Discontinuous Galerkin method for a distributed optimal control problem governed by a time fractional diffusion equation *

Tao Wang[†1], Binjie Li[‡2], and Xiaoping Xie[§2]

[1]South China Research Center for Applied Mathematics and Interdisciplinary Studies, South China Normal University, Guangzhou 510631, China

[2]School of Mathematics, Sichuan University, Chengdu 610064, China

Abstract

This paper is devoted to the numerical analysis of a control constrained distributed optimal control problem subject to a time fractional diffusion equation with non-smooth initial data. The solutions of state and co-state are decomposed into singular and regular parts, and some growth estimates are obtained for the singular parts. By following the variational discretization concept, a full discretization is applied to the corresponding state and co-state equations by using linear conforming finite element method in space and piecewise constant discontinuous Galerkin method in time. Error estimates are derived by employing the growth estimates. In particular, graded temporal grids are adopted to obtain the first-order temporal accuracy. Finally, numerical experiments are performed to verify the theoretical results.

Keywords: distributed optimal control, time fractional diffusion equation, growth estimate, finite element, discontinuous Galerkin method, error estimate.

1 Introduction

Let $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ be a convex polytope, and assume that $0 < \alpha < 1$, $-\infty < u_* < u^* < \infty$, and $0 < \nu, T < \infty$. We consider the following distributed optimal control problem:

$$\min_{\substack{u \in U_{\rm ad} \\ y \in L^2(0,T;L^2(\Omega))}} J(y, u) = \frac{1}{2}\|y - y_d\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\nu}{2}\|u\|^2_{L^2(0,T;L^2(\Omega))}, \tag{1}$$

*Tao Wang was supported in part by the China Postdoctoral Science Foundation (2019M66294), Binjie Li was supported in part by the National Natural Science Foundation of China (11901410), and Xiaoping Xie was supported in part by the National Natural Science Foundation of China (11771312).

†tao_wang@m.scnu.edu.cn; wangtao5233@hotmail.com

‡Corresponding author: libinjie@scu.edu.cn

§xpxie@scu.edu.cn

subject to the state equation

$$\begin{cases} \big(\mathrm{D}_{0+}^{\alpha}(y-y_0)\big)(t) - \Delta y(t) = u(t), \quad 0 < t \leqslant T, \\ y(0) = y_0. \end{cases} \tag{2}$$

Here, Δ is the realization of the Laplace operator with homogeneous Dirichlet boundary condition in $L^2(\Omega)$, D_{0+}^{α} is the Riemann-Liouville fractional differential operator of order α, $y_0 \in L^2(\Omega)$ and $y_d \in L^2(0,T;L^2(\Omega))$ are given, and

$$U_{\mathrm{ad}} := \left\{ v \in L^2(0,T;L^2(\Omega)) : \ u_* \leqslant v \leqslant u^* \text{ a.e. in } \Omega \times (0,T) \right\}.$$

The optimal control problem (1) subject to an elliptic or heat equation is a classic problem, which has been thoroughly studied both in theoretical and numerical aspects; see, e.g. [11, 24, 43] for theoretical analysis and [9, 18, 21, 45, 46] for finite element analysis. In general, there are mainly two discretization concepts for this problem ([38, 11]): direct and variational discretizations. The difference between these two concepts is that in the variational discretization, the control is implicitly discretized by the L^2 projection of the discrete co-state into the admissible set U_{ad}. Since the control may has singularity near the boundary of active set, the variational discretization is easier to obtain high accuracy than the direct discretization. However, it should be pointed out that the resultant discrete system of variational discretization is generally more difficult to solve [12, 39].

The state equation (2) is a fractional diffusion equation, which is used to model some physical processes like subdiffusion [3] and water movement in soils [41]. There are many methods to solve this equation, including finite difference methods [7, 13, 17, 23, 26, 49, 50], spectral methods [22, 47], finite element methods [20, 29, 32, 33, 34, 35] and so on.

In recent years, fractional optimal control problems have attracted more and more research interest [1, 2, 10, 16, 25, 48, 51, 52, 53]. For problem (1), Zhou and Gong [52] employed the conforming linear finite element method and L1 scheme for spatial and temporal discretizations, respectively, and obtained optimal convergence results for the spatially semi-discrete approximation. By using the conforming linear finite element method in space and the L1 scheme/backward Euler convolution quadrature in time, Jin et al. [16] gave the first error estimate of the fully discrete scheme with $y_0 = 0$, which is nearly optimal with respect to the regularity. In [10, 51] error estimates were derived for fully discrete finite element approximations for problem (1) with a variant state equation like

$$y' - \mathrm{D}_{0+}^{1-\alpha} \Delta y = f + u.$$

We note that the above works [10, 16, 51, 52] all focus on the case $y_0 = 0$, and their analyses are based on uniform or quasi-uniform temporal grids. However, the situation will be quite subtle when considering nonvanishing y_0. In fact, the nonvanishing initial value may cause essential singularities (cf. Theorem 3.3), which can not be handled well by using uniform or quasi-uniform temporal grids. On the other hand, the non-vanishing $y_d(T)$ may also cause singularities. Fortunately, all these singularities can be dealt with by using special graded temporal grids (cf. Theorem 4.1).

In this paper, for a full discretization using the conforming linear finite element method in space and the piecewise constant discontinuous Galerkin method in time, we provide the first numerical analysis of problem (1) with nonvanishing y_0. Moreover, for the case with $y_0 \in \dot H^{2r}(\Omega)$ with $0 < r < \min\{1, \frac{1-\alpha}{\alpha}\}$ and $y_d \in H^1(0,T;L^2(\Omega))$, we have the following decompositions of the control u, the state y, and the co-state p:

$$u = u_1 + u_2, \quad y = y_1 + y_2, \quad p = p_1 + p_2,$$

with the regularity estimates

$$\begin{aligned}
&\|u_1\|_{{}_0H^1(0,T;L^2(\Omega))} \leqslant C,\\
&\|u_2'(t)\|_{L^2(\Omega)} \leqslant C\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big),\\
&\|y_1\|_{{}_0H^{1+\alpha}(0,T;L^2(\Omega))} + \|y_1\|_{{}_0H^1(0,T;\dot H^2(\Omega)} \leqslant C,\\
&\|p_1\|_{{}^0H^{1+\alpha}(0,T;L^2(\Omega))} + \|p_1\|_{{}^0H^1(0,T;\dot H^2(\Omega))} \leqslant C,\\
&\|y_2'(t)\|_{L^2(\Omega)} \leqslant C\big(t^{\alpha r-1} + \omega_2(T-t)\big), \quad 0<t<T,\\
&\|y_2'(t)\|_{\dot H^1(\Omega)} \leqslant C\big(t^{\alpha r-\alpha/2-1} + \omega_1(T-t)\big), \quad 0<t<T,\\
&\|p_2'(t)\|_{L^2(\Omega)} \leqslant C\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big), \quad 0<t<T,\\
&\|p_2'(t)\|_{\dot H^1(\Omega)} \leqslant C\big(t^{\alpha r+\alpha/2-1} + (T-t)^{\alpha/2-1}\big), \quad 0<t<T,
\end{aligned}$$

where

$$\omega_1(t) = \begin{cases} 1 + \frac{t^{3\alpha/2-1}}{|\alpha(2-3\alpha)|} & \text{if } \alpha \neq 2/3,\\ |\ln t| & \text{if } \alpha = 2/3,\end{cases}$$
$$\omega_2(t) = \begin{cases} 1 + \frac{t^{2\alpha-1}}{|\alpha(1-2\alpha)|} & \text{if } \alpha \neq 1/2,\\ |\ln t| & \text{if } \alpha = 1/2,\end{cases}$$

and C is a generic positive constant depending only on α, r, ν, u_*, u^*, y_0, y_d, T, and Ω. By the above estimates, we obtain first-order temporal accuracy and $\min\{2, 1/\alpha + 2r\}$-order spatial accuracy on graded temporal grids.

The rest of this paper is organized as follows. Section 2 introduces several Sobolev spaces and the Riemann-Liouville fractional calculus operators. Section 3 investigates the regularity of problem (1). Section 4 carries out the convergence analysis for the discontinuous Galerkin method. Finally, Section 5 provides several numerical experiments to confirm the theoretical results.

2 Preliminaries

In this paper, we introduce the following conventions: if $D \subset \mathbb{R}^l (l = 1,2,3,4)$ is Lebesgue measurable, then define $(v,w)_D := \int_D v \cdot w$ for scaler or vector valued functions v and w, and if X is a Banach space, then $\langle\cdot,\cdot\rangle_X$ means the duality pairing between X^* (the dual space of X) and X; the notation $C_\times$ means a positive constant depending only on its subscript(s), and its value may differ at each occurrence. Let $H^\gamma(D)$ $(-\infty < \gamma < \infty)$ and $H_0^\gamma(D)$ $(0 < \gamma < \infty)$ be the usual γ-th order Sobolev spaces on D with norm $\|\cdot\|_{H^\gamma(D)}$ and seminorm $|\cdot|_{H^\gamma(D)}$. In particular, $H^0(D) = L^2(D)$.

Sobolev Spaces. Assume that $-\infty < a < b < \infty$ and X is a Hilbert space.

For each $m \in \mathbb{N}$ and $1 \leqslant q \leqslant \infty$, define

$$
\begin{aligned}
{}_0W^{m,q}(a,b;X) &:= \left\{ v \in W^{m,q}(a,b;X) : \ v^{(k)}(a) = 0, \ 0 \leqslant k < m \right\}, \\
{}^0W^{m,q}(a,b;X) &:= \left\{ v \in W^{m,q}(a,b;X) : \ v^{(k)}(b) = 0, \ 0 \leqslant k < m \right\},
\end{aligned}
$$

where $W^{m,q}(a,b;X)$ is the usual vector valued Sobolev space and $v^{(k)}$ is the k-th weak derivative of v. We equip the above two spaces with the norms

$$
\begin{aligned}
\|v\|_{{}_0W^{m,q}(a,b;X)} &:= \|v^{(m)}\|_{L^q(a,b;X)} \quad \forall v \in {}_0W^{m,q}(a,b;X), \\
\|v\|_{{}^0W^{m,q}(a,b;X)} &:= \|v^{(m)}\|_{L^q(a,b;X)} \quad \forall v \in {}^0W^{m,q}(a,b;X),
\end{aligned}
$$

respectively. For any $m \in \mathbb{N}_{>0}$ and $0 < \theta < 1$, define

$$
\begin{aligned}
W^{m-1+\theta,q}(a,b;X) &:= \big(W^{m-1,q}(a,b;X), W^{m,q}(a,b;X)\big)_{\theta,q}, \\
{}_0W^{m-1+\theta,q}(a,b;X) &:= \big({}_0W^{m-1,q}(a,b;X), \ {}_0W^{m,q}(a,b;X)\big)_{\theta,q}, \\
{}^0W^{m-1+\theta,q}(a,b;X) &:= \big({}^0W^{m-1,q}(a,b;X), \ {}^0W^{m,q}(a,b;X)\big)_{\theta,q},
\end{aligned}
$$

where $(A,B)_{\theta,q}$ denotes the real interpolation space of two Banach spaces, A and B, constructed by the K-method [42]. In addition, for $q = 2$ and $0 \leqslant \beta < \infty$, denote $H^\beta(a,b;X) := W^{\beta,2}(a,b;X)$, ${}_0H^\beta(a,b;X) := {}_0W^{\beta,2}(a,b;X)$, and ${}^0H^\beta(a,b;X) := {}^0W^{\beta,2}(a,b;X)$. We also need the space

$$
W^{1,\infty}_{\text{loc}}(a,b;X) := \left\{ v : (a,b) \to X : \ v \in W^{1,\infty}(c,d;X) \ \text{ for all } a < c < d < b \right\}.
$$

Remark 2.1. *If $0 < \theta < 1$ and $1 \leqslant q < \infty$ satisfy $\theta q < 1$, then*

$$
W^{\theta,q}(a,b;X) = {}_0W^{\theta,q}(a,b;X) = {}^0W^{\theta,q}(a,b;X)
$$

with equivalent norms.

