Received: 28 November 2023 | Revised: 4 April 2024 | Accepted: 17 July 2024

DOI: 10.1002/num.23134

RESEARCH ARTICLE

WILEY

Optimal control of variably distributed-order time-fractional diffusion equation: Analysis and computation

Xiangcheng Zheng[1] | **Huan Liu**[2] | **Hong Wang**[3] | **Xu Guo**[4]

[1]School of Mathematics, Shandong University, Jinan, China
[2]School of Mathematics, Shanghai University of Finance and Economics, Shanghai, China
[3]Department of Mathematics, University of South Carolina, Columbia, South Carolina, USA
[4]Geotechnical and Structural Engineering Research Center, Shandong University, Jinan, China

Correspondence
Huan Liu, School of Mathematics, Shanghai University of Finance and Economics, Shanghai 200433, China.
Email: liuhuan@sufe.edu.cn
Hong Wang, Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA.
Email: hwang@math.sc.edu

Funding information
Taishan Scholars Program of Shandong Province of China, Grant/Award Numbers: tsqn201909044, tsqn202306083; National Natural Science Foundation of China, Grant/Award Numbers: 12001337, 12301555, 12271303; National Science Foundation, Grant/Award Numbers: DMS-2012291, DMS-2245097; Major Fundamental Research Project of Shandong Province of China, Grant/Award Number: ZR2023ZD33; Fundamental Research Funds for the Central Universities, Grant/Award Number: 2022110089.

Abstract

Fractional diffusion equations exhibit competitive capabilities in modeling many challenging phenomena such as the anomalously diffusive transport and memory effects. We prove the well-posedness and regularity of an optimal control of a variably distributed-order fractional diffusion equation with pointwise constraints, where the distributed-order operator accounts for, for example, the effect of uncertainties. We accordingly develop and analyze a fully-discretized finite element approximation to the optimal control without any artificial regularity assumption of the true solution. Numerical experiments are also performed to substantiate the theoretical findings.

KEYWORDS

error estimate, finite element method, optimal control, regularity, variably distributed-order fractional diffusion equation, well-posedness

1 | INTRODUCTION

Optimal control governed by classical integer-order partial differential equations (PDEs) have been widely used in engineering design, manufacturing, biology and medicine, and a variety of other applications. Its mathematical analysis and numerical approximation have been well studied [6, 13, 23, 28, 43, 47]. In applications such as solute transport in porous medium and additive manufacturing, the integer-order diffusion PDEs, derived assuming the existence of a mean free path and a mean waiting time of the underlying particle movements [44], accurately model the diffusive transport of solute in homogeneous media when the solute plumes in field applications were observed to exhibit Gaussian exponential decays [8].

The diffusive transport of solutes in heterogeneous media exhibits power-law decays that indicates why integer-order diffusion PDEs tend to yield less accurate approximations [9, 35, 42, 44]. In contrast, fractional PDEs derived via continuous time random walk under the assumption that their solutions (i.e., the probability density function of the underlying particle jumps) have power-law decays [36, 42, 44], can accurately model the anomalously diffusive transport in heterogeneous porous materials. In recent years, optimal control [5, 6, 12, 15, 25, 32, 46] of fractional PDEs has attracted intensive investigations [4, 14, 17, 22, 26, 29, 45, 48, 55, 56, 59].

A pointwise optimal control problem of a space-time fractional PDE, in which a spectral fractional Laplacian operator was adopted to model the superdiffusive transport while a Caputo time fractional differential operator was used to model the subdiffusive transport, was analyzed rigorously [4]. The regularities of the solution in some weighted Sobolev spaces were proved and the Caffarelli-Silvestre extension [10] and the L1 discretization [37, 47] were employed for numerical computation. A pointwise optimal control of a time-fractional PDE was analyzed, in which a discrete-in-space but continuous-in-time finite element method (FEM) approximation was considered [59]. Based on the proved regularity of the solutions, the error estimate of the numerical approximations was proved, and the stability and truncation error of the fully-discretized FEM were analyzed. The pointwise-in-time error estimate of a fully-discretized FEM approximation to the subdiffusion optimal control problem was proved with the regularity of the solutions proved via resolvent estimates [29], in which both the L1 temporal discretization and the backward Euler convolution quadrature were analyzed. The well-posedness and smoothing properties of the solutions to the optimal control problem governed by a different type of time-fractional PDE were proved, and an optimal-order error estimate of a fully discretized FEM approximation with the convolution quadrature discretization in time was proved [26].

