



Finite Element Approximation of Optimal Control Problem Governed by Space Fractional Equation

Zhaojie Zhou¹ · Zhiyu Tan²

Received: 2 May 2018 / Revised: 2 September 2018 / Accepted: 10 September 2018 /
Published online: 21 September 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

In this paper we investigate finite element approximation of optimal control problem governed by space fractional diffusion equation with control constraints. The control variable is approximated by piecewise constant. Regularity estimate for the control problem is proved based on the first order optimality system and a priori error estimates for the state, the adjoint state and the control variables are derived. Due to the nonlocal property of fractional derivative, which will leads to a full stiff matrix, we develop a fast primal dual active set algorithm for the control problem. Numerical examples are given to illustrate the theoretical findings and the efficiency of the fast algorithm.

Keywords Finite element method · Optimal control problem · Space fractional equation · Primal dual active set algorithm · Fast algorithm.

1 Introduction

Our main goal in this paper is to study finite element discretization of optimal control problem governed by 1-D space fractional diffusion equation. We consider the following control problem:

$$\min_{u \in U_{ad}} J(y, u) := \frac{1}{2} \|y - y_d\|_0^2 + \frac{\gamma}{2} \|u\|_0^2 \quad (1.1)$$

The research was supported by Natural Science Foundation of Shandong Province (No. ZR2016JL004, ZR2017MA020) and National Natural Science Foundation of China (No. 11301311, 11471196).

✉ Zhiyu Tan
zhiyutan@amss.ac.cn

Zhaojie Zhou
zhouzhaojie@sdnu.edu.cn

¹ School of Mathematics and Statistics, Shandong Normal University, Jinan, China

² Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong

subject to

$$\begin{cases} \mathcal{L}_r^\alpha y = f + u & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma, \end{cases} \quad (1.2)$$

where $\Omega = (0, 1)$, $\Gamma = \partial\Omega$, $y_d \in L^2(\Omega)$ is the desired state, $\gamma > 0$ is the regularization parameter, $f \in L^2(\Omega)$ is a given function and the control constraint

$$U_{ad} = \{u \in L^2(\Omega) : u_a \leq u(x) \leq u_b \text{ a.e. in } \Omega \text{ with } u_a, u_b \in \mathbb{R} \text{ and } u_a \leq u_b\}.$$

Here the fractional differential operator \mathcal{L}_r^α is defined by

$$\mathcal{L}_r^\alpha y := -D(r {}_0D_x^{-(2-\alpha)} + (1-r)_xD_1^{-(2-\alpha)})Dy.$$

The parameters $r \in [0, 1]$ and $\alpha \in (1, 2)$. More details about fractional derivative and integral will be specified later.

In the past decades, lots of researches show that anomalous diffusion phenomena widely exists in real world applications, for example, contaminant transport in groundwater flow. In [1,2], it was shown that solutes moving through aquifers do not generally follow a classical second-order Fickian diffusion equation. The heavy tail behavior of the transport processes can be accurately described by Levy distribution, which can be viewed as a probability description of fractional diffusion equations. Due to the self-similarity the plume spreads faster than a traditional Brownian motion. The traditional dispersion equation would seriously underestimate the risk of downstream contamination if the plume represent a pollutant heading to a drinking water well. The stable density that solves the fractional diffusion equation can capture the super-diffusive spreading observed in the data. Motivated by above facts, in this paper we mainly focus on optimal control problem governed by a space fractional diffusion equation described in (1.2).

In recent years, the research of optimal control problem governed by fractional PDEs forms a hot topic not only in model problems, but also in numerical methods. For the model problems, a distributed optimal control problem governed by time fractional diffusion equation with Riemann–Liouville derivative was discussed in [3] and the first order optimality system was derived there. In [4], the authors studied the optimal control problem governed by time fractional diffusion equation with state constraints and proved the well-posedness of the control problem. The controllability of the time fractional diffusion equation was discussed in [5]. A new type of identification problems was studied in [6], where the fractional order in a nonlocal evolution equation was identified and the well-posedness of the identification problem was proved.

On the numerical method aspect, spectral approximation of time fractional optimal control problems were studied in [7–9] and error estimates were derived under the assumption that both state and control variables are sufficiently smooth. In [10], Legendre pseudo-spectral method combined with $L1$ scheme was applied to approximate optimal control problem governed by a time-fractional diffusion equation. Numerical simulation of distributed-order fractional optimal control problems was investigated in [11]. A fully spectral collocation scheme was proposed. Besides, finite element approximation and finite difference approximation of fractional optimal control problems are also widely studied. In [12–15], optimal control problems governed by fractional Laplacian were investigated, where the fractional Laplacian operator was characterized as fractional powers of the Dirichlet Laplace operator in the sense of spectral theory. By using the Caffarelli–Silvestre extension, the fractional Laplacian equation was realized as the Dirichlet-to-Neumann map for a nonuniformly elliptic problem posed on a semi-infinite cylinder in one more spatial dimension, which overcame the

nonlocality of the fractional Laplacian operator. Finite element discretization of the control problem was investigated, and a priori as well as a posteriori error estimate were derived. In [16], the authors investigated the controllability of a one-dimensional heat equation involving the fractional Laplacian operator both from theoretical and numerical aspects. Finite element approximation of time fractional optimal control problems was studied in [17], where the well-posedness of control problem and optimal-order error estimates for the space semidiscrete approximation were proved. Pointwise-in-time error estimates for an optimal control problem with subdiffusion constraint was derived in [18] for $L1$ and back Euler discretization of time. A fast projected gradient algorithm for optimal control problem governed by space fractional diffusion equation was developed in [19] based on finite difference discretization of the state equation.

The present work is devoted to develop a rigorous error estimate for finite element approximation of optimal control problem governed by space fractional diffusion equation. The control space is approximated by the piecewise constant finite element space. The regularity estimates for the state, the adjoint state and the control variables are derived based on the first order optimality system. Note that the fractional derivative is nonlocal, which leads to a full matrix in finite element or finite difference discretization of the state equation as well as the adjoint equation. To reduce the computational cost, we develop a fast primal dual active set algorithm based on the Toeplitz structure of the coefficient matrix ([20]) in the discrete state and adjoint state equations. Finally, numerical examples are given to illustrate the theoretical results.

Our paper is organized as follows. In next section, we present some preliminary knowledge about the fractional derivative and integral, first. Then we derive the first order optimality system and prove some results with respect to the regularity of the solutions. In Sect. 3, we consider the finite element approximation of control problem and prove a priori error estimates for the state variable, the adjoint state variable and the control variable. To save computational cost, a fast algorithm based on the primal dual active set strategy is developed in Sect. 4. In Sect. 5, numerical example is given to confirm our theoretical findings. Finally, we draw some concluding remarks in Sect. 6.

2 Optimal Control Problem

In this section, we begin with a brief review of the definition of fractional integral, derivative and the related Sobolev space.

For a function u defined on the interval Ω and $\beta > 0$, we have the left and right fractional integrals of order β defined by:

$${}_0D_x^{-\beta}u(x) := \frac{1}{\Gamma(\beta)} \int_0^x \frac{u(s)}{(x-s)^{1-\beta}} ds,$$

$${}_xD_1^{-\beta}u(x) := \frac{1}{\Gamma(\beta)} \int_x^1 \frac{u(s)}{(s-x)^{1-\beta}} ds.$$

The left and right Caputo fractional derivatives of order $1 - \sigma$ with $0 \leq \sigma < 1$ are defined by:

$${}_0^C D_x^{1-\sigma}u(x) = \frac{1}{\Gamma(\sigma)} \int_0^x \frac{u'(s)}{(x-s)^{1-\sigma}} ds = {}_0D_x^{-\sigma} Du(x),$$

$${}_x^C D_1^{1-\sigma}u(x) = -\frac{1}{\Gamma(\sigma)} \int_x^1 \frac{u'(s)}{(s-x)^{1-\sigma}} ds = -{}_xD_1^{-\sigma} Du(x).$$

Similarly, the left and right Riemann–Liouville fractional derivatives of order $1 - \sigma$ are defined by:

$$\begin{aligned}
 {}_0D_x^{1-\sigma}u(x) &= \frac{1}{\Gamma(\sigma)} \frac{d}{dx} \int_0^x \frac{u(s)}{(x-s)^{1-\sigma}} ds = D_0D_x^{-\sigma}u(x), \\
 {}_xD_1^{1-\sigma}u(x) &= -\frac{1}{\Gamma(\sigma)} \frac{d}{dx} \int_x^1 \frac{u(s)}{(s-x)^{1-\sigma}} ds = -D_xD_1^{-\sigma}u(x).
 \end{aligned}$$

For $s \geq 0$, let $H^s(\Omega)$ denote the Sobolev space of order s on the interval Ω and $\tilde{H}^s(\Omega)$ denote the set of functions in $H^s(\Omega)$ whose extension by 0 are in $H^s(R)$. For u defined on Ω and \tilde{u} its extension by zero, $\tilde{H}^s(\Omega)$ is the closure of $C_0^\infty(\Omega)$ with the norm $\|u\|_{H^s(\Omega)} := \|\tilde{u}\|_{\tilde{H}^s(R)}$.