Let Δ be the realization of the Laplace operator with homogeneous Dirichlet boundary condition in $L^2(\Omega)$. For any $-\infty < r < \infty$, define

$$
\dot{H}^r(\Omega) := \{(-\Delta)^{-r/2} v : \ v \in L^2(\Omega)\}
$$

and endow this space with the norm

$$
\|v\|_{\dot{H}^r(\Omega)} := \|(-\Delta)^{r/2} v\|_{L^2(\Omega)} \quad \forall v \in \dot{H}^\beta(\Omega).
$$

Remark 2.2. *For $r \in [0,1] \setminus \{0.5\}$, $\dot{H}^r(\Omega) = H_0^r(\Omega)$ holds with equivalent norms, and for $1 < r \leqslant 2$, the space $\dot{H}^r(\Omega)$ is continuously embedded into $H^r(\Omega)$.*

Fractional calculus operators. For $\gamma > 0$, the left-sided and right-sided Riemann-Liouville fractional integral operators of order γ are defined respectively by

$$
\begin{aligned}
\big(\mathrm{D}_{0+}^{-\gamma} v\big)(t) &:= \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} v(s) \, \mathrm{d}s, \quad 0 < t < T, \\
\big(\mathrm{D}_{T-}^{-\gamma} v\big)(t) &:= \frac{1}{\Gamma(\gamma)} \int_t^T (s-t)^{\gamma-1} v(s) \, \mathrm{d}s, \quad 0 < t < T,
\end{aligned}
$$

for all $v \in L^1(0,T;X)$, where $\Gamma(\cdot)$ is the gamma function. In addition, let D_{0+}^0 and D_{T-}^0 be the identity operator on $L^1(0,T;X)$. Then for $0 < \gamma \leqslant 1$, the left-sided and right-sided Riemann-Liouville fractional differential operators of order γ are defined respectively by

$$\mathrm{D}_{0+}^{\gamma}\, v := \mathrm{D}\,\mathrm{D}_{0+}^{\gamma-1}\, v,$$
$$\mathrm{D}_{T-}^{\gamma}\, v := -\,\mathrm{D}\,\mathrm{D}_{T-}^{\gamma-1}\, v,$$

for all $v \in L^1(0,T;X)$, where D is the first-order differential operator in the distribution sense.

Lemma 2.1 ([6]). *If $v \in H^{\alpha/2}(0,T)$, then*

$$(\mathrm{D}_{0+}^{\alpha/2}\, v, \mathrm{D}_{T-}^{\alpha/2}\, v)_{(0,T)} \geqslant C_{\alpha,T} \|v\|_{H^{\alpha/2}(0,T)}^2.$$

Moreover, if $v, w \in H^{\alpha/2}(0,T)$, then

$$(\mathrm{D}_{0+}^{\alpha/2}\, v, \mathrm{D}_{T-}^{\alpha/2}\, w)_{(0,T)} \leqslant C_{\alpha,T} \|v\|_{H^{\alpha/2}(0,T)} \|w\|_{H^{\alpha/2}(0,T)},$$
$$(\mathrm{D}_{0+}^{\alpha}\, v, w)_{H^{\alpha/2}(0,T)} = (\mathrm{D}_{0+}^{\alpha/2}\, v, \mathrm{D}_{T-}^{\alpha/2}\, w)_{(0,T)} = (\mathrm{D}_{T-}^{\alpha}\, w, v)_{H^{\alpha/2}(0,T)}.$$

3 Regularity

For any $g \in L^q(0,T;L^2(\Omega))$ with $1 < q < \infty$, define Sg (cf. Appendix A) such that $(\mathrm{D}_{0+}^{\alpha} - \Delta)Sg = g$. From Lemma A.2 we summarize several properties of S as follows: (cf. Lemma A.2):

- for any $g \in {}_0W^{\beta,q}(0,T;L^2(\Omega))$ with $\beta \in (0,2] \setminus \{1-\alpha, 2-\alpha\}$ and $1 < q < \infty$,

$$\|Sg\|_{{}_0W^{\alpha+\beta,q}(0,T;L^2(\Omega))} + \|Sg\|_{{}_0W^{\beta,q}(0,T;\dot{H}^2(\Omega))} \leqslant C_{\alpha,\beta,q} \|g\|_{{}_0W^{\beta,q}(0,T;L^2(\Omega))}; \quad (3)$$

- for any $g \in {}_0H^{\beta}(0,T;L^2(\Omega))$ with $0 \leqslant \beta < \infty$,

$$\|Sg\|_{{}_0H^{\alpha+\beta}(0,T;L^2(\Omega))} + \|Sg\|_{{}_0H^{\beta}(0,T;\dot{H}^2(\Omega))} \leqslant C_{\alpha,\beta} \|g\|_{{}_0H^{\beta}(0,T;L^2(\Omega))}. \quad (4)$$

Symmetrically, for any $g \in L^q(0,T;L^2(\Omega))$ with $1 < q < \infty$, define S^*g such that $(\mathrm{D}_{T-}^{\alpha} - \Delta)S^*g = g$. Similar to S, there hold following properties of S^*:

- for any $g \in {}^0W^{\beta,q}(0,T;L^2(\Omega))$ with $\beta \in (0,2] \setminus \{1-\alpha, 2-\alpha\}$ and $1 < q < \infty$,

$$\|S^*g\|_{{}^0W^{\alpha+\beta,q}(0,T;L^2(\Omega))} + \|S^*g\|_{{}^0W^{\beta,q}(0,T;\dot{H}^2(\Omega))} \leqslant C_{\alpha,\beta,q} \|g\|_{{}^0W^{\beta,q}(0,T;L^2(\Omega))}; \quad (5)$$

- for any $g \in {}^0H^{\beta}(0,T;L^2(\Omega))$ with $0 \leqslant \beta < \infty$,

$$\|S^*g\|_{{}^0H^{\alpha+\beta}(0,T;L^2(\Omega))} + \|S^*g\|_{{}^0H^{\beta}(0,T;\dot{H}^2(\Omega))} \leqslant C_{\alpha,\beta} \|g\|_{{}^0H^{\beta}(0,T;L^2(\Omega))}. \quad (6)$$

In addition, by the definitions of S and S^*, (4), (6) and Lemma 2.1, we obtain that, for any $v, w \in L^2(0,T;L^2(\Omega))$,

$$\begin{aligned}
(Sv, w)_{\Omega\times(0,T)} &= \big(Sv, (\mathrm{D}_{T-}^{\alpha} - \Delta)S^*w\big)_{\Omega\times(0,T)} \\
&= \big((\mathrm{D}_{0+}^{\alpha} - \Delta)Sv, S^*w\big)_{\Omega\times(0,T)} \\
&= (v, S^*w)_{\Omega\times(0,T)}. \quad (7)
\end{aligned}$$

Assuming that $y_0 \in L^2(\Omega)$ and $y_d \in L^2(0,T;L^2(\Omega))$, we call $u \in U_{\rm ad}$ a solution to problem (1) if u solves the minimization problem

$$\min_{u\in U_{\rm ad}} J(u) = \frac{1}{2}\|S(u + \mathrm{D}_{0+}^{\alpha}\, y_0) - y_d\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\nu}{2}\|u\|^2_{L^2(0,T;L^2(\Omega))}. \tag{8}$$

By (7), a routine argument gives the following theorem (cf. [11, 43]).

Theorem 3.1. *Problem (1) admits a unique solution $u \in U_{ad}$, and*

$$\left(S^*\big(S(u + \mathrm{D}_{0+}^{\alpha}\, y_0) - y_d\big) + \nu u, v - u\right)_{\Omega\times(0,T)} \geqslant 0 \tag{9}$$

for all $v \in U_{ad}$.

In the rest of this paper, we use u to denote the solution of problem (1) and use

$$y := S(u + \mathrm{D}_{0+}^{\alpha}\, y_0) \quad \text{and} \quad p := S^*\big(y - y_d\big) \tag{10}$$

to denote the corresponding state and co-state, respectively.

The main task of this section is to prove the following two theorems.

Theorem 3.2. *Assume that $y_0 \in \dot H^{2r}(\Omega)$ with $0 < r < \min\{1, \frac{1-\alpha}{\alpha}\}$, and $y_d \in H^1(0,T;L^2(\Omega))$. There exists a decomposition*

$$u = u_1 + u_2,$$

where

$$\|u_1\|_{{}_0H^1(0,T;L^2(\Omega))} \leqslant C \tag{11}$$

and $u_2 \in W^{1,\infty}_{loc}(0,T;L^2(\Omega))$ satisfies that

$$\|u_2'(t)\|_{L^2(\Omega)} \leqslant C\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big), \quad a.e.\ 0<t<T, \tag{12}$$

where C is a positive constant depending only on α, r, ν, u_*, u^*, y_0, y_d, T and Ω.

Theorem 3.3. *Assume that $y_0 \in \dot H^{2r}(\Omega)$ with $0 < r < \min\{1, \frac{1-\alpha}{\alpha}\}$, and $y_d \in H^1(0,T;L^2(\Omega))$. There exist decompositions*

$$y = y_1 + y_2, \quad p = p_1 + p_2,$$

where

$$y_1 \in {}_0H^{1+\alpha}(0,T;L^2(\Omega)) \bigcap {}_0H^1(0,T;\dot H^2(\Omega)), \tag{13}$$

$$p_1 \in {}^0H^{1+\alpha}(0,T;L^2(\Omega)) \bigcap {}^0H^1(0,T;\dot H^2(\Omega)) \tag{14}$$

$$\text{and} \quad y_2, p_2 \in C^1((0,T);\dot H^1(\Omega)). \tag{15}$$

Moreover,

$$\|y_1\|_{{}_0H^{1+\alpha}(0,T;L^2(\Omega))} + \|y_1\|_{{}_0H^1(0,T;\dot H^2(\Omega)} \leqslant C, \tag{16}$$

$$\|p_1\|_{{}^0H^{1+\alpha}(0,T;L^2(\Omega))} + \|p_1\|_{{}^0H^1(0,T;\dot H^2(\Omega))} \leqslant C, \tag{17}$$

$$\|y_2'(t)\|_{\dot H^1(\Omega)} \leqslant C\big(t^{\alpha r-\alpha/2-1} + \omega_1(T-t)\big), \quad 0<t<T, \tag{18}$$

$$\|y_2'(t)\|_{L^2(\Omega)} \leqslant C\big(t^{\alpha r-1} + \omega_2(T-t)\big), \quad 0<t<T, \tag{19}$$

$$\|p_2'(t)\|_{\dot H^1(\Omega)} \leqslant C\big(t^{\alpha r+\alpha/2-1} + (T-t)^{\alpha/2-1}\big), \quad 0<t<T, \tag{20}$$

$$\|p_2'(t)\|_{L^2(\Omega)} \leqslant C\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big), \quad 0<t<T. \tag{21}$$

The above C is a positive constant depending only on α, r, ν, u_*, u^*, y_0, y_d, T and Ω, and for any $t > 0$,

$$\omega_1(t) := \begin{cases} 1 + \frac{t^{3\alpha/2-1}}{|\alpha(2-3\alpha)|} & \text{if } \alpha \neq 2/3, \\ |\ln t| & \text{if } \alpha = 2/3, \end{cases} \tag{22}$$

$$\omega_2(t) := \begin{cases} 1 + \frac{t^{2\alpha-1}}{|\alpha(1-2\alpha)|} & \text{if } \alpha \neq 1/2, \\ |\ln t| & \text{if } \alpha = 1/2. \end{cases} \tag{23}$$

Remark 3.1. *The results of Theorems 3.2 and 3.3 can be easily extended to the case $y_0 \in \dot{H}^{2r}(\Omega)$ with $r \geqslant \min\{1, \frac{1-\alpha}{\alpha}\}$.*

3.1 Proofs of Theorems 3.2 and 3.3

For $g \in L^1(0,T;L^2(\Omega))$, we have that [15]

$$(Sg)(t) = \int_0^t E(s)g(t-s)\,\mathrm{d}s, \qquad \text{a.e. } 0 < t < T, \tag{24}$$

$$(S^*g)(t) = \int_t^T E(s-t)g(s)\,\mathrm{d}s, \qquad \text{a.e. } 0 < t < T, \tag{25}$$

where, for each $0 < s \leqslant T$,

$$E(s) := \frac{1}{2\pi i}\int_0^\infty e^{-rs}\big((r^\alpha e^{-i\alpha\pi} - \Delta)^{-1} - (r^\alpha e^{i\alpha\pi} - \Delta)^{-1}\big)\,\mathrm{d}r.$$

Lemma 3.1 ([15])**.** *The function E is an $\mathcal{L}(L^2(\Omega), \dot{H}^1(\Omega))$-valued analytic function on $(0,\infty)$, and*

$$\|E(t)\|_{\mathcal{L}(L^2(\Omega))} + t^{\alpha/2}\|E(t)\|_{\mathcal{L}(L^2(\Omega),\dot{H}^1(\Omega))} \leqslant C_\alpha t^{\alpha-1}, \quad t > 0.$$

In the rest of this subsection, for convenience we will always assume that $y_d \in H^1(0,T;L^2(\Omega))$ and $y_0 \in \dot{H}^{2r}(\Omega)$, where $0 < r < \min\{1, \frac{1-\alpha}{\alpha}\}$.