The motivation of this paper is as follows: The time-fractional PDE (tFDE) yields solutions with nonphysical initial weak singularity [40, 49, 51], because it was derived as a stochastic limit when the number of particle jumps tends to infinity and hence holds only for large time [44]. Recently, a two time-scale mobile-immobile tFDE, which contains an additional $\partial_t u$ term with a partition coefficient k in front of the time-fractional derivative term, was presented [50]. The model describes dynamic mass exchange between the $1/(1+k)$ portion of the solute mass in the mobile phase and the $k/(1+k)$ portion of the solute mass in the immobile phase absorbed to the porous materials, exhibits the Fickian diffusion behavior initially and naturally switches to the subdiffusive transport behavior as time evolves, and so is valid on the entire interval including the initial time without the initial weak singularity of the conventional tFDE [53].

Furthermore, the order of fractional PDEs is related to the fractal dimension of porous materials via the Hurst index [42]. For highly heterogeneous porous materials, a scalar Hurst index does not necessarily suffice to quantify its fractional dimension. The distributed-order fractional differential operator [11, 24, 38]

$$\mathbb{D}_t^{\nu} g := \int_0^1 \nu(\alpha)\, \partial_t^{\alpha} g d\alpha, \quad \partial_t^{\alpha} g :={}_0 I_t^{1-\alpha} \partial_t g, \quad {}_0I_t^{\alpha} g := \int_0^t \frac{g(s)}{\Gamma(\alpha)(t-s)^{1-\alpha}} ds, \tag{1}$$

with $\Gamma(\alpha)$ being the Gamma function, was introduced to integrate the accumulated impact of a wide spectrum of fractional differential operators modeling complex phenomena. Moreover, in subsurface solute transport and hydrocarbon or gas recovery [7, 20], the geological information is limited and often polluted with noises and an accurate determination of a scalar fractional order is impossible [19]. The distributed-order fractional differential operator provides a feasible approach to quantify the uncertainties. Distributed-order FDEs have attracted extensive research on their application, analysis and discretization [16, 21, 30, 34, 40, 41]. In applications such as bioclogging [7] and nonconventional gas and oil recovery [20], the structure of porous materials may evolve in time so the probability density function $\nu(\alpha)$ in (1) may depend on the time t too. All these factors lead to the two time-scale variably distributed-order FDE [53]

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

Here $\Omega \subset \mathbb{R}^d$, $d = 1,2,3$, is a simply-connected bounded domain with the piecewise smooth boundary $\partial\Omega$ and convex corners, $\mathbb{D}_t^{\omega} g := \int_0^1 \omega(\alpha,t)\, \partial_t^{\alpha} g d\alpha$, $\boldsymbol{x} := (x_1, \ldots, x_d)$, $\mathcal{B} := -\nabla \cdot (\boldsymbol{K}(\boldsymbol{x})\nabla)$ with $\nabla := (\partial/\partial x_1, \ldots, \partial/\partial x_d)^{\top}$ and $\boldsymbol{K}(\boldsymbol{x}) := (k_{ij}(\boldsymbol{x}))_{i,j=1}^d$ is the symmetric diffusivity tensor. The partition coefficient k is a positive constant, f and c denote the source and the control variable, respectively.

Let u_d be the target function, in this paper we discretize and analyze the optimal control

$$\min_{c\in\Lambda(l,r)} J(u,c) = \frac{1}{2}\|u-u_d\|^2_{L^2(L^2(\Omega))} + \frac{\gamma}{2}\|c\|^2_{L^2(L^2(\Omega))}, \tag{3}$$

which is governed by the variably distributed-order tFDE (2). For $l \le r$, the admissible set $\Lambda(l,r)$ is defined by

$$\Lambda(l,r) := \{c \in L^2(0,T;L^2(\Omega)) : l \le c(\boldsymbol{x},t) \le r \text{ a.e. in } \Omega \times [0,T]\}. \tag{4}$$

Let $L^p(\Omega)$ and $W^{m,p}(\Omega)$, $1 \le p \le \infty$, $m \in \mathbb{N}$, be the spaces of p-th Lebesgue integrable functions and those with p-th Lebesgue integrable derivatives up to order m in Ω, respectively, and $H^m(\Omega) := W^{m,2}(\Omega)$. For $s > 0$ the fractional Sobolev space $H^s(\Omega)$ is defined by interpolation [1]. All the spaces can also be defined on an interval. Given a Banach space S, $L^p(a,b;S)$ and $H^1(a,b;S)$ consist of functions f such that $\|f\|_S$ and $\|\partial_t f\|_S$ in $L^p(a,b)$ and $L^2(a,b)$, respectively [1, 18]. All the spaces are equipped with the standard norms [1, 18]. For $g \in H^1(0,T;L^2(\Omega))$, the following estimates hold for the cut-off projection [31]

$$\begin{aligned} &\mathcal{P}g := \max\{l, \min\{g,r\}\} \in H^1(0,T;L^2(\Omega)), \\ &\|\mathcal{P}g\|_{H^1(0,T;L^2(\Omega))} \le \|g\|_{H^1(0,T;L^2(\Omega))}, \\ &\|\mathcal{P}g\|_{H^{\mu}(\Omega)} \le \|g\|_{H^{\mu}(\Omega)}, \quad \forall g \in H^{\mu}(\Omega),\ 0 \le \mu \le 1. \end{aligned} \tag{5}$$