Let ${}_2F_1$ denote the Gaussian three-parameter hypergeometric function, which is defined by

$$\begin{aligned}
 {}_2F_1(a; b; c; x) &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 z^{b-1}(1-z)^{c-b-1}(1-zx)^{-a} dz \\
 &= \sum_{n=0}^\infty \frac{(a)_n(b)_n x^n}{(c)_n n!}.
 \end{aligned}$$

Here $(\cdot)_n$ denotes the rising Pochhammer symbol. This function will be used in the kernel function of fractional operator.

Theorem 2.1 *Let (y, u) be the solution of control problem (1.1)–(1.2). Then the following first order optimality system holds*

$$\begin{cases} -D(r_0D_x^{-(2-\alpha)} + (1-r)_xD_1^{-(2-\alpha)})Dy = f + u \text{ in } \Omega, \\ y = 0 \text{ on } \Gamma, \end{cases} \tag{2.1}$$

$$\begin{cases} -D(r_xD_1^{-(2-\alpha)} + (1-r)_0D_x^{-(2-\alpha)})Dz = y - y_d \text{ in } \Omega, \\ z = 0 \text{ on } \Gamma \end{cases} \tag{2.2}$$

and

$$\int_\Omega (\gamma u + z)(\chi - u) dx \geq 0, \forall \chi \in U_{ad}. \tag{2.3}$$

Proof To derive the first order optimality system, we introduce the following reduced optimization problem:

$$\min_{u \in U_{ad}} \hat{J}(u) := J(y(u), u),$$

where $y(u)$ is the solution of the state equation. Then the following optimality condition holds

$$\hat{J}'(u)(\chi - u) \geq 0, \forall \chi \in U_{ad}.$$

Note that the objective functional is strictly convex, so the above condition is sufficient and necessary.

By simple calculation, we have

$$\hat{J}'(u)(\chi - u) = \int_\Omega (y(u) - y_d)[y'(u)(\chi - u)] dx + \gamma \int_\Omega u(\chi - u) dx.$$

In order to simplify the above inequality, we need to calculate $y'(u)(\chi - u)$, According to the state equation, we have

$$\begin{cases} -D(r_0D_x^{-(2-\alpha)} + (1-r)_xD_1^{-(2-\alpha)})D[y'(u)(\chi - u)] = \chi - u & \text{in } \Omega, \\ y'(u)(\chi - u) = 0 & \text{on } \Gamma. \end{cases}$$

We introduce the adjoint state equation:

$$\begin{cases} -D(r_xD_1^{-(2-\alpha)} + (1-r)_0D_x^{-(2-\alpha)})Dz = y(u) - y_d & \text{in } \Omega, \\ z = 0 & \text{on } \Gamma. \end{cases}$$

Then by using integration by parts, we deduce

$$\begin{aligned} \int_{\Omega} (y(u) - y_d)[y'(u)(\chi - u)]dx &= \int_{\Omega} \left(-D(r_xD_1^{-(2-\alpha)} + (1-r)_0D_x^{-(2-\alpha)})Dz \right) \cdot [y'(u)(\chi - u)]dx \\ &= \int_{\Omega} (r_xD_1^{-(2-\alpha)} + (1-r)_0D_x^{-(2-\alpha)})Dz \cdot D[y'(u)(\chi - u)]dx \\ &= \int_{\Omega} Dz(x) \cdot \left((r_0D_x^{-(2-\alpha)} + (1-r)_xD_1^{-(2-\alpha)})D[y'(u)(\chi - u)] \right)dx \\ &= - \int_{\Omega} z(x) \cdot D(r_0D_x^{-(2-\alpha)} + (1-r)_xD_1^{-(2-\alpha)})D[y'(u)(\chi - u)]dx \\ &= \int_{\Omega} z(\chi - u)dx. \end{aligned}$$

Combing the above equations leads to

$$\hat{J}'(u)(\chi - u) = \int_{\Omega} (\gamma u + z(x))(\chi - u)dx \geq 0.$$

□

Let

$$P_{U_{ad}}(u) = \max\{u_a, \min\{u, u_b\}\}$$

denote the pointwise projection onto the admissible set U_{ad} . Then (2.3) is equivalent to

$$u = P_{U_{ad}}\left(-\frac{1}{\gamma}z\right).$$

In the following, we are going to show the regularity of the control problem.

Theorem 2.2 *Suppose that (y, z, u) is the solution of optimality system (2.1)–(2.3). Then we have the following regularity results*

$$y, z \in H^{v+\frac{3}{2}-\epsilon}(\Omega), u \in \begin{cases} H^{v+\frac{3}{2}-\epsilon}(\Omega), & v + \frac{3}{2} \leq 1, \\ H^1(\Omega), & v + \frac{3}{2} > 1. \end{cases}$$

Here $v = \min\{p, q\}$, $\forall \epsilon > 0$ and p, q are constants satifying $\alpha - 2 \leq p, q < 0$ and the following relation:

$$p + q = \alpha - 2, \quad r \sin(\pi(-q)) = (1 - r) \sin(\pi(-p)).$$

Proof According to [21], the kernel function of operator \mathcal{L}_r^α is given by $\ker(\mathcal{L}_r^\alpha) = \text{Span}\{1, K(x)\}$, where $K(x) = \frac{1}{p+1}x^{p+1}{}_2F_1(-q, p+1, p+2; x)$ with ${}_2F_1$ being the Gaussian three-parameter hypergeometric function and $\alpha - 2 \leq p, q < 0$ satisfying the following relation:

$$p + q = \alpha - 2, \quad r \sin(\pi(-q)) = (1 - r) \sin(\pi(-p)).$$

This implies that the state y belongs to the space $H^{v+\frac{3}{2}-\epsilon}(\Omega)$ with $v = \min\{p, q\}$ and any constant $\epsilon > 0$.

By similar argument, we can obtain that the adjoint state z has the same regularity as that of the state y . According to [22], the projection operator $P_{U_{ad}}$ satisfies the following property: if $v \in H^s(\Omega), \forall s \in [0, 1]$, then $P_{U_{ad}}(v) \in H^s(\Omega)$. Combining this property with the regularity of z implies that $u \in \begin{cases} H^{v+\frac{3}{2}-\epsilon}(\Omega), & v + \frac{3}{2} \leq 1, \\ H^1(\Omega), & v + \frac{3}{2} > 1. \end{cases}$ □

Remark 2.3 In the case $r = \frac{1}{2}$, we have $p = q = \frac{\alpha}{2} - 1$. This leads to $y, z \in H^{\frac{\alpha}{2}+\frac{1}{2}-\epsilon}(\Omega), u \in H^1(\Omega)$.

In the case $r \rightarrow 1$, we have $q \rightarrow 0, p \rightarrow \alpha - 2$. This gives $y, z \in H^{\alpha-\frac{1}{2}-\epsilon}(\Omega)$, which implies that $u \in H^1(\Omega)$, for $\alpha \in (\frac{3}{2}, 2)$ and $u \in H^{\alpha-\frac{1}{2}-\epsilon}(\Omega)$, for $\alpha \in (1, \frac{3}{2}]$.

3 Finite Element Approximation

In this section, we will investigate finite element approximation of the control problem. For simplicity, we set $\beta = 1 - \frac{\alpha}{2}$.

Following [21,23], let

$$A(y, v) := r({}_0D_x^{-\beta} Dy, {}_x D_1^{-\beta} Dv) + (1 - r)({}_x D_1^{-\beta} Dy, {}_0D_x^{-\beta} Dv)$$

denote the the bilinear form, which satisfies

$$A(y, y) \geq C_0 \|y\|_{\tilde{H}^{\frac{\alpha}{2}}(\Omega)}^2$$

and

$$A(y, v) \leq C_1 \|y\|_{\tilde{H}^{\frac{\alpha}{2}}(\Omega)} \|v\|_{\tilde{H}^{\frac{\alpha}{2}}(\Omega)},$$

where C_0 and C_1 are positive constants.

Then the weak formulation of the control problem reads as:

$$\min_{u \in U_{ad}} J(y, u) \tag{3.1}$$

subject to

$$A(y, v) = (f + u, v), \forall v \in \tilde{H}^{\frac{\alpha}{2}}(\Omega). \tag{3.2}$$

Define the Lagrange functional

$$\mathcal{L}(y, z, u) = J(y, u) + (f + u, z) - A(y, z).$$

Then we can derive the first order optimality system of the above variational problem:

$$A(y, v) = (f + u, v), \forall v \in \tilde{H}^{\frac{\alpha}{2}}(\Omega), \tag{3.3}$$

$$A(w, z) = (y - y_d, w), \forall w \in \tilde{H}^{\frac{\alpha}{2}}(\Omega), \tag{3.4}$$

and

$$\int_{\Omega} (\gamma u + z)(\chi - u) \geq 0, \forall \chi \in U_{ad}. \tag{3.5}$$

To define the finite element scheme, we introduce an uniform partition of the interval Ω with the mesh parameter $h = \frac{1}{N}$, where $N > 0$ is an integer. Let V_h denote the continuous finite element space consisting of piecewise linear polynomial on each interval $I_i = [x_i, x_{i+1}]$, where the nodes $x_i = ih, i = 0, 1, 2, \dots, N$. Since $\frac{\alpha}{2} \in (\frac{1}{2}, 1)$, we have $V_h \subset \tilde{H}^{\frac{\alpha}{2}}(\Omega)$. Following [24], the space V_h satisfies the approximation property: if the partition is quasi-uniform and $\frac{\alpha}{2} \leq \kappa \leq 2$, for $v \in H^{\kappa}(\Omega) \cap \tilde{H}^{\frac{\alpha}{2}}(\Omega)$, then

$$\inf_{v_h \in V_h} \|v - v_h\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq Ch^{\kappa - \frac{\alpha}{2}} \|v\|_{H^{\kappa}(\Omega)}. \tag{3.6}$$

For the discretization of control variable, we introduce a piecewise constant finite element space $U_h \subset L^2(\Omega)$. Let $U_{ad}^h = U_h \cap U_{ad}$.