Lemma 3.2. *Assume that $g \in W^{1,\infty}_{loc}(0,T;L^2(\Omega))$ and A is a positive constant. If*

$$\|g'(t)\|_{L^2(\Omega)} \leqslant A\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big), \quad \text{a.e. } 0 < t < T, \tag{26}$$

then $Sg \in C^1((0,T);\dot{H}^1(\Omega))$ and, for any $0 < t < T$,

$$\|(Sg)'(t)\|_{L^2(\Omega)} \leqslant C_{\alpha,T}\left(A + \|g(0)\|_{L^2(\Omega)}\right)\left(t^{\alpha-1} + \omega_2(T-t)\right), \tag{27}$$

$$\|(Sg)'(t)\|_{\dot{H}^1(\Omega)} \leqslant C_{\alpha,T}\left(A + \|g(0)\|_{L^2(\Omega)}\right)\left(t^{\alpha/2-1} + \omega_1(T-t)\right). \tag{28}$$

If

$$\|g'(t)\|_{L^2(\Omega)} \leqslant A\big(t^{\alpha r-1} + \omega_2(T-t)\big), \quad \text{a.e. } 0 < t < T,$$

then $S^*g \in C^1((0,T);\dot{H}^1(\Omega))$ and, for any $0 < t < T$,

$$\|(S^*g)'(t)\|_{L^2(\Omega)} \leqslant C_{\alpha,T}\big(A + \|g(T)\|_{L^2(\Omega)}\big)\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big),$$

$$\|(S^*g)'(t)\|_{\dot{H}^1(\Omega)} \leqslant C_{\alpha,T}\big(A + \|g(T)\|_{L^2(\Omega)}\big)\big(t^{\alpha r+\alpha/2-1} + (T-t)^{\alpha/2-1}\big).$$

The above ω_1 and ω_2 are defined by (22) and (23), respectively.

Proof. By (24), (26) and Lemma 3.1, a straightforward computation gives

$$(Sg)'(t) = E(t)g(0) + \int_0^t E(s)g'(t-s)\,\mathrm{d}s, \quad 0 < t < T, \tag{29}$$

and, by the techniques in the proof of [4, Theorem 2.6], it is easy to verify that $Sg \in C^1((0,T); \dot{H}^1(\Omega))$. Furthermore, by (26), (29) and Lemma 3.1,

$$\begin{aligned}
\|(Sg)'(t)\|_{L^2(\Omega)} &\leqslant C_\alpha\Big(t^{\alpha-1}\|g(0)\|_{L^2(\Omega)} + A\int_0^t s^{\alpha-1}\big((t-s)^{\alpha r+\alpha-1} + (T-t+s)^{\alpha-1}\big)\,\mathrm{d}s\Big) \\
&= C_\alpha\Big(t^{\alpha-1}\|g(0)\|_{L^2(\Omega)} + At^{\alpha r+2\alpha-1}\int_0^1 x^{\alpha-1}(1-x)^{\alpha r+\alpha-1}\mathrm{d}x \\
&\qquad + A(T-t)^{2\alpha-1}\int_0^{t/(T-t)} x^{\alpha-1}(1+x)^{\alpha-1}\,\mathrm{d}x\Big) \\
&\leqslant C_{\alpha,T}\big(A + \|g(0)\|_{L^2(\Omega)}\big)\Big(t^{\alpha-1} + \omega_2(T-t)\Big)
\end{aligned}$$

and

$$\begin{aligned}
\|(Sg)'(t)\|_{\dot{H}^1(\Omega)} &\leqslant C_\alpha\Big(t^{\alpha/2-1}\|g(0)\|_{L^2(\Omega)} + A\int_0^t s^{\alpha/2-1}\big((t-s)^{\alpha r+\alpha-1} + (T-t+s)^{\alpha-1}\big)\mathrm{d}s\Big) \\
&= C_\alpha\Big(t^{\alpha/2-1}\|g(0)\|_{L^2(\Omega)} + At^{\alpha r+3\alpha/2-1}\int_0^1 x^{\alpha/2-1}\big((1-x)^{\alpha r+\alpha-1}\mathrm{d}x \\
&\qquad + A(T-t)^{3\alpha/2-1}\int_0^{t/(T-t)} x^{\alpha/2-1}(1+x)^{\alpha-1}\,\mathrm{d}x\Big) \\
&\leqslant C_{\alpha,T}\big(A + \|g(0\|_{L^2(\Omega)}\big)\big(t^{\alpha/2-1} + \omega_1(T-t)\big)
\end{aligned}$$

for all $0 < t < T$. This proves estimates (27) and (28). Since the rest of this lemma can be proved analogously, this completes the proof. ■

From (9) it follows that

$$u = f(p), \tag{30}$$

where $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$f(v) := \begin{cases} u^* & \text{if } v < -\nu u^*, \\ -v/\nu & \text{if } -\nu u^* \leqslant v \leqslant -\nu u_*, \\ u_* & \text{if } v > -\nu u_*. \end{cases}$$

We set

$$f'(r) := \begin{cases} 0 & \text{if } r \leqslant -\nu u^*, \\ -1/\nu & \text{if } -\nu u^* < r < \nu u_*, \\ 0 & \text{if } r \geqslant -\nu u_*. \end{cases}$$

For any $v \in W^{1,1}(0,T; L^2(\Omega))$, from [8, Thoerem 7.8] we conclude that $f(v) \in W^{1,1}(0,T; L^2(\Omega))$ and

$$(f(v))'(t) = f'(v(t))v'(t), \quad 0 < t < T.$$

Furthermore, applying [42, Lemma 28.1] yields the following interpolation result.

Lemma 3.3. *If $v \in W^{\beta,q}(0,T; L^2(\Omega))$ with $0 \leqslant \beta \leqslant 1$ and $1 < q < \infty$, then*

$$\|f(v)\|_{W^{\beta,q}(0,T;L^2(\Omega))} \leqslant C_{\nu,u_*,u^*,q,T,\Omega}\big(1 + \|v\|_{W^{\beta,q}(0,T;L^2(\Omega))}\big).$$

Lemma 3.4. *If $u \in W^{\beta,q_0}(0,T;L^2(\Omega))$ with $\beta \in (0,1-\alpha) \setminus \{1-2\alpha\}$ and $1 < q_0 < 1/(1-\alpha r)$, then*

$$\begin{aligned} &\|u\|_{W^{\min\{2\alpha+\beta,1\},q_0}(0,T;L^2(\Omega))} \\ &\leqslant C_{\alpha,\beta,q_0,r,\nu,u_*,u^*,T,\Omega}\big(1 + \|y_0\|_{\dot H^{2r}(\Omega)} + \|y_d\|_{H^1(0,T;L^2(\Omega))} + \|u\|_{W^{\beta,q_0}(0,T;L^2(\Omega))}\big). \end{aligned} \tag{31}$$

Proof. We only prove the case $\beta < 1-2\alpha$, since the other cases can be proved analogously. For simplicity, we denote by C in this proof a generic positive constant depending only on α, β, q_0, r, ν, u_*, u^*, T and Ω, and its value may differ in different places. Some straightforward calculations give

$$\|S\,\mathrm{D}_{0+}^{\alpha}\, y_0\|_{W^{1,q_0}(0,T;L^2(\Omega))} \leqslant C\|y_0\|_{\dot H^{2r}(\Omega)} \quad \text{(by (94))}$$

and

$$\begin{aligned} \|Su\|_{{}_0W^{\alpha+\beta,q_0}(0,T;L^2(\Omega))} &\leqslant C\|u\|_{{}_0W^{\beta,q_0}(0,T;L^2(\Omega))} \quad \text{(by (3))} \\ &\leqslant C\|u\|_{W^{\beta,q_0}(0,T;L^2(\Omega))} \quad \text{(by the fact } \beta q_0 < 1), \end{aligned}$$

so that

$$\begin{aligned} \|y\|_{W^{\alpha+\beta,q_0}(0,T;L^2(\Omega))} &= \|S(u + \mathrm{D}_{0+}^{\alpha}\, y_0)\|_{W^{\alpha+\beta,q_0}(0,T;L^2(\Omega))} \\ &\leqslant C\big(\|u\|_{W^{\beta,q_0}(0,T;L^2(\Omega))} + \|y_0\|_{\dot H^{2r}(\Omega)}\big). \end{aligned}$$

By (5) and the fact $(\alpha+\beta)q_0 < 1$,

$$\begin{aligned} &\|p\|_{{}^0W^{2\alpha+\beta,q_0}(0,T;L^2(\Omega))} = \|S^*(y-y_d)\|_{{}^0W^{2\alpha+\beta,q_0}(0,T;L^2(\Omega))} \\ &\leqslant C\|y-y_d\|_{{}^0W^{\alpha+\beta,q_0}(0,T;L^2(\Omega))} \leqslant C\|y-y_d\|_{W^{\alpha+\beta,q_0}(0,T;L^2(\Omega))} \\ &\leqslant C\big(\|y\|_{W^{\alpha+\beta,q_0}(0,T;L^2(\Omega))} + \|y_d\|_{H^1(0,T;L^2(\Omega))}\big). \end{aligned}$$

In addition,

$$\begin{aligned} \|u\|_{W^{2\alpha+\beta,q_0}(0,T;L^2(\Omega))} &= \|f(p)\|_{W^{2\alpha+\beta,q_0}(0,T;L^2(\Omega))} \quad \text{(by (30))} \\ &\leqslant C\big(1 + \|p\|_{W^{2\alpha+\beta,q_0}(0,T;L^2(\Omega))}\big) \quad \text{(by Lemma 3.3)}. \end{aligned}$$

Finally, combining the above three estimates proves (31) and hence this lemma. ■

Lemma 3.5. *Assume that $1 < q \leqslant 2$, $0 < A < \infty$, and $u(t) = u_1(t) + u_2(t)$ for each $0 < t < T$, with $u_1 \in {}_0W^{1,q}(0,T;L^2(\Omega))$ and $u_2 \in W^{1,\infty}_{loc}(0,T;L^2(\Omega))$. If*

$$\|u_2'(t)\|_{L^2(\Omega)} \leqslant A\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big), \quad \text{a.e. } 0 < t < T,$$

then there exists a decomposition

$$u(t) = \widetilde u_1(t) + \widetilde u_2(t), \quad 0 < t < T, \tag{32}$$

such that

$$\|\widetilde u_1\|_{{}_0W^{1,q+\alpha/(1-\alpha)}(0,T;L^2(\Omega))} \leqslant C\big(\|u_1\|_{{}_0W^{1,q}(0,T;L^2(\Omega))} + \|y_d\|_{H^1(0,T;L^2(\Omega))}\big) \tag{33}$$

and

$$\|\widetilde{u}_2'(t)\|_{L^2(\Omega)} \leqslant C\Big(A + \|u(0)\|_{L^2(\Omega)} + \|y_0\|_{\dot H^{2r}(\Omega)} + \|y_d\|_{H^1(0,T;L^2(\Omega))} + \|u_1\|_{{}_0W^{1,q}(0,T;L^2(\Omega))}\Big)\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big) \tag{34}$$

for almost all $0<t<T$*, where* C *is a positive constant depending only on* α*,* r*,* ν*,* u_**,* u^**,* q*,* T *and* Ω*.*

Proof. A simple calculation gives, by (10), that

$$p = \mathbb{I}_1 + \mathbb{I}_2, \tag{35}$$

where

$$\mathbb{I}_1 := S^*\big(Su_1 - y_d - (Su_1 - y_d)(T)\big),$$
$$\mathbb{I}_2 := S^*\big(S(u_2 + \mathrm{D}_{0+}^\alpha\, y_0) + (Su_1 - y_d)(T)\big).$$

