Journal of Scientific Computing (2022) 91:56
https://doi.org/10.1007/s10915-022-01795-x

Discretization and Analysis of an Optimal Control of a Variable-Order Time-Fractional Diffusion Equation with Pointwise Constraints

Xiangcheng Zheng[1] · Hong Wang[2]

Received: 8 June 2021 / Revised: 27 September 2021 / Accepted: 6 February 2022 /
Published online: 4 April 2022

Abstract
We prove the well-posedness and regularity of an optimal control model with pointwise constraints governed by a variable-order Caputo time-fractional diffusion equation (tFDE), in which the adjoint equation reduces to a Riemann–Liouville tFDE with a different type of variable-order fractional differential operator. We develop and analyze a finite element discretization to the optimal control model without any regularity assumptions of the true solution. Numerical experiments are performed to substantiate the theoretical findings.

Keywords Optimal control · Variable-order time-fractional diffusion equation · Well-posedness · Regularity · Finite element method · Error estimate

1 Introduction

Optimal control [4, 14, 15, 20, 30] governed by tFDEs is getting popular because of its modeling abilities of challenging phenomena such as anomalously diffusive transport, memory effect and long-range interactions [3, 6–10, 13, 18, 21, 29, 32, 33, 35, 44]. In applications (e.g., bioclogging [5], hydrofracturing in gas/oil recovery [12], viscoelastic materials [32]), the structure of porous materials may evolve in time and so the fractal dimension of the material does, leading to a variable-order tFDE via the Hurst index [23, 25, 40]. Further, the classical tFDE just contains the fractional time derivative $\partial_t^\alpha u$ and was derived as the

This work was partially funded by ARO MURI Grant W911NF-15-1-0562, by the National Science Foundation under Grant DMS-2012291, by the China Postdoctoral Science Foundation 2021TQ0017 and 2021M700244, and by the International Postdoctoral Exchange Fellowship Program (Talent-Introduction Program) YJ20210019.

✉ Hong Wang
hwang@math.sc.edu

Xiangcheng Zheng
zhengxch@math.pku.edu.cn

1 School of Mathematical Sciences, Peking University, Beijing 100871, China

2 Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

diffusion limit of a continuous time random walk as the number of particle jumps tends to infinity [28], and so holds only for $t \gg 1$. That is why it has initial weak singularity [34]. Motivated by these discussions, we analyze and discretize the optimal control problem

$$\min_{c\in U(a,b)} J(u,c) = \frac{1}{2}\|u-u_d\|^2_{L^2(L^2(\Omega))} + \frac{\gamma}{2}\|c\|^2_{L^2(L^2(\Omega))}, \tag{1}$$

which is governed by the two time-scale variable-order Caputo tFDE Model that not only resolves the initial singularity but captures the long-term subdiffusive transport behavior [22, 23, 39–41]

$$\begin{aligned} &\partial_t u + k\,\partial_t^{\alpha(t)} u + \mathcal{B}u = f(\boldsymbol{x},t) + c(\boldsymbol{x},t), \;\; (\boldsymbol{x},t)\in\Omega\times(0,T]; \\ &u(\boldsymbol{x},0) = u_0(\boldsymbol{x}),\;\; \boldsymbol{x}\in\Omega; \quad u(\boldsymbol{x},t)=0,\;\; (\boldsymbol{x},t)\in\partial\Omega\times[0,T]. \end{aligned} \tag{2}$$

Here $\Omega \subset \mathbb{R}^d$ $(d = 1,2,3)$ is a convex polygonal domain, $\mathcal{B} := -\nabla\cdot\big(\boldsymbol{K}(\boldsymbol{x})\nabla\big)$ with $\nabla := (\partial/\partial x_1,\cdots,\partial/\partial x_d)^\top$ and $\boldsymbol{K}(\boldsymbol{x}) := (k_{ij}(\boldsymbol{x}))_{i,j=1}^d$ is the symmetric diffusivity tensor. $k\in\mathbb{R}^+$, and u_0, f, u_d and c denote the initial value, source, the target function and the control variable, respectively. $\partial_t^{\alpha(t)}$ and $U(a,b)$ are defined by [23, 40]

$$\begin{aligned} &\partial_t^{\alpha(t)} g := {}_0I_t^{1-\alpha(t)}\partial_t g, \qquad {}_0I_t^{1-\alpha(t)} g := \frac{1}{\Gamma(1-\alpha(t))}\int_0^t \frac{g(s)}{(t-s)^{\alpha(t)}}ds, \\ &U(a,b) := \big\{c\in L^2(0,T;L^2(\Omega)) : a\le c(\boldsymbol{x},t)\le b \text{ a.e. in } \Omega\times[0,T],\; a,b\in\mathbb{R}\big\}. \end{aligned} \tag{3}$$

The optimal control problem (1)–(2) presents mathematical and numerical challenges that are not common in classical constant-order tFDEs: (i) For constant-order tFDEs, the adjoint equation of a Caputo tFDE is a Riemann–Liouville tFDE; and the singularity of their solutions limits the accuracy of their numerical approximations [34, 40]. For problem (1)–(2), the adjoint equation yields a different type of variable-order Riemann–Liouville tFDE (11), in which α assumes its value $\alpha(s)$ for each $s\in[0,t]$ inside the integral (13) [23, 42]. (ii) There is no analytical result on the variable-order Riemann–Liouville tFDE (11) in the literature, except for the one for its Caputo analogue [42] that requires much smoother data and does not apply here. Unlike constant-order fractional derivatives, the variable-order Riemann–Liouville and Caputo derivatives are not equivalent even for a zero initial data. (iii) The coupling of the variable-order tFDE (2), the adjoint tFDE (11), and the pointwise constraint (3) limit the regularity of the solution to the optimal control model (1)–(2), so those proved for linear variable-order tFDEs are no longer valid [38, 42]. (iv) The coefficients of the L1 discretization to the Riemann–Liouville variable-order fractional derivative (13) lose the key properties in their error estimates. (v) Error estimates of the discretization of the optimal control problem are to be proved to accommodate its reduced regularity since those for linear variable-order tFDEs [42] no longer hold.

In a recent work a variable-order time-fractional optimal control problem with integral constraints, in which the variable-order fractional derivative in the state equation is different from that in (2), was investigated in [43]. The solutions were estimated in weighted Sobolev norms, based on which the numerical analysis was performed. Compared with [43], the current paper has the following features:

- The Caputo variable-order fractional derivative in (2) is the type of $\alpha = \alpha(t)$, which is different from the $\alpha = \alpha(s)$ type Caputo derivative in [43] due to the hidden-memory effect of the later. Although the adjoint of (2) has the $\alpha = \alpha(s)$ type Riemann–Liouville derivative, it is not equivalent to its Caputo analogue as in the case of constant fractional-order problems even if the function has a zero initial (or terminal) value, due to the incommutability of the integer-order derivative and the variable-order fractional integral operators. Therefore, the model in this paper is indeed different from that in [43].

- Due to the differences of the models, different and improved techniques are also required in the current work compared with [43]. For instance, in the proof of Lemma 1 we need to perform the integration by parts for the $\alpha = \alpha(s)$ type fractional integral before differentiation, which is quite complicated and requires technical treatments due to the dependence of α on s. However, this difficulty is not encountered in [43], in which the α of the variable-order Riemann–Liouville fractional derivative depends on t and thus the corresponding kernel could be integrated exactly.
- The current work studies a pointwise constraint problem, which is in general more difficult than the integral constraint case investigated in [43]. As [43] adopts the spectral decomposition approach in theoretical analysis, which is established under the L^2 framework, the high-order derivatives of the solutions are only estimated in weighted L^2 norms. In this work we alternatively prove the regularity estimates via resolvent estimates to obtain sharper results in non-weighted L^p spaces. Furthermore, a critical step (cf. the estimates (31)–(32)) of this work is proved in a simper and clearer manner than that in [43] via properties of solution operators.
- Since the regularity results in this paper and [43] are analyzed in different spaces, the analysis of the truncation errors, which is based on the smoothing properties of the solutions, are thus different. Furthermore, due to the differences of the models in these two works caused by the non-equivalence of the variable-order Riemann–Liouville and Caputo fractional operators, the techniques in [43] do not apply directly for the current problem and need to be modified and improved, which distinguish this work from [43] from the aspect of numerical analysis.

The rest of the paper is as follows: In Sect. 2 we derive a formulation from the optimality condition. In Sect. 3 we prove the well-posedness and regularity of the adjoint equation under a weaker condition from the control problem. In Sect. 4 we prove the well-posedness of the optimal control problem (1)–(2) and (11) and the regularity of its solution. In Sect. 5 we discretize the optimal control model. In Sect. 6 we prove the stability and optimal-order error estimate of the discretization without any regularity assumption of the true solution. In Sect. 7 we perform numerical experiments to substantiate the theoretical findings. In Sect. 8 we prove auxiliary lemmas.

2 Preliminaries and the Optimality Condition

For a postive integer m, let $C^m(\Omega)$ and $H^m(\Omega)$ be the spaces of continuous functions with continuous derivatives up to order m and the Lebesgue square integrable functions with square integrable weak derivatives up to order m in Ω. For a non-integer real $s > 0$, $1 \leq p \leq \infty$, and a Banach space S, the fractional Sobolev space $H^s(\Omega)$ is defined by the real method of interpolation between $H^m(\Omega)$ and $H^{m+1}(\Omega)$ with $m = \lfloor s \rfloor$ [1] and $L^p(a, b; S)$ and $H^1(a, b; S)$ denote the spaces of functions f such that $\|f\|_S$ and $\|\partial_t f\|_S$ in $L^p(a, b)$ and $L^2(a, b)$, respectively. All the spaces are equipped with the standard norms [1, 11]. The eigenfunctions $\{\phi_i\}_{i=1}^{\infty}$ of the Sturm-Liouville problem $\mathcal{B}\phi_i(\boldsymbol{x}) = \lambda_i\phi_i(\boldsymbol{x})$ for $\boldsymbol{x} \in \Omega$ with $\phi_i(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \partial\Omega$ form an orthonormal basis in $L^2(\Omega)$ and $0 < \lambda_i \uparrow \infty$ [11].

Throughout this paper we assume that $\boldsymbol{K}$ is symmetric and positive definite uniformly in $\boldsymbol{x} \in \Omega$ and $k_{i,j} \in C^1(\overline{\Omega})$ for $1 \le i, j \le d$. Define the Sobolev space with $\gamma \ge 0$ by [36]

$$\check{H}^\gamma(\Omega) := \Big\{v \in L^2(\Omega) : \|v\|^2_{\check{H}^\gamma(\Omega)} := \big(\mathcal{B}^\gamma v, v\big) = \sum_{i=1}^\infty \lambda_i^\gamma (v, \phi_i)^2 < \infty\Big\}, \tag{4}$$

and $\check{H}^0(\Omega) = L^2(\Omega)$ and $\check{H}^2(\Omega) = H^2(\Omega) \cap H_0^1(\Omega)$ [1, 36]. We also denote $\dot{H}^1(0,T) := \{g \in H^1(0,T) : g(0) = 0\}$.

Let Q denote a positive constant that may assume different values at different occurrences, drop the subscript L^2 in $(\cdot,\cdot)_{L^2}$ and $\|\cdot\|_{L^2}$, and Ω in the Sobolev spaces and norms, and write $\|\cdot\|_{L^p(S)}$ for $\|\cdot\|_{L^p(0,T;S)}$ when no confusion occurs. We use the convention that a sum vanishes if its upper limit is smaller than its lower limit.

For $g \in H^1(L^2)$, the cut-off $\mathcal{P}g := \max\{a, \min\{g, b\}\} \in H^1(L^2)$ satisfies [19]

$$\|\mathcal{P}\bar{g}\|_{H^\mu} \le \|\bar{g}\|_{H^\mu}, \quad \forall \bar{g} \in H^\mu,\ 0 \le \mu \le 1. \tag{5}$$

Assumption A: (a) $\alpha \in W^{2,\infty}[0,T]$ and $0 \le \alpha(t) \le \alpha^* := \|\alpha\|_{C[0,T]} < 1$ on $[0,T]$. (b) $f, u_d \in H^1(L^2)$ with $\mathcal{B}u_0, f(\boldsymbol{x},0), u_d(\boldsymbol{x},0) \in \check{H}^1$.

Remark 1 In this paper we will reduce the initial value of u to 0 by the transformation $u \to u - u_0$, cf. Section 2.2, then $\mathcal{B}u_0$ appears at the right-hand side of the state equation and serves as part of f. Therefore, $\mathcal{B}u_0$ needs the same regularity as f in spatial direction. In particular, $f(\boldsymbol{x},0) \in \check{H}^1$ leads to the condition $\mathcal{B}u_0 \in \check{H}^1$.