Then the finite element approximation of control problem (3.1)–(3.2) is to find $(y_h, u_h) \in V_h \times U_{ad}^h$ such that

$$\min_{u_h \in U_{ad}^h} J(y_h, u_h) \tag{3.7}$$

subject to

$$A(y_h, v_h) = (f + u_h, v_h), \forall v_h \in V_h. \tag{3.8}$$

Analogous to the continuous case, we can derive the discrete first order optimality system

$$A(y_h, v_h) = (f + u_h, v_h), \forall v_h \in V_h, \tag{3.9}$$

$$A(w_h, z_h) = (y_h - y_d, w_h), \forall w_h \in V_h, \tag{3.10}$$

and

$$\int_{\Omega} (\gamma u_h + z_h)(\chi_h - u_h) dx \geq 0, \forall \chi_h \in U_{ad}^h. \tag{3.11}$$

In the following analysis, we are going to derive a priori error estimate for the control problem. For this purpose, we introduce some auxiliary problems:

$$A(y_h(u), v_h) = (f + u, w_h), \quad \forall w_h \in V_h, \tag{3.12}$$

$$A(w_h, z_h(u)) = (y_h(u) - y_d, w_h), \quad \forall w_h \in V_h, \tag{3.13}$$

$$A(w_h, z_h(y)) = (y - y_d, w_h), \quad \forall w_h \in V_h. \tag{3.14}$$

It is easy to see that $y_h(u)$ and $z_h(y)$ are the finite element approximations of the state y and the adjoint state z , respectively. Therefore, following [21] and the regularity of the state and the adjoint state variables, we have

$$\|y - y_h(u)\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq Ch^{v + \frac{3}{2} - \frac{\alpha}{2} - \epsilon} \|y\|_{H^{v + \frac{3}{2} - \epsilon}(\Omega)}, \tag{3.15}$$

$$\|y - y_h(u)\|_{L^2(\Omega)} \leq Ch^{2(v + \frac{3}{2} - \frac{\alpha}{2} - \epsilon)} \|y\|_{H^{v + \frac{3}{2} - \epsilon}(\Omega)} \tag{3.16}$$

and

$$\|z - z_h(y)\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq Ch^{v + \frac{3}{2} - \frac{\alpha}{2} - \epsilon} \|z\|_{H^{v + \frac{3}{2} - \epsilon}(\Omega)}, \tag{3.17}$$

$$\|z - z_h(y)\|_{L^2(\Omega)} \leq Ch^{2(v+\frac{3}{2}-\frac{\alpha}{2}-\epsilon)} \|z\|_{H^{v+\frac{3}{2}-\epsilon}(\Omega)}. \tag{3.18}$$

By the coercivity of the bilinear form $A(\cdot, \cdot)$, we can derive

$$\|z_h(y) - z_h(u)\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq C \|y_h(u) - y\|_{L^2(\Omega)}, \tag{3.19}$$

$$\|z_h(y) - z_h\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq C \|y_h - y\|_{L^2(\Omega)} \tag{3.20}$$

and

$$\|y_h - y_h(u)\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq C \|u - u_h\|_{L^2(\Omega)}, \tag{3.21}$$

$$\|z_h - z_h(u)\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq C \|u - u_h\|_{L^2(\Omega)}. \tag{3.22}$$

Note that the estimate of the state and the adjoint state depends on the estimate of the control. Therefore, we show the error estimate of the control firstly.

Lemma 3.1 *Let (y, z, u) and (y_h, z_h, u_h) be the solutions of (2.1)–(2.3) and (3.9)–(3.11), respectively. Then the following estimate holds*

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^{2(v+\frac{3}{2}-\frac{\alpha}{2}-\epsilon)}.$$

Proof Let

$$(\widehat{J}'_h(u), \chi) = (\gamma u + z_h(u), \chi).$$

Then we can derive

$$(\widehat{J}'_h(u), u - u_h) - (\widehat{J}'_h(u_h), u - u_h) \geq \gamma \|u - u_h\|_{L^2(\Omega)}^2 + (z_h(u) - z_h, u - u_h).$$

By (3.8) and (3.12), we obtain

$$\begin{aligned} (z_h(u) - z_h, u - u_h) &= A(y_h(u) - y_h, z_h(u) - z_h) \\ &= (y_h(u) - y_h, y_h(u) - y_h) \geq 0. \end{aligned}$$

This implies

$$(\widehat{J}'_h(u), u - u_h) - (\widehat{J}'_h(u_h), u - u_h) \geq \gamma \|u - u_h\|_{L^2(\Omega)}^2.$$

Then we have

$$\begin{aligned} &\gamma \|u - u_h\|_{L^2(\Omega)}^2 \\ &\leq (\widehat{J}'_h(u), u - u_h) - (\widehat{J}'_h(u_h), u - u_h) \\ &= \int_{\Omega} (\gamma u + z_h(u))(u - u_h) dx - \int_{\Omega} (\gamma u_h + z_h)(u - u_h) dx \\ &= \int_{\Omega} (\gamma u + z)(u - u_h) dx + \int_{\Omega} (z_h(u) - z)(u - u_h) dx - \int_{\Omega} (\gamma u_h + z_h)(u - u_h) dx. \end{aligned}$$

Let $\Pi_h u \in U_{ad}^h$ be the L^2 projection of u defined by

$$\Pi_h u|_{I_i} = \frac{\int_{I_i} u}{\int_{I_i} 1}, \quad \forall I_i.$$

It is easy to see that

$$\|u - \Pi_h u\|_{L^2(\Omega)} \leq Ch^s \|u\|_{H^s(\Omega)}, \quad 0 < s \leq 1. \tag{3.23}$$

Therefore, it follows from (2.3), (3.11), (3.22) and the Young’s inequality that

$$\begin{aligned}
 & \gamma \|u - u_h\|_{L^2(\Omega)}^2 \\
 &= \int_{\Omega} (\gamma u + z)(u - u_h) dx + \int_{\Omega} (z_h(u) - z)(u - u_h) dx - \int_{\Omega} (\gamma u_h + z_h)(u - \Pi_h u) dx \\
 &\quad - \int_{\Omega} (\gamma u_h + z_h)(\Pi_h u - u_h) dx \\
 &\leq 0 + \int_{\Omega} (z_h(u) - z)(u - u_h) dx - \int_{\Omega} (\gamma u_h + z_h)(u - \Pi_h u) dx + 0 \\
 &= \int_{\Omega} (z_h(u) - z)(u - u_h) dx - \int_{\Omega} (\gamma u \\
 &\quad + z - \gamma u_h - z_h)(u - \Pi_h u) dx + \int_{\Omega} (\gamma u + z)(u - \Pi_h u) dx \\
 &= \int_{\Omega} (z_h(u) - z)(u - u_h) dx + \gamma \int_{\Omega} (u_h - u)(u - \Pi_h u) dx \\
 &\quad + \int_{\Omega} (\gamma u + z)(u - \Pi_h u) dx \\
 &\quad + \int_{\Omega} (z_h(u) - z)(u - \Pi_h u) dx + \int_{\Omega} (z_h(u) - z_h)(\Pi_h u - u) dx \\
 &\leq \delta \|u - u_h\|_{L^2(\Omega)}^2 + C(\delta) \|u - \Pi_h u\|_{L^2(\Omega)}^2 \\
 &\quad + C(\delta) \|z_h(u) - z\|_{L^2(\Omega)}^2 + \int_{\Omega} (\gamma u + z)(u - \Pi_h u) dx.
 \end{aligned}$$

Here δ is an arbitrary positive constant. Furthermore, by the definition of $\Pi_h u$ and the regularity of u and z , we have

$$\begin{aligned}
 \int_{\Omega} (\gamma u + z)(u - \Pi_h u) &= \int_{\Omega} (\gamma u + z - \Pi_h(\gamma u + z))(u - \Pi_h u) dx \\
 &\leq \|\gamma u + z - \Pi_h(\gamma u + z)\|_{L^2(\Omega)} \|u - \Pi_h u\|_{L^2(\Omega)} \\
 &\leq Ch^{2\min\{1, v+\frac{3}{2}-\epsilon\}}.
 \end{aligned}$$

Combining the above estimates and setting δ small enough lead to

$$\begin{aligned}
 \|u - u_h\|_{L^2(\Omega)}^2 &\leq Ch^{2\min\{1, v+\frac{3}{2}-\epsilon\}} + Ch^{4(v+\frac{3}{2}-\frac{\alpha}{2}-\epsilon)} \\
 &\leq Ch^{4(v+\frac{3}{2}-\frac{\alpha}{2}-\epsilon)}.
 \end{aligned}$$

where the following fact is used

$$2\left(v + \frac{3}{2} - \frac{\alpha}{2} - \epsilon\right) \leq 2\left(\frac{\alpha}{2} - 1 + \frac{3}{2} - \frac{\alpha}{2} - \epsilon\right) = 2\left(\frac{1}{2} - \epsilon\right) < 1.$$

□

Finally, by using the above estimate for the control variable, we can derive a priori error estimates for the state and the adjoint state variables.