We have

$$\begin{aligned}
&\|\mathbb{I}_1\|_{{}^0W^{1+\alpha,q}(0,T;L^2(\Omega))} \\
&\leqslant C\|Su_1 - y_d - (Su_1-y_d)(T)\|_{{}^0W^{1,q}(0,T;L^2(\Omega))} \quad \text{(by (5))} \\
&\leqslant C\big(\|Su_1\|_{{}_0W^{1,q}(0,T;L^2(\Omega))} + \|y_d\|_{H^1(0,T;L^2(\Omega))}\big) \\
&\leqslant C\big(\|u_1\|_{{}_0W^{1,q}(0,T;L^2(\Omega))} + \|y_d\|_{H^1(0,T;L^2(\Omega))}\big) \quad \text{(by (3))}.
\end{aligned}$$

As ${}^0W^{1+\alpha,q}(0,T;L^2(\Omega))$ is continuously embedded into ${}^0W^{1,q+\alpha/(1-\alpha)}(0,T;L^2(\Omega))$, it holds

$$\|\mathbb{I}_1\|_{{}^0W^{1,q+\alpha/(1-\alpha)}(0,T;L^2(\Omega))} \leqslant C\big(\|u_1\|_{{}_0W^{1,q}(0,T;L^2(\Omega))} + \|y_d\|_{H^1(0,T;L^2(\Omega))}\big). \tag{36}$$

From (27) it follows that

$$\|(Su_2)'(t)\|_{L^2(\Omega)} \leqslant C(A + u_2(0))\big(t^{\alpha-1} + \omega_2(T-t)\big).$$

Therefore, the fact that

$$\|(Su_1)(T)\|_{L^2(\Omega)} \leqslant C\|Su_1\|_{{}_0W^{1+\alpha/2,q}(0,T;L^2(\Omega))} \leqslant C\|u_1\|_{{}_0W^{1-\alpha/2,q}(0,T;L^2(\Omega))} \quad \text{(by (3))}$$

and

$$\|(S\,\mathrm{D}_{0+}^\alpha\, y_0)'(t)\| \leqslant Ct^{\alpha r-1}\|y_0\|_{\dot H^{2r}(\Omega)} \quad \text{(by (94))},$$

together with Lemma 3.2 and (92), yields

$$\|\mathbb{I}_2'(t)\|_{L^2(\Omega)} \leqslant C\Big(A + \|u(0)\|_{L^2(\Omega)} + \|y_0\|_{\dot H^{2r}(\Omega)} + \|u_1\|_{{}_0W^{1,q}(0,T;L^2(\Omega))} + \|y_d\|_{H^1(0,T;L^2(\Omega))}\Big)\big(t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1}\big) \tag{37}$$

for all $0<t<T$.

Finally, letting

$$\widetilde{u}_1(t) := \int_0^t f'(p(s))\mathbb{I}_1'(s)\,\mathrm{d}s, \quad 0<t<T, \tag{38}$$

$$\widetilde{u}_2(t) := u(0) + \int_0^t f'(p(s))\mathbb{I}_2'(s)\,\mathrm{d}s, \quad 0<t<T, \tag{39}$$

by (30) and (35) we obtain

$$\widetilde{u}_1(t)+\widetilde{u}_2(t) = u(0) + f(p(t)) - f(p(0)) = u(t), \quad 0 < t \leqslant T,$$

which proves (32). Furthermore, (33) follows from (36) and (38), and (34) follows from (37) and (39). This completes the proof. ■

Finally, we are in a position to prove Theorems 3.2 and 3.3.

Proof of Theorem 3.2. By the fact $u \in U_{ad} \subset L^2(0,T;L^2(\Omega))$, a similar proof as that of Lemma 3.4 gives

$$\|u\|_{W^{\alpha,q_0}(0,T;L^2(\Omega))} \leqslant C,$$

and then applying Lemma 3.4 several times yields

$$\|u\|_{W^{1,q_0}(0,T;L^2(\Omega))} \leqslant C,$$

where $q_0 := \frac{1}{2}(1+1/(1-\alpha r)) < 1/(1-\alpha r)$ and C is a positive constant depending only on α, r, ν, u_*, u^*, y_0, y_d, T and Ω. Letting

$$m := \min\{n \in \mathbb{N} : \; q_0 + n\alpha/(1-\alpha) \geqslant 2\}$$

and applying Lemma 3.5 m time(s) then prove Theorem 3.2 ($u_2 = 0$ for the first time). ■

Proof of Theorem 3.3. Let $y_1 := Su_1$ and $y_2 := S(u_2 + \mathrm{D}_{0+}^{\alpha}\, y_0)$, where u_1 and u_2 are defined in Theorem 3.2. By (4) and (11) we have

$$\|y_1\|_{{}_0H^{1+\alpha}(0,T;L^2(\Omega))} + \|y_1\|_{{}_0H^1(0,T;\dot{H}^2(\Omega))} \leqslant C, \tag{40}$$

and by (12), (27), (28) and (94) we have that

$$y_2 \in C([0,T];L^2(\Omega)) \bigcap C^1((0,T);\dot{H}^1(\Omega))$$

and

$$\begin{aligned} &\|y_2'(t)\|_{L^2(\Omega)} \leqslant C\big(t^{\alpha r-1} + \omega_2(T-t)\big), \quad 0 < t < T, \\ &\|y_2'(t)\|_{\dot{H}^1(\Omega)} \leqslant C\big(t^{(r-1/2)\alpha-1} + \omega_1(T-t)\big), \quad 0 < t < T, \end{aligned}$$

where C is a positive constant depending only on α, r, ν, u_*, u^*, y_0, y_d, T and Ω. Hence, $y = y_1 + y_2$ is the desired decomposition in Theorem 3.3. Since the rest of this theorem can be proved analogously, this concludes the proof. ■

4 Discretization

Assume that $0 < r < 1$, $M > 1$ is an integer and

$$\begin{cases} \sigma_1 > \max\left\{1, \frac{2-\alpha}{(2r-1)\alpha+1}\right\}, \\ \sigma_2 > \max\left\{1, \frac{2-\alpha}{\alpha+1}\right\}. \end{cases} \tag{41}$$

Let $0 = t_0 < t_1 < \ldots < t_{2M} = T$ be a graded partition of the temporal interval $[0, T]$ with

$$t_j := \begin{cases} \left(\frac{j}{M}\right)^{\sigma_1} \frac{T}{2}, & \text{if } \ 0 \leqslant j \leqslant M, \\ T - \left(2 - \frac{j}{M}\right)^{\sigma_2} \frac{T}{2}, & \text{if } \ M < j \leqslant 2M. \end{cases}$$

Define $\tau_j := t_j - t_{j-1}$ for each $1 \leqslant j \leqslant 2M$ and set $\tau := \max\{\tau_j : 1 \leqslant j \leqslant 2M\}$. Let $\mathcal{K}_h$ be a conventional conforming and shape-regular triangulation of Ω consisting of d-simplexes with mesh size $h := \max_{K \in \mathcal{K}_h}\{\text{diameter of } K\}$. Then we introduce the following finite element spaces:

$$\begin{aligned} \mathcal{V}_h &:= \left\{ v_h \in \dot{H}^1(\Omega) : \ v_h \text{ is linear on } K \text{ for each } K \in \mathcal{K}_h \right\}, \\ \mathcal{W}_{h\tau} &:= \left\{ V \in L^2(0,T;\mathcal{V}_h) : \ V \text{ is constant on } (t_{j-1}, t_j) \text{ for each } 1 \leqslant j \leqslant 2M \right\}. \end{aligned}$$

For any $g \in L^2(0,T;L^2(\Omega))$, define $S_{h\tau} g$, $S_{h\tau}^* g \in \mathcal{W}_{h\tau}$ respectively by that

$$\left(\mathrm{D}_{0+}^{\alpha/2}\, S_{h\tau} g, \mathrm{D}_{T-}^{\alpha/2}\, V \right)_{\Omega\times(0,T)} + \left(\nabla S_{h\tau} g, \nabla V \right)_{\Omega\times(0,T)} = (g, V)_{\Omega\times(0,T)}, \tag{42}$$

$$\left(\mathrm{D}_{T-}^{\alpha/2}\, S_{h\tau}^* g, \mathrm{D}_{0+}^{\alpha/2}\, V \right)_{\Omega\times(0,T)} + \left(\nabla S_{h\tau}^* g, \nabla V \right)_{\Omega\times(0,T)} = (g, V)_{\Omega\times(0,T)}, \tag{43}$$

for all $V \in \mathcal{W}_{h\tau}$. It is evident that

$$(S_{h\tau} g_1, g_2)_{\Omega\times(0,T)} = (g_1, S_{h\tau}^* g_2)_{\Omega\times(0,T)} \tag{44}$$

for all $g_1, g_2 \in \mathcal{W}_{h\tau}$.

Remark 4.1. *By Lemma 2.1 and the Lax-Milgram theorem, it is easy to verify that the operators $S_{h\tau}$ and $S_{h\tau}^*$ are well defined.*

With the above two operators, we consider the following optimal control problem:

$$\min_{U \in U_{\mathrm{ad}}} J(U) = \frac{1}{2} \| S_{h\tau}(U + \mathrm{D}_{0+}^{\alpha}\, y_0) - y_d \|_{L^2(0,T;L^2(\Omega))}^2 + \frac{\nu}{2} \|U\|_{L^2(0,T;L^2(\Omega))}^2. \tag{45}$$

Similar to the continuous case, there exists a unique discrete control $U \in U_{\mathrm{ad}}$ such that

$$\left(S_{h\tau}^* \left(S_{h\tau}(U + \mathrm{D}_{0+}^{\alpha}\, y_0) - y_d \right) + \nu U, \, v - U \right)_{\Omega\times(0,T)} \geqslant 0, \quad \forall v \in U_{ad}. \tag{46}$$

The corresponding discrete state Y and co-state P are defined respectively by

$$Y := S_{h\tau}(U + \mathrm{D}_{0+}^{\alpha}\, y_0) \quad \text{and} \quad P := S_{h\tau}^*(Y - y_d). \tag{47}$$

Remark 4.2. *Since the co-state P is in the finite dimensional space $\mathcal{W}_{h\tau}$, the control u is indirectly discretized by the projection*

$$U = Q_{U_{ad}}(-P/\nu),$$

where $Q_{U_{ad}}$ is the L^2 projection onto the admissible set U_{ad}. This is the key point of the variational discretization concept (cf. [11]).

In this section, we use $\bar{a} \lesssim \bar{b}$ to denote $\bar{a} \leqslant C\bar{b}$, where C is a positive constant independent of h and M. The main result of this section is the following theorem.