2.1 Resolvent Estimates

For $\theta \in (\frac{\pi}{2}, \pi)$, $\delta > 0$, let Γ_θ be the contour

$$\Gamma_\theta := \big\{z \in \mathbb{C} : |\arg(z)| = \theta, |z| \ge \delta\big\} \cup \big\{z \in \mathbb{C} : |\arg(z)| \le \theta, |z| = \delta\big\}.$$

The following inequalities hold for $0 < \mu \le 1$, $t \in (0,T]$, and $Q = Q(\theta, \mu)$ [2, 17, 24]

$$\int_{\Gamma_\theta} |z|^{\mu-1}|e^{tz}|\,|dz| \le Qt^{-\mu}, \quad \Big\|\int_{\Gamma_\theta} z^\mu (z+\mathcal{B})^{-1} e^{tz}\,dz\Big\|_{L^2} \le Qt^{-\mu}. \tag{6}$$

For $g \in L_{loc}(a,b)$, let $\tilde{g}(t)$ be its zero extension outside (a,b) [32]

$$\begin{gathered}\mathcal{L}g(z) := \int_0^\infty \tilde{g}(t)e^{-tz}dt, \quad \mathcal{L}^{-1}(\mathcal{L}g(z)) := \frac{1}{2\pi i}\int_{\Gamma_\theta} e^{tz}\mathcal{L}g(z)dz = g(t),\\ \mathcal{L}\big({}^R\partial_t^\mu g(t)\big) = z^\mu \mathcal{L}\big(g(t)\big),\ 0 \le \mu < 1; \quad {}^R\partial_t^\mu g := \partial_t\, {}_0I_t^{1-\mu} g.\end{gathered} \tag{7}$$

Let $e^{-t\mathcal{B}}$ be the semigroup of operators defined by $\partial_t\, e^{-t\mathcal{B}}g + \mathcal{B}\, e^{-t\mathcal{B}}g = 0$ with $e^{-t\mathcal{B}}g = 0$ for $\boldsymbol{x} \in \partial\Omega$ and $e^{-t\mathcal{B}}g\big|_{t=0} = g$ for $\boldsymbol{x} \in \Omega$. The solution $u(\boldsymbol{x},t)$ to the homogeneous initial-boundary-value problem of

$$\partial_t u(\boldsymbol{x},t) + \mathcal{B}u(\boldsymbol{x},t) = f(\boldsymbol{x},t) \tag{8}$$

can be expressed as

$$\begin{gathered}u(\boldsymbol{x},t) = \int_0^t e^{-(t-s)\mathcal{B}} f(\boldsymbol{x},s)ds, \quad e^{-t\mathcal{B}}g(\boldsymbol{x}) = \sum_{i=1}^\infty e^{-\lambda_i t}(g,\phi_i)\phi_i(\boldsymbol{x}),\\ e^{-t\mathcal{B}}g(\boldsymbol{x}) = \frac{1}{2\pi \mathrm{i}}\int_{\Gamma_\theta} e^{zt}(z+\mathcal{B})^{-1}g(\boldsymbol{x})\,dz, \quad q \in L^2.\end{gathered} \tag{9}$$

The following estimates hold for $s \geq r \geq -1$ and for any $t > 0$ [36]

$$\|e^{-t\mathcal{B}}\|_{L^2\to L^2} \leq Q; \quad \|e^{-t\mathcal{B}}g\|_{\check{H}^s} \leq Qt^{-(s-r)/2}\|g\|_{\check{H}^r}, \quad g \in \check{H}^r. \tag{10}$$

2.2 The First-Order Optimality Condition

Use $w = u - u_0$ to rewrite problem (2) in terms of w with $u_0 = 0$, u and f replaced by w and $f - \mathcal{B}u_0$, respectively. By Assumption A, the data in the new problem retains the regularity of the original. Thus we again use (2) with $u_0 = 0$, denote $f - \mathcal{B}u_0$ by f, and use u for the unknown solution in the rest of the paper.

Theorem 1 *Under Assumption A the optimal control problem (1)–(2) admits a unique solution (u, c) such that $u \in H^1(L^2) \cap L^2(\check{H}^2)$. There exists an adjoint state z such that (u, c, z) satisfies tFDE (2), and*

$$\begin{aligned} -\partial_t z + k\,{}^R\hat{\partial}_t^{\alpha(t)} z + \mathcal{B}z &= u(\boldsymbol{x}, t; c) - u_d(\boldsymbol{x}, t), \ (\boldsymbol{x}, t) \in \Omega \times [0, T); \\ z(\boldsymbol{x}, T) = 0, \ \boldsymbol{x} \in \Omega; &\quad z(\boldsymbol{x}, t) = 0, \ (\boldsymbol{x}, t) \in \partial\Omega \times [0, T], \end{aligned} \tag{11}$$

$$\int_0^T \int_\Omega (\gamma c + z)(v - c)d\boldsymbol{x}dt \geq 0, \qquad \forall v \in U(a, b). \tag{12}$$

The backward Riemann–Liouville differential operator ${}^R\hat{\partial}_t^{\alpha(t)}$ is [23, 42]

$${}^R\hat{\partial}_t^{\alpha(t)} g := -\partial_t\, {}_t\hat{I}_T^{1-\alpha(t)} g, \quad {}_t\hat{I}_T^{1-\alpha(t)} g := \int_t^T \frac{g(s)}{\Gamma(1 - \alpha(s))(s - t)^{\alpha(s)}} ds. \tag{13}$$

Proof We prove that (1)–(2) admits a unique solution (cf. [37] or [44,Theorem 3.3]). Since $\hat{J}(c) := J(u(c), c) \geq 0$, there exists a sequence $\{c^{(l)}\}_{l=1}^\infty \subset U(a, b)$ with $J_{inf} \leq \hat{J}(c^{(l+1)}) \leq \hat{J}(c^{(l)})$ such that $\lim_{l\to\infty} \hat{J}(c^{(l)}) = J_{inf} := \inf_{c\in U(a,b)} \hat{J}(c)$. Let $u^{(l)}$ be the solution to (2) with c being replaced by $c^{(l)}$. We have $\|c^{(l)}\|_{L^2(L^2)} \leq Q_0$ for some $Q_0 > 0$ and all the l. We apply this estimate and (29) (that is, $\|u^{(l)}\|_{H^1(L^2)} + \|u^{(l)}\|_{L^2(\check{H}^2)} \leq Q(\|f\|_{L^2(L^2)} + \|c^{(l)}\|_{L^2(L^2)})$) to obtain $\|u^{(l)}\|_{H^1(L^2)} + \|u^{(l)}\|_{L^2(\check{H}^2)} \leq Q\big(\|f\|_{L^2(L^2)} + Q_0\big)$, which implies $\|\partial_t^{\alpha(t)} u^{(l)}\|_{L^2(L^2)} \leq Q\big(\|f\|_{L^2(L^2)} + Q_0\big)$. Then there exist weakly convergent subsequences $\{u^{(l_j)}\}_{j=1}^\infty$ and $\{c^{(l_j)}\}_{j=1}^\infty$ such that $\{u^{(l_j)}\}_{j=1}^\infty$, $\{\partial_t^{\alpha(t)} u^{(l_j)}\}_{j=1}^\infty$ and $\{c^{(l_j)}\}_{j=1}^\infty$ converge weakly to u_*, $\partial_t^{\alpha(t)} u_*$ and c_* in $H^1(L^2) \cap L^2(\check{H}^2)$, $L^2(L^2)$ and $L^2(L^2)$, respectively. Since $U(a, b)$ is closed and convex, $U(a, b)$ is weakly closed. Thus, $c^* \in U(a, b)$. We multiply (2) associated with c_* by any $\phi \in C^\infty(\Omega \times [0, T])$ with $\phi(\boldsymbol{x}, T) = 0$ for $\boldsymbol{x} \in \Omega$ and $\phi(\boldsymbol{x}, t) = 0$ on $\partial\Omega \times [0, T]$, integrate the equation by parts and apply the adjoint property of $\partial_t^{\alpha(t)}$ and ${}^R\hat{\partial}_t^{\alpha(t)}$ [31,Theorem 3.2] to get

$$\begin{aligned} \int_0^T \int_\Omega (f + c_*)\phi d\boldsymbol{x}dt &= \lim_{j\to\infty} \int_0^T \int_\Omega \big(f + c^{(l_j)}\big)\phi d\boldsymbol{x}dt \\ &= \lim_{j\to\infty} \int_0^T \int_\Omega \big(\partial_t + k\partial_t^{\alpha(t)} + \mathcal{B}\big)u^{(l_j)} \cdot \phi d\boldsymbol{x}dt \\ &= \lim_{j\to\infty} \int_0^T \int_\Omega u^{(l_j)} \cdot \big(-\partial_t + k\,{}^R\hat{\partial}_t^{\alpha(t)} + \mathcal{B}\big)\phi d\boldsymbol{x}dt \\ &= \int_0^T \int_\Omega u_* \cdot \big(-\partial_t + k\,{}^R\hat{\partial}_t^{\alpha(t)} + \mathcal{B}\big)\phi d\boldsymbol{x}dt = \int_0^T \int_\Omega \big(\partial_t + k\partial_t^{\alpha(t)} + \mathcal{B}\big)u_* \cdot \phi d\boldsymbol{x}dt. \end{aligned}$$

Thus, u_* solves Eq. (2) associated with c_*. By the weakly lower semi-continuity of $\hat{J}(c)$, $J_{inf} \geq \liminf_{j\to\infty} \hat{J}(c^{(l_j)}) \geq \hat{J}(c_*) \geq J_{inf}$. So (u_*, c_*) is a solution to problem (1)–(2). The uniqueness follows from the strict convexity of J. For any $v \in U(a,b)$, $0 < \varepsilon \ll 1$, $\delta c := v - c$, $c + \varepsilon\delta c \in U(a,b)$. $\delta u(c) := \big(u(c+\varepsilon\delta c) - u(c)\big)/\varepsilon$ satisfies the homogeneous initial-boundary value problem of

$$\partial_t \delta_\varepsilon u + k\partial_t^{\alpha(t)}\delta_\varepsilon u + \mathcal{B}\delta_\varepsilon u = \delta c(\boldsymbol{x},t), \quad (\boldsymbol{x},t) \in \Omega \times (0,T] \tag{14}$$

which has ε-independent coefficients and right-hand side. Hence, the solution $\delta_\varepsilon u = \delta u$ is independent of ε by the uniqueness of the solution to problem (14). Use $u(c+\varepsilon\delta c) - u(c) = \varepsilon\delta u$ and the adjoint property of $\partial_t^{\alpha(t)}$ and ${}^R\hat{\partial}_t^{\alpha(t)}$ to obtain

$$\begin{aligned} 0 \leq \partial_c \hat{J}(c)\delta c &= \lim_{\varepsilon\to 0^+} (\hat{J}(c+\varepsilon\delta c) - \hat{J}(c))/\varepsilon \\ &= \int_0^T\int_\Omega (u(c) - u_d)\delta u \, d\boldsymbol{x}dt + \int_0^T\int_\Omega \gamma c\delta c \, d\boldsymbol{x}dt \\ &= \int_0^T\int_\Omega \big(-\partial_t z + k\,{}^R\hat{\partial}_t^{\alpha(t)} z + \mathcal{B}z\big)\delta u \, d\boldsymbol{x}dt + \int_0^T\int_\Omega \gamma c\delta c \, d\boldsymbol{x}dt \\ &= \int_0^T\int_\Omega \big(\gamma c + z\big)\delta c \, d\boldsymbol{x}dt = \int_0^T\int_\Omega \big(\gamma c + z\big)(v - c)\, d\boldsymbol{x}dt. \end{aligned} \tag{15}$$

□

Remark 2 Inequality (12) implies that c is given by the projection $\mathcal{P}$ [13, 44]

$$c(\boldsymbol{x},t) = \mathcal{P}\big(-z(\boldsymbol{x},t)/\gamma\big). \tag{16}$$

3 Adjoint tFDE

We analyze the well-posedness and regularity of the forward-in-time analogue of problem (11) with homogeneous initial and boundary values

$$\big(\partial_t + k\,{}^R\bar{\partial}_t^{\alpha(t)} + \mathcal{B}\big)z = p, \quad {}^R\bar{\partial}_t^{\alpha(t)} g := \partial_t\, {}_0\bar{I}_t^{1-\alpha(t)} g = \partial_t \int_0^t \frac{(t-s)^{-\alpha(s)} g(s)}{\Gamma(1-\alpha(s))} ds. \tag{17}$$

Lemma 1 *Under assumption (a), $g \in \dot{H}^1(0,T)$ has the estimate*

$$\big|{}^R\bar{\partial}_t^{\alpha(t)} g\big| \leq Q\int_0^t |g'(s)|(t-s)^{-\alpha^*} ds, \quad 0 \leq t \leq T; \quad Q = Q(\|\alpha\|_{W^{1,\infty}}, T, \alpha^*).$$