Theorem 3.2 Assume that (y, z, u) and (y_h, z_h, u_h) are the solutions of the optimality system (2.1)–(2.3) and the discrete counterpart. Then the following error estimates hold

$$\|y - y_h\|_{H^{\frac{\alpha}{2}}(\Omega)} + \|z - z_h\|_{H^{\frac{\alpha}{2}}(\Omega)} \leq Ch^{v+\frac{3}{2}-\frac{\alpha}{2}-\epsilon}$$

and

$$\|y - y_h\|_{L^2(\Omega)} + \|z - z_h\|_{L^2(\Omega)} \leq Ch^{2(v+\frac{3}{2}-\frac{\alpha}{2}-\epsilon)}.$$

Proof Note that

$$\begin{aligned} \|y - y_h\|_{H^{\frac{\alpha}{2}}(\Omega)} &\leq \|y - y_h(u)\|_{H^{\frac{\alpha}{2}}(\Omega)} + \|y_h(u) - y_h\|_{H^{\frac{\alpha}{2}}(\Omega)} \\ &\leq \|y - y_h(u)\|_{H^{\frac{\alpha}{2}}(\Omega)} + \|u - u_h\|_{L^2(\Omega)} \end{aligned}$$

and

$$\begin{aligned} \|y - y_h\|_{L^2(\Omega)} &\leq \|y - y_h(u)\|_{L^2(\Omega)} + \|y_h(u) - y_h\|_{L^2(\Omega)} \\ &\leq \|y - y_h(u)\|_{L^2(\Omega)} + \|u - u_h\|_{L^2(\Omega)}. \end{aligned}$$

Then the estimate of the state follows from (3.15), (3.16) and Lemma 3.1. In an analogue way,

$$\begin{aligned} \|z - z_h\|_{H^{\frac{\alpha}{2}}(\Omega)} &\leq \|z - z_h(y)\|_{H^{\frac{\alpha}{2}}(\Omega)} + \|z_h(y) - z_h\|_{H^{\frac{\alpha}{2}}(\Omega)} \\ &\leq \|z - z_h(y)\|_{H^{\frac{\alpha}{2}}(\Omega)} + \|y - y_h\|_{L^2(\Omega)} \end{aligned}$$

and

$$\begin{aligned} \|z - z_h\|_{L^2(\Omega)} &\leq \|z - z_h(y)\|_{L^2(\Omega)} + \|z_h(y) - z_h\|_{L^2(\Omega)} \\ &\leq \|z - z_h(y)\|_{L^2(\Omega)} + \|y - y_h\|_{L^2(\Omega)}. \end{aligned}$$

Then the estimate of the adjoint state follows from (3.17), (3.18) and the estimate of the state. □

4 Numerical Algorithm

In this section, we will present a numerical algorithm for the above control problem based on the primal dual active set strategy. The primal dual active set strategy has been developed in [25] for the elliptic optimal control problem with control constraints.

4.1 Primal–Dual Active Set Algorithm

In order to present the primal–dual active set algorithm, we reformulate the optimality system (2.1)–(2.3) into the following equivalent formulation

$$\begin{cases} -D(r_0D_x^{-(2-\alpha)} + (1-r)_xD_1^{-(2-\alpha)})Dy = f + u & \text{in } \Omega, \\ -D(r_xD_1^{-(2-\alpha)} + (1-r)_0D_x^{-(2-\alpha)})Dz = y - y_d & \text{in } \Omega, \\ \gamma u + z = -\lambda & \text{in } \Omega. \end{cases} \tag{4.1}$$

Here λ is defined as follows, for $\mu > 0$

$$\lambda = \mu \max \left\{ 0, u + \frac{\lambda}{\mu} - u_b \right\} - \mu \min \left\{ 0, u + \frac{\lambda}{\mu} - u_b \right\} \tag{4.2}$$

or

$$\lambda = \max\{0, \lambda + \mu(u - u_b)\} - \min\{0, \lambda + \mu(u - u_a)\}. \tag{4.3}$$

The primal dual active set algorithm is based on utilizing (4.2) or (4.3) as a prediction strategy. For given (u^{n-1}, λ^{n-1}) , the active set and the inactive set for the next iteration are chosen as

$$\begin{aligned} \mathcal{A}_a^n &= \{x \in \Omega \mid \lambda^{n-1}(x) + \mu(u^{n-1}(x) - u_a) < 0\} \\ \mathcal{A}_b^n &= \{x \in \Omega \mid \lambda^{n-1}(x) + \mu(u^{n-1}(x) - u_b) > 0\} \\ \mathcal{I}^n &= \Omega \setminus (\mathcal{A}_a^n \cup \mathcal{A}_b^n). \end{aligned}$$

Then the primal–dual active set algorithm for the continuous problem is specified as follows

Algorithm 1: Primal-dual active set algorithm for continuous problem

1. Choose the initial value u^0, y^0, λ^0 and $\mu > 0$. Set $n = 1$.
2. Compute the following subsets $\mathcal{A}_a^n, \mathcal{A}_b^n$ and \mathcal{I}^n .
3. Solving the following equations for (y^n, z^n)

$$\begin{cases} -D(r_0 D_x^{-(2-\alpha)} + (1-r)_x D_1^{-(2-\alpha)}) D y^n = f + u^n & \text{in } \Omega, \\ -D(r_x D_1^{-(2-\alpha)} + (1-r)_0 D_x^{-(2-\alpha)}) D z^n = y^n - y_d & \text{in } \Omega, \end{cases} \tag{4.4}$$

where

$$u^n = \begin{cases} u_a, & \text{on } \mathcal{A}_a^n \\ u_b, & \text{on } \mathcal{A}_b^n \end{cases}$$

and

$$\lambda^n = 0, \text{ on } \mathcal{I}^n.$$

4. End loop, if $n \geq 2$, $\mathcal{A}_a^n = \mathcal{A}_a^{n-1}$ and $\mathcal{A}_b^n = \mathcal{A}_b^{n-1}$. Otherwise, update $n = n + 1$ and goto step 2.
-

In the following, we are going to discuss the algorithm in the discrete level. Let

$$\begin{aligned} V_h &= \text{Span}\{\phi_1, \dots, \phi_{N-1}\}, U_h = \text{Span}\{\psi_1, \dots, \psi_N\}, \mathbf{y} = (y_1, \dots, y_{N-1})^T, \\ \mathbf{z} &= (z_1, \dots, z_{N-1})^T. \end{aligned}$$

Then the finite element scheme for the state equation with right hand term g can be expressed as the following matrix form

$$\mathcal{A}\mathbf{y} = \mathbf{g}, \tag{4.5}$$

where $\mathbf{g} = (g_1, \dots, g_{N-1})^T$ with $g_k = (g, \phi_k(x))$, $k = 1, \dots, N - 1$. The entries of matrix \mathcal{A} are calculated by

$$\begin{aligned} a_{ij} &= A(\phi_j, \phi_i) \\ &= r({}_0D_x^{-\beta} D\phi_j(x), {}_x D_1^{-\beta} D\phi_i(x)) + (1-r)({}_x D_1^{-\beta} D\phi_j(x), {}_0D_x^{-\beta} D\phi_i(x)). \end{aligned}$$

Note that the finite element space V_h consists of piecewise linear polynomial on each interval $[x_i, x_{i+1}]$. Then by simple calculation, matrix \mathcal{A} takes the following form ([26,27])

$$\mathcal{A} = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \ddots & a_{3-N} & a_{2-N} \\ a_1 & a_0 & a_{-1} & a_{-2} & \ddots & a_{3-N} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & a_1 & a_0 & a_{-1} & a_{-2} \\ a_{N-3} & \ddots & \ddots & \ddots & a_0 & a_{-1} \\ a_{N-2} & a_{N-3} & \ddots & \ddots & a_1 & a_0 \end{pmatrix} \tag{4.6}$$

where

$$\begin{cases} a_0 = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} \left(r(-4 + 2^{3-\alpha}) + (1-r)(-4 + 2^{3-\alpha}) \right), \\ a_1 = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} \left(r(3^{3-\alpha} - 4 \cdot 2^{3-\alpha} + 6) + (1-r) \right), \\ a_i = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} r \left((i+2)^{3-\alpha} - 4 \cdot (i+1)^{3-\alpha} + 6 \cdot i^{3-\alpha} - 4 \cdot (i-1)^{3-\alpha} + (i-2)^{3-\alpha} \right), \\ \quad i = 2, 3, 4, \dots, N-2, \\ a_{-1} = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} \left(r + (1-r)(3^{3-\alpha} - 4 \cdot 2^{3-\alpha} + 6) \right), \\ a_{-i} = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} (1-r) \left((i+2)^{3-\alpha} - 4 \cdot (i+1)^{3-\alpha} + 6 \cdot i^{3-\alpha} - 4 \cdot (i-1)^{3-\alpha} + (i-2)^{3-\alpha} \right), \\ \quad i = 2, 3, 4, \dots, N-2. \end{cases}$$