We make the following *assumptions* throughout the paper.

- **(a)** $\omega(\alpha,t) \ge 0$ on $[0,1]\times[0,T]$ satisfies $\int_0^1 \omega(\alpha,t)d\alpha = 1$ for any $t \in [0,T]$. There exist some $0 < \alpha^* < 1$ and $Q_0 > 0$ such that supp $\omega \subset [0,\alpha^*]\times[0,T]$, and $|\partial_\alpha^m \omega|, |\partial_t^m \omega|, |\partial_\alpha^2 \partial_t \omega| \le Q_0$ on supp ω for $m = 0,1,2$.
- **(b)** $0 < K_* \le \boldsymbol{\xi}^T \boldsymbol{K}\boldsymbol{\xi} \le K^* < \infty$, $\boldsymbol{\xi} \in \mathbb{R}^d, |\boldsymbol{\xi}| = 1$, $k_{ij} \in C^1(\overline{\Omega})$, $1 \le i,j \le d$.
- **(c)** $f, u_d \in H^1(0,T;L^2(\Omega)) \cap L^2(0,T;\check{H}^2(\Omega))$.

To date, the optimal control of distributed-order tFDEs of the form (1) was analyzed assuming that the exact solution is sufficiently smooth [54]. A numerical approximation to a forward variably distributed-order tFDE with a space-dependent probability density function was analyzed assuming the exact solution is sufficiently smooth in both space and time [3]. However, the fact that tFDEs exhibit initial weak singularity [40, 49, 51] does not seem to justify these assumptions. The well-posedness of the variably distributed-order tFDE (2) and its smoothing properties were proved and based on which a numerical approximation was analyzed [53].

Up to now, no result on the optimal control of variably distributed-order tFDEs was reported in the literature. This motivates the study of the variably distributed-order fractional optimal control model (2) in this paper. Compared to the analysis of the variably distributed-order tFDE (2) [53], in the context of the optimal control problem, we have to address the following issues:

- The coupling of the state Equation (2), the adjoint state Equation (13), and the pointwise constraint (4) reduces the regularity of the exact solution to the optimal control model (3) and (2), the well-posedness and smoothing properties of which were not analyzed even in the conventional distributed-order tFDEs. In particular, the high-order temporal derivatives of the solutions may not be bounded in the L^2 norm in time and could only be analyzed in the L^1 norm (cf. Section 4). Consequently, the results proved for the linear variably distributed-order tFDE [53], which provided pointwise-in-time estimates of high-order temporal derivatives of the solutions based on the high-order regularity assumptions of the data, are no longer valid. Indeed, to find the aforementioned L^1 estimates, the spectral decomposition method used in the literature [53], which is particularly suitable for L^2 estimates, does not apply and we instead employ the method of resolvent estimates in this work in the analysis to circumvent this issue.
- Due to the low regularity of the exact solution to optimal control problem (3) and (2), the analysis of the truncation errors and consequently the convergence estimates of the discretization to the state Equation (2), which are based on the pointwise-in-time estimates of the high-order temporal derivatives of the solutions [53], does not hold any longer and need to be refined carefully.
- Due to the time dependence of the $\omega(\alpha, t)$, the distributed-order fractional differential operator in the adjoint state Equation (13) acts on the product (ωz) instead of z. Hence, the adjoint of the Caputo-type distributed-order differential operator $\mathbb{D}_t^{\nu}$ is no longer its Riemann-Liouville analogue (cf. Section 2.2), which brings new difficulties in its analysis and discretization. For instance, due to the strong coupling of the weight ω and the solution z in the adjoint state equation, the coefficients of the L1 discretization lose their monotonicity that was critical in the error estimates of the numerical discretization.