Proof We write $g(s) = \int_0^s g'(y)dy$ and interchange the order of integration to get

$$\begin{aligned} {}^R\bar{\partial}_t^{\alpha(t)} g &= \partial_t \int_0^t \frac{\int_0^s g'(y)dy\, ds}{\Gamma(1-\alpha(s))(t-s)^{\alpha(s)}} = \partial_t \int_0^t g'(y)\int_y^t \frac{(t-s)^{-\alpha(s)}}{\Gamma(1-\alpha(s))} ds dy \\ &= \int_0^t g'(y)\bigg[\partial_t \int_y^t \frac{(t-s)^{-\alpha(s)}}{\Gamma(1-\alpha(s))} ds\bigg] dy, \end{aligned} \tag{18}$$

because the inner integral on the right end of the first row tends to 0 as $y \to t$ since α is strictly less than 1. We integrate the term in the square bracket by parts to obtain

$$\begin{aligned} &\partial_t \int_y^t \frac{(t-s)^{-\alpha(s)}}{\Gamma(1-\alpha(s))} ds = -\partial_t \int_y^t \frac{(t-s)^{\alpha(t)-\alpha(s)}}{\Gamma(1-\alpha(s))} d\bigg(\frac{(t-s)^{1-\alpha(t)}}{1-\alpha(t)}\bigg) \\ &= \partial_t \bigg(\frac{(t-y)^{1-\alpha(y)}}{\Gamma(1-\alpha(y))(1-\alpha(t))}\bigg) + \partial_t \int_y^t \frac{(t-s)^{1-\alpha(t)}}{1-\alpha(t)} \partial_s\Big(\frac{(t-s)^{\alpha(t)-\alpha(s)}}{\Gamma(1-\alpha(s))}\Big) ds. \end{aligned}$$

We bound the (leading terms of the) two terms on the right-hand side by

$$
\begin{aligned}
&\left|\partial_t\left(\frac{(t-y)^{1-\alpha(y)}}{\Gamma(1-\alpha(y))(1-\alpha(t))}\right)\right| \le \frac{Q(t-y)^{-\alpha(y)}}{\Gamma(-\alpha(y))(1-\alpha(t))} \le \frac{Q}{(t-y)^{\alpha(y)}} \le \frac{Q}{(t-y)^{\alpha^*}},\\
&\left|\partial_t \int_y^t \frac{(t-s)^{1-\alpha(s)}\left(\alpha'(s)\ln(t-s)+\frac{\alpha(t)-\alpha(s)}{t-s}+\frac{\Gamma'(1-\alpha(s))\alpha'(s)}{\Gamma(1-\alpha(s))}\right)ds}{(1-\alpha(t))\Gamma(1-\alpha(s))}\right|\\
&\le \int_y^t \frac{|\alpha'(s)\ln(t-s)|\,(t-s)^{-\alpha(s)}(1+Q(t-s))}{(1-\alpha(t))\Gamma(-\alpha(s))}ds \le Q\int_y^t \frac{|\ln(t-s)|ds}{(t-s)^{\alpha^*}} \le Q.
\end{aligned}
\tag{19}
$$

To relax $\alpha \in W^{2,\infty}$ to $\alpha \in W^{1,\infty}$, we rebound the second term in the bracket

$$
\begin{aligned}
&\left|\int_y^t \frac{(t-s)^{1-\alpha(s)}\partial_t\left(\frac{\alpha(t)-\alpha(s)}{t-s}\right)ds}{(1-\alpha(t))\Gamma(1-\alpha(s))}\right| = \left|\int_y^t \frac{(t-s)^{1-\alpha(s)}\partial_s\left(\frac{\alpha(t)-\alpha(s)}{t-s}\right)ds}{(1-\alpha(t))\Gamma(1-\alpha(s))}\right|\\
&= \left|\int_y^t \partial_s\left(\frac{(t-s)^{1-\alpha(s)}}{(1-\alpha(t))\Gamma(1-\alpha(s))}\right)\frac{\alpha(t)-\alpha(s)}{t-s}ds\right.\\
&\quad\left.+\frac{(t-y)^{1-\alpha(y)}}{(1-\alpha(t))\Gamma(1-\alpha(y))}\frac{\alpha(t)-\alpha(y)}{t-y}\right| \le Q(\alpha^*, \|\alpha\|_{W^{1,\infty}}, T).
\end{aligned}
$$

□

Theorem 2 *Suppose assumption (a) holds and $p \in L^2(L^2)$, problem (17) has a unique solution $z \in H^1(L^2) \cap L^2(\check{H}^2)$ and there is a $Q = Q(\alpha^*, \|\alpha\|_{W^{1,\infty}}, k, T)$ with*

$$
\|z\|_{H^1(L^2)} + \|z\|_{L^2(\check{H}^2)} \le Q\|p\|_{L^2(L^2)}.
$$

Proof The proof could be performed following that of [43,Theorem 3.1] based on Lemma 1, and is thus omitted. □

Theorem 3 *Suppose assumption (a) holds, $p(\boldsymbol{x}, 0) \in \check{H}^1$, $p \in H^1(L^2)$. For $1 \le q \le 2$ with $q < 1/\alpha(0)$, the solution z to problem (17) has the estimate*

$$
\left\|\partial_t^2 z\right\|_{L^q(L^2)} + \left\|\partial_t^2\left({}_0\bar{I}_t^{1-\alpha(t)} z\right)\right\|_{L^q(L^2)} + \left\|\partial_t z\right\|_{L^q(\check{H}^2)} \le Q\left(\|p\|_{H^1(L^2)} + \|p(\cdot,0)\|_{\check{H}^1}\right).
$$

Here $Q = Q(\alpha^*, \|\alpha\|_{W^{2,\infty}}, k, T)$.

Proof By Theorem 2 $z \in H^1(L^2) \cap L^2(\check{H}^2)$. Move the second term on the left side of (17) to the right side and use (9) to express z as follows

$$
z(\boldsymbol{x}, t) = \int_0^t e^{-(t-s)\mathcal{B}} p(\boldsymbol{x}, s)ds - k\int_0^t e^{-(t-s)\mathcal{B}}\, {}^R\bar{\partial}_s^{\alpha(s)} z(\boldsymbol{x}, s)ds =: L_1 - kL_2. \tag{20}
$$

We differentiate L_1 twice in time to obtain

$$
\begin{aligned}
\partial_t^2 L_1 = &-\sum_{i=1}^{\infty}\lambda_i e^{-\lambda_i t}(p(\cdot,0),\phi_i)\phi_i - \sum_{i=1}^{\infty}(\partial_t p(\cdot,t),\phi_i)\phi_i\\
&+\int_0^t \sum_{i=1}^{\infty}\lambda_i e^{-\lambda_i(t-s)}(\partial_s p(\cdot,s),\phi_i)\phi_i ds.
\end{aligned}
$$

By Young's inequality, $\|\partial_t^2 L_1\|_{L^2(L^2)} \le \|p(\cdot,0)\|_{\check{H}^1} + 2\|p\|_{H^1(L^2)}$. By (20) we have

$$
\partial_t^2 L_2 = -\int_0^t e^{-(t-s)\mathcal{B}}\mathcal{B}\left(\partial_s\, {}^R\bar{\partial}_s^{\alpha(s)} z(\boldsymbol{x}, s)\right)ds + \partial_t\, {}^R\bar{\partial}_t^{\alpha(t)} z(\boldsymbol{x}, t). \tag{21}
$$

Like (18)–(19) in the proof of Lemma 1, the leading terms of ${}^R\bar\partial_t^{\alpha(t)} z$ are

$$\int_0^t \frac{\partial_y z(\boldsymbol{x}, y)(t-y)^{-\alpha(y)}dy}{\Gamma(-\alpha(y))(1-\alpha(t))}, \quad \int_0^t \partial_y z(\boldsymbol{x}, y)\int_y^t \frac{\alpha'(s)(t-s)^{-\alpha(s)}\ln(t-s)dsdy}{(1-\alpha(t))\Gamma(-\alpha(s))}. \tag{22}$$

To bound $\partial_t {}^R\bar\partial_t^{\alpha(t)} z$, we just need to bound the derivative of (22). We express $\partial_y z(\boldsymbol{x}, y) = \int_0^y \partial_\theta^2 z(\boldsymbol{x}, \theta)d\theta + \partial_y z(\boldsymbol{x}, 0)$ to rewrite the first term in (22) as

$$\int_0^t \partial_\theta^2 z(\boldsymbol{x}, \theta)\int_\theta^t \frac{(t-y)^{-\alpha(y)}dy}{\Gamma(-\alpha(y))(1-\alpha(t))}d\theta + \partial_t z(\boldsymbol{x}, 0)\int_0^t \frac{(t-y)^{-\alpha(y)}dy}{\Gamma(-\alpha(y))(1-\alpha(t))}. \tag{23}$$

Like (18)–(19), we bound the derivative of (23) (and the first term in (22)) by

$$\begin{aligned}&\left|\partial_t\int_0^t \partial_\theta^2 z(\boldsymbol{x}, \theta)\int_\theta^t \frac{(t-y)^{-\alpha(y)}dy}{\Gamma(-\alpha(y))(1-\alpha(t))}d\theta + \partial_t z(\boldsymbol{x}, 0)\partial_t\int_0^t \frac{(t-y)^{-\alpha(y)}dy}{\Gamma(-\alpha(y))(1-\alpha(t))}\right|\\ &\quad\le Q\int_0^t \frac{|\partial_\theta^2 z(\boldsymbol{x}, \theta)|}{(t-\theta)^{\alpha(\theta)}}d\theta + Q|\partial_t z(\boldsymbol{x}, 0)|t^{-\alpha(0)}.\end{aligned}$$

Similarly to (23), we rewrite the second term in (22) as

$$\begin{aligned}&\int_0^t \partial_\theta^2 z(\boldsymbol{x}, \theta)\int_\theta^t\int_0^{t-y} \frac{\alpha'(t-s)s^{-\alpha(t-s)}\ln s}{(1-\alpha(t))\Gamma(-\alpha(t-s))}dsdyd\theta\\ &\qquad+\partial_t z(\boldsymbol{x}, 0)\int_0^t\int_0^{t-y} \frac{\alpha'(t-s)s^{-\alpha(t-s)}\ln s}{(1-\alpha(t))\Gamma(-\alpha(t-s))}dsdy.\end{aligned}$$

We use the estimate

$$\begin{aligned}&\left|\partial_t\int_\theta^t\int_0^{t-y} \frac{\alpha'(t-s)s^{-\alpha(t-s)}\ln s}{(1-\alpha(t))\Gamma(-\alpha(t-s))}dsdy\right| = \left|\int_\theta^t \frac{\alpha'(y)(t-y)^{-\alpha(y)}\ln(t-y)}{(1-\alpha(t))\Gamma(-\alpha(y))}dy\right.\\ &\qquad\left.+\int_\theta^t\int_0^{t-y}\partial_t\left(\frac{\alpha'(t-s)s^{-\alpha(t-s)}\ln s}{(1-\alpha(t))\Gamma(-\alpha(t-s))}\right)dsdy\right| \le Q\end{aligned} \tag{24}$$

to bound the derivative of the second term in (22) by

$$\left|\partial_t\int_0^t \partial_y z(\boldsymbol{x}, y)\int_y^t \frac{\alpha'(s)(t-s)^{-\alpha(s)}\ln(t-s)}{(1-\alpha(t))\Gamma(-\alpha(s))}dsdy\right| \le Q\Big(\int_0^t |\partial_\theta^2 z(\boldsymbol{x}, \theta)|d\theta + |\partial_t z(\boldsymbol{x}, 0)|\Big)$$

Passing the limit $t \to 0^+$ in (17) yields $\partial_t z(\boldsymbol{x}, 0) = p(\boldsymbol{x}, 0)$. Thus we incorporate the preceding estimates for the two terms in (22) to conclude that

$$\left|\partial_t {}^R\bar\partial_t^{\alpha(t)} z\right| \le Q\int_0^t \frac{|\partial_\theta^2 z(\boldsymbol{x}, \theta)|}{(t-\theta)^{\alpha^*}}d\theta + Q|p(\boldsymbol{x}, 0)|t^{-\alpha(0)}. \tag{25}$$

Use (25) to bound ${}^R\partial_t^\varepsilon\, \partial_t {}^R\bar\partial_t^{\alpha(t)} z$ for $0 < \varepsilon < 1-\alpha^*$. By Lemma 1, ${}^R\bar\partial_t^{\alpha(t)} z|_{t=0} = 0$. Use ${}_0I^{1-\varepsilon}\partial_t g = \partial_t\, {}_0I^{1-\varepsilon} g$ if $g(0) = 0$ to obtain ${}^R\partial_t^\varepsilon\, \partial_t {}^R\bar\partial_t^{\alpha(t)} z = \partial_t {}^R\bar\partial_t^{\alpha(t)+\varepsilon} z$, which, together with (25), yields the estimate

$$\left|{}^R\partial_t^\varepsilon\, \partial_t {}^R\bar\partial_t^{\alpha(t)} z\right| \le Q\int_0^t \frac{|\partial_\theta^2 z(\boldsymbol{x}, \theta)|}{(t-\theta)^{\alpha^*+\varepsilon}}d\theta + Q\big|p(\boldsymbol{x}, 0)\big|t^{-\alpha(0)-\varepsilon}. \tag{26}$$