In an analogous way, the coefficient matrix for the adjoint state equation can be expressed as follows

$$\mathcal{B} = \begin{pmatrix} \tilde{a}_0 & \tilde{a}_{-1} & \tilde{a}_{-2} & \ddots & \tilde{a}_{3-N} & \tilde{a}_{2-N} \\ \tilde{a}_1 & \tilde{a}_0 & \tilde{a}_{-1} & \tilde{a}_{-2} & \ddots & \tilde{a}_{3-N} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \tilde{a}_1 & \tilde{a}_0 & \tilde{a}_{-1} & \tilde{a}_{-2} \\ \tilde{a}_{N-3} & \ddots & \ddots & \ddots & \tilde{a}_0 & \tilde{a}_{-1} \\ \tilde{a}_{N-2} & \tilde{a}_{N-3} & \ddots & \ddots & \tilde{a}_1 & \tilde{a}_0 \end{pmatrix} \tag{4.7}$$

where

$$\begin{cases} \tilde{a}_0 = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} \left((1-r)(-4 + 2^{3-\alpha}) + r(-4 + 2^{3-\alpha}) \right), \\ \tilde{a}_1 = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} \left((1-r)(3^{3-\alpha} - 4 \cdot 2^{3-\alpha} + 6) + r \right), \\ \tilde{a}_i = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} (1-r) \left((i+2)^{3-\alpha} - 4 \cdot (i+1)^{3-\alpha} + 6 \cdot i^{3-\alpha} - 4 \cdot (i-1)^{3-\alpha} + (i-2)^{3-\alpha} \right), \\ \quad i = 2, 3, 4, \dots, N-2, \\ \tilde{a}_{-1} = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} \left((1-r) + r(3^{3-\alpha} - 4 \cdot 2^{3-\alpha} + 6) \right), \\ \tilde{a}_{-i} = -\frac{h^{1-\alpha}}{\Gamma(4-\alpha)} r \left((i+2)^{3-\alpha} - 4 \cdot (i+1)^{3-\alpha} + 6 \cdot i^{3-\alpha} - 4 \cdot (i-1)^{3-\alpha} + (i-2)^{3-\alpha} \right), \\ \quad i = 2, 3, 4, \dots, N-2. \end{cases}$$

Let $N_h = \{1, 2, \dots, N\}$. $\mathbf{F}, Y_d \in \mathbb{R}^{N-1}$, where $\mathbf{F} = ((f, \phi_i))_{(N-1) \times 1}$ and $Y_d = ((y_d, \phi_i))_{(N-1) \times 1}$. $\mathbf{A}, \mathbf{u}, \mathbf{u}_d, \mathbf{u}_b \in \mathbb{R}^N$ are vectors corresponding to the discrete counterpart. $M \in \mathbb{R}^{(N-1) \times (N-1)}$ and $M_0 \in \mathbb{R}^{N \times N}$ are the mass matrix for space V_h and U_h , respectively.

$M_1 \in \mathbb{R}^{(N-1) \times N}$ denotes the matrix whose entry is calculated by $m_{ij} = (\psi_j(x), \phi_i(x))$, $i = 1, \dots, N-1, j = 1, 2, \dots, N$. We should point out that M_0 is a diagonal matrix in this case.

Similar to the continuous case, we can define the active set and the inactive set of the n -th step as follows

$$\begin{aligned} \tilde{\mathcal{A}}_a^n &= \{i \in N_h \mid (\Lambda^{n-1} + \mu(\mathbf{u}^{n-1} - \mathbf{u}_a))_i < 0\}, \\ \tilde{\mathcal{A}}_b^n &= \{i \in N_h \mid (\Lambda^{n-1} + \mu(\mathbf{u}^{n-1} - \mathbf{u}_b))_i > 0\}, \\ \tilde{\mathcal{I}}^n &= N_h \setminus (\tilde{\mathcal{A}}_a^n \cup \tilde{\mathcal{A}}_b^n), \end{aligned}$$

where $\Lambda^n = \gamma M_0 \mathbf{u}^n + M_1^T \mathbf{z}^n$.

Algorithm 2: Primal-dual active set algorithm for discrete problem

1. Choose the initial value $\mathbf{u}^0, \mathbf{y}^0, \Lambda^0$ and $\mu > 0$. Set $n = 1$.
2. Compute the following subsets $\tilde{\mathcal{A}}_a^n, \tilde{\mathcal{A}}_b^n$ and $\tilde{\mathcal{I}}^n$.
3. Solving the following system

$$\begin{cases} \mathcal{A}\mathbf{y}^n - M_1\mathbf{u}^n = \mathbf{F}, \\ \mathcal{B}\mathbf{z}^n - M\mathbf{y}^n = -Y_d. \end{cases} \tag{4.8}$$

Here

$$\mathbf{u}^n = \begin{cases} \mathbf{u}_a, & \text{on } \tilde{\mathcal{A}}_a^n \\ \mathbf{u}_b, & \text{on } \tilde{\mathcal{A}}_b^n \end{cases}$$

and

$$\Lambda^n = 0, \text{ on } \tilde{\mathcal{I}}^n.$$

4. End loop, if $n \geq 2, \tilde{\mathcal{A}}_a^n = \tilde{\mathcal{A}}_a^{n-1}$ and $\tilde{\mathcal{A}}_b^n = \tilde{\mathcal{A}}_b^{n-1}$. Otherwise, update $n = n + 1$ and goto step 2.
-

It is easy to see that in each iteration of the above algorithm, we need to solve the following equation

$$\begin{pmatrix} \mathcal{A} & \frac{1}{\gamma} \tilde{M}_1 M_0^{-1} \tilde{M}_1^T \\ -M & \mathcal{B} \end{pmatrix} \begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} \mathbf{F} + M_1(\tilde{\mathbf{u}}_a + \tilde{\mathbf{u}}_b) \\ -Y_d \end{pmatrix} \tag{4.9}$$

where \tilde{M}_1 consists of the $\tilde{\mathcal{I}}^n$ columns of the matrix M_1 , the $\tilde{\mathcal{A}}_a^n$ entries of $\tilde{\mathbf{u}}_a$ is the same as that of \mathbf{u}_a and the other entries are zero, the $\tilde{\mathcal{A}}_b^n$ entries of $\tilde{\mathbf{u}}_b$ is the same as that of \mathbf{u}_b and the other entries are zero.

4.2 Fast Algorithm

Note that the fractional differential operator is nonlocal, which leads to a full coefficient matrix. In the following part, we will develop a fast algorithm based on the Toeplitz structure of the coefficient matrix.

The Eq. (4.9) can be solved by some Krylov subspace methods, for example, BiCGStab method. In the above algorithm, the main computational cost comes from the matrix-vector multiplication. Note that M is tri-diagonal sparse matrix, which needs $O(N)$ operations and $O(N)$ storage. However, \mathcal{A} and \mathcal{B} are full, which usually cost $O(N^2)$ operations and $O(N^2)$ storage.

It is easy to observe that the stiff matrices \mathcal{A} and \mathcal{B} are two $(N - 1)$ -by- $(N - 1)$ Toeplitz matrices. In the computation, we only need to store the first row vector and the first column vector which requires $O(N)$ storage. Note that a Toeplitz matrix T_N can be embedded into a circulant matrix C_{2N} . According to [28,29], a circulant matrix C_{2N} can be decomposed as follows

$$C_{2N} = F_{2N}^{-1} \text{diag}(F_{2N} \mathbf{d}) F_{2N}, \tag{4.10}$$

where \mathbf{d} is the first column vector of C_{2N} and F_{2N} is the $2N \times 2N$ discrete Fourier transform matrix in which the (j, l) -entry $F_{2N}(j, l)$ of the matrix F_{2N} is defined by

$$F_{2N}(j, l) = \frac{1}{\sqrt{2N}} \exp\left(-\frac{2\pi i j l}{2N}\right), \quad 0 \leq j, l \leq 2N - 1$$

with $i = \sqrt{-1}$. It is well known that the matrix-vector multiplication $F_{2N} W_{2N}$ for $W_{2N} \in R^{2N}$ can be carried out by fast Fourier transform (FFT). Thus the matrix-vector multiplication $C_{2N} W_{2N}$ can be calculated by $O(N \log N)$ operations.

