequality and (4.7) that

$$\|P_hy-y_h\|_{\boldsymbol{\Xi}} \leq C\left(k^2+h^2\right). \tag{4.8}$$

Analogously, according to (3.1), (4.4)-(4.5), (4.8), and the Poincaré inequality, we derive

$$\|P_hz-z_h\|_{\boldsymbol{\Xi}} \leq C\left(k^2+h^2\right). \tag{4.9}$$

Then (4.1) follows from (4.8)-(4.9). □

5. Numerical Experiments

We perform some experiments to support the previous theoretical analysis in this section. Let $T=1$, $\Omega=(0,1)\times(0,1)$, $\boldsymbol{A}(\boldsymbol{x})$ to be the identity matrix and $c(\boldsymbol{x})=1$. The numerical example is realized by AFEPack [27]. The discretization scheme is as described in (2.8)-(2.12).

Example 1. The data is given by:

$$\begin{aligned}&u_*=-0.25,\ u^*=0.5,\\ &y(t,\boldsymbol{x})=t\sin(2\pi x_1)\sin(2\pi x_2),\\ &z(t,\boldsymbol{x})=(1-t)\sin(2\pi x_1)\sin(2\pi x_2),\\ &u(t,\boldsymbol{x})=\max\{u_*,\min\{u^*,-z(t,\boldsymbol{x})\}\},\\ &f(t,\boldsymbol{x})=-\mathrm{div}(\nabla y(t,\boldsymbol{x}))+y_t(t,\boldsymbol{x})+y(t,\boldsymbol{x})-u(t,\boldsymbol{x}),\\ &y_d(t,\boldsymbol{x})=z_t(t,\boldsymbol{x})+\mathrm{div}(\nabla z(t,\boldsymbol{x}))-z(t,\boldsymbol{x})+y(t,\boldsymbol{x}).\end{aligned}$$

Fixed $h=\frac{1}{100}$, errors $\|u-u_h\|_{\boldsymbol{\Theta}}$, $\|y-y_h\|_{\boldsymbol{\Gamma}}$, $\|z-z_h\|_{\boldsymbol{\Gamma}}$, $\|y-y_h\|_{\boldsymbol{\Xi}}$, $\|z-z_h\|_{\boldsymbol{\Xi}}$, $\|P_hy-y_h\|_{\boldsymbol{\Xi}}$ and $\|P_hz-z_h\|_{\boldsymbol{\Xi}}$ based on $k=\frac{1}{10},\frac{1}{20},\frac{1}{40},\frac{1}{80}$ are presented in Table 1. Temporal error convergent rates are reported in Figure 1. Fixed $k=\frac{1}{100}$, errors based on $h=\frac{1}{10},\frac{1}{20},\frac{1}{40},\frac{1}{80}$ are provided in Table 2. Spatial error convergent rates are reported in Figure 2.

TABLE 1. Errors of Example 1 with $h=\frac{1}{100}$.

k	1/10	1/20	1/40	1/80
$\|u-u_h\|_{\boldsymbol{\Theta}}$	4.3561e-02	1.1092e-02	2.7730e-03	6.9325e-04
$\|y-y_h\|_{\boldsymbol{\Gamma}}$	3.5145e-02	8.8053e-03	2.2013e-03	5.5033e-04
$\|z-z_h\|_{\boldsymbol{\Gamma}}$	4.0137e-02	1.1034e-02	2.7585e-03	6.8963e-04
$\|y-y_h\|_{\boldsymbol{\Xi}}$	8.8053e-02	2.2135e-02	5.5338e-03	1.3834e-03
$\|z-z_h\|_{\boldsymbol{\Xi}}$	9.2405e-02	2.3285e-02	5.8213e-03	1.4553e-03
$\|P_hy-y_h\|_{\boldsymbol{\Xi}}$	7.3163e-02	1.8307e-02	4.5768e-03	1.1442e-03
$\|P_hz-z_h\|_{\boldsymbol{\Xi}}$	8.5268e-02	2.1542e-02	5.3855e-03	1.3464e-03