Use (9) and the Laplace transform of the first term on the right side of $\partial_t^2 L_2$ in (21) to conclude that for $0 < \varepsilon < 1-\alpha^*$

$$\begin{aligned}&\mathcal{L}\Big[-\int_0^t e^{-(t-s)\mathcal{B}}\mathcal{B}\big(\partial_s\, {}^R\bar\partial_s^{\alpha(s)} z(\boldsymbol{x}, s)\big)ds\Big]\\ &\quad= z(z+\mathcal{B})^{-1}\mathcal{L}\big(\partial_t {}^R\bar\partial_t^{\alpha(t)} z(\boldsymbol{x}, t)\big) = \big(z^{1-\varepsilon}(z+\mathcal{B})^{-1}\big)\big(z^\varepsilon\mathcal{L}(\partial_t {}^R\bar\partial_t^{\alpha(t)} z(\boldsymbol{x}, t))\big).\end{aligned}$$

Apply (7) and the Laplace transform of the convolution formula to obtain

$$
\begin{aligned}
&-\int_0^t e^{-(t-s)\mathcal{B}}\mathcal{B}\left(\partial_s\, {}^R\bar{\partial}_s^{\alpha(s)} z(\boldsymbol{x},s)\right)ds \\
&=\int_0^t \left[\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_\theta} z^{1-\varepsilon}(z+\mathcal{B})^{-1}e^{z(t-s)}\mathrm{d}z\right]\left({}^R\partial_s^{\varepsilon}\partial_s\, {}^R\bar{\partial}_s^{\alpha(s)} z(\boldsymbol{x},s)\right)ds.
\end{aligned}
$$

Use (6) to bound the term in the square brackets and (26) to bound the first term on the right side of (21)

$$
\begin{aligned}
&\left\|\int_0^t e^{-(t-s)\mathcal{B}}\mathcal{B}\left(\partial_s\, {}^R\bar{\partial}_s^{\alpha(s)} z(\boldsymbol{x},s)\right)ds\right\| \\
&\le Q\int_0^t \frac{\left\|{}^R\partial_s^{\varepsilon}\partial_s\, {}^R\bar{\partial}_s^{\alpha(s)} z(\cdot,s)\right\|ds}{(t-s)^{1-\varepsilon}} \\
&\le Q\int_0^t \frac{1}{(t-s)^{1-\varepsilon}}\left(\int_0^s \frac{\|\partial_{\theta\theta}z(\cdot,\theta)\|}{(s-\theta)^{\alpha_*+\varepsilon}}d\theta+\|p(\cdot,0)\|s^{-\varepsilon-\alpha(0)}\right)ds \\
&\le Q\int_0^t \frac{\|\partial_{\theta\theta}z(\cdot,\theta)\|}{(t-\theta)^{\alpha_*}}d\theta+Q\|p(\cdot,0)\|t^{-\alpha(0)}.
\end{aligned}
\tag{27}
$$

Apply (25) to bound the L^2 norm of the second term on the right side of (21) by the right side of (27). Multiply (21) by $e^{-\sigma t}$ and take the $\|\cdot\|_{L^q(0,T)}$ on both sides of the equation and use Young's convolution inequality to get

$$
\begin{aligned}
\left\|e^{-\sigma t}\partial_t^2 L_2\right\|_{L^q(L^2)} &\le Q\left\|\int_0^t \frac{e^{-\sigma(t-\theta)}}{(t-\theta)^{\alpha_*}}e^{-\sigma\theta}\|\partial_{\theta\theta}z(\cdot,\theta)\|d\theta\right\|_{L^q(0,T)}+Q\|p(\cdot,0)\| \\
&\le Q\sigma^{\alpha^*-1}\|e^{-\sigma t}\partial_t^2 z\|_{L^q(L^2)}+Q\|p(\cdot,0)\|.
\end{aligned}
\tag{28}
$$

Differentiate (20) twice, take the $\|\cdot\|_{L^q(L^2)}$ on both sides of the resulting equation multiplied by $e^{-\lambda t}$ and use the estimate of $\|\partial_t^2 L_1\|_{L^2(L^2)}$ and (28) to obtain

$$
\left\|e^{-\sigma t}\partial_t^2 z\right\|_{L^q(L^2)} \le Q\sigma^{\alpha^*-1}\|e^{-\sigma t}\partial_t^2 z\|_{L^q(L^2)}+Q\left(\|p\|_{H^1(L^2)}+\|p(\cdot,0)\|_{\check{H}^1}\right).
$$

We set σ large enough to get $\|\partial_t^2 z\|_{L^q(L^2)} \le Q\left(\|p\|_{H^1(L^2)}+\|p(\cdot,0)\|_{\check{H}^1}\right)$, which, together with (25), yields the estimate of $\partial_t^2\left({}_0\bar{I}_t^{1-\alpha(t)}z\right)$. We finally use (4), (17) and preceding two estimates to obtain the bound of $\|\partial_t z\|_{L^q(\check{H}^2)}$. □

Remark 3 In (a) the α is supposed to belong to $W^{2,\infty}[0,T]$, which is required in such estimates as (24), in which the α'' is encountered. The inherent reason of imposing this strong condition lies in the difficulties of estimating the complicated derivatives of variable-order operators. How to relax this condition is an important but challenging topic and we will carry out investigations in the near future.

4 Optimal Control Model

We prove the well-posedness and regularity of the solutions to the optimal control model (1)–(2) based on the results in previous sections.

Theorem 4 *If assumption (a) holds and $f, c \in L^2(L^2)$, problem (2) has a unique solution $u \in H^1(L^2) \cap L^2(\check{H}^2)$ and there is a $Q = Q(\alpha^*, \|\alpha\|_{W^{1,\infty}}, k, T)$*

$$\|u\|_{H^1(L^2)} + \|u\|_{L^2(\check{H}^2)} \le Q\|f + c\|_{L^2(L^2)} \tag{29}$$

If $f, u_d \in H^1(L^2)$, $f(\boldsymbol{x}, 0), u_d(\boldsymbol{x}, 0) \in \check{H}^1$, $c \in H^1(L^2)$, then for $1 \le q < \min\{\frac{4}{3}, \frac{1}{\alpha(0)}\}$

$$\begin{aligned}\|\partial_t^2 u\|_{L^q(L^2)} + \|\partial_t u\|_{L^q(\check{H}^2)} \le Q\big(&\|f\|_{H^1(L^2)} + \|f(\cdot,0)\|_{\check{H}^1} + \|c\|_{H^1(L^2)} + \|u_d\|_{H^1(L^2)} \\ &+\|u_d(\cdot,0)\|_{\check{H}^1}\big), \quad Q = Q(\alpha^*, \|\alpha\|_{W^{2,\infty}}, k, T).\end{aligned} \tag{30}$$

Proof We prove (29) as in Theorem 3. However, the proof does not carry over to (30) as the regularity of $\partial_t^2 u$ requires $c(\boldsymbol{x}, 0) \in \check{H}^1$ that contradicts (16). By Theorems 2 and 3, $z \in H^1(L^2) \cap L^2(\check{H}^2)$ and the estimates in the theorems hold. To bound $\partial_t^2 u$, we refine the proof of Theorem 3 by re-estimating $\partial_t^2 L_1$ with L_1 given by (20) and p replaced by c. We use the second equation in (9) to get

$$\begin{aligned}\partial_t \int_0^t e^{-(t-s)\mathcal{B}} c(\boldsymbol{x}, s)ds &= \partial_t \int_0^t e^{-y\mathcal{B}} c(\boldsymbol{x}, t-y)dy \\ &= e^{-t\mathcal{B}} c(\boldsymbol{x}, 0) + \int_0^t e^{-y\mathcal{B}} \partial_t c(\boldsymbol{x}, t-y)dy \\ &= e^{-t\mathcal{B}} c(\boldsymbol{x}, 0) + \int_0^t e^{-(t-s)\mathcal{B}} \partial_s c(\boldsymbol{x}, s)ds.\end{aligned}$$

We differentiate the equation with respect to t to find

$$\partial_t^2 \int_0^t e^{-(t-s)\mathcal{B}} c(\boldsymbol{x}, s)ds = -e^{-t\mathcal{B}}\mathcal{B}c(\boldsymbol{x}, 0) + \partial_t c(\boldsymbol{x}, t) + \partial_t \int_0^t e^{-(t-s)\mathcal{B}} \partial_s c(\boldsymbol{x}, s)ds. \tag{31}$$

We use Young's convolution inequality to bound the last term on the right-hand side

$$\begin{aligned}\left\|\partial_t \int_0^t e^{-(t-s)\mathcal{B}} \partial_s c(\boldsymbol{x}, s)ds\right\|_{L^2(L^2)} &= \left[\sum_{i=1}^\infty \left\|\int_0^t \lambda_i e^{-\lambda_i(t-s)} (\partial_s c(\cdot, s), \phi_i)ds\right\|^2_{L^2(0,T)}\right]^{\frac{1}{2}} \\ &\le \left[\sum_{i=1}^\infty \int_0^T (\partial_s c(\cdot, s), \phi_i)^2 ds\right]^{\frac{1}{2}} \le \|c\|_{H^1(L^2)}.\end{aligned}$$

We use (5), (10), (16), Theorem 3, and the equivalence between $\check{H}^{\frac{1}{2}-\varepsilon}$ and $H^{\frac{1}{2}-\varepsilon}$ to bound the first term on the right side of (31) for $0 < \varepsilon \ll 1$ by

$$\begin{aligned}\big\|e^{-t\mathcal{B}}\mathcal{B}c(\boldsymbol{x}, 0)\big\|_{L^2} &= \big\|e^{-t\mathcal{B}}c(\boldsymbol{x}, 0)\big\|_{\check{H}^2} \le Qt^{-(\frac{3}{4}+\frac{\varepsilon}{2})}\|c(\boldsymbol{x}, 0)\|_{\check{H}^{\frac{1}{2}-\varepsilon}} \\ &\le Qt^{-(\frac{3}{4}+\frac{\varepsilon}{2})}\|c(\boldsymbol{x}, 0)\|_{H^{\frac{1}{2}-\varepsilon}} \le Qt^{-(\frac{3}{4}+\frac{\varepsilon}{2})}\|z(\boldsymbol{x}, 0)\|_{H^{\frac{1}{2}-\varepsilon}} \\ &\le Qt^{-(\frac{3}{4}+\frac{\varepsilon}{2})}\|\partial_t z\|_{L^1(H^{\frac{1}{2}-\varepsilon})} \\ &\le Qt^{-(\frac{3}{4}+\frac{\varepsilon}{2})}\big(\|u - u_d\|_{H^1(L^2)} + \|u_d(\cdot, 0)\|_{\check{H}^1}\big).\end{aligned} \tag{32}$$

So for $1 \le q < (\frac{3}{4} + \frac{\varepsilon}{2})^{-1}$, $\|e^{-t\mathcal{B}}\mathcal{B}c(\boldsymbol{x}, 0)\|_{L^q(L^2)} \le Q\big(\|u - u_d\|_{H^1(L^2)} + \|u_d\|_{H^1}\big)$. The remaining analysis can be carried out as in Theorem 3 and is omitted. □

Theorem 5 *Under Assumption A, there is a* $Q = Q(\alpha^*, \|\alpha\|_{W^{2,\infty}}, k, T)$ *such that the solution* (u, z, c) *of the optimal control problem has the regularity estimate*

$$\begin{aligned}
&\|c\|_{H^1(L^2)} \le Q\big(\|u_d\|_{L^2(L^2)} + \|f\|_{L^2(L^2)} + \|c\|_{L^2(L^2)}\big),\\
&\|u\|_{H^1(L^2)} + \|u\|_{L^2(\check{H}^2)} + \|\partial_t^2 u\|_{L^q(L^2)} + \|\partial_t u\|_{L^q(\check{H}^2)} \le Q\big(\|u_d\|_{H^1(L^2)} + \|u_d(\cdot,0)\|_{\check{H}^1}\\
&\quad + \|f\|_{H^1(L^2)} + \|f(\cdot,0)\|_{\check{H}^1} + \|c\|_{L^2(L^2)}\big),\ 1 \le q < \min\{\tfrac{4}{3}, \tfrac{1}{\alpha(0)}\},\\
&\|z\|_{H^1(L^2)} + \|z\|_{L^2(\check{H}^2)} + \big\|\partial_t^2 z\big\|_{L^q(L^2)} + \big\|\partial_t^2\big({}_t\hat{I}_T^{1-\alpha(t)} z\big)\big\|_{L^q(L^2)} + \big\|\partial_t z\big\|_{L^q(\check{H}^2)}\\
&\quad \le Q\big(\|u_d\|_{H^1(L^2)} + \|u_d(\cdot,0)\|_{\check{H}^1} + \|f\|_{L^2(L^2)} + \|c\|_{L^2(L^2)}\big),\ 1 \le q \le 2,\ q < \tfrac{1}{\alpha(0)}.
\end{aligned} \tag{33}$$

Proof By Assumption A, $c \in L^2(L^2)$, and Theorem 4, $u \in H^1(L^2) \cap L^2(\check{H}^2)$. By Theorem 2 Eq. (11) has a unique solution $z \in H^1(L^2) \cap L^2(\check{H}^2)$ and estimates in Theorems 2–3 hold. We combine these estimates with (29) to prove the estimate of z in (33). Use estimates (29)–(30) to get the estimate of u in (33). The estimate of c in (33) is a consequence of (5), (16), Theorem 2, and (29). □

5 Discretization

We develop a numerical scheme to the fractional optimal control model (1)–(2) by discretizing the first-order optimality condition derived in Sect. 2.2.