Indeed, the $(N - 1) \times (N - 1)$ Toeplitz matrix \mathcal{A} can be embedded into a $(2N - 2) \times (2N - 2)$ circulant matrix C_{2N-2} as follows

$$C_{2N-2} = \begin{pmatrix} \mathcal{A} & \widehat{\mathcal{A}} \\ \widehat{\mathcal{A}} & \mathcal{A} \end{pmatrix},$$

where

$$\widehat{\mathcal{A}} = \begin{pmatrix} 0 & a_{N-2} & a_{N-3} & \cdots & a_2 & a_1 \\ a_{2-N} & 0 & a_{N-2} & a_{N-3} & \cdots & a_2 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & a_{2-N} & 0 & a_{N-2} & a_{N-3} \\ a_{-2} & \ddots & \ddots & \ddots & 0 & a_{N-2} \\ a_{-1} & a_{-2} & \ddots & \ddots & a_{2-N} & 0 \end{pmatrix}.$$

Define a $2N - 2$ vector $\mathbf{v}_{2N-2} = (\mathbf{v}_{N-1}, \mathbf{0})^T$, then we have

$$C_{2N-2} \mathbf{v}_{2N-2} = \begin{pmatrix} \mathcal{A} \mathbf{v}_{N-1} \\ \widehat{\mathcal{A}} \mathbf{v}_{N-1} \end{pmatrix}.$$

Thus, the first part of matrix-vector products $C_{2N-2} \mathbf{v}_{2N-2}$ leads to the matrix-vector products $\mathcal{A} \mathbf{v}_{N-1}$. Therefore, the computational cost of $\mathcal{A} \mathbf{v}_{N-1}$ is $O(N \log N)$ operations. For $\mathcal{B} \mathbf{v}_{N-1}$ we can treat in the same way. Applying above skill to the BiCGStab method leads to a fast algorithm for the optimal control problem.

5 Numerical Example

In this section, some numerical experiments are carried out to show the performance of our finite element scheme and the iterative algorithm.

Table 1 L^2 error and $H^{\frac{\alpha}{2}}$ error versus mesh size h and orders of convergence for Example 1

h	$e_{0,h}$	Order	$e_{\frac{\alpha}{2},h}$	Order
1/64	3.6093E-2	–	4.0813E-1	–
1/128	1.7990E-2	1.0045	2.8176E-1	0.5346
1/256	8.9215E-3	1.0119	1.9682E-1	0.5176
1/512	4.4369E-3	1.0077	1.3832E-1	0.5089
1/1024	2.2092E-3	1.0061	9.7508E-2	0.5044
1/2048	1.0981E-3	1.0085	6.8843E-2	0.5022

Table 2 L^2 error of u , y and z versus mesh size h and orders of convergence for Example 1

h	$\ u - u_h\ _0$	Order	$\ y - y_h\ _0$	Order	$\ z - z_h\ _0$	Order
1/64	2.9094E-2	–	4.2734E-4	–	6.5717E-3	–
1/128	1.4700E-2	0.9849	1.9709E-4	1.1166	3.0932E-3	1.0872
1/256	7.3477E-3	1.0004	8.8477E-5	1.1555	1.4853E-4	1.0584
1/512	3.6713E-3	1.0010	4.1732E-5	1.0841	7.2392E-4	1.0368
1/1024	1.8339E-3	1.0014	1.9812E-5	1.0748	3.5547E-4	1.0261
1/2048	9.1714E-4	0.9997	9.4673E-6	1.0653	1.7147E-4	1.0518

Table 3 $H^{\frac{\alpha}{2}}$ error of y and z versus mesh size h and orders of convergence for Example 1

h	$\ y - y_h\ _{\frac{\alpha}{2}}$	Order	$\ z - z_h\ _{\frac{\alpha}{2}}$	Order
1/64	3.7230E-2	–	3.7090E-1	–
1/128	2.5647E-2	0.5377	2.5611E-1	0.5342
1/256	1.7905E-2	0.5184	1.7892E-1	0.5175
1/512	1.2579E-2	0.5093	1.2574E-1	0.5088
1/1024	8.8659E-3	0.5047	8.8642E-1	0.5044
1/2048	6.2590E-3	0.5023	6.2584E-1	0.5022

In the following part, we denote $e_{0,h} = \gamma \|u - u_h\|_0 + \|y - y_h\|_0 + \|z - z_h\|_0$, $e_{\frac{\alpha}{2},h} = \|y - y_h\|_{\frac{\alpha}{2}} + \|z - z_h\|_{\frac{\alpha}{2}}$. The $\|w\|_{\frac{\alpha}{2}}$ data presented in the tables denotes the norm

$$\|w\|_{\frac{\alpha}{2}} = \left(\|w\|_0^2 + \int_0^1 \int_0^1 \frac{|w(x) - w(y)|^2}{|x - y|^{1+\alpha}} dx dy \right)^{\frac{1}{2}}.$$

Example 1 In this example, we consider problem (1.1)–(1.2) with the exact solutions defined by

$$\begin{aligned} y &= x^{\frac{\alpha}{2}}(1 - x)^{\frac{\alpha}{2}}, \\ z &= sx^{\frac{\alpha}{2}}(1 - x)^{\frac{\alpha}{2}}, \\ u &= \max \left\{ u_a, \min \left\{ u_b, -\frac{1}{\gamma}z \right\} \right\}. \end{aligned}$$

where $s \in \mathbb{R}$ is a constant and if we take $\alpha = 1.6$, $r = 0.5$, the corresponding f and y_d are

$$\begin{aligned} f &= -\Gamma(1 + \alpha) \cos(\pi\alpha/2) - u, \\ y_d &= y + s\Gamma(1 + \alpha) \cos(\pi\alpha/2). \end{aligned}$$

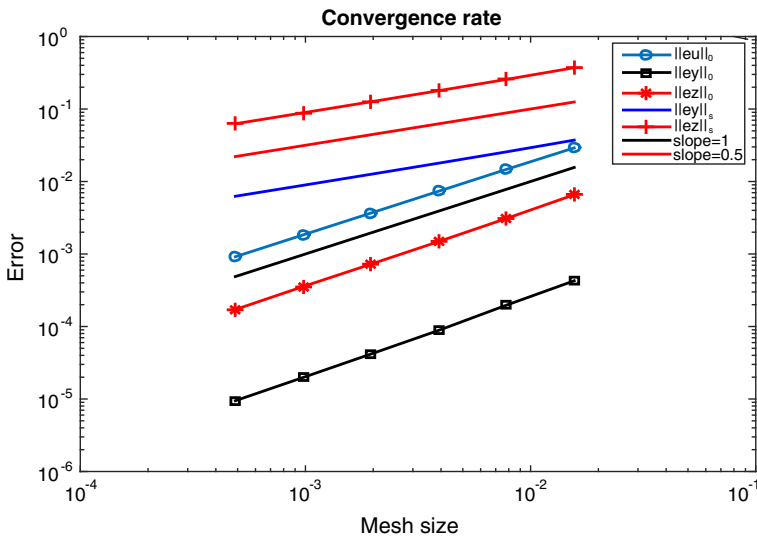


Fig. 1 The convergence rate of Example 1

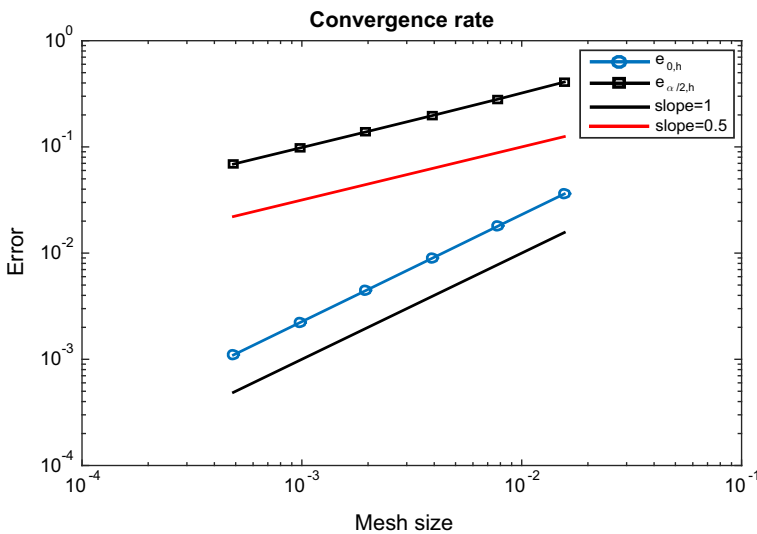


Fig. 2 The convergence rate of Example 1

In the first numerical test, we take $\gamma = 1$, $s = 10$, $u_a = -2.5$ and $u_b = -0.5$. We refine the mesh uniformly in the test. The L^2 error, the $H^{\frac{\alpha}{2}}$ error and the orders of convergence with respect to the mesh size are listed in Tables 1, 2 and 3, respectively. The Figs. 1 and 2 show the convergence rate. The Fig. 3 shows the numerical solutions and the exact solutions of the optimal control problem with the mesh size $h = 1/64$. According to these results, we know that the orders of convergence of the L^2 error and the $H^{\frac{\alpha}{2}}$ error are 1 and 0.5, respectively, which are just the same as the theoretical analysis.