Theorem 4.1. *Assume that $0 < r < \min\{1/2, (1-\alpha)/\alpha\}$. Let u and y be the control and state of (1) respectively, and let U and Y be the control and state of (45) respectively. If $y_0 \in \dot{H}^{2r}(\Omega)$ and $y_d \in H^1(0,T;L^2(\Omega))$, then*

$$\|u - U\|_{L^2(0,T;L^2(\Omega))} + \|y - Y\|_{L^2(0,T;L^2(\Omega))} \lesssim h^{\min\{1/\alpha+2r,2\}} + M^{-1}. \tag{48}$$

Remark 4.3. *By following a similar routine in the proof of Theorem 4.1, we can show that the estimate (48) also holds for $0 < r < 1$ with $r \neq \frac{1}{\alpha} - 1, \frac{1}{\alpha} - \frac{1}{2}$.*

Remark 4.4. *Note that, under the condition that $y_0 = 0$ and $y_d \in H^1(0,T;L^2(\Omega))$, Jin et al. [16, Theorem 3.10] derived temporal accuracy $O(\tau^{1/2+\min\{1/2,\alpha-\epsilon\}})$, where $\epsilon > 0$ can be arbitrarily small.*

4.1 Proof of Theorem 4.1

Throughout this subsection, u, y and p are the control, state and co-state of (1), respectively. By (9), (46) and the standard technique in [12], we obtain

$$\begin{aligned} &\|u - U\|_{L^2(0,T;L^2(\Omega))} + \|y - Y\|_{L^2(0,T;L^2(\Omega))} \\ \lesssim\ &\|y - \widetilde{Y}\|_{L^2(0,T;L^2(\Omega))} + \|p - \widetilde{P}\|_{L^2(0,T;L^2(\Omega))}, \end{aligned}$$

where

$$\widetilde{Y} := S_{h\tau}(u + \mathrm{D}_{0+}^{\alpha}\, y_0), \tag{49}$$

$$\widetilde{P} := S_{h\tau}^{*}(y - y_d). \tag{50}$$

Therefore, to conclude the proof of Theorem 4.1, it remains to prove

$$\|y - y_h\|_{L^2(0,T;L^2(\Omega))} + \|p - p_h\|_{L^2(0,T;L^2(\Omega))} \lesssim h^{\min\{1/\alpha+2r,2\}}, \tag{51}$$

$$\|y_h - \widetilde{Y}\|_{L^2(0,T;L^2(\Omega))} + \|p_h - \widetilde{P}\|_{L^2(0,T;L^2(\Omega))} \lesssim M^{-1}, \tag{52}$$

where y_h is the solution of the equation

$$\begin{cases} \mathrm{D}_{0+}^{\alpha}(y_h - Q_h y_0) - \Delta_h y_h = Q_h u, \\ y_h(0) = Q_h y_0, \end{cases} \tag{53}$$

and p_h is the solution of the equation

$$\begin{cases} (\mathrm{D}_{T-}^{\alpha} - \Delta_h) p_h = Q_h(y - y_d), \\ p_h(T) = 0. \end{cases} \tag{54}$$

Here Q_h is the $L^2(\Omega)$-orthogonal projection operator onto $\mathcal{V}_h$ and $\Delta_h : \mathcal{V}_h \to \mathcal{V}_h$ is the discrete Laplace operator defined by

$$(\Delta_h v_h, w_h)_\Omega = -(\nabla v_h, \nabla w_h)_\Omega \quad \forall v_h, w_h \in \mathcal{V}_h. \tag{55}$$

Let us first prove estimate (51) in the following lemma.

Lemma 4.1. *Under the condition of Theorem 4.1, we have*

$$\|y - y_h\|_{L^2(0,T;L^2(\Omega))} \lesssim h^{\min\{1/\alpha+2r,2\}}, \tag{56}$$
$$\|p - p_h\|_{L^2(0,T;L^2(\Omega))} \lesssim h^2. \tag{57}$$

Proof. By (4) and the fact $u \in U_{\mathrm{ad}}$, we have

$$Su \in {}_0H^\alpha(0,T;L^2(\Omega)) \cap L^2(0,T;\dot{H}^2(\Omega)),$$

so that by interpolation we obtain

$$Su \in {}_0H^{\alpha/2}(0,T;\dot{H}^1(\Omega)).$$

In addition, Lemma A.4 implies

$$S\,\mathrm{D}_{0+}^\alpha\, y_0 \in {}_0H^{\alpha/2}\big(0,T;\dot{H}^{\min\{1/\alpha+2r-1,1\}}(\Omega)\big) \cap L^2\big(0,T;\dot{H}^{\min\{1/\alpha+2r,2\}}(\Omega)\big).$$

Consequently, we conclude from the fact $y = S(u + \mathrm{D}_{0+}^\alpha\, y_0)$ that

$$y \in {}_0H^{\alpha/2}\big(0,T;\dot{H}^{\min\{1/\alpha+2r-1,1\}}(\Omega)\big) \cap L^2\big(0,T;\dot{H}^{\min\{1/\alpha+2r,2\}}(\Omega)\big).$$

A routine energy argument (cf. [20]) then yields (56). Since (57) can be derived analogously, this completes the proof. ■

Then let us prove estimate (52). Similar to the properties of y and p presented in Theorem 3.3, under the condition of Theorem 4.1, we have the following properties: there exist decompositions

$$y_h = y_{h,1} + y_{h,2}, \quad p_h = p_{h,1} + p_{h,2}, \tag{58}$$

where

$$\|y_{h,1}\|_{{}_0H^{1+\alpha}(0,T;L^2(\Omega))} + \|\Delta_h y_{h,1}\|_{{}_0H^1(0,T;L^2(\Omega))} \lesssim 1, \tag{59}$$
$$\|p_{h,1}\|_{{}^0H^{1+\alpha}(0,T;L^2(\Omega))} + \|\Delta_h p_{h,1}\|_{{}^0H^1(0,T;L^2(\Omega))} \lesssim 1, \tag{60}$$
$$\|y_{h,2}'(t)\|_{L^2(\Omega)} \lesssim t^{\alpha r-1} + \omega_2(T-t) \quad \forall 0<t<T, \tag{61}$$
$$\|y_{h,2}'(t)\|_{\dot{H}^1(\Omega)} \lesssim t^{\alpha r-\alpha/2-1} + \omega_1(T-t) \quad \forall 0<t<T, \tag{62}$$
$$\|p_{h,2}'(t)\|_{L^2(\Omega)} \lesssim t^{\alpha r+\alpha-1} + (T-t)^{\alpha-1} \quad \forall 0<t<T, \tag{63}$$
$$\|p_{h,2}'(t)\|_{\dot{H}^1(\Omega)} \lesssim t^{\alpha r+\alpha/2-1} + (T-t)^{\alpha/2-1} \quad \forall 0<t<T. \tag{64}$$

Here ω_1 and ω_2 are defined by (22) and (23), respectively. For any $v \in L^2(0,T;L^2(\Omega))$, space, define $Q_\tau v \in L^\infty(0,T;L^2(\Omega))$ by

$$(Q_\tau v)|_{(t_{j-1},t_j)} := \frac{1}{\tau_j}\int_{t_{j-1}}^{t_j} v(t)\,\mathrm{d}t, \quad \forall 1 \leqslant j \leqslant 2M.$$

By (59) and (60) it is standard that

$$\|(I-Q_\tau)y_{h,1}\|_W \lesssim M^{-1+\alpha/2}, \tag{65}$$
$$\|(I-Q_\tau)p_{h,1}\|_W \lesssim M^{-1+\alpha/2}, \tag{66}$$

where

$$\|\cdot\|_W := \|\cdot\|_{{}_0H^{\alpha/2}(0,T;L^2(\Omega))} + \|\cdot\|_{L^2(0,T;\dot{H}^1(\Omega))}.$$

Lemma 4.2. *Under the condition of Theorem 4.1, we have*

$$\|(I - Q_\tau) y_h\|_W \lesssim M^{-1+\alpha/2}, \tag{67}$$

$$\|(I - Q_\tau) p_h\|_W \lesssim M^{-1+\alpha/2}. \tag{68}$$

Proof. We only prove (68), since (67) can be derived analogously. By (58) and (66) it remains to prove

$$\|(I - Q_\tau) p_{h,2}\|_{{}_0H^{\alpha/2}(0,T;L^2(\Omega))} \lesssim M^{-1+\alpha/2}, \tag{69}$$

$$\|(I - Q_\tau) p_{h,2}\|_{L^2(0,T;\dot{H}^1(\Omega))} \lesssim M^{-1+\alpha/2}. \tag{70}$$

We will give the proof of (69), the proof of (70) being easier. To this end, we proceed as follows. Let

$$g_1(t) := \begin{cases} (p_{h,2} - Q_\tau p_{h,2})(t) & \text{if } t \in (0, T/2), \\ 0 & \text{if } t \in (-\infty, 0) \cup (T/2, \infty), \end{cases}$$

$$g_2(t) := \begin{cases} (p_{h,2} - Q_\tau p_{h,2})(t) & \text{if } t \in (T/2, T), \\ 0 & \text{if } t \in (-\infty, T/2) \cup (T, \infty). \end{cases}$$

By [42, Lemma 16.3] we obtain

$$\|g_1\|^2_{H^{\alpha/2}(-\infty,\infty;L^2(\Omega))} \lesssim \mathbb{I}_1 + \mathbb{I}_2 + \mathbb{I}_3 + \mathbb{I}_4, \tag{71}$$

where

$$\begin{aligned}
\mathbb{I}_1 &:= \int_0^{t_1} \int_0^{t_1} \frac{\|g_1(s) - g_1(t)\|^2_{L^2(\Omega)}}{|s-t|^{1+\alpha}} \, \mathrm{d}s \, \mathrm{d}t, \\
\mathbb{I}_2 &:= \sum_{j=2}^{M} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t_j} \frac{\|g_1(s) - g_1(t)\|^2_{L^2(\Omega)}}{|s-t|^{1+\alpha}} \, \mathrm{d}s \, \mathrm{d}t, \\
\mathbb{I}_3 &:= \int_0^{t_1} \|g_1(t)\|^2_{L^2(\Omega)} \big((t_1 - t)^{-\alpha} + t^{-\alpha}\big) \, \mathrm{d}t, \\
\mathbb{I}_4 &:= \sum_{j=2}^{M} \int_{t_{j-1}}^{t_j} \|g_1(t)\|^2_{L^2(\Omega)} \big((t_j - t)^{-\alpha} + (t - t_{j-1})^{-\alpha}\big) \, \mathrm{d}t.
\end{aligned}$$

By (21), an elementary calculation gives the following four estimates:

$$\begin{aligned}
\mathbb{I}_1 &= 2 \int_0^{t_1} \int_t^{t_1} \frac{\|g_1(s) - g_1(t)\|^2_{L^2(\Omega)}}{|s-t|^{1+\alpha}} \, \mathrm{d}s \, \mathrm{d}t \\
&\lesssim \int_0^{t_1} \int_t^{t_1} \frac{(s^{\alpha r+\alpha} - t^{\alpha r+\alpha})^2}{(s-t)^{1+\alpha}} \, \mathrm{d}s \, \mathrm{d}t \\
&= \int_0^{t_1} \int_t^{t_1} \left(\left(\frac{s}{t}\right)^{\alpha r+\alpha} - 1 \right)^2 \left(\frac{s-t}{t}\right)^{-(1+\alpha)} t^{2\alpha r+\alpha-1} \, \mathrm{d}s \, \mathrm{d}t \\
&\lesssim \int_0^{t_1} t^{2\alpha r+\alpha} \int_1^{t_1/t} \big(x^{\alpha r+\alpha} - 1\big)^2 (x-1)^{-(1+\alpha)} \, \mathrm{d}x \, \mathrm{d}t \\
&\lesssim \int_0^{t_1} t^{2\alpha r+\alpha} \left(1 + \left(\frac{t_1}{t}\right)^{2\alpha r+\alpha} \right) \mathrm{d}t \\
&\lesssim t_1^{2\alpha r+\alpha+1},
\end{aligned}$$

$$
\begin{aligned}
\mathbb{I}_2 &= 2\sum_{j=2}^{M}\int_{t_{j-1}}^{t_j}\int_{t}^{t_j}\frac{\|g_1(s)-g_1(t)\|_{L^2(\Omega)}^2}{(s-t)^{1+\alpha}}\,\mathrm{d}s\,\mathrm{d}t \\
&\lesssim \sum_{j=2}^{M}\int_{t_{j-1}}^{t_j}\int_{t}^{t_j}\frac{(s-t)^2\|p_{h,2}'\|_{L^\infty(s,t;L^2(\Omega))}^2}{(s-t)^{1+\alpha}}\,\mathrm{d}s\,\mathrm{d}t \\
&\lesssim \sum_{j=2}^{M}\big(1+t_{j-1}^{2(\alpha r+\alpha-1)}\big)\int_{t_{j-1}}^{t_j}\int_{t}^{t_j}(s-t)^{1-\alpha}\,\mathrm{d}s\,\mathrm{d}t \\
&\lesssim \sum_{j=2}^{M}\big(1+t_{j-1}^{2(\alpha r+\alpha-1)}\big)(t_j-t_{j-1})^{3-\alpha},
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{I}_3 &\lesssim \|g_1\|_{L^\infty(0,t_1;L^2(\Omega))}^2\int_0^{t_1}\big((t_1-t)^{-\alpha}+t^{-\alpha}\big)\,\mathrm{d}t \\
&\lesssim t_1^{2(\alpha r+\alpha)}\int_0^{t_1}\big((t_1-t)^{-\alpha}+t^{-\alpha}\big)\,\mathrm{d}t \\
&\lesssim t_1^{2\alpha r+\alpha+1}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{I}_4 &\lesssim \sum_{j=2}^{M}(t_j-t_{j-1})^2\|p_{h,2}'\|_{L^\infty(t_{j-1},t_j;L^2(\Omega))}^2\int_{t_{j-1}}^{t_j}\big((t_j-t)^{-\alpha}+(t-t_{j-1})^{-\alpha}\big)\,\mathrm{d}t \\
&\lesssim \sum_{j=2}^{M}\big(1+t_{j-1}^{2(\alpha r+\alpha-1)}\big)(t_j-t_{j-1})^{3-\alpha}.
\end{aligned}
$$