5.1 The tFDEs (2) and (11)

Let $t_n := n\tau$ for $\tau := T/N$ and $0 \le n \le N$, $f_n := f(\boldsymbol{x}, t_n)$, $c_n := c(\boldsymbol{x}, t_n)$, and $u_n := u(\boldsymbol{x}, t_n)$. Discretize $\partial_t u$ and $\partial_t^{\alpha(t)} u$ by

$$\begin{aligned}
&\partial_t u(\boldsymbol{x}, t_n) = \delta_\tau u_n + E_n := \frac{u_n - u_{n-1}}{\tau} + \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \partial_t^2 u(\boldsymbol{x}, t)(t - t_{n-1})dt,\\
&\partial_t^{\alpha(t_n)} u(\boldsymbol{x}, t_n) = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \frac{\partial_s u(\boldsymbol{x}, s)ds}{\Gamma(1-\alpha(t_n))(t_n - s)^{\alpha(t_n)}} = \delta_\tau^{\alpha(t_n)} u_n + R_n
\end{aligned} \tag{34}$$

for $1 \le n \le N$. Here $\delta_\tau^{\alpha(t_n)} u_n$ and R_n are defined by

$$\begin{aligned}
&\delta_\tau^{\alpha(t_n)} u_n := \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \frac{(t_n - s)^{-\alpha(t_n)} \delta_\tau u_k\, ds}{\Gamma(1-\alpha(t_n))} = b_{n,n} u_n + \sum_{k=1}^{n-1} (b_{n,k} - b_{n,k+1}) u_k,\\
&R_n = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \frac{1}{\tau \Gamma(1-\alpha(t_n))(t_n - s)^{\alpha(t_n)}} \bigg[\int_{t_{k-1}}^{t_k} \int_z^s \partial_\theta^2 u(\boldsymbol{x}, \theta) d\theta dz\bigg] ds
\end{aligned} \tag{35}$$

where $b_{n,k} := \big[(t_n - t_{k-1})^{1-\alpha(t_n)} - (t_n - t_k)^{1-\alpha(t_n)}\big]/\big[\Gamma(2-\alpha(t_n))\tau\big]$. We plug (34) into (2) and integrate the equation multiplied by $\chi \in H_0^1(\Omega)$ on Ω to obtain the following for Eq. (2) for any $\chi \in H_0^1$ and $n = 1, 2, \ldots, N$

$$\big(\delta_\tau u_n + k\delta_\tau^{\alpha(t_n)} u_n + \mathcal{B} u_n, \chi\big) = \big(f_n + c_n, \chi\big) - \big(kR_n + E_n, \chi\big). \tag{36}$$

We discretize $-\partial_t z$ and ${}^R\hat{\partial}_t^{\alpha(t)} z$ backward in time for $n = N, N-1, \ldots, 1$ by

$$\begin{aligned} -\partial_t z(\boldsymbol{x}, t_{n-1}) &= -\delta_\tau z_n + \hat{E}_{n-1} := \frac{z_{n-1} - z_n}{\tau} + \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \partial_t^2 z(\boldsymbol{x}, t)(t_n - t)dt, \\ {}^R\hat{\partial}_t^{\alpha(t)} z(\boldsymbol{x}, t)\big|_{t=t_{n-1}} &= \frac{1}{\tau}\big({}_t\hat{I}_T^{1-\alpha(t)} z\big)\Big|_{t=t_n}^{t=t_{n-1}} + \hat{F}_{n-1} = \hat{\delta}_\tau^{\alpha(t_{n-1})} z_{n-1} + \hat{F}_{n-1} + \hat{R}_{n-1}. \end{aligned} \tag{37}$$

Here $\hat{\delta}_\tau^{\alpha(t_{n-1})} z_{n-1}$, $\hat{F}_{n-1}$ and $\hat{R}_{n-1}$ are defined by

$$\begin{aligned} \hat{\delta}_\tau^{\alpha(t_{n-1})} z_{n-1} &:= \sum_{k=n}^{N} b_{k,n} z_{k-1} - \sum_{k=n}^{N-1} b_{k+1,n+1} z_k, \\ \hat{F}_{n-1} &:= \hat{\partial}_t^{\alpha(t)} z_{n-1} - \frac{1}{\tau}\big({}_t\hat{I}_T^{1-\alpha(t)} z(\boldsymbol{x}, t)\big)\Big|_{t=t_n}^{t=t_{n-1}}, \\ \hat{R}_{n-1} &:= \frac{1}{\tau}\Big[\sum_{k=n}^{N}\int_{t_{k-1}}^{t_k} \frac{z(\boldsymbol{x}, s)(s - t_{n-1})^{-\alpha(s)}}{\Gamma(1-\alpha(s))} - \frac{z_{k-1}(s - t_{n-1})^{-\alpha(t_k)}}{\Gamma(1-\alpha(t_k))} ds \\ &\quad - \sum_{k=n}^{N-1}\int_{t_k}^{t_{k+1}} \frac{z(\boldsymbol{x}, s)(s - t_n)^{-\alpha(s)}}{\Gamma(1-\alpha(s))} - \frac{z_k(s - t_n)^{-\alpha(t_{k+1})}}{\Gamma(1-\alpha(t_{k+1}))} ds\Big]. \end{aligned} \tag{38}$$

Plug (37) in (11) to get the formulation: For $n = N, N-1, \ldots, 1$, find $z(\boldsymbol{x}, t_{n-1})$

$$\begin{aligned} &\big(-\delta_\tau z_n + k\,\hat{\delta}_\tau^{\alpha(t_{n-1})} z_{n-1} + \mathcal{B} z_{n-1}, \chi\big) \\ &\quad = \big(u_{n-1} - u_d(\cdot, t_{n-1}), \chi\big) - \big(k(\hat{F}_{n-1} + \hat{R}_{n-1}) + \hat{E}_{n-1}, \chi\big), \quad \forall \chi \in H_0^1. \end{aligned} \tag{39}$$

5.2 The Optimal Control

Define a quasi-uniform mesh on Ω with the mesh diameter h. Let $S_h(\Omega)$ be continuous and piecewise-linear finite element space with respect to the mesh and $U^\tau(a, b)$ be a time discretization of $U(a, b)$ (3)

$$U^\tau(a, b) := \big\{\boldsymbol{C} := C_{n-1}(\boldsymbol{x}),\ t \in [t_{n-1}, t_n) :\ a \le C_{n-1} \le b,\ 1 \le n \le N\big\}.$$

We now define a discretized optimal control model

$$\min_{\boldsymbol{C} \in U^\tau(a,b)} J^\tau(\boldsymbol{U}, \boldsymbol{C}) := \frac{\tau}{2}\sum_{n=1}^{N}\big(\|U_n - u_d(\cdot, t_n)\|^2 + \gamma\|C_{n-1}\|^2\big), \tag{40}$$

in which $\boldsymbol{U} := \{U_n\}_{n=1}^N \subset S_h$ and U_n satisfies the following equation with $U_0 = 0$

$$\big(\delta_\tau U_n + k\delta_\tau^{\alpha(t_n)} U_n + \mathcal{B}_h U_n, \chi\big) = (f_n + C_{n-1}, \chi), \quad \chi \in S_h, \quad 1 \le n \le N. \tag{41}$$

Here $\mathcal{B}_h : S_h \to S_h$ is defined via $(\mathcal{B}_h V, W) = (\boldsymbol{K}\nabla V, \nabla W)$ for any $V, W \in S_h$.

Theorem 6 *The discrete system (40)–(41) admits a unique solution* $(\boldsymbol{U}, \boldsymbol{C})$*, and an adjoint state* $\boldsymbol{Z} = \{Z_n\}_{n=0}^{N-1} \subset S_h$ *with* $Z_N = 0$ *such that for any* $\chi \in S_h$

$$\big(-\delta_\tau Z_n + k\hat{\delta}_\tau^{\alpha(t_{n-1})} Z_{n-1} + \mathcal{B}_h Z_{n-1}, \chi\big) = \big(U_n - u_d(\cdot, t_n), \chi\big), \quad n = N, \ldots, 1, \tag{42}$$

$$\big(\gamma C_{n-1} + Z_{n-1}, v - C_{n-1}\big) \ge 0, \quad \forall v \in L^2 \text{ with } a \le v \le b. \tag{43}$$

Proof The convexity of J^τ implies the system (40)–(41) admits a unique solution $(\boldsymbol{U}, \boldsymbol{C})$. Let $\delta\boldsymbol{C} := \boldsymbol{V} - \boldsymbol{C}$ for any $\boldsymbol{V} \in U^\tau(a, b)$, $\boldsymbol{C} + \varepsilon\delta\boldsymbol{C} \in U^\tau(a, b)$ for $0 < \varepsilon \ll 1$. Then $\delta_\varepsilon \boldsymbol{U} := \big(\boldsymbol{U}(\boldsymbol{C} + \varepsilon\delta\boldsymbol{C}) - \boldsymbol{U}(\boldsymbol{C})\big)/\varepsilon$ satisfies

$$(\delta_\tau\delta_\varepsilon U_n + k\delta_\tau^{\alpha(t_n)}\delta_\varepsilon U_n + \mathcal{B}_h\delta_\varepsilon U_n, \chi\big) = (\delta C_{n-1}, \chi), \quad \chi \in S_h, \quad 1 \le n \le N \tag{44}$$

By (35) for $\delta_\tau^{\alpha(t_n)}$ and (38) for $\hat{\delta}_\tau^{\alpha(t_{n-1})}$ we observe that

$$\sum_{n=1}^{N}(\delta_\tau^{\alpha(t_n)} U_n, Z_{n-1}) = \sum_{n=1}^{N}(U_n, \hat{\delta}_\tau^{\alpha(t_{n-1})} Z_{n-1}). \tag{45}$$

Use (45) to differentiate $\hat{J}^\tau(\boldsymbol{C}) = J^\tau(\boldsymbol{U}(\boldsymbol{C}), \boldsymbol{C})$ as in Theorem 1 to get

$$0 \le \lim_{\varepsilon\to 0^+} \frac{\hat{J}^\tau(\boldsymbol{C} + \varepsilon\delta\boldsymbol{C}) - \hat{J}^\tau(\boldsymbol{C})}{\varepsilon} = \tau\sum_{m=1}^{N}\int_\Omega \big(\gamma C_{m-1} + Z_{m-1}\big)\delta C_{m-1}\, d\boldsymbol{x}.$$

$\boldsymbol{V} - \boldsymbol{C} = (0, \cdots, 0, v - C_{n-1}, 0, \cdots, 0)$ for $\forall v \in L^2$ with $a \le v \le b$ gives (43). □

As in Remark 2, (43) implies

$$C_{n-1}(\boldsymbol{x}) = \mathcal{P}\big(-Z_{n-1}(\boldsymbol{x})/\gamma\big), \quad 1 \le n \le N. \tag{46}$$

We follow [13, 16, 44] to use (46) to compute C_{n-1}, which has a proved second-order accuracy [13, 44] but generally is not in S_h. An alternative approach is to discretize the control variable in S_h, which is computationally efficient but with a proved sub-optimal convergence rate [26, 27] and will be investigated in the future.

6 Stability and Error Estimates

We prove the stability and error estimates of the numerical schemes (41)–(43) to the optimal control model.

6.1 Stability Estimates of Finite Element Schemes (41) and (42)

We first prove properties of the discretization coefficients of the variable-order fractional derivatives in the follow lemma.