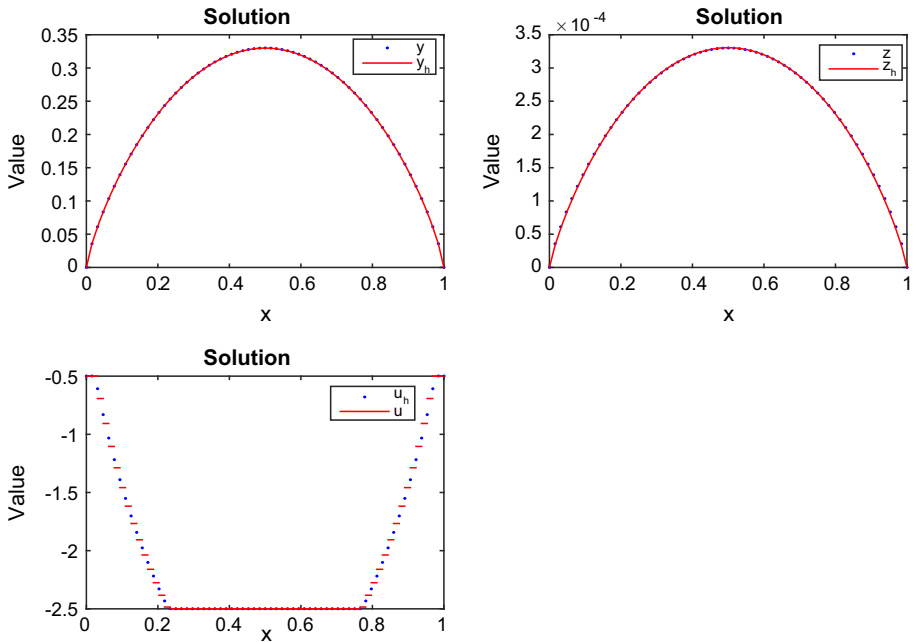


Fig. 3 The numerical solutions and the exact solutions of Example 1

Table 4 L^2 error and $H^{\frac{\alpha}{2}}$ error versus mesh size h and orders of convergence for Example 1

h	$e_{0,h}$	Order	$e_{\frac{\alpha}{2},h}$	Order
1/64	3.0865E-2	–	3.7388E-2	–
1/128	1.5621E-2	0.9824	2.5723E-2	0.5395
1/256	7.7955E-3	1.0028	1.7940E-2	0.5199
1/512	3.8906E-3	1.0027	1.2598E-2	0.5100
1/1024	1.9412E-3	1.0030	8.8769E-3	0.5051
1/2048	9.7103E-4	0.9994	6.2660E-3	0.5025

In the second numerical test, we take $\gamma = 10^{-4}$, $s = 0.001$, $u_a = -2.5$ and $u_b = -0.5$. We do the same test and list the results in Tables 4, 5 and 6. The Figs. 4 and 5 show the convergence rate. The Fig. 6 shows the numerical solutions and the exact solutions of the optimal control problem with the mesh size $h = 1/64$. These results are in accordance with the theoretical analysis.

Meanwhile, we also test the performance of our fast algorithm combing with the BiCGStab method for solving the linear system. We denote the original one by “PDAS” and the fast algorithm by “FFT PDAS”. We compare the time consuming between these two algorithms. We list the test results in Table 7, where N is the number of the element, “T” is the total time consuming of the algorithm. We list the time consuming and the iteration number for solving the corresponding linear system, in each iteration of the primal dual active set method, by the BiCGStab method. For example, “1.054E-1(115.5)” means that the time consuming of solving the linear system is “1.054E-1s” and the iteration number of BiCGStab method is

Table 5 L^2 error of u , y and z versus mesh size h and orders of convergence for Example 1

h	$\ u - u_h\ _0$	Order	$\ y - y_h\ _0$	Order	$\ z - z_h\ _0$	Order
1/64	3.0464E-2	–	3.9954E-4	–	1.3186E-6	–
1/128	1.5450E-2	0.9795	1.7080E-4	1.2260	6.1349E-7	1.1039
1/256	7.7205E-3	1.0008	7.4648E-5	1.1942	3.0067E-7	1.0289
1/512	3.8569E-3	1.0012	3.3485E-5	1.1566	1.4760E-7	1.0265
1/1024	1.9257E-3	1.0021	1.5449E-5	1.1160	7.4906E-8	0.9785
1/2048	9.6372E-4	0.9987	7.2695E-6	1.0876	4.1548E-8	0.8503

Table 6 $H^{\frac{\alpha}{2}}$ error of y and z versus mesh size h and orders of convergence for Example 1

h	$\ y - y_h\ _{\frac{\alpha}{2}}$	Order	$\ z - z_h\ _{\frac{\alpha}{2}}$	Order
1/64	3.7351E-2	–	3.7561E-5	–
1/128	2.5698E-2	0.5195	2.5305E-5	0.5698
1/256	1.7923E-2	0.5199	1.7693E-5	0.5162
1/512	1.2585E-2	0.5100	1.2483E-5	0.5033
1/1024	8.8681E-3	0.5051	8.8260E-6	0.5001
1/2048	6.2597E-3	0.5025	6.2432E-6	0.4995

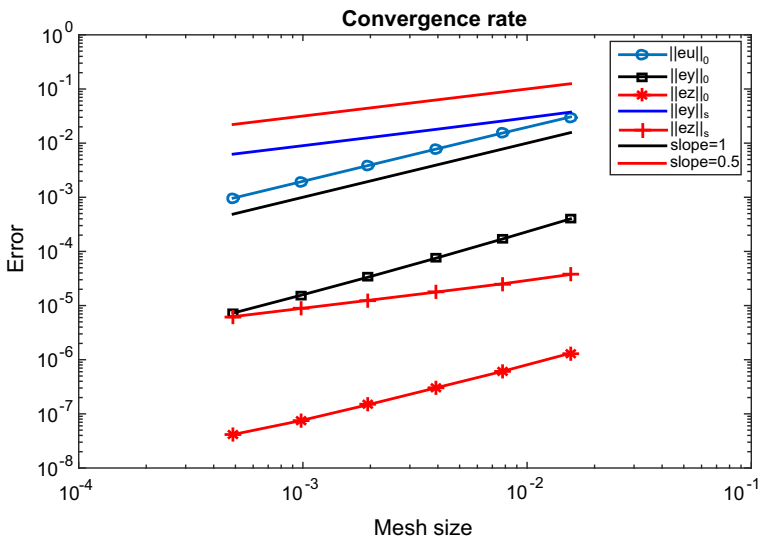


Fig. 4 The convergence rate of Example 1

“115.5”. According to these results, we can see that the fast algorithm can speed up the primal dual active method for solving the constraint optimal control problem combining with the BiCGStab method.