The rest of the paper is organized as follows: In Section 2 we derive the optimality condition. In Section 3 we prove the well-posedness and regularity of the adjoint state equation under a weaker condition from the control problem. In Section 4 we prove the well-posedness of the optimal control problem (3), (2) and (13) and the regularity of its solution. In Section 5 we discretize the optimal control model. In Section 6 we prove the stability and optimal-order error estimate of the numerical approximation without any artificial assumption of the true solution. In Section 7 we carry out numerical experiments to substantiate the theoretical findings. In Section 8 we prove auxiliary lemmas.

2 | OPTIMALITY CONDITION

We go over some results in the literature and derive an optimality condition. Throughout this paper we use $Q_0 - Q_2$ to denote fixed positive constants and Q to denote a generic positive constant that may assume different values at different occurrences. We may drop the subscript L^2 in $(\cdot, \cdot)_{L^2}$ and $\|\cdot\|_{L^2}$, and the domain 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 also follow the convention that a sum equals zero if its upper limit is smaller than its lower limit.

2.1 | Solution representation and resolvent estimates

It is well known that the eigenfunctions $\{\phi_i\}_{i=1}^{\infty}$ of the Sturm-Liouville problem

$$\mathcal{B}\phi_i(\boldsymbol{x}) = \lambda_i\phi_i(\boldsymbol{x}),\ \boldsymbol{x} \in \Omega; \quad \phi_i(\boldsymbol{x}) = 0,\ \boldsymbol{x} \in \partial\Omega \tag{6}$$

form an orthonormal basis in $L^2(\Omega)$ and the eigenvalues $\{\lambda_i\}_{i=1}^{\infty}$ form a positive increasing sequence going to ∞ [18]. For any $\gamma \geq 0$, define the Sobolev space by

$$\check{H}^{\gamma}(\Omega) := \{v \in L^2(\Omega) : |v|^2_{\check{H}^{\gamma}(\Omega)} := (\mathcal{B}^{\gamma}v, v) = \sum_{i=1}^{\infty}\lambda_i^{\gamma}(v,\phi_i)^2 < \infty\} \tag{7}$$

with $\|g\|_{\check{H}^{\gamma}} := \left(\|g\|^2_{L^2} + |g|^2_{\check{H}^{\gamma}}\right)^{1/2}$, and $\check{H}^0 = L^2$ and $\check{H}^2 = H^2 \cap H_0^1$ [1, 49, 52].

For $\theta \in (\pi/2, \pi)$ and $\delta > 0$, let Γ_θ be the contour in the complex plane

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

The following inequalities hold for $0 < \mu \leq 1$ and $t \in (0, T]$ [2, 28, 39]

$$\int_{\Gamma_\theta} |z|^{\mu-1}|e^{tz}|\ |dz| \leq Qt^{-\mu}, \left\| \int_{\Gamma_\theta} z^{\mu}(z+\mathcal{B})^{-1}e^{tz}\ dz\right\|_{L^2} \leq Qt^{-\mu} \tag{8}$$

where $|dz|$ denotes the arc length element on the contour Γ_θ and $Q = Q(\theta,\mu)$.

For $g \in L_{loc}(a,b)$, the space of locally integrable functions on (a,b), the Laplace transform $\mathcal{L}$ of its extension $\tilde{g}(t)$ with compact support on (a,b) and the corresponding inverse transform $\mathcal{L}^{-1}$ are denoted by

$$\mathcal{L}g(z) := \int_0^{\infty} g(t)e^{-tz}dt, \quad \mathcal{L}^{-1}(\mathcal{L}g(z)) := \frac{1}{2\pi i}\int_{\Gamma_\theta} e^{tz}\mathcal{L}g(z)dz = g(t).$$

$\mathcal{L}g$ is always interpreted as the Laplace transform of $\tilde{g}$. It is known that [48]

$$\mathcal{L}\left({}^R\partial_t^{\mu} g(t)\right) = z^{\mu}\mathcal{L}(g(t)), \quad 0 \leq \mu < 1; \qquad {}^R\partial_t^{\mu} g := \partial_t\, {}_0I_t^{1-\mu} g. \tag{9}$$

The solution $u(\boldsymbol{x},t)$ to the heat equation $\partial_t u(\boldsymbol{x},t) + \mathcal{B}u(\boldsymbol{x},t) = f(\boldsymbol{x},t)$ with zero initial and boundary conditions can be expressed as

$$u(\boldsymbol{x},t) = \int_0^t e^{-(t-s)\mathcal{B}}f(\boldsymbol{x},s)ds. \tag{10}$$

Here $e^{-t\mathcal{B}}$ is the semigroup of operators generated by

$$\partial_t\, e^{-t\mathcal{B}}g + \mathcal{B}\, e^{-t\mathcal{B}}g = 0; \quad e^{-t\mathcal{B}}g = 0,\ \boldsymbol{x} \in \partial\Omega; \quad e^{-t\mathcal{B}}g\Big|_{t=0} = g,\ \boldsymbol{x} \in \Omega.$$