Since

$$
\begin{aligned}
\sum_{j=2}^{M}(t_j-t_{j-1})^{3-\alpha} &\lesssim M^{(\alpha-3)\sigma_1}\sum_{j=2}^{M}\big(j^{\sigma_1}-(j-1)^{\sigma_1}\big)^{3-\alpha} \\
&\lesssim M^{(\alpha-3)\sigma_1}\sum_{j=2}^{M}j^{(\sigma_1-1)(3-\alpha)} \\
&\lesssim M^{\alpha-2}
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{j=2}^{M}t_{j-1}^{2(\alpha r+\alpha-1)}(t_j-t_{j-1})^{3-\alpha} \\
&\lesssim M^{-(2\alpha r+\alpha+1)\sigma_1}\sum_{j=2}^{M}(j-1)^{2(\alpha r+\alpha-1)\sigma_1}\big(j^{\sigma_1}-(j-1)^{\sigma_1}\big)^{3-\alpha} \\
&\lesssim M^{-(2\alpha r+\alpha+1)\sigma_1}\sum_{j=2}^{M}j^{2(\alpha r+\alpha-1)\sigma_1}j^{(\sigma_1-1)(3-\alpha)} \\
&= M^{-(2\alpha r+\alpha+1)\sigma_1}\sum_{j=2}^{M}j^{(2\alpha r+\alpha+1)\sigma_1+\alpha-3} \\
&\lesssim M^{\alpha-2}\quad\text{(by (41))},
\end{aligned}
$$

combining the above estimates for $\mathbb{I}_2$ and $\mathbb{I}_4$ yields

$$\mathbb{I}_2 + \mathbb{I}_4 \lesssim M^{\alpha-2}.$$

In addition, combining the above estimate for $\mathbb{I}_1$ and $\mathbb{I}_3$ yields

$$\mathbb{I}_1 + \mathbb{I}_3 \lesssim t_1^{2\alpha r+\alpha+1} \lesssim M^{-(2\alpha r+\alpha+1)\sigma_1} \lesssim M^{\alpha-2} \quad \text{(by (41))}.$$

Consequently, we conclude from (71) that

$$\|g_1\|_{H^{\alpha/2}(-\infty,\infty;L^2(\Omega))} \lesssim M^{\alpha-2}.$$

A similar argument yields

$$\|g_2\|_{H^{\alpha/2}(-\infty,\infty;L^2(\Omega))} \lesssim M^{\alpha-2}.$$

Therefore, (69) follows from the estimate

$$\begin{aligned}\|g_1 + g_2\|_{{}_0H^{\alpha/2}(0,T;L^2(\Omega))} &\lesssim \|g_1 + g_2\|_{H^{\alpha/2}(-\infty,\infty;L^2(\Omega))} \\ &\lesssim \|g_1\|_{H^{\alpha/2}(-\infty,\infty;L^2(\Omega))} + \|g_2\|_{H^{\alpha/2}(-\infty,\infty;L^2(\Omega))}.\end{aligned}$$

This completes the proof. ∎

Lemma 4.3. *Under the condition of Theorem 4.1, we have*

$$\|y_h - \widetilde{Y}\|_{L^2(0,T;L^2(\Omega))} \lesssim M^{-1}, \tag{72}$$

$$\|p_h - \widetilde{P}\|_{L^2(0,T;L^2(\Omega))} \lesssim M^{-1}. \tag{73}$$

Proof. By an energy argument (cf. [20]), we obtain

$$\begin{aligned}\|y_h - \widetilde{Y}\|_{L^2(0,T;L^2(\Omega))} &\lesssim M^{-\alpha/2}\|(I - Q_\tau)y_h\|_W, \\ \|p_h - \widetilde{P}\|_{L^2(0,T;L^2(\Omega))} &\lesssim M^{-\alpha/2}\|(I - Q_\tau)p_h\|_W,\end{aligned}$$

so that (72) and (73) follows from (67) and (68), respectively. This completes the proof. ∎

5 Numerical results

This section provides three numerical experiments to verify the theoretical results. We use uniform grids for the spatial discretization and employ a fixed point method [11] to solve the discrete system. The convergent condition is that the difference of the discrete control (in l^2 norm) between two steps is less than 1e-13. We adopt the following setting:

$$\begin{aligned}&\alpha = 0.4 \text{ or } 0.8; \quad r = 0 \text{ or } 0.25; \\ &\nu = 1; \quad T = 1; \quad \Omega = (0,1); \quad u_* = -0.1; \ u^* = 0.1; \\ &y_0(x) := x^{2r-0.49}(1-x) \quad \text{for all} \quad 0 \leqslant x \leqslant 1; \\ &y_d(x,t) := x^{-0.49}(1-x) \quad \text{for all} \quad 0 < x \leqslant 1 \text{ and } 0 \leqslant t \leqslant T.\end{aligned}$$

Let $U^{m,n}$ be the numerical solution of (46) with the mesh parameters $M = 2^m$, $h = 1/n$ and

$$\begin{cases} \sigma_1 = \max\left\{1, \frac{2-\alpha}{(2r-1)\alpha+1}\right\}, \\ \sigma_2 = \max\left\{1, \frac{2-\alpha}{\alpha+1}\right\}. \end{cases} \tag{74}$$

Define the discrete state and co-state respectively as

$$\begin{aligned} Y^{m,n} &:= S_{h\tau}(U^{m,n} + \mathrm{D}_{0+}^{\alpha}\, y_0), \\ P^{m,n} &:= S_{h\tau}^{*}(S_{h\tau}(U^{m,n} + \mathrm{D}_{0+}^{\alpha}\, y_0) - y_d). \end{aligned}$$

Throughout this section, $\|\cdot\|_{L^2(0,T;L^2(\Omega))}$ is abbreviated to $\|\cdot\|$ for convenience.

Experiment 1. This experiment verifies the spatial accuracy. The reference solutions are $U^{14,512}$, $Y^{14,512}$ and $P^{14,512}$. Tables 1 and 2 demonstrate that the accuracies of state are close to $\mathcal{O}(h^{\min\{2,1/\alpha+2r\}})$, and this agrees well with Theorem 4.1. In particular, it is observed that the convergence orders of the co-state and control are higher than the state in the case $\alpha = 0.8$.

	n	$\|Y^{14,512} - Y^{14,n}\|$		$\|P^{14,512} - P^{14,n}\|$		$\|U^{14,512} - U^{14,n}\|$	
$\alpha = 0.4$	10	2.12e-3	Order	1.60e-3	Order	1.50e-3	Order
	20	5.94e-4	1.84	4.38e-4	1.87	4.16e-4	1.85
	30	2.78e-5	1.87	2.03e-4	1.90	1.94e-4	1.89
	40	1.61e-5	1.90	1.17e-4	1.92	1.12e-4	1.91
	50	1.05e-5	1.99	7.57e-5	1.94	7.26e-5	1.94
$\alpha = 0.8$	10	1.01e-2	Order	1.60e-3	Order	1.49e-3	Order
	20	4.23e-3	1.25	4.38e-4	1.87	4.13e-4	1.85
	30	2.53e-3	1.27	2.03e-4	1.90	1.93e-4	1.88
	40	1.74e-3	1.30	1.17e-4	1.92	1.11e-4	1.90
	50	1.30e-3	1.31	7.57e-5	1.94	7.22e-5	1.94

Table 1: *Convergence history with* $r = 0$.

	n	$\|Y^{14,512} - Y^{14,n}\|$		$\|P^{14,512} - P^{14,n}\|$		$\|U^{14,512} - U^{14,n}\|$	
$\alpha = 0.4$	10	5.77e-4	Order	1.56e-3	Order	1.49e-3	Order
	20	1.45e-4	2.00	4.30e-4	1.86	4.14e-4	1.85
	30	6.43e-5	2.01	2.00e-4	1.89	1.93e-4	1.89
	40	3.58e-5	2.03	1.15e-4	1.91	1.11e-4	1.90
	50	2.28e-5	2.03	7.46e-5	1.94	7.23e-5	1.94
$\alpha = 0.8$	10	1.70e-3	Order	1.57e-3	Order	1.48e-3	Order
	20	5.07e-4	1.74	4.31e-3	1.86	4.12e-3	1.85
	30	2.46e-4	1.78	2.00e-3	1.89	1.92e-4	1.88
	40	1.45e-4	1.83	1.15e-3	1.91	1.11e-4	1.91
	50	9.58e-5	1.86	7.47e-4	1.94	7.21e-5	1.93

Table 2: *Convergence history with* $r = 0.25$.

Experiment 2. This experiment investigates the temporal accuracy with graded temporal grids. The reference solutions are $U^{14,512}$, $Y^{14,512}$ and $P^{14,512}$. Tables 3 and 4 illustrate that the temporal accuracy of the numerical control, state and co-state are close to $\mathcal{O}(M^{-1})$, which agrees well with Theorem 4.1.

	m	$\|Y^{m,512}-Y^{14,512}\|$		$\|P^{m,512}-P^{14,512}\|$		$\|U^{m,512}-U^{14,512}\|$	
	8	3.06e-4	Order	2.29e-4	Order	2.28e-4	Order
	9	1.54e-4	0.99	1.36e-4	0.75	1.36e-4	0.75
$\alpha=0.4$	10	7.72e-5	1.00	7.87e-5	0.79	7.86e-5	0.79
	11	3.85e-5	1.00	4.40e-5	0.84	4.39e-5	0.84
	12	1.88e-5	1.03	2.33e-5	0.91	2.33e-5	0.91
	8	9.13e-4	Order	2.02e-4	Order	1.97e-4	Order
	9	4.53e-4	1.01	1.01e-4	0.99	9.91e-5	0.99
$\alpha=0.8$	10	2.25e-4	1.01	5.03e-5	1.01	4.92e-5	1.01
	11	1.11e-4	1.02	2.46e-5	1.03	2.40e-5	1.03
	12	5.39e-5	1.04	1.17e-5	1.07	1.14e-5	1.07

Table 3: *Convergence history with* $r = 0$.

	m	$\|Y^{m,512}-Y^{14,512}\|$		$\|P^{m,512}-P^{14,512}\|$		$\|U^{m,512}-U^{14,512}\|$	
	8	4.30e-4	Order	2.23e-4	Order	2.22e-4	Order
	9	2.33e-4	0.89	1.33e-4	0.75	1.32e-4	0.75
$\alpha=0.4$	10	1.25e-4	0.90	7.67e-5	0.79	7.66e-5	0.79
	11	6.57e-5	0.92	4.29e-5	0.84	4.28e-5	0.84
	12	3.37e-5	0.96	2.28e-5	0.91	2.28e-5	0.91
	8	1.30e-3	Order	1.95e-4	Order	1.94e-4	Order
	9	7.31e-4	0.84	9.82e-5	0.99	9.76e-5	0.99
$\alpha=0.8$	10	4.05e-4	0.85	4.88e-5	1.01	4.85e-5	1.01
	11	2.19e-4	0.88	2.38e-5	1.03	2.38e-5	1.03
	12	1.14e-4	0.95	1.13e-5	1.08	1.13e-5	1.08

Table 4: *Convergence history with* $r = 0.25$.