Lemma 2 *The following estimates hold for* $1 \le n \le N-1$

$$\sum_{k=n+1}^{N}(b_{k,n+1} - b_{k,n}) \le b_{n,n} + Q_0, \quad Q_0 = Q_0(\alpha^*, \|\alpha\|_{W^{1,\infty}}, T).$$

Proof For $1 \le n \le N-1$, we have

$$\sum_{k=n+1}^{N}(b_{k,n+1} - b_{k,n}) = b_{n,n} + \sum_{k=n}^{N-1}(b_{k+1,n+1} - b_{k,n}) - b_{N,n}.$$

For $n \le k \le N-1$ and $0 < \varepsilon < 1-\alpha^*$, we apply the mean-value theorem to obtain

$$
\begin{aligned}
b_{k+1,n+1} - b_{k,n} &= \frac{1}{\tau}\int_{t_n}^{t_{n+1}} \frac{(t_{k+1}-s)^{-\alpha(t_{k+1})}}{\Gamma(1-\alpha(t_{k+1}))}ds - \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \frac{(t_k-s)^{-\alpha(t_k)}}{\Gamma(1-\alpha(t_k))}ds\\
&= \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \frac{(t_k-s)^{-\alpha(t_{k+1})}}{\Gamma(1-\alpha(t_{k+1}))} - \frac{(t_k-s)^{-\alpha(t_k)}}{\Gamma(1-\alpha(t_k))}ds\\
&= \frac{\alpha(t_{k+1})-\alpha(t_k)}{\tau}\int_{t_{n-1}}^{t_n} \frac{-(t_k-s)^{-\zeta}\ln(t_k-s)}{\Gamma(1-\zeta)} + \frac{(t_k-s)^{-\zeta}\Gamma'(1-\zeta)}{\Gamma^2(1-\zeta)}ds\\
&\le Q\int_{t_{n-1}}^{t_n} \frac{1}{(t_k-s)^{\alpha^*+\varepsilon}}ds = Q\frac{(t_{k+1}-t_n)^{1-\alpha^*-\varepsilon} - (t_k-t_n)^{1-\alpha^*-\varepsilon}}{1-\alpha_*-\varepsilon}, \quad (47)
\end{aligned}
$$

$$
\sum_{k=n}^{N-1}(b_{k+1,n+1}-b_{k,n}) \le \frac{Q(t_N-t_n)^{1-\alpha^*-\varepsilon}}{1-\alpha_*-\varepsilon} \le Q\frac{\max\{1,T\}}{1-\alpha_*-\varepsilon}.
$$

□

Theorem 7 *Schemes (41) and (42) are stable*

$$
\begin{aligned}
&|||\boldsymbol{U}|||_{L^\infty(L^2)} \le Q\big(|||\boldsymbol{f}|||_{L^1(L^2)} + |||\boldsymbol{C}|||_{L^1(L^2)}\big), \quad |||\boldsymbol{Z}|||_{L^\infty(L^2)} \le Q|||\boldsymbol{U}-\boldsymbol{u}_d|||_{L^1(L^2)},\\
&|||\boldsymbol{V}|||_{L^\infty(L^2)} := \max_{1\le n\le N}\|V_n\|, \quad |||\boldsymbol{V}|||_{L^p(L^2)} := \Big[\tau\sum_{n=1}^N \big\|V_n\big\|^p\Big]^{\frac{1}{p}}, \;\; p=1,2
\end{aligned} \quad (48)
$$

with the norms of $\boldsymbol{C}$ *and* $\boldsymbol{Z}$ *being defined for n from* 0 *to* $N-1$.

Proof We set $\chi = Z_{n-1}$ in (42) and use expression (38) for $\hat{\delta}_\tau^{\alpha(t_{n-1})}$ to get

$$
\begin{aligned}
&\big(1+k\tau b_{n,n}\big)\|Z_{n-1}\|^2 + \big(\mathcal{B}_h Z_{n-1}, Z_{n-1}\big)\\
&\quad = (Z_n, Z_{n-1}) + k\tau\sum_{k=n+1}^N (b_{k,n+1}-b_{k,n})(Z_{k-1}, Z_{n-1}) + \tau(U_n - u_d(\cdot,t_n), Z_{n-1}).
\end{aligned}
$$

Use the coercivity of $\mathcal{B}_h$ and $b_{k,n+1} > b_{k,n}$ to cancel $\|Z_{n-1}\|$ on both sides

$$
\big(1+k\tau b_{n,n}\big)\|Z_{n-1}\| \le \|Z_n\| + k\tau\sum_{k=n+1}^N \big(b_{k,n+1}-b_{k,n}\big)\|Z_{k-1}\| + \tau\|U_n - u_d(\cdot,t_n)\|.
$$

We then prove the following relation by induction

$$
\|Z_{m-1}\| \le A_m\tau\sum_{j=m}^N \big\|U_j - u_d(\cdot,t_j)\big\|, \quad A_m := \big(1+Q_0k\tau\big)^{N-m+1} \quad (49)
$$

where Q_0 is defined in Lemma 2. It is clear that (49) holds for $n = N$. Suppose (49) holds for $n+1 \le m \le N$, we apply $A_1 > A_2 > \cdots > A_N > 1$ and Lemma 2 to obtain

$$
\begin{aligned}
&(1+k\tau b_{n,n})\|Z_{n-1}\|\\
&\quad \le A_{n+1}\tau\sum_{j=n}^N \big\|U_j - u_d(\cdot,t_j)\big\| + k\tau\Big[A_{n+1}\tau\sum_{j=n+1}^N \big\|U_j - u_d(\cdot,t_j)\big\|\Big]\big(b_{n,n}+Q_0\big)\\
&\quad \le \big(1+k\tau b_{n,n}+Q_0k\tau\big)A_{n+1}\tau\sum_{j=n}^N \big\|U_j - u_d(\cdot,t_j)\big\|.
\end{aligned}
$$

We divide both sides by $(1 + k\tau b_{n,n})$ to conclude that (49) holds for $m = n$ and for all $1 \le m \le N$ by induction. We have proved the second estimate in (48). The first one can be proved in a similar manner. □

6.2 An Optimal-Order Error Estimate

We prove optimal-order error estimates of the schemes (41)–(43). The Ritz projection $\Pi_h : H_0^1 \to S_h$ defined by $(\boldsymbol{K}\nabla(v - \Pi_h v), \nabla\chi) = 0$ for any $\chi \in S_h$ and $v \in H_0^1$ satisfies [36]

$$\left\| v - \Pi_h v \right\| \le Q h^2 \|v\|_{H^2}, \quad \forall v \in H^2 \cap H_0^1. \tag{50}$$

Theorem 8 *Suppose Assumption A holds. Then*

$$|||u - U|||_{L^\infty(L^2)} + |||z - Z|||_{L^\infty(L^2)} + |||c - C|||_{L^\infty(L^2)} \le QM\big(\tau + h^2\big) \tag{51}$$

with $M := \|u_d\|_{H^1(L^2)} + \|u_d(\cdot,0)\|_{\check{H}^1} + \|f\|_{H^1(L^2)} + \|f(\cdot,0)\|_{\check{H}^1} + \|c\|_{L^2(L^2)}$.

Proof Subtract Eq. (36) from scheme (41) and decompose $U_n - u_n = \xi_n + \eta_n$ with $\xi_n = U_n - \Pi_h u_n \in S_h$ and $\eta(\boldsymbol{x},t) := \Pi_h u(\boldsymbol{x},t) - u(\boldsymbol{x},t)$ to obtain

$$\begin{aligned} &\big(\delta_\tau \xi_n + k\delta_\tau^{\alpha(t_n)}\xi_n, \chi\big) + \big(\boldsymbol{K}\nabla\xi_n, \nabla\chi\big) \\ &\quad = -\big(\delta_\tau \eta_n + k\delta_\tau^{\alpha(t_n)}\eta_n, \chi\big) + \big(C_{n-1} - c_n, \chi\big) + \big(kR_n + E_n, \chi\big). \end{aligned} \tag{52}$$

Apply Theorem 7 to Eq. (52) to arrive at the following estimate

$$\begin{aligned} |||\xi|||_{L^\infty(L^2)} \le Q\big(&|||\delta_\tau \eta + k\delta_\tau^{\alpha(t)}\eta|||_{L^1(L^2)} + |||kR + E|||_{L^1(L^2)} \\ &+ |||C - c|||_{L^1(L^2)}\big) + Q\tau \sum_{n=1}^{N} \|c_n - c_{n-1}\|. \end{aligned} \tag{53}$$

Utilize estimates of u and c in 33, (50), and Lemmas 3–5 to get

$$\begin{aligned} &|||\eta|||_{L^\infty(L^2)} \le Qh^2 \|u\|_{L^\infty(H^2)} \le Qh^2 \|u\|_{W^{1,1}(H^2)} \le QMh^2, \\ &|||\delta_\tau^{\alpha(t)}\eta|||_{L^1(L^2)} + |||kR + E|||_{L^1(L^2)} + |||C - c|||_{L^1(L^2)} \le QM\big(\tau + h^2\big), \\ &\tau \sum_{n=1}^{N} \|c_n - c_{n-1}\| \le Q\tau \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\partial_t c\| dt \le Q\tau \|c\|_{H^1(L^2)} \le QM\tau, \\ &|||\delta_\tau \eta|||_{L^1(L^2)} = \sum_{n=1}^{N} \left\| \int_{t_{n-1}}^{t_n} \partial_t \eta dt \right\| \le Q \|\Pi_h \partial_t u - \partial_t u\|_{L^1(L^2)} \le QMh^2. \end{aligned} \tag{54}$$

Combine estimates (53)–(54) and Lemma 5 to get the estimate (51) for $U - u$.

A similar derivation from (39) and (42) yields an equation in terms of $\hat{\xi}_n = Z_n - \Pi_h z_n \in S_h$ and $\hat{\eta}(\boldsymbol{x},t) := \Pi_h z(\boldsymbol{x},t) - z(\boldsymbol{x},t)$, and we apply Theorem 7 to the resulting equation and analyze the truncation errors like (54) and the estimate of $|||U - u|||_{L^1(L^2)}$ to get that for $Z - z$ in (51).

To bound $|||C - c|||_{L^\infty(L^2)}$, we use (16) and (46) to find that

$$|c_{n-1} - C_{n-1}| = |\mathcal{P}(-z_{n-1}/\gamma) - \mathcal{P}(-Z_{n-1}/\gamma)|, \quad 1 \le n \le N. \tag{55}$$

Table 1 Accuracy of the discretization in Sect. 7.1 with $\alpha(0) = 0.2$, $\alpha(T) = 0.8$

τ	1/8	1/16	1/32	1/64	υ
$\lvert\lvert\lvert c - C \rvert\rvert\rvert_{L^\infty(L^2)}$	2.08E−02	9.64E−03	6.35E−03	4.16E−03	1.07
$\lvert\lvert\lvert u - U \rvert\rvert\rvert_{L^\infty(L^2)}$	2.24E−02	1.15E−02	7.58E−03	5.00E−03	1.00
$\lvert\lvert\lvert z - Z \rvert\rvert\rvert_{L^\infty(L^2)}$	2.38E−02	1.21E−02	8.12E−03	5.38E−03	0.99
h	1/60	1/72	1/90	1/120	ι
$\lvert\lvert\lvert c - C \rvert\rvert\rvert_{L^\infty(L^2)}$	4.66E−04	3.71E−04	2.11E−04	1.15E−04	2.08
$\lvert\lvert\lvert u - U \rvert\rvert\rvert_{L^\infty(L^2)}$	1.33E−03	9.27E−04	5.91E−04	3.22E−04	2.04
$\lvert\lvert\lvert z - Z \rvert\rvert\rvert_{L^\infty(L^2)}$	9.59E−04	6.72E−04	4.31E−04	2.35E−04	2.03

If both $-z_{n-1}(\boldsymbol{x})/\gamma, -Z_{n-1}(\boldsymbol{x})/\gamma \in [a, b]$, then $|c_{n-1}(\boldsymbol{x}) - C_{n-1}(\boldsymbol{x})| = |z_{n-1}(\boldsymbol{x}) - Z_{n-1}(\boldsymbol{x})|/\gamma$. Otherwise, say $-z_{n-1}/\gamma \le a$ and $-Z_{n-1}/\gamma \ge b$, we have from (55) $|c_{n-1}(\boldsymbol{x}) - C_{n-1}(\boldsymbol{x})| = b - a \le |z_{n-1}(\boldsymbol{x}) - Z_{n-1}(\boldsymbol{x})|/\gamma$. We similarly bound $|c_{n-1}(\boldsymbol{x}) - C_{n-1}(\boldsymbol{x})|$ by $|z_{n-1}(\boldsymbol{x}) - Z_{n-1}(\boldsymbol{x})|/\gamma$ for other cases, which gives rise to the estimate $\lvert\lvert\lvert C - c \rvert\rvert\rvert_{L^\infty(L^2)} \le Q \lvert\lvert\lvert Z - z \rvert\rvert\rvert_{L^\infty(L^2)} \le QM(\tau + h^2)$. □

7 Numerical Experiments

We study the performance of the discretization schemes (41)–(43) by several numerical experiments.

7.1 One Dimensional Test

The data are as follows: $\Omega = (0, 1)$, $[0, T] = [0, 1]$, $k = 1$, $K = 0.01$, $a = 0.2$, $b = 0.3$, $\gamma = 1$, $f = 1$, $u_d = 1$, and $\alpha(t) = \alpha(T) + (\alpha(0) - \alpha(T))\big(1 - t/T - \sin\big(2\pi(1 - t/T)\big)/(2\pi)\big)$. As the closed-form analytical solution to the problem is not available, we use the numerical solution computed with $\tau_f = 1/720$ and $h_f = 1/360$ as the reference solution. We compute the spatial convergence rate ι measured in the norms of $\lvert\lvert\lvert c - C \rvert\rvert\rvert_{L^\infty(L^2)}$, $\lvert\lvert\lvert u - U \rvert\rvert\rvert_{L^\infty(L^2)}$ and $\lvert\lvert\lvert z - Z \rvert\rvert\rvert_{L^\infty(L^2)}$, in which we choose a fine time step of $\tau_f = 1/720$. We similarly compute the temporal convergence rate υ measured in the norms of $\lvert\lvert\lvert c - C \rvert\rvert\rvert_{L^\infty(L^2)}$, $\lvert\lvert\lvert u - U \rvert\rvert\rvert_{L^\infty(L^2)}$ and $\lvert\lvert\lvert z - Z \rvert\rvert\rvert_{L^\infty(L^2)}$, in which we choose a fine space mesh size of $h_f = 1/360$. We present the numerical results in Tables 1 and 2 and observe the second-order accuracy in space and first-order convergence in time as proved in Theorem 8.