Moreover, $e^{-t\mathcal{B}}$ has the following expressions for any $g \in L^2(\Omega)$

$$e^{-t\mathcal{B}}g(\boldsymbol{x}) = \frac{1}{2\pi \mathrm{i}}\int_{\Gamma_\theta} e^{zt}(z+\mathcal{B})^{-1}g(\boldsymbol{x})\ dz, \quad e^{-t\mathcal{B}}g(\boldsymbol{x}) = \sum_{i=1}^{\infty}e^{-\lambda_i t}(g,\phi_i)\phi_i(\boldsymbol{x}). \tag{11}$$

The following estimates hold for any $t > 0$ [52]

$$\|e^{-t\mathcal{B}}\|_{L^2\to L^2} \leq Q; \quad \|e^{-t\mathcal{B}}g\|_{\check{H}^s} \leq Qt^{-(s-r)/2}\|g\|_{\check{H}^r}, \quad g \in \check{H}^r,\ s \geq r \geq -1. \tag{12}$$

2.2 | The first order optimality condition

Theorem 1. *Under assumptions (a)–(c) the optimal control problem* (3) *and* (2) *admits a unique solution* (u,c). *There exists an adjoint state* z *such that* (u,c,z) *satisfies state*

Equation (2), the adjoint state equation

$$
\begin{aligned}
&-\partial_t z + k\, {}^R\mathbb{D}_t^{\alpha^*}(\omega z) + \mathcal{B}z = u(\boldsymbol{x}, t; c) - u_d(\boldsymbol{x}, t), \quad (\boldsymbol{x}, t) \in \Omega \times [0, T);\\
&z(\boldsymbol{x}, T) = 0, \quad \boldsymbol{x} \in \Omega; \qquad z(\boldsymbol{x}, t) = 0, \quad (\boldsymbol{x}, t) \in \partial\Omega \times [0, T]
\end{aligned} \tag{13}
$$

and the variational inequality

$$
\int_0^T \int_\Omega (\gamma c + z)(v - c)\, d\boldsymbol{x} dt \geq 0, \qquad \forall v \in \Lambda(l, r). \tag{14}
$$

The distributed-order Riemann-Liouville fractional differential operator ${}^R\mathbb{D}_t^{\alpha^*}$ *is*

$$
{}^R\mathbb{D}_t^{\alpha^*} g := \int_0^{\alpha^*} {}^R\hat{\partial}_t^\alpha g\, d\alpha, \;\; {}^R\hat{\partial}_t^\alpha g := -\partial_{tt} I_T^{1-\alpha} g, \;\; {}_t I_T^\alpha g := \int_t^T \frac{g(s)}{\Gamma(\alpha)(s-t)^{1-\alpha}} ds. \tag{15}
$$

Proof. The proof follows almost the same procedure as, for example, [57, theorem 2.1] or [58, theorem 1] and is thus omitted. ∎

Remark 1. The variational inequality (14) implies [26, 59]

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

The uniqueness of the minimizer of $\hat{J}(c)$ shows that (14) is equivalent to (16).

3 | ANALYSIS OF THE ADJOINT STATE FRACTIONAL PDE

For convenience we analyze a forward-in-time analogue of problem (13)

$$
\begin{aligned}
&\partial_t z + k\, {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z) + \mathcal{B}z = p(\boldsymbol{x}, t), \quad (\boldsymbol{x}, t) \in \Omega \times (0, T];\\
&z(\boldsymbol{x}, 0) = 0, \;\; \boldsymbol{x} \in \Omega; \quad z(\boldsymbol{x}, t) = 0, \quad (\boldsymbol{x}, t) \in \partial\Omega \times [0, T].
\end{aligned} \tag{17}
$$

Here ${}^R\overline{\mathbb{D}}_t^{\alpha^*} g := \int_0^{\alpha^*} {}^R\partial_t^\alpha g\, d\alpha$ is the forward-in-time analogue of (15).