Experiment 3. This experiment investigates the temporal accuracy with uniform temporal grids. The reference solutions are $U^{14,512}$, $Y^{14,512}$ and $P^{14,512}$, and we use $\widetilde{U}^{m,512}$, $\widetilde{Y}^{m,512}$ and $\widetilde{P}^{m,512}$ to denote the corresponding numerical solutions of (46) with the mesh parameters $M = 2^m$, $h = 512$ and $\sigma_1 = \sigma_2 = 1$. From Tables 5 and 6, it is easy to see that the errors are generally larger than the cases with graded temporal grids.

	m	$\|\widetilde{Y}^{m,512}-Y^{14,512}\|$		$\|\widetilde{P}^{m,512}-P^{14,512}\|$		$\|\widetilde{U}^{m,512}-U^{14,512}\|$	
	8	4.66e-3	Order	3.88e-4	Order	3.64e-4	Order
	9	3.35e-3	0.48	2.49e-4	0.64	2.36e-4	0.63
$\alpha=0.4$	10	2.40e-3	0.48	1.56e-4	0.67	1.49e-4	0.66
	11	1.71e-3	0.48	9.55e-5	0.71	9.19e-5	0.70
	12	1.22e-3	0.49	5.72e-5	0.74	5.53e-5	0.73
	8	1.11e-2	Order	2.09e-4	Order	1.97e-4	Order
	9	7.86e-3	0.49	1.05e-4	0.99	9.97e-5	0.99
$\alpha=0.8$	10	5.57e-3	0.50	5.21e-5	1.01	4.95e-5	1.01
	11	3.94e-3	0.50	2.55e-5	1.03	2.42e-5	1.03
	12	2.78e-4	0.50	1.22e-5	1.06	1.15e-5	1.07

Table 5: *Convergence history with* $r = 0$.

	m	$\|\widetilde{Y}^{m,512} - Y^{14,512}\|$		$\|\widetilde{P}^{m,512} - P^{14,512}\|$		$\|\widetilde{U}^{m,512} - U^{14,512}\|$	
	8	2.14e-3	Order	3.60e-4	Order	3.53e-4	Order
	9	1.45e-3	0.56	2.33e-4	0.63	2.29e-4	0.62
$\alpha = 0.4$	10	9.68e-4	0.58	1.47e-4	0.67	1.45e-4	0.66
	11	6.42e-4	0.59	9.02e-5	0.70	8.93e-5	0.70
	12	4.24e-5	0.60	5.42e-5	0.73	5.37e-5	0.73
	8	2.45e-3	Order	1.97e-4	Order	1.94e-4	Order
	9	1.51e-3	0.70	9.92e-5	0.99	9.78e-5	0.99
$\alpha = 0.8$	10	9.24e-4	0.71	4.92e-5	1.01	4.85e-5	1.01
	11	5.62e-4	0.72	2.41e-5	1.03	2.37e-5	1.03
	12	3.39e-4	0.73	1.14e-5	1.07	1.13e-5	1.07

Table 6: *Convergence history with* $r = 0.25$.

A Regularity of a fractional diffusion equation

A.1 Regularity in interpolation spaces

We first introduce the interpolation space theory (cf. [27, Chapter 2, pp. 54–55]). Assume that (X, Y) is an interpolation couple of complex Banach spaces. For any $0 < \theta_1 < \theta_2 < 1$, $0 < \theta < 1$ and $1 \leqslant q \leqslant \infty$,

$$([X, Y]_{\theta_1}, [X, Y]_{\theta_2})_{\theta,q} = (X, Y)_{(1-\theta)\theta_1+\theta\theta_2,q} \tag{75}$$

with equivalent norms, where $[\cdot, \cdot]_\theta$ and $(\cdot, \cdot)_{\theta,q}$ denote the interpolation spaces defined by the complex method and the real method, respectively. For each $w \in [X, Y]_\theta$ with $0 < \theta < 1$,

$$K(t, w) \leqslant 2t^\theta \|w\|_{[X,Y]_\theta}, \quad t > 0, \tag{76}$$

where

$$K(t, w) := \inf_{w=x+y,x\in X,y\in Y} \|x\|_X + t\|y\|_Y.$$

If Y is continuously embedded into X, then

$$[X, Y]_{\theta_2} \text{ is continuously embedded into } [X, Y]_{\theta_1} \tag{77}$$

for any $0 < \theta_1 < \theta_2 < 1$.

Lemma A.1. *If* $0 < r < s < 1$ *and* $1 \leqslant q < \infty$*, then*

$$\|w\|_{(X,Y)_{r,q}} \leqslant C\|w\|_{[X,Y]_s} \tag{78}$$

for all $w \in [X, Y]_s$*, where* C *is a positive constant independent of* w*.*

Proof. A straightforward calculation gives that

$$\begin{aligned}
&\|w\|^q_{(X,Y)_{r,q}} = \int_0^\infty |t^{-r}K(t, w)|^q \frac{\mathrm{d}t}{t} \\
&\leqslant 2^q \int_0^1 \left(t^{s-r}\|w\|_{[X,Y]_s}\right)^q \frac{\mathrm{d}t}{t} + 2^q \int_1^\infty \left(t^{-r/2}\|w\|_{[X,Y]_{r/2}}\right)^q \frac{\mathrm{d}t}{t} \quad (\text{by } (76)) \\
&= \frac{2^q}{(s-r)q}\|w\|^q_{[X,Y]_s} + \frac{2^{q+1}}{qr}\|w\|^q_{[X,Y]_{r/2}} \\
&\leqslant C\|w\|^q_{[X,Y]_s} \quad (\text{by } (77)),
\end{aligned}$$

where C is a positive constant independent of w. This proves (78) and hence this lemma. ∎

For $m \in \mathbb{N}_{>0}$, $0 < \theta < 1$ and $1 \leqslant q < \infty$, define

$$ {}_0H^{m\theta,q}(0,T;X) := \big[L^q(0,T;X),\, {}_0W^{m,q}(0,T;X)\big]_\theta, $$

where X is a Hilbert space. We use ${}_0H^{0,q}(0,T;X)$ to denote the space $L^q(0,T;X)$. For $0 < \beta < \infty$ and $1 \leqslant q < \infty$, we have the following properties: if $\beta \in \mathbb{N}_{>0}$ then

$$ {}_0H^{\beta,q}(0,T;X) = {}_0W^{\beta,q}(0,T;X) \quad \text{with equivalent norms;} \tag{79} $$

if $q = 2$ then (cf. [27, Corollary 4.37])

$$ {}_0H^{\beta,q}(0,T;X) = {}_0W^{\beta,q}(0,T;X) \quad \text{with equivalent norms.} \tag{80} $$

By [37, Theorem 4.5.15], for any $g \in {}_0H^{\beta,q}(0,T;L^2(\Omega))$ with $0 \leqslant \beta < \infty$ and $1 < q < \infty$, there exists a unique $Sg \in {}_0H^{\alpha+\beta,q}(0,T;L^2(\Omega)) \bigcap {}_0H^{\beta,q}(0,T;\dot{H}^2(\Omega))$ such that

$$ (\mathrm{D}_{0+}^\alpha - \Delta)Sg = g; $$

moreover,

$$ \|Sg\|_{{}_0H^{\alpha+\beta,q}(0,T;L^2(\Omega))} + \|Sg\|_{{}_0H^{\beta,q}(0,T;\dot{H}^2(\Omega))} \leqslant C_{\alpha,\beta,q}\|g\|_{{}_0H^{\beta,q}(0,T;L^2(\Omega))}. \tag{81} $$

Lemma A.2. *Assume that $1 < q < \infty$. If $g \in L^q(0,T;L^2(\Omega))$, then*

$$ \|Sg\|_{{}_0W^{\beta,q}(0,T;L^2(\Omega))} \leqslant C_{\alpha,\beta,q}\|g\|_{L^q(0,T;L^2(\Omega))} \tag{82} $$

for all $0 < \beta < \alpha$. If $g \in {}_0W^{\beta,q}(0,T;L^2(\Omega))$ with $0 < \beta < \infty$, then

$$ \|Sg\|_{{}_0W^{\beta,q}(0,T;\dot{H}^2(\Omega))} \leqslant C_{\alpha,\beta,q}\|g\|_{{}_0W^{\beta,q}(0,T;L^2(\Omega))}. \tag{83} $$

If $g \in {}_0W^{\beta,q}(0,T;L^2(\Omega))$ with $0 < \beta < \infty$ and $\alpha + \beta \notin \mathbb{N}$, then

$$ \|Sg\|_{{}_0W^{\alpha+\beta,q}(0,T;L^2(\Omega))} \leqslant C_{\alpha,\beta,q}\|g\|_{{}_0W^{\beta,q}(0,T;L^2(\Omega))}. \tag{84} $$

Proof. Estimate (82) follows from Lemma A.1 and (81), and estimate (83) follows from (79), (81) and [27, Theorem 1.6]. Let us proceed to prove (84) for the case $\beta \in (0,1] \setminus \{1-\alpha\}$. By (81) it holds that

$$ \|Sw\|_{{}_0H^{\alpha,q}(0,T;L^2(\Omega))} \leqslant C_{\alpha,q}\|w\|_{L^q(0,T;L^2(\Omega))} $$

for all $w \in L^q(0,T;L^2(\Omega))$ and that

$$ \|Sw\|_{{}_0H^{\alpha+1,q}(0,T;L^2(\Omega))} \leqslant C_{\alpha,q}\|w\|_{{}_0W^{1,q}(0,T;L^2(\Omega))} $$

for all $w \in {}_0W^{1,q}(0,T;L^2(\Omega))$. Hence, applying (75) and the real interpolation of type (β,q) yields that

$$ \|Sw\|_{{}_0W^{\alpha+\beta,q}(0,T;L^2(\Omega))} \leqslant C_{\alpha,\beta,q}\|w\|_{{}_0W^{\beta,q}(0,T;L^2(\Omega))} $$

for all $w \in {}_0W^{\beta,q}(0,T;L^2(\Omega))$. For $1 < \beta < \infty$ and $\alpha + \beta \notin \mathbb{N}$, (84) can be proved analogously. This completes the proof. ∎

Remark A.1. *In the above proof of* (84)*, we have used the following well-known result (cf. [27, Proposition 3.8]): if $j, k, m, n \in \mathbb{N}$ and $0 < r, s < 1$ satisfy that $j < k$, $m < n$ and $(1-r)j + rk = (1-s)m + sn \notin \mathbb{N}$, then*

$$\begin{aligned} &\big({}_0W^{j,q}(0,T;L^2(\Omega)),\, {}_0W^{k,q}(0,T;L^2(\Omega))\big)_{r,q} \\ &= \big({}_0W^{m,q}(0,T;L^2(\Omega)),\, {}_0W^{n,q}(0,T;L^2(\Omega))\big)_{s,q}, \end{aligned}$$

with equivalent norms, where $1 \leqslant q < \infty$.

A.2 Regularity from the Mittag-Leffler function

For any $\beta, \gamma > 0$, define the Mittag-Leffler function $E_{\beta,\gamma}$ by that

$$E_{\beta,\gamma}(z) := \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(k\beta + \gamma)}, \quad z \in \mathbb{C}. \tag{85}$$

This function has a useful growth estimate [36]:

$$|E_{\beta,\gamma}(-t)| \leqslant \frac{C_{\beta,\gamma}}{1+t}, \quad t > 0. \tag{86}$$

Moreover, the Mittag-Leffler function admits the asymptotic expansion (cf. [5, pp. 207]):

$$E_{\beta,1}(-t) = \sum_{k=1}^{N} \frac{(-1)^{k+1} t^{-k}}{\Gamma(1-k\beta)} + O(t^{-N-1}), \quad \text{as } t \to \infty.$$

A straightforward calculation gives that [40], for any $v \in L^2(\Omega)$,

$$(Sv)(t) = t^{\alpha} \sum_{k=0}^{\infty} E_{\alpha,1+\alpha}(-\lambda_k t^{\alpha})(v, \phi_k)_{\Omega}\, \phi_k, \tag{87}$$

$$(S\,\mathrm{D}_{0+}^{\alpha}\, v)(t) = \sum_{k=0}^{\infty} E_{\alpha,1}(-\lambda_k t^{\alpha})(v, \phi_k)_{\Omega}\, \phi_k, \tag{88}$$

for each $0 \leqslant t \leqslant T$. Here

$$\{\phi_k : k \in \mathbb{N}\} \subset H_0^1(\Omega) \cap H^2(\Omega) \tag{89}$$

is an orthonormal basis of $L^2(\Omega)$ such that for all $k \in \mathbb{N}$,

$$-\Delta \phi_k = \lambda_k \phi_k \quad \text{in } \Omega, \tag{90}$$

where $\{\lambda_k : \ k \in \mathbb{N}\} \subset \mathbb{R}_{>0}$ is a non-decreasing sequence.

By the above properties of Mittag-Leffler function, [28, Lemma 3.4] and some techniques in [31], a few straightforward calculations yield the following two lemmas.