7.2 Two Dimensional Test

The data are as follows: $\Omega = (0, 1)^2$, $[0, T] = [0, 1]$, $k = 1$, $\boldsymbol{K} = \mathrm{diag}(0.01, 0.01)$, $\alpha(t) = \alpha(T) + (\alpha(0) - \alpha(T))(1 - t/T)$. The solution is chosen to be $u(\boldsymbol{x}, t) = t^{2-\alpha(0)} \sin(\pi x_1) \sin(\pi x_2)$, $z(\boldsymbol{x}, t) = (T - t)^{2-\alpha(T)} \sin(\pi x_1) \sin(\pi x_2)$, $c(\boldsymbol{x}, t) = \max\{a, \min\{-z(\boldsymbol{x}, t)/\gamma, b\}\}$ with $a = -0.2$ and $b = -0.1$, $\gamma = 1$, and f and u_d calculated accordingly. A mesh size of $h = 1/64$ is used to measure the temporal convergence rate, while $\tau = 2h^2$ is used to measure the spatial convergence rate. We present the numerical

Table 2 Accuracy of the discretization in Sect. 7.1 with $\alpha(0) = 0.9$, $\alpha(T) = 0.7$

τ	1/8	1/16	1/32	1/64	υ
$\|\|\|c - C\|\|\|_{L^\infty(L^2)}$	1.66E−02	8.83E−03	6.36E−03	4.15E−03	0.91
$\|\|\|u - U\|\|\|_{L^\infty(L^2)}$	7.48E−03	3.96E−03	2.71E−03	1.83E−03	0.93
$\|\|\|z - Z\|\|\|_{L^\infty(L^2)}$	2.27E−02	1.17E−02	7.77E−03	5.13E−03	0.99
h	1/60	1/72	1/90	1/120	ι
$\|\|\|c - C\|\|\|_{L^\infty(L^2)}$	4.61E−04	2.67E−04	2.12E−04	9.58E−05	2.15
$\|\|\|u - U\|\|\|_{L^\infty(L^2)}$	1.27E−03	8.88E−04	5.67E−04	3.09E−04	2.04
$\|\|\|z - Z\|\|\|_{L^\infty(L^2)}$	9.38E−04	6.60E−04	4.23E−04	2.31E−04	2.02

Table 3 Accuracy of the discretization in Sect. 7.2 with $\alpha(0) = 0.3$, $\alpha(T) = 0.6$

τ	1/80	1/96	1/120	1/144	υ
$\|\|\|c - C\|\|\|_{L^\infty(L^2)}$	1.81E−03	1.52E−03	1.22E−03	1.03E−03	0.96
$\|\|\|u - U\|\|\|_{L^\infty(L^2)}$	2.47E−03	2.06E−03	1.65E−03	1.38E−03	1.00
$\|\|\|z - Z\|\|\|_{L^\infty(L^2)}$	2.89E−03	2.42E−03	1.94E−03	1.62E−03	0.98
h	1/60	1/72	1/90	1/120	ι
$\|\|\|c - C\|\|\|_{L^\infty(L^2)}$	1.15E−03	5.15E−04	2.97E−04	1.92E−04	1.95
$\|\|\|u - U\|\|\|_{L^\infty(L^2)}$	1.48E−03	6.61E−04	3.72E−04	2.38E−04	2.00
$\|\|\|z - Z\|\|\|_{L^\infty(L^2)}$	1.74E−03	7.87E−04	4.46E−04	2.87E−04	1.97

Table 4 Accuracy of the discretization in Sect. 7.2 with $\alpha(0) = 0$, $\alpha(T) = 0.5$

τ	1/80	1/96	1/120	1/144	υ
$\|\|\|c - C\|\|\|_{L^\infty(L^2)}$	2.17E−03	1.80E−03	1.46E−03	1.23E−03	0.96
$\|\|\|u - U\|\|\|_{L^\infty(L^2)}$	3.23E−03	2.69E−03	2.15E−03	1.79E−03	1.00
$\|\|\|z - Z\|\|\|_{L^\infty(L^2)}$	3.41E−03	2.84E−03	2.28E−03	1.90E−03	0.99
h	1/60	1/72	1/90	1/120	ι
$\|\|\|c - C\|\|\|_{L^\infty(L^2)}$	1.37E−03	6.10E−04	3.49E−04	2.25E−04	1.97
$\|\|\|u - U\|\|\|_{L^\infty(L^2)}$	1.95E−03	8.68E−04	4.89E−04	3.13E−04	2.00
$\|\|\|z - Z\|\|\|_{L^\infty(L^2)}$	2.06E−03	9.23E−04	5.21E−04	3.34E−04	1.98

results in Tables 3 and 4 and again observe the second-order accuracy in space and first-order convergence in time as proved in Theorem 8.

Acknowledgements The authors would like to express their most sincere thanks to the referees for their very constructive and helpful comments and suggestions, which greatly improved the quality of this paper.

Data Availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors declare no conflict of interest.

8 Appendix

We prove Lemmas 3–5 used in the proof of Theorem 8.

Lemma 3 *Under Assumption A, E_n, R_n, $\hat{E}_n$, $\hat{F}_n$, $\hat{R}_n$ in (35)–(38) satisfy*

$$|||E + kR|||_{L^1(L^2)} + |||\hat{E} + k(\hat{F} + \hat{R})|||_{L^1(L^2)} \leq QM\tau. \tag{56}$$

Here M and $|||\cdot|||_{L^1(L^2)}$ were given in (51) and (48), and Q is independent of N.

Proof Use estimates (33) to bound $\hat{E}_{k-1}$ by $|||\hat{E}|||_{L^1(L^2)} \leq \tau\|\partial_t^2 z\|_{L^1(L^2)} \leq QM\tau$. Use (13) and Taylor expansion to similarly bound $\hat{F}$

$$\begin{aligned}\|\hat{F}_{k-1}\| &= \Big\|\hat{\partial}_t^{\alpha(t)} z_{k-1} - \frac{1}{\tau}\big({}_t\hat{I}_T^{1-\alpha(t)} z(\boldsymbol{x}, t)\big)\Big|_{t=t_k}^{t=t_{k-1}}\Big\| \\ &= \Big\|\frac{1}{\tau}\int_{t_{k-1}}^{t_k} \partial_t^2({}_t\hat{I}_T^{1-\alpha(t)} z)(t_k - t)dt\Big\| \leq \big\|\partial_t^2({}_t\hat{I}_T^{1-\alpha(t)} z)\big\|_{L^1(t_{k-1},t_k;L^2)},\end{aligned}$$

yielding $|||\hat{F}|||_{L^1(L^2)} \leq Q\tau\big\|\partial_t^2({}_t\hat{I}_T^{1-\alpha(t)} z)\big\|_{L^1(L^2)} \leq QM\tau$ by (33). Split $\hat{R}_{k-1}$ as

$$\begin{aligned}\hat{R}_{k-1} &= \frac{1}{\tau}\sum_{m=k}^{N-1}\bigg[\int_{t_{m-1}}^{t_m} \frac{z(\boldsymbol{x}, s)(s - t_{k-1})^{-\alpha(s)}}{\Gamma(1 - \alpha(s))} - \frac{z_{m-1}(s - t_{k-1})^{-\alpha(t_{m-1})}}{\Gamma(1 - \alpha(t_{m-1}))}ds \\ &\quad + \int_{t_{m-1}}^{t_m} z_{m-1}\Big(\frac{(s - t_{k-1})^{-\alpha(t_{m-1})}}{\Gamma(1 - \alpha(t_{m-1}))} - \frac{(s - t_{k-1})^{-\alpha(t_m)}}{\Gamma(1 - \alpha(t_m))}\Big)ds\bigg] \\ &\quad - \frac{1}{\tau}\sum_{m=k}^{N-1}\bigg[\int_{t_{m-1}}^{t_m} \frac{z(\boldsymbol{x}, s + \tau)(s - t_{k-1})^{-\alpha(s+\tau)}}{\Gamma(1 - \alpha(s + \tau))} - \frac{z_m(s - t_{k-1})^{-\alpha(t_m)}}{\Gamma(1 - \alpha(t_m))}ds \\ &\quad + \int_{t_{m-1}}^{t_m} z_m\Big(\frac{(s - t_{k-1})^{-\alpha(t_m)}}{\Gamma(1 - \alpha(t_m))} - \frac{(s - t_{k-1})^{-\alpha(t_{m+1})}}{\Gamma(1 - \alpha(t_{m+1}))}\Big)ds\bigg] \\ &\quad + \frac{1}{\tau}\int_{t_{N-1}}^{t_N} \frac{z(\boldsymbol{x}, s)(s - t_{k-1})^{-\alpha(s)}}{\Gamma(1 - \alpha(s))} - \frac{z_{N-1}(s - t_{k-1})^{-\alpha(t_N)}}{\Gamma(1 - \alpha(t_N))}ds \\ &=: (J_1^k + J_2^k) - (J_3^k + J_4^k) + J_5^k.\end{aligned}$$

$$\begin{aligned}\big|J_1^k - J_3^k\big| &= \Big|\frac{1}{\tau}\sum_{m=k}^{N-1}\int_{t_{m-1}}^{t_m}\Big(\int_{t_{m-1}}^{s} - \int_{t_m}^{s+\tau}\Big)\partial_y\Big(\frac{z(\boldsymbol{x}, y)(s - t_{k-1})^{-\alpha(y)}}{\Gamma(1 - \alpha(y))}\Big)dyds\Big| \\ &= \Big|\frac{1}{\tau}\sum_{m=k}^{N-1}\int_{t_{m-1}}^{t_m}\int_{t_{m-1}}^{s}\int_{y}^{y+\tau}\partial_\theta^2\Big(\frac{z(\boldsymbol{x}, \theta)(s - t_{k-1})^{-\alpha(\theta)}}{\Gamma(1 - \alpha(\theta))}\Big)d\theta dyds\Big|.\end{aligned}$$

Direct calculations show that for $0 < \varepsilon < 1 - \alpha^*$

$$\Big|\partial_\theta^2\Big(\frac{z(\boldsymbol{x}, \theta)(s - t_{k-1})^{-\alpha(\theta)}}{\Gamma(1 - \alpha(\theta))}\Big)\Big| \leq \frac{Q\big(|z| + |\partial_\theta z| + |\partial_\theta^2 z|\big)}{(s - t_{k-1})^{\alpha^*+\varepsilon}}, \tag{57}$$

which implies

$$|J_1^k - J_3^k| \leq Q\sum_{m=k}^{N-1}\int_{t_{m-1}}^{t_{m+1}}\big(|z| + |\partial_\theta z| + |\partial_\theta^2 z|\big)d\theta\big((t_m - t_{k-1})^{\alpha^*+\varepsilon} - (t_m - t_k)^{\alpha^*+\varepsilon}\big).$$

Apply (33) to obtain

$$\tau \sum_{k=1}^{N} \|J_1^k - J_3^k\| \le Q\tau \sum_{m=1}^{N-1} \int_{t_{m-1}}^{t_{m+1}} \big(\|z\| + \|\partial_\theta z\| + \|\partial_\theta^2 z\|\big) d\theta$$
$$\times \sum_{k=1}^{m} \big((t_m - t_{k-1})^{\alpha^*+\varepsilon} - (t_m - t_k)^{\alpha^*+\varepsilon}\big) \le Q\tau \|z\|_{W^{2,1}(L^2)} \le QM\tau. \tag{58}$$

A similar derivation as (57)–(58) gives $\tau \sum_{k=1}^{N} \|J_2^k - J_4^k\| \le QM\tau$, and $\tau \sum_{k=1}^{N} \|J_5^k\|$ bounded as $\tau \sum_{k=1}^{N} \|J_1^k - J_3^k\|$, completing the estimate of $\|\hat{R}\|_{L^1(L^2)}$. The rest terms in (56) are bounded similarly and the proofs are omitted. □

Lemma 4 *Under Assumption A, the following estimate holds*

$$\tau \sum_{k=1}^{N} \big(\|\delta_\tau^{\alpha(t_k)} \eta_k\| + \|\hat{\delta}_\tau^{\alpha(t_{k-1})} \hat{\eta}_{k-1}\|\big) \le QMh^2. \tag{59}$$