Theorem 2. *Suppose that assumptions (a) and (b) hold 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*

$$
\|z\|_{H^1(L^2)} + \|z\|_{L^2(\check{H}^2)} \leq Q\|p\|_{L^2(L^2)}, \quad Q = Q(\alpha^*, Q_0, k, T).
$$

Proof. The proof could be performed following that of [57, theorem 3.1] based on the estimate below: For $w \in H^1(0, T)$ with $w(0) = 0$, one could apply $w(t) = \int_0^t w'(s) ds$, $(t-s)^{-\alpha} = (t-s)^{-\alpha^*}(t-s)^{\alpha^*-\alpha} \leq \max\{1, T\}(t-s)^{-\alpha^*}$ and $\partial_t^\alpha w = {}^R\partial_t^\alpha w$ to get

$$
\begin{aligned}
\left|{}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega w)\right| &\leq \int_0^{\alpha^*} \left|{}_0 I_t^{1-\alpha} \partial_t(\omega w)\right| d\alpha\\
&\leq \int_0^{\alpha^*} \int_0^t \frac{|w(s)\partial_s \omega(\alpha, s) + \omega(\alpha, s) w'(s)|}{\Gamma(1-\alpha)(t-s)^\alpha} ds d\alpha \leq Q \int_0^t \frac{|w(s)| + |w'(s)|}{(t-s)^{\alpha^*}} ds\\
&\leq Q \int_0^t \frac{\int_0^s |w'(y)| dy + |w'(s)|}{(t-s)^{\alpha^*}} ds \leq Q \int_0^t \frac{|w'(s)|}{(t-s)^{\alpha^*}} ds.
\end{aligned} \tag{18}
$$

Thus the proof is omitted. ∎

Theorem 3. *Suppose assumptions (a) and (b) hold and* $p \in H^1(L^2) \cap L^2(\check{H}^2)$. *Then the solution* z *to problem* (17) *has the estimate*

$$\left\|\partial_t^2 z\right\|_{L^1(L^2)} + \left\|\partial_t z\right\|_{L^1(\check{H}^2)} \le Q\left(\|p\|_{H^1(L^2)} + \|p\|_{L^2(\check{H}^2)}\right), \quad Q = Q(\alpha^*, Q_0, k, T).$$

Proof. By Theorem 2 problem (17) has a unique solution $z \in H^1(L^2) \cap L^2(\check{H}^2)$. Move $k\, {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z)$ in (17) to the right side and use (10) and (11) to express z as

$$z(\boldsymbol{x}, t) = \int_0^t e^{-(t-s)\mathcal{B}} p(\boldsymbol{x}, s) ds - k\int_0^t e^{-(t-s)\mathcal{B}}\, {}^R\overline{\mathbb{D}}_s^{\alpha^*}(\omega z(\boldsymbol{x}, s)) ds =: G_1 - kG_2. \tag{19}$$

$$\partial_t G_1 = \partial_t \int_0^t \sum_{i=1}^{\infty} e^{-\lambda_i(t-s)}(p(\cdot, s), \phi_i)\phi_i ds = p - \int_0^t \sum_{i=1}^{\infty} \lambda_i e^{-\lambda_i(t-s)}(p(\cdot, s), \phi_i)\phi_i ds,$$

$$\partial_t^2 G_1 = \partial_t p - \sum_{i=1}^{\infty} \lambda_i (p(\cdot, t), \phi_i)\phi_i + \int_0^t \sum_{i=1}^{\infty} \lambda_i^2 e^{-\lambda_i(t-s)}(p(\cdot, s), \phi_i)\phi_i ds.$$

Application of the Young's inequality yields

$$\begin{aligned} \|\partial_t^2 G_1\|_{L^2(L^2)} \le &\left[\sum_{i=1}^{\infty} \left\| \int_0^t \lambda_i^2 e^{-\lambda_i(t-s)}(p(\cdot, s), \phi_i) ds \right\|_{L^2(0,T)}^2\right]^{1/2} \\ &+ \|p\|_{H^1(L^2)} + \|p\|_{L^2(\check{H}^2)} \le \|p\|_{H^1(L^2)} + 2\|p\|_{L^2(\check{H}^2)}. \end{aligned} \tag{20}$$

We use the expression of G_2 in (19) to directly evaluate

$$\partial_t G_2 = -\int_0^t e^{-(t-s)\mathcal{B}}\, \mathcal{B}\, {}^R\overline{\mathbb{D}}_s^{\alpha^*}(\omega z(\boldsymbol{x}, s)) ds + {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z(\boldsymbol{x}, t)). \tag{21}$$