Lemma A.3. *Assume that $0 < t < T$ and $v \in L^2(\Omega)$ then*

$$t^{-\alpha} \|(Sv)(t)\|_{L^2(\Omega)} + t^{1-\alpha} \|(Sv)'(t)\|_{L^2(\Omega)} \leqslant C_\alpha \|v\|_{L^2(\Omega)}, \tag{91}$$

$$(T-t)^{-\alpha} \|(S^*v)(t)\|_{L^2(\Omega)} + (T-t)^{1-\alpha} \|(S^*v)'(t)\|_{L^2(\Omega)} \leqslant C_\alpha \|v\|_{L^2(\Omega)}. \tag{92}$$

Moreover, if $v \in \dot{H}^{2r}(\Omega)$ *with* $0 < r < 1$ *then*

$$\|(S\,\mathrm{D}_{0+}^{\alpha}\, v)(t)\|_{L^2(\Omega)} \leqslant C_\alpha t^{\alpha r} \|v\|_{\dot{H}^{2r}(\Omega)}. \tag{93}$$

$$\|(S\,\mathrm{D}_{0+}^{\alpha}\, v)'(t)\|_{L^2(\Omega)} + t^{\alpha/2} \|(S\,\mathrm{D}_{0+}^{\alpha}\, v)'\|_{\dot{H}^1(\Omega)} \leqslant C_\alpha t^{\alpha r-1} \|v\|_{\dot{H}^{2r}(\Omega)}. \tag{94}$$

Lemma A.4. *Assume that* $v \in \dot{H}^{2r}(\Omega)$ *with* $0 < r < 1$*. If* $0 < \alpha < 1/3$*, then*

$$\|S\,\mathrm{D}_{0+}^{\alpha}\, v\|_{{}_0H^{\alpha/2}(0,T;\dot{H}^{2+2r}(\Omega))} \leqslant C_{\alpha,r,\Omega,T} \|v\|_{\dot{H}^{2r}(\Omega)}.$$

If $\alpha = 1/3$*, then for any* $0 < \epsilon < 1/2$*,*

$$\|S\,\mathrm{D}_{0+}^{\alpha}\, v\|_{{}_0H^{\alpha/2}(0,T;\dot{H}^{2+2r-\epsilon}(\Omega))} \leqslant C_{\alpha,r,\Omega,T}\, \epsilon^{-1/2} \|v\|_{\dot{H}^{2r}(\Omega)}.$$

If $1/3 < \alpha < 1$*, then*

$$\|S\,\mathrm{D}_{0+}^{\alpha}\, v\|_{{}_0H^{\alpha/2}(0,T;\dot{H}^{1/\alpha+2r-1}(\Omega))} \leqslant C_{\alpha,r,\Omega,T} \|v\|_{\dot{H}^{2r}(\Omega)}.$$

If $0 < \alpha < 1/2$*, then*

$$\|S\,\mathrm{D}_{0+}^{\alpha}\, v\|_{L^2(0,T;\dot{H}^{2+2r}(\Omega))} \leqslant C_{\alpha,r,\Omega,T} \|v\|_{\dot{H}^{2r}(\Omega)}.$$

If $\alpha = 1/2$*, then for any* $0 < \epsilon < 1/2$*,*

$$\|S\,\mathrm{D}_{0+}^{\alpha}\, v\|_{L^2(0,T;\dot{H}^{2+2r-\epsilon}(\Omega))} \leqslant C_{\alpha,r,\Omega,T}\, \epsilon^{-1/2} \|v\|_{\dot{H}^{2r}(\Omega)}.$$

If $1/2 < \alpha < 1$*, then*

$$\|S\,\mathrm{D}_{0+}^{\alpha}\, v\|_{L^2(0,T;\dot{H}^{2r+1/\alpha}(\Omega))} \leqslant C_{\alpha,r,\Omega,T} \|v\|_{\dot{H}^{2r}(\Omega)}.$$

References

[1] H. Antil and E. Otárola. A FEM for an optimal control problem of fractional powers of elliptic operators. *SIAM Journal on Control and Optimization*, 53(6):3432–3456, 2015.

[2] H. Antil, E. Otárola and A. J. Salgado. A space-time fractional optimal control problem: analysis and discretization. *SIAM Journal on Control and Optimization*, 54(3):1295–1328, 2016.

[3] J. Bouchaud and A. Georges. Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications. *Physics Reports*, 195(4–5):127–293, 1990.

[4] K. Diethelm. *The analysis of fractional differential equations*. Springer Berlin, 2010.

[5] A. Erdelyi. *Higher transcendental functions Volume 3*. McGraw-Hill, New York, 1955.

[6] V. Ervin and J. Roop. Variational formulation for the stationary fractional advection dispersion equation. *Numerical Methods for Partial Differential Equations*, 22(3):558–576, 2006.

[7] G. Gao and Z. Sun. A compact finite difference scheme for the fractional sub-diffusion equations. *Journal of Computational Physics*, 230(3):586–595, 2011.

[8] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer, Berlin, 2001.

[9] W. Gong and N. Yan. Adaptive finite element method for elliptic optimal control problems: convergence and optimality. *Numerische Mathematik*, 135(4):1121–1170, 2017.

[10] M. Gunzburger and J. Wang. Error analysis of fully discrete finite element approximations to an optimal control problem governed by a time-fractional PDE. *SIAM Journal on Control and Optimization*, 57(1):241–263, 2019.

[11] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich. *Optimization with PDE Constraints*. Springer Dordrecht, 2009.

[12] M. Hinze and M. Vierling. The semi-smooth Newton method for variationally discretized control constrained elliptic optimal control problems; implementation, convergence and globalization. *Optimization Methods & Software*, 27(6):933–950, 2012.

[13] B. Jin, R. Lazarov and Z. Zhou. Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data. *SIAM Journal of Scientific Computing*, 38(1):A146–A170, 2016.

[14] B. Jin, B. Li and Z. Zhou. Discrete maximal regularity of time-stepping schemes for fractional evolution equations. *Numerische Mathematik*, 138(1):101–131, 2018.

[15] B. Jin, B. Li and Z. Zhou. Numerical analysis of nonlinear subdiffusion equations. *SIAM Journal on Numerical Analysis,* 56(1):1–23, 2018.

[16] B. Jin, B. Li and Z. Zhou. Pointwise-in-time error estimates for an optimal control problem with subdiffusion constraint. *IMA Journal of Numerical Analysis*, 40:377–404, 2020.

[17] T. Langlands and B. Henry. The accuracy and stability of an implicit solution method for the fractional diffusion equation. *Journal of Computational Physic*, 205(2):719–736, 2005.

[18] D. Leykekhman and B. Vexler. Optimal a priori error estimates of parabolic optimal control problems with pointwise control. *SIAM Journal on Numerical Analysis*, 5(51):2797–2821, 2013.

[19] B. Li, H. Luo and X. Xie. A time-spectral algorithm for fractional wave problems. *Journal of Scientific Computing*, 7(2):1164–1184, 2018.

[20] B. Li, H. Luo and X. Xie. Analysis of a time-stepping scheme for time fractional diffusion problems with nonsmooth data. *SIAM Journal on Numerical Analysis*, 57(2):779–798, 2019.

[21] R. Li, W. Liu, H. Ma and T. Tang. Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM Journal on Control and Optimization*, 41(5):1321–1349, 2002.

[22] X. Li and C. Xu. A space-time spectral method for the time fractional diffusion equation. *SIAM Journal on Numerical Analysis*, 47(3):2108–2131, 2009.

[23] Y. Lin and C. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *Journal of Computational Physics*, 225(2):1533–1552, 2007.

[24] J. L. Lions. *Optimal control of systems governed by partial differential equations*. Springer-Verlag, Berlin, 1971.

[25] A. Lotfi, M. Dehghan, S. A. Yousefi. A numerical technique for solving fractional optimal control problems. *Computers & Mathematics with Applications*, 62(3):1055-1067, 2011.

[26] C. Lubich, I. Sloan and V. Thomée. Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term. *Mathematics of Computation*, 65(213):1–17, 1996.

[27] A. Lunardi. *Interpolation theory*. Edizioni della Normale, Pisa, 2018.

[28] H. Luo, B. Li and X. Xie. Convergence analysis of a Petrov–Galerkin method for fractional wave problems with nonsmooth data. *Journal of Scientific Computing*, 80(2):957–992, 2019.

[29] W. McLean and K. Mustapha. Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation. *Numerical Algorithms*, 52(1):69–88, 2009.

[30] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1–77, 2000.

[31] W. McLean. Regularity of solutions to a time-fractional diffusion equation. *Anziam Journal*, 52(2):123–138, 2010.

[32] K. Mustapha and W. McLean. Discontinuous Galerkin method for an evolution equation with a memory term of positive type. *Mathematics of Computation*, 78(268):1975–1995, 2009.

[33] K. Mustapha and W. McLean. Piecewise-linear, discontinuous Galerkin method for a fractional diffusion equation. *Numerical Algorithms*, 56(2):159–184, 2011.

[34] K. Mustapha and W. McLean. Uniform convergence for a discontinuous Galerkin, time-stepping method applied to a fractional diffusion equation. *IMA J. Numer. Anal.*, 32(3):906–925, 2012.

[35] K. Mustapha, B. Abdallah, and K. Furati. A discontinuous Petrov-Galerkin method for time-fractional diffusion equations. *SIAM Journal on Numerical Analysis*, 52(5):2512–2529, 2014.

[36] I. Podlubny. *Fractional differential equations.* Academic Press, 1998.

[37] J. Prüss and G. Simonett. *Moving interfaces and quasilinear parabolic evolution equations.* Birkhäuser Basel, 2016.

[38] A. Rösch and B. Vexler. Optimal control of the stokes equations: A priori error analysis for finite element discretization with postprocessing. *SIAM Journal on Numerical Analysis*, 44(5):1903–1920, 2006.

[39] A. Rösch and G. Wachsmuth. Mass lumping for the optimal control of elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 55(3):1412–1436, 2017.

[40] K. Sakamoto and M. Yamamoto. Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. *Journal of Mathematical Analysis and Applications*, 55(3):1412–1436, 2017.

[41] R. Schumer, M. M. Meerschaert and B. Baeumer. Fractional advection-dispersion equations for modeling transport at the earth surface. *Journal of Geophysical Research Earth Surface*, 114(F4):F00A07, 2009.

[42] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces.* Springer Berlin, 2007.

[43] F. Tröltzsch. Optimal control of partial differential equations. *Graduate studies in mathematics*, 2010.

[44] V. R. Voller. On a fractional derivative form of the Green-Ampt infiltration model. *Advances in Water Resources*, 34(2):257–262, 2011.

[45] T. Wang, C. Yang, and X. Xie. A Nitsche-eXtended finite element method for distributed optimal control problems of elliptic interface equations. *Computational Methods in Applied Mathematics*, doi:10.1515/cmam-2018-0256; arXiv:1810.02271; 2019.

[46] T. Wang, C. Yang, and X. Xie. Extended finite element methods for optimal control problems governed by Poisson equation in non-convex domains. *Science China Mathematics*. https://doi.org/10.1007/s11425-018-9487-4, 2019.

[47] Y. Yang, Y. Chen, Y. Huang, and H. Wei. Spectral collocation method for the time-fractional diffusion-wave equation and convergence analysis. *Computers & Mathematics with Applications*, 73(6):1218–1232, 2017.

[48] X. Ye and C. Xu. A space-time spectral method for the time fractional diffusion optimal control problems. *Advances in Difference Equations*, 156 (2015), 2015.

[49] S. Yuste and L. Acedo. An explicit finite difference method and a new von Neumann-type stability analysis for fractional diffusion equations. *SIAM Journal on Numerical Analysis*, 42(5):1862–1874, 2005.

[50] F. Zeng, C. Li, F. Liu and I. Turner. The use of finite difference/element approaches for solving the time-fractional subdiffusion equation. *SIAM Journal on Scientific Computing*, 35(6):2976–3000, 2013.

[51] C. Zhang, H. Liu and Z. Zhou. A priori error analysis for time-stepping discontinuous Galerkin finite element approximation of time fractional optimal control problem. *Journal of Scientific Computing*, 80(2):993–1018, 2019.

[52] Z. Zhou and W. Gong. Finite element approximation of optimal control problems governed by time fractional diffusion equation. *Computers & Mathematics with Applications*, 71(1):301–318, 2016.

[53] Z. Zhou and C. Zhang. Time-stepping discontinuous Galerkin approximation of optimal control problem governed by time fractional diffusion equation. *Numerical Algorithms*, 79:437–455, 2018.