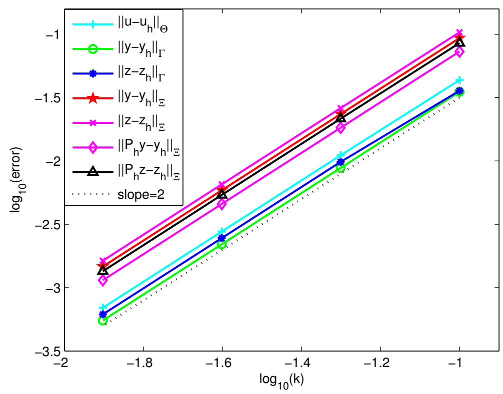

FIGURE 1. Temporal error convergence rates of Example 1 with $h = \frac{1}{100}$.

TABLE 2. Errors of Example 1 with $k = \frac{1}{100}$.

h	$1/10$	$1/20$	$1/40$	$1/80$
$\|u - u_h\|_{\mathbf{\Theta}}$	5.3078e-02	1.3145e-02	3.2863e-03	8.2156e-04
$\|y - y_h\|_{\mathbf{\Gamma}}$	3.9275e-02	9.8187e-03	2.4547e-03	6.1367e-04
$\|z - z_h\|_{\mathbf{\Gamma}}$	5.1434e-02	1.2858e-02	3.2146e-03	8.0365e-04
$\|y - y_h\|_{\mathbf{\Xi}}$	2.8646e-02	1.4323e-02	7.1614e-03	3.5807e-03
$\|z - z_h\|_{\mathbf{\Xi}}$	2.9263e-02	1.4631e-02	7.3156e-03	3.6578e-03
$\|P_h y - y_h\|_{\mathbf{\Xi}}$	1.4912e-02	3.7257e-03	9.3141e-04	2.3285e-04
$\|P_h z - z_h\|_{\mathbf{\Xi}}$	1.6535e-02	4.1086e-03	1.0272e-03	2.5679e-04

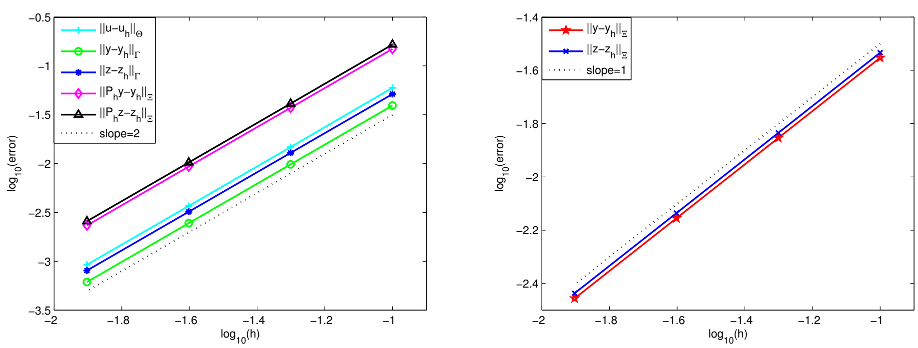

FIGURE 2. Spatial error convergence rates of Example 1 with $k = \frac{1}{100}$.

Example 2. The data is given by:

$$
\begin{aligned}
&u_* = -0.75,\ u^* = 0.75,\\
&y(t,\boldsymbol{x}) = t^2 x_1(x_1 - 1)(1 - x_2)x_2,\\
&z(t,\boldsymbol{x}) = (1-t)^2 x_1(x_1 - 1)(1 - x_2)x_2,\\
&u(t,\boldsymbol{x}) = \max\{u_*, \min\{u^*, -z(t,\boldsymbol{x})\}\},\\
&f(t,\boldsymbol{x}) = -\mathrm{div}(\nabla y(t,\boldsymbol{x})) + y_t(t,\boldsymbol{x}) + y(t,\boldsymbol{x}) - u(t,\boldsymbol{x}),\\
&y_d(t,\boldsymbol{x}) = z_t(t,\boldsymbol{x}) + \mathrm{div}(\nabla z(t,\boldsymbol{x})) - z(t,\boldsymbol{x}) + y(t,\boldsymbol{x}).
\end{aligned}
$$