We utilize the following relation

$$\begin{aligned} \partial_t \int_0^t e^{-(t-s)\mathcal{B}} \mathcal{B}\, {}^R\overline{\mathbb{D}}_s^{\alpha^*}(\omega z(\boldsymbol{x}, s)) ds &= \partial_t \int_0^t e^{-s\mathcal{B}} \mathcal{B} \left({}^R\overline{\mathbb{D}}_y^{\alpha^*}(\omega z(\boldsymbol{x}, y))\right)\Big|_{y=t-s} ds \\ = \int_0^t e^{-s\mathcal{B}} \mathcal{B}\, \partial_t \left(\left({}^R\overline{\mathbb{D}}_y^{\alpha^*}(\omega z(\boldsymbol{x}, y))\right)\Big|_{y=t-s}\right) ds &= \int_0^t e^{-(t-s)\mathcal{B}} \mathcal{B}\, (\partial_s {}^R\overline{\mathbb{D}}_s^{\alpha^*}(\omega z(\boldsymbol{x}, s))) ds \end{aligned}$$

and (21) to evaluate $\partial_t^2 G_2$ as follows

$$\partial_t^2 G_2 = -\int_0^t e^{-(t-s)\mathcal{B}} \mathcal{B} \left(\partial_s\, {}^R\overline{\mathbb{D}}_s^{\alpha^*}(\omega z(\boldsymbol{x}, s))\right) ds + \partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z(\boldsymbol{x}, t)). \tag{22}$$

We integrate ${}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z)$ by parts to reformulate $\partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z)$ as

$$\begin{aligned} \partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z) &= \partial_t \int_0^{\alpha^*} \left[\frac{\partial_t(\omega z)|_{t=0} t^{1-\alpha}}{\Gamma(2-\alpha)} + \int_0^t \frac{(t-s)^{1-\alpha}}{\Gamma(2-\alpha)} \partial_s^2(\omega(\alpha, s) z(\boldsymbol{x}, s)) ds\right] d\alpha \\ &= \int_0^{\alpha^*} \left[\frac{\partial_t(\omega z)|_{t=0} t^{-\alpha}}{\Gamma(1-\alpha)} + \int_0^t \frac{(t-s)^{-\alpha}}{\Gamma(1-\alpha)} \partial_s^2(\omega(\alpha, s) z(\boldsymbol{x}, s)) ds\right] d\alpha. \end{aligned} \tag{23}$$

Passing the limit $t \to 0^+$ in (17) yields $\partial_t z(\boldsymbol{x}, 0) = p(\boldsymbol{x}, 0)$, which in turn leads to $\partial_t(\omega z)|_{t=0} = \omega(\alpha, 0) p(\boldsymbol{x}, 0)$. We plug them into (23) to conclude that

$$\left|\partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z)\right| \le Q \int_0^t \frac{|\partial_s^2 z(\boldsymbol{x}, s)|}{(t-s)^{\alpha^*}} ds + Q|p(\boldsymbol{x}, 0)| t^{-\alpha^*}. \tag{24}$$

We similarly bound ${}^R\partial_t^\varepsilon\ \partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z)$ for any $0 < \varepsilon < 1-\alpha^*$. By estimate (18), ${}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z)|_{t=0} = 0$. We use the substitution $s = y + (t-y)\theta$ to obtain

$$
\begin{aligned}
{}^R\partial_t^\varepsilon \partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z) &= \partial_t\ {}_0I_t^{1-\varepsilon}\ \partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z) = \partial_t^2\ {}_0I_t^{1-\varepsilon}\, {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z)\\
&= \int_0^{\alpha^*} \partial_t^3\ {}_0I_t^{1-\varepsilon}\,{}_0I_t^{1-\alpha}(\omega z)d\alpha\\
&= \int_0^{\alpha^*} \partial_t^3\left[\int_0^t \frac{\omega(\alpha,y)z(\boldsymbol{x},y)}{\Gamma(1-\varepsilon)\Gamma(1-\alpha)}\left(\int_y^t (t-s)^{-\varepsilon}(s-y)^{-\alpha}ds\right)dy\right]d\alpha\\
&= \int_0^{\alpha^*}\partial_t^3\int_0^t \frac{\omega(\alpha,y)z(\boldsymbol{x},y)(t-y)^{1-\alpha-\varepsilon}B(1-\varepsilon,1-\alpha)}{\Gamma(1-\varepsilon)\Gamma(1-\alpha)}dyd\alpha\\
&= \int_0^{\alpha^*}\partial_t^2\int_0^t \frac{\omega(\alpha,y)z(\boldsymbol{x},y)(t-y)^{-\alpha-\varepsilon}}{\Gamma(1-\alpha-\varepsilon)}dyd\alpha = \int_0^{\alpha^*} {}^R\partial_t^{\alpha+\varepsilon}\left(\partial_t(\omega z)\right)d\alpha,
\end{aligned}
$$

which, together with estimates like (23) and (24), yields

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

We use (11) and the Laplace transform to evaluate the first term on the right-hand side of $\partial_t^2 G_2$ in (22) to conclude that for $0<\varepsilon<1-\alpha^*$