Proof We use the expression of $\hat{\delta}_\tau^{\alpha(t_{k-1})}$ in (38) and $\hat{\eta}_N = 0$ to obtain

$$\hat{\delta}_\tau^{\alpha(t_{k-1})} \hat{\eta}_{k-1} = \sum_{m=k}^{N} b_{m,k}(\hat{\eta}_{m-1} - \hat{\eta}_m) + \sum_{m=k}^{N-1} (b_{m,k} - b_{m+1,k+1})\hat{\eta}_m =: \sum_{i=1}^{2} I_i^k.$$

Interchange the order of summation to get

$$\tau \sum_{k=1}^{N} \|I_1^k\| \le \textstyle\sum_{k=1}^{N} \sum_{m=k}^{N} \int_{t_{m-1}}^{t_m} \|\partial_t \hat{\eta}\| dt \int_{t_{k-1}}^{t_k} (t_m - s)^{-\alpha^*} ds$$
$$\le Qh^2 \sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|\partial_t z\|_{H^2} dt \sum_{k=1}^{m} \big[(t_m - t_{k-1})^{1-\alpha(t_{k-1})} - (t_m - t_k)^{1-\alpha(t_{k-1})}\big]$$
$$\le Qh^2 \|z\|_{W^{1,1}(H^2)} \le QMh^2.$$

Use the mean-value theorem as (47) to find that $|b_{m,k} - b_{m+1,k+1}| \le Q \int_{t_{k-1}}^{t_k} (t_m - s)^{-(\alpha^*+\varepsilon)} ds$. Combine this estimate with (33) and $z(\boldsymbol{x}, T) = 0$ to get

$$\tau \sum_{k=1}^{N} \|I_2^k\| \le Qh^2 \tau \|z\|_{L^\infty(H^2)} \sum_{k=1}^{N} \sum_{m=k}^{N-1} \int_{t_{k-1}}^{t_k} \frac{1}{(t_m - s)^{\alpha^*+\varepsilon}} ds \le Qh^2 \|z\|_{W^{1,1}(H^2)}.$$

Use the preceding estimates to bound the second term on the left-hand side of (59). The first term can be estimated similarly and is omitted. □

Lemma 5 *Under Assumption A,* $\|C - c\|_{L^2(L^2)} \le QM(\tau + h^2)$.

Proof Let $U(c)$ satisfy $U_0(c) = 0$ and (41) with C_{n-1} replaced by c_{n-1} and $Z(U(c))$ satisfy $Z_N(U(c)) = 0$ and (42) with U_n replaced by $U_n(c)$. Then an application of $(\gamma c_{n-1} + z_{n-1}, c_{n-1} - C_{n-1}) \le 0$ leads to

$$\gamma |||c - C|||^2_{L^2(L^2)} \le \tau \sum_{n=1}^{N} \Big[\big(Z_{n-1} - Z_{n-1}(U(c)), c_{n-1} - C_{n-1} \big) + \big(Z_{n-1}(U(c)) - z_{n-1}, c_{n-1} - C_{n-1} \big) \Big]. \tag{60}$$

We incorporate $\sum_{n=1}^{N} \Big[\big(Z_{n-1} - Z_{n-1}(U(c)), c_{n-1} - C_{n-1} \big) = \sum_{n=1}^{N} \big(U_n(c) - U_n, U_n - U_n(c) \big) \le$ 0 with (60) and use Cauchy inequality to conclude that

$$\gamma |||c - C|||_{L^2(L^2)} \le |||Z(U(c)) - z|||_{L^2(L^2)} \le |||z - Z(u)||| + |||Z(u) - Z(U(c))|||. \tag{61}$$

Here $Z(u)$ satisfies $Z_N(u) = 0$ and (42) with U_n replaced by u_n. Split $Z_n(u) - z_n = \check{\xi}_n + \hat{\eta}_n$ for $n = 0, 1, \ldots, N-1$ with $\check{\xi}_n = Z_n(u) - \Pi_h z_n \in S_h$ and $\hat{\eta}(\boldsymbol{x}, t) := \Pi_h z(\boldsymbol{x}, t) - z(\boldsymbol{x}, t)$, which leads to

$$\begin{aligned} &\big(-\delta_\tau \check{\xi}_n + k \hat{\delta}_\tau^{\alpha(t_{n-1})} \check{\xi}_{n-1}, \chi \big) + \big(\boldsymbol{K} \nabla \check{\xi}_{n-1}, \nabla \chi \big) \\ &= -\big(-\delta_\tau \hat{\eta}_n + k \hat{\delta}_\tau^{\alpha(t_{n-1})} \hat{\eta}_{n-1}, \chi \big) + \big(u_n - u_{n-1}, \chi \big) + \big(u_d(\cdot, t_{n-1}) - u_d(\cdot, t_n), \chi \big) \\ &\quad + \big(k(\hat{F}_{n-1} + \hat{R}_{n-1}) + \hat{E}_{n-1}, \chi \big), \quad \forall \chi \in S_h, \quad n = N, N-1, \ldots, 1. \end{aligned}$$

We apply Theorem 7 to the equation and employ similar estimates as (54) and Lemmas 3–4 to get $|||z - Z(u)|||_{L^\infty(L^2)} \le QM(\tau + h^2)$. We estimate the second right-hand side term in (61) by Theorem 7 to get $|||Z(u) - Z(U(c))|||_{L^\infty(L^2)} \le Q|||u - U(c)|||_{L^1(L^2)}$, which completes the proof. □

References

1. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Elsevier, San Diego (2003)
2. Akrivis, G., Li, B., Lubich, C.: Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. Math. Comp. **86**, 1527–1552 (2017)
3. Antil, H., Otarola, E., Salgado, A.: A space–time fractional optimal control problem: analysis and discretization. SIAM J. Control. Optim. **54**, 1295–1328 (2016)
4. Athans, M., Falb, P.L.: Optimal Control: An Introduction to the Theory and its Applications. Dover Publications, Mineola (2006)
5. Baveye, P., Vandevivere, P., Hoyle, B.L., DeLeo, P.C., de Lozada, D.S.: Environmental impact and mechanisms of the biological clogging of saturated soils and aquifer materials. Crit. Rev. Environ. Sci. Technol. **28**, 123–191 (2006)
6. Chen, S., Shen, J., Zhang, Z., Zhou, Z.: A spectrally accurate approximation to subdiffusion equations using the log orthogonal functions. SIAM J. Sci. Comput. **42**, A849–A877 (2020)
7. D'Elia, M., Glusa, C., Otarola, E.: A priori error estimates for the optimal control of the integral fractional Laplacian. SIAM J. Control Optim. **57**, 2775–2798 (2019)
8. Deng, W., Hesthaven, J.: Local discontinuous Galerkin methods for fractional diffusion equations. ESAIM: M2AN **47**(6), 1845–1864
9. Diethelm, K., Ford, N.: Analysis of fractional differential equations. J. Math. Anal. Appl. **265**, 229–248 (2002)
10. Du, Q., Toniazzi, L., Zhou, Z.: Stochastic representation of solution to nonlocal-in-time diffusion. Stoch. Process. Appl. **130**(4), 2058–2085
11. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, V 19. American Mathematical Society, Rhode Island (1998)
12. Gandossi, L., Von Estorff, U.: An overview of hydraulic fracturing and other formation stimulation technologies for shale gas production. In: Scientific and Technical Research Reports, Joint Research Centre of the European Commission. Publications Office of the European Union (2015). https://doi.org/10.2790/379646
13. Gunzburger, M., Wang, J.: Error analysis of fully discrete finite element approximations to an optimal control problem governed by a time-fractional PDE. SIAM J. Control Optim. **57**, 241–263 (2019)

14. Gunzburger, M., Kunoth, A.: Space–time adaptive wavelet methods for control problems constrained by parabolic evolution equations. SIAM J. Control Optim. **49**, 1150–1170 (2011)
15. Herzog, R., Stadler, G., Wachsmuth, G.: Directional sparsity in optimal control of partial differential equations. SIAM J. Control Optim. **50**, 943–963 (2012)
16. Hinze, M.: A variational discretization concept in control constrained optimization: the linear quadratic case. J. Comput. Optim. Appl. **30**, 45–63 (2005)
17. Jin, B., Li, B., Zhou, Z.: Correction of high-order BDF convolution quadrature for fractional evolution equations. SIAM J. Sci. Comput. **39**, A3129–A3152 (2017)
18. Jin, B., Li, B., Zhou, Z.: Pointwise-in-time error estimates for an optimal control problem with subdiffusion constraint. IMA J. Numer. Anal. **40**, 377–404 (2020)
19. Kunisch, K., Vexler, B.: Constrained Dirichlet boundary control in L^2 for a class of evolution equations. SIAM J. Control Optim. **46**, 1726–1753 (2007)
20. Kunoth, A., Schwab, C.: Analytic regularity and GPC approximation for control problems constrained by linear parametric elliptic and parabolic PDEs. SIAM J. Control Optim. **51**, 2442–2471 (2013)
21. Liao, H., Tang, T., Zhou, T.: A second-order and nonuniform time-stepping maximum-principle preserving scheme for time-fractional Allen–Cahn equations. J. Comput. Phys. **414**, 109473 (2020)
22. Li, B., Wang, H., Wang, J.: Well-posedness and numerical approximation of a fractional diffusion equation with a nonlinear variable order. ESAIM: M2AN **55**, 171–207 (2021)
23. Lorenzo, C.F., Hartley, T.T.: Variable order and distributed order fractional operators. Nonlinear Dyn. **29**, 57–98 (2002)
24. Lubich, C., Sloan, T., Thomée, V.: Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term. Math. Comput. **65**, 1–17 (1996)
25. Meerschaert, M.M., Sikorskii, A.: Stochastic Models for Fractional Calculus. De Gruyter Studies in Mathematics (2011)
26. Meidner, D., Vexler, B.: A priori error estimates for space–time finite element discretization of parabolic optimal control problems. Part I: problems without control constraints. SIAM J. Control Optim. **47**, 1150–1177 (2008)
27. Meidner, D., Vexler, B.: A priori error estimates for space–time finite element discretization of parabolic optimal control problems. Part II: problems with control constraints. SIAM J. Control Optim. **47**, 1301–1329 (2008)
28. Metzler, R., Klafter, J.: The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys. Rep. **339**, 1–77 (2000)
29. Mophou, G.: Optimal control of fractional diffusion equation. Comput. Math. Appl. **61**, 68–78 (2011)
30. Neittaanmaki, P., Tiba, D., Dekker, M.: Optimal Control of Nonlinear Parabolic Systems: Theory. Algorithms and Applications. Marcel Dekker, New York (1994)
31. Odzijewicz, T., Malinowska, A., Torres, D.: Fractional variational calculus of variable order. In: Almeida, A., Castro, L., Speck, F.O. (eds.) Advances in Harmonic Analysis and Operator Theory. Operator Theory: Advances and Applications, vol. 229. Birkhäuser, Basel (2013)
32. Podlubny, I.: Fractional Differential Equations. Academic Press, New York (1999)
33. Shen, J., Sheng, C.: An efficient space–time method for time fractional diffusion equation. J. Sci. Comput. **81**, 1088–1110 (2019)
34. Stynes, M., O'Riordan, E., Gracia, J.L.: Error analysis of a finite difference method on graded mesh for a time-fractional diffusion equation. SIAM J.. Numer. Anal. **55**, 1057–1079 (2017)
35. Sweilam, N., Al-Ajami, T., Hoppe, R.: Numerical solution of some types of fractional optimal control problems. Sci. World J. **2013**, 306237 (2013)
36. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems. Lecture Notes in Mathematics, vol. 1054. Springer, New York (1984)
37. Troltzsch, F.: Optimal Control of Partial Differential Equations: Theory, Methods and Applications. Graduate Studies in Mathematics. AMS, Ann Arbor (2010)
38. Wang, H., Zheng, X.: Wellposedness and regularity of the variable-order time-fractional diffusion equations. J. Math. Anal. Appl. **475**, 1778–1802 (2019)
39. Zeng, F., Zhang, Z., Karniadakis, G.: A generalized spectral collocation method with tunable accuracy for variable-order fractional differential equations. SIAM J. Sci. Comput. **37**, A2710–A2732 (2015)
40. Zheng, X., Wang, H.: Optimal-order error estimates of finite element approximations to variable-order time-fractional diffusion equations without regularity assumptions of the true solutions. IMA J. Numer. Anal. **41**, 1522–1545 (2021)
41. Zheng, X., Wang, H.: A time-fractional partial differential equation with a space–time dependent hidden-memory variable order: analysis and approximation. BIT Numer. Math. (2021). https://doi.org/10.1007/s10543-021-00861-4

42. Zheng, X., Wang, H.: Wellposedness and smoothing properties of history-state based variable-order time-fractional diffusion equations. Zeitschrift für Angewandte Mathematik und Physik (ZAMP) **71**, 34 (2020)
43. Zheng, X., Wang, H.: A hidden-memory variable-order fractional optimal control model: analysis and approximation. SIAM J. Control Optim. **59**, 1851–1880 (2021)
44. Zhou, Z., Gong, W.: Finite element approximation of optimal control problems governed by time fractional diffusion equation. Comput. Math. Appl. **71**, 301–318 (2016)