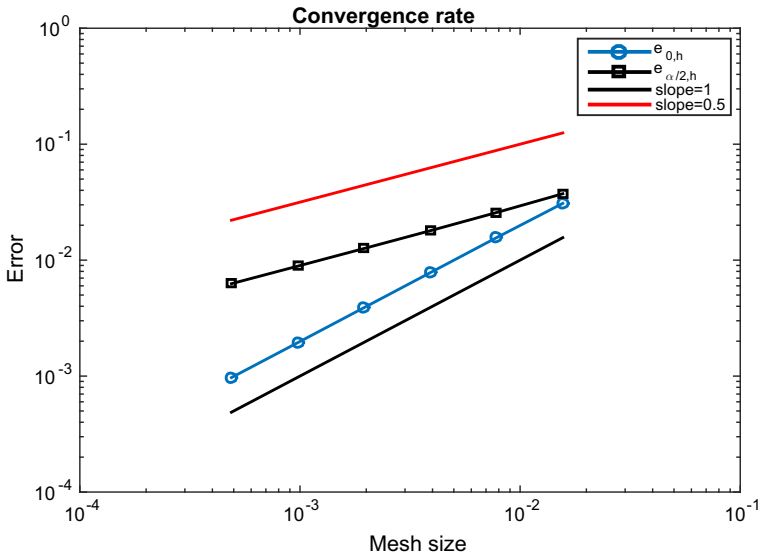


Fig. 5 The convergence rate of Example 1

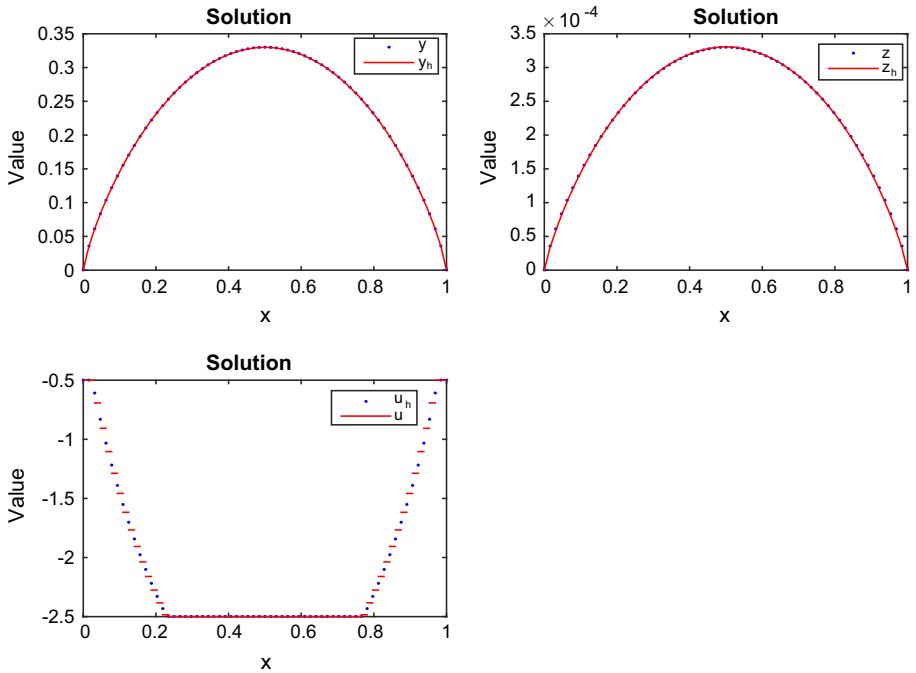


Fig. 6 The numerical solution and the exact solution of Example 1

Table 7 The time consuming of PDAS method for $\gamma = 1.0$

N	256	512	1024	2048	4096	8192
PDAS(s)	1.054E-1(115.5)	2.760E-1(198.5)	1.603E+0(354.5)	1.118E+1(621.5)	8.321E+1(1079.5)	6.597E+2(2177.5)
	1.156E-1(115.5)	3.300E-1(192.5)	1.609E+0(356)	1.135E+1(631.5)	8.289E+1(1056.5)	7.921E+2(2619.5)
	1.115E-1(114.5)	3.507E-1(205.5)	1.566E+0(344.5)	1.121E+1(623.5)	8.443E+1(1099.5)	6.394E+2(2106.5)
T(s)	6.500E-1	1.410E+0	5.690E+0	4.392E+1	2.784E+2	2.258E+3
FFT PDAS(s)	1.085E-1(120.5)	1.738E-1(202)	2.676E-1(341.5)	1.325E+0(640)	2.405E+0(1234.5)	1.782E+1(2079.5)
	9.446E-2(109.5)	1.832E-1(201.5)	3.153E-1(342.5)	1.349E+0(635.5)	2.301E+0(1179.5)	2.033E+1(2417.5)
	9.958E-2(118.5)	1.764E-1(192.5)	3.175E-1(341.5)	1.453E+0(686)	2.481E+0(1275.5)	1.761E+1(1275.5)
T(s)	5.200E-2	7.900E-1	1.304E+0	9.250E+0	1.901E+1	1.084E+2

6 Conclusion

In this paper, finite element approximation of optimal control governed by space fractional equation is investigated. A priori error estimate is derived. A fast primal dual active set algorithm based on the Toeplitz structure of coefficient matrix of discrete equations is designed. Numerical experiments are carried out to show the performance of the numerical method and the fast algorithm. There are many issues in this field can be addressed, for example, optimal control problem governed by fractional Laplacian operator, identification of the fractional order in a nonlocal equation and other type optimal control problem ([30]).

Acknowledgements The authors would like to thank the referees for their careful reviews and many valuable suggestions which have led to a considerably improved paper.

References

- Benson, D.A., Wheatcraft, S.W., Meerschaert, M.M.: The fractional order governing equations of Levy motion. *Water Resour. Res.* **36**, 1413–1423 (2000)
- Meerschaert, M.M., Sikorskii, A.: *Stochastic Models for Fractional Calculus*, De Gruyter Studies in Mathematics, vol. 43. Walter de Gruyter, Berlin (2012)
- Mophou, G.: Optimal control of fractional diffusion equation. *Comput. Math. Appl.* **61**, 68–78 (2011)
- Mophou, G., N'Guérékata, G.M.: Optimal control of fractional diffusion equation with state constraints. *Comput. Math. Appl.* **62**, 1413–1426 (2011)
- Fujishiro, K., Yamamoto, M.: Approximate controllability for fractional diffusion equations by interior control. *Appl. Anal.* **93**(9), 1793–1810 (2014)
- Sprekels, J., Valdinoci, E.: A new type of identification problems: optimizing the fractional order in a nonlocal evolution equation. *SIAM J. Control. Optim.* **55**, 70–93 (2017)
- Ye, X.Y., Xu, C.J.: A spectral method for optimal control problem governed by the abnormal diffusion equation with integral constraint on the state. *Sci. Sin. Math.* **46**, 1053–1070 (2016)
- Ye, X.Y., Xu, C.J.: Spectral optimization methods for the time fractional diffusion inverse problem. *Numer. Math. Theory Methods Appl.* **6**(3), 499–519 (2013)
- Ye, X.Y., Xu, C.J.: A space-time spectral method for the time fractional diffusion optimal control problems. *Adv. Differ. Equ.* **2015**, 156 (2015)
- Li, S.Y., Zhou, Z.J.: Legendre pseudo-spectral method for optimal control problem governed by a time-fractional diffusion equation. *Int. J. Comput. Math.* **95**(6–7), 1308–1325 (2018)
- Zaky, M.A., Machado, J.A.T.: On the formulation and numerical simulation of distributed-order fractional optimal control problems. *Commun. Nonlinear Sci. Numer. Simul.* **52**, 177–189 (2017)
- Antil, H., Otárola, E.: A FEM for an optimal control problem of fractional powers of elliptic operators. *SIAM J. Control Optim.* **53**(6), 3432–3456 (2015)
- Antil, H., Otárola, E., Salgado, A.J.: A space–time fractional optimal control problem: analysis and discretization. *SIAM J. Control Optim.* **54**(3), 1295–1328 (2016)
- Antil, H., Otárola, E.: An a posteriori error analysis for an optimal control problem involving the fractional Laplacian. *IMA J. Numer. Anal.* **38**(1), 198–226 (2017)
- Antil, H., Otárola, E., Salgado, A.J.: Optimization with respect to order in a fractional diffusion model: analysis, approximation and algorithmic aspects. *J. Sci. Comput.* (2018). <https://doi.org/10.1007/s10915-018-0703-0>
- Biccari, U., Hernández-Santamaría, V.: Controllability of a one-dimensional fractional heat equation: theoretical and numerical aspects. *IMA J. Math. Control Inform.* (2018). <https://doi.org/10.1093/imamci/dny025>
- Zhou, Z.J., Gong, W.: Finite element approximation of optimal control problems governed by time fractional diffusion equation. *Comput. Math. Appl.* **71**, 301–318 (2016)
- Jin, B.T., Li, B.Y., Zhou, Z.: Pointwise-in-time error estimates for an optimal control problem with subdiffusion constraint, [arXiv:1707.08808](https://arxiv.org/abs/1707.08808)
- Du, N., Wang, H., Liu, W.B.: A fast gradient projection method for a constrained fractional optimal control. *J. Sci. Comput.* **68**, 1–20 (2016)
- Wang, H., Wang, K.X., Sircar, T.: A direct $O(N \log N)$ finite difference method for fractional diffusion equations. *J. Comput. Phys.* **229**(21), 8095–8104 (2010)

21. Ervin, V.J., Heuer, N., Roop, J.P.: Regularity of the solution to 1-D fractional order diffusion equations. *Math. Comp.* **87**, 2273–2294 (2018)
22. Kunisch, K., Vexler, B.: Constrained Dirichlet boundary control in L^2 for a class of evolution equations. *SIAM J. Control Optim.* **46**(5), 1726–1753 (2007)
23. Ervin, V.J., Roop, J.P.: Variational formulation for the stationary fractional advection dispersion equation. *Numer. Methods Partial Differ. Equ.* **22**, 558–576 (2006)
24. Jin, B.T., Lazarov, R., Pasciak, J., Rundell, W.: Variational formulation of problems involving fractional order differential operators. *Math. Comp.* **84**, 2665–2700 (2015)
25. Bergounioux, M., Ito, K., Kunisch, K.: Primal–dual strategy for constrained optimal control problems. *SIAM J. Control Optim.* **37**(4), 1176–1194 (1999)
26. Li, Y.S., Chen, H.Z., Wang, H.: A mixed-type Galerkin variational formulation and fast algorithms for variable-coefficient fractional diffusion equations. *Math. Methods Appl. Sci.* **40**(14), 5018–5034 (2017)
27. Jia, L.L., Chen, H.Z., Wang, H.: Mixed-type Galerkin variational principle and numerical simulation for a generalized nonlocal elastic model. *J. Sci. Comput.* **71**(2), 660–681 (2017)
28. Davis, P.J.: *Circulant Matrices*. Wiley, New York (1979)
29. Gary, R.M.: Toeplitz and circulant matrices: a review. *Found Trends Commun. Inf. Theory* **2**, 155–239 (2001)
30. Gong, W., Hinze, M., Zhou, Z.J.: A priori error analysis for finite element approximation of parabolic optimal control problems with pointwise control. *SIAM J. Control Optim.* **52**(1), 97–119 (2014)