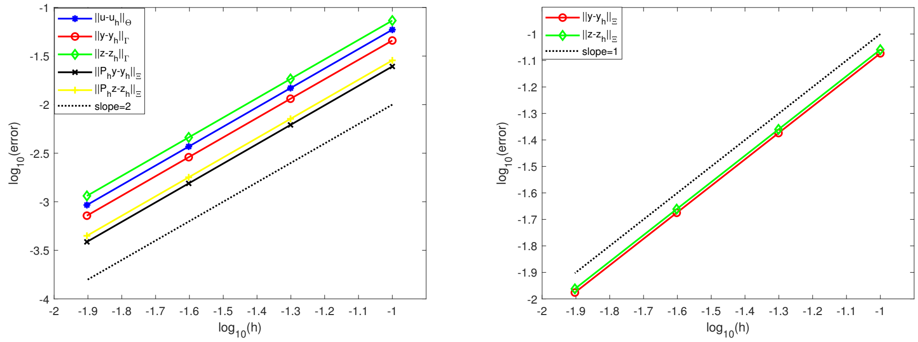

FIGURE 3. Convergence rates of Example 2.

TABLE 3. Errors of Example 2.

$h=k$	$1/10$	$1/20$	$1/40$	$1/80$
$\|u-u_h\|_{\Theta}$	5.9105e-02	1.4806e-02	3.7015e-03	9.2536e-04
$\|y-y_h\|_{\Gamma}$	4.5738e-02	1.1504e-02	2.8760e-03	7.1901e-04
$\|z-z_h\|_{\Gamma}$	5.8304e-02	1.4606e-02	3.6515e-03	9.1287e-04
$\|y-y_h\|_{\Xi}$	3.5882e-02	1.7804e-02	8.9010e-03	4.4485e-03
$\|z-z_h\|_{\Xi}$	3.7560e-02	1.8709e-02	9.5245e-03	4.7515e-03
$\|P_hy-y_h\|_{\Xi}$	2.4728e-02	6.1820e-03	1.5455e-03	3.8637e-04
$\|P_hz-z_h\|_{\Xi}$	2.5457e-02	6.3643e-03	1.5911e-03	3.9765e-04

In this example, we take gradually decreasing mesh sizes $h=\frac{1}{10},\frac{1}{20},\frac{1}{40},\frac{1}{80}$ and the time step size k is taken as $k=h$. Numerical errors and their convergence rates for different h are displayed in Table 3 and Figure 3. It is clear that numerical results are in good agreement with the theoretical results in Theorem 3.3 and Theorem 4.1.

Funding

This work was supported by the Scientific Research Foundation of Hunan Provincial Department of Education (20A211) and the Natural Science Foundation of Hunan Province (2020JJ4323).

REFERENCES

[1] J. Lions, Optimal Control of System Governed by Partial Differential Equations, Springer, Berlin, 1971.

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

[3] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, Optimization with PDE constraints, Springer Science & Business Media, 2008.

[4] Y. Chen, Z. Lu, High Efficient and Accuracy Numercial Methods for Optimal Control Problems, Science Press, Beijing, 2015.

[5] A. Lapin, S. Zhang, S. Lapin, Numerical solution of a parabolic optimal control problem arising in economics and management, Appl. Math. Comput. 361 (2019) 715-729.

[6] A. Borzì, Multigrid methods for parabolic distributed optimal control problems, J. Comput. Appl. Math. 157 (2003) 365-382.

[7] B. Li, J. Liu, M. Xiao, A new multigrid method for unconstrained parabolic optimal control problems, J. Comput. Appl. Math. 326 (2017) 358-373.

[8] Y. Tang, Y. Hua, Superconvergence of fully discrete finite elements for parabolic control problems with integral constraints, East Asian J. Appl. Math. 3 (2013) 138-153.
[9] D. Liang, W. Gong, X. Xie, Finite element error estimation for parabolic optimal control problems with pointwise observations, Numer. Math. Theor. Meth. Appl. 15 (2022) 165-199.
[10] T. Hou, C. Liu, H. Chen, Fully discrete H^1-Galerkin mixed finite element methods for parabolic optimal control problems, Numer. Math. Thero. Meth. Appl. 12 (2019) 134-153.
[11] J. Wang, Z. Lu, F. Cai, Y. Feng, Fully discrete interpolation coefficients mixed finite element methods for semi-linear parabolic optimal control problem, IEEE Access 10 (2022) 54291-54300.
[12] Y. Hua, Y. Tang, Z. Chen, Interpolated coefficient characteristic mixed finite element method for semilinear convection-diffusion optimal control problems, J. Nonlinear Funct. Anal. 2024 (2024) 12.
[13] Z. Zhang, D. Liang, Q. Wang, Immersed finite element method and its analysis for parabolic optimal control problems with interfaces, Appl. Numer. Math. 147 (2020) 174-195.
[14] F. Thomas, K, Michael, Space-time least-squares finite element methods for parabolic distributed optimal control problems, Comput. Meth. Appl. Math. 24 (2024) 673-691.
[15] X. Luo, Y. Chen, Y. Huang and T. Hou, Some error estimates of finite volume element method for parabolic optimal control problems, Optim. Control Appl. Meth. 35 (2014) 145-165.
[16] Z. Lu, R. Xu, C. Hou, L. Xing, A priori error estimates of finite volume element method for bilinear parabolic optimal control problem, AIMS Math. 8 (2023) 19374-19390.
[17] J. Zhou, D. Yang, Legendre-Galerkin spectral methods for optimal control problems with integral constraint for state in one dimension, Comput. Optim. Appl. 61 (2015) 135-158.
[18] Y. Chen, X. Lin and Y. Huang, Error analysis of Galerkin spectral methods for nonlinear optimal control problems with integral control constraint, Commun. Math. Sci. 20 (2022) 1659-1683.
[19] C. Meyer, A. Rösch, Superconvergence properties of optimal control problems, SIAM J. Control Optim. 43 (2004) 970-985.
[20] M. Hinze, A variational discretization conception in control constrained optimization: The linear case, Comput. Optim. Appl. 30 (2005) 45-61.
[21] C. Yang, T. Sun, Crank-Nicolson finite difference schemes for parabolic optimal Dirichlet boundary control problems, Math. Meth. Appl. Sci. 45(12) (2022) 7346-7363.
[22] X. Zhang, J. Zhao, R. Hou, A priori error estimates of Crank-Nicolson finite element method for parabolic optimal control problems, Comput. Math. Appl. 140 (2023) 274-289.
[23] Y. Tang, Y. Hua, Crank-Nicolson splitting positive definite mixed element discretization of parabolic control problems, J. Nonlinear Funct. Anal. 2024 (2024) 1.
[24] C. Yang, T. Sun, BDF2 schemes for optimal parameter control problems governed by bilinear parabolic equations, Optim. Control Appl. Meth. 44 (2023) 2055-2081.
[25] P. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amstterdam, 1978.
[26] J. Liu, M. Xiao, A leapfrog semi-smooth Newton-multigrid method for semilinear parabolic optimal control problems, Comput. Optim. Appl. 63 (2016) 69-95.
[27] R. Li, W. Liu, N. Yan, A posteriori error estimates of recovery type for distributed convex optimal control problems, J. Sci. Comput. 33 (2007) 155-182.