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

Invert the Laplace transform, use (9) and the Laplace transform of convolution to get

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

We use (8) to bound the integral in the square brackets and use (25) to bound the first term on the right-hand side of $\partial_t^2 G_2$ in (22)

$$
\begin{aligned}
\left\|\int_0^t e^{-(t-s)\mathcal{B}}\mathcal{B}\ \left(\partial_s\ {}^R\overline{\mathbb{D}}_s^{\alpha^*}(\omega z(\boldsymbol{x},s))\right)ds\right\| &\le Q\int_0^t \frac{\left\|{}^R\partial_s^\varepsilon\partial_s {}^R\overline{\mathbb{D}}_s^{\alpha^*}(\omega z(\cdot,s))\right\|ds}{(t-s)^{1-\varepsilon}}\\
&\le Q\int_0^t \frac{1}{(t-s)^{1-\varepsilon}}\left(\int_0^s \frac{\|\partial_\theta^2 z(\cdot,\theta)\|}{(s-\theta)^{\alpha_*+\varepsilon}}d\theta + \|p(\cdot,0)\|s^{-\varepsilon-\alpha^*}\right)ds\\
&\le Q\int_0^t \frac{\|\partial_\theta^2 z(\cdot,\theta)\|}{(t-\theta)^{\alpha_*}}d\theta + Q\|p(\cdot,0)\|t^{-\alpha^*}.
\end{aligned} \tag{26}
$$

Use (24) to bound the L^2 norm of the second term on the right side of (22) by the right side of (26). Multiply (22) by $e^{-\sigma t}$ and bound the $\|\cdot\|_{L^q(0,T)}$ on both sides of the equation by Young's inequality to obtain

$$\begin{aligned}\left\|e^{-\sigma t}\partial_t^2 G_2\right\|_{L^1(L^2)} &\le Q\left\|\int_0^t \frac{e^{-\sigma(t-\theta)}}{(t-\theta)^{\alpha_*}}e^{-\sigma\theta}\|\partial_\theta^2 z(\cdot,\theta)\|d\theta\right\|_{L^1(0,T)} + Q\|p(\cdot,0)\| \\ &\le Q\sigma^{\alpha^*-1}\|e^{-\sigma t}\partial_t^2 z\|_{L^1(L^2)} + Q\|p(\cdot,0)\|.\end{aligned} \tag{27}$$

Differentiate (19) twice in time, bound the $\|\cdot\|_{L^q(0,T)}$ norm on both sides of the equation multiplied by $e^{-\lambda t}$ and invoke (20) and (27) to obtain

$$\begin{aligned}\left\|e^{-\sigma t}\partial_t^2 z\right\|_{L^1(L^2)} &\le \|e^{-\sigma t}\partial_t^2 G_1\|_{L^1(L^2)} + \|e^{-\sigma t}\partial_t^2 G_2\|_{L^1(L^2)} \\ &\le Q\sigma^{\alpha^*-1}\|e^{-\sigma t}\partial_t^2 z\|_{L^1(L^2)} + Q\left(\|p\|_{H^1(L^2)} + \|p\|_{L^2(\check{H}^2)}\right).\end{aligned}$$

Set σ large enough to get $\|\partial_t^2 z\|_{L^1(L^2)} \le Q\left(\|p\|_{H^1(L^2)} + \|p\|_{L^2(\check{H}^2)}\right)$, which, together with (7), (17) and (24), yields

$$\begin{aligned}\left\|\partial_t z\right\|_{L^1(\check{H}^2)} = \left\|\partial_t \mathcal{B} z\right\|_{L^1(L^2)} &= \left\|\partial_t^2 z + k\partial_t {}^R\overline{\mathbb{D}}_t^{\alpha^*}(\omega z) - \partial_t p\right\|_{L^1(L^2)} \\ &\le Q\left(\|p\|_{H^1(L^2)} + \|p\|_{L^2(\check{H}^2)}\right).\end{aligned}$$

We thus complete the proof. ■

4 | ANALYSIS OF THE VARIABLY DISTRIBUTED-ORDER FRACTIONAL OPTIMAL CONTROL MODEL

We restate Theorems 2 and 3 for the adjoint state Equation (13) in the corollary.

Corollary 1. *If assumptions (a) and (b) hold and $u, u_d \in L^2(L^2)$, the adjoint state Equation (13) has a unique solution $z \in H^1(L^2) \cap L^2(\check{H}^2)$ and*

$$\|z\|_{H^1(L^2)} + \|z\|_{L^2(\check{H}^2)} \le Q\|u - u_d\|_{L^2(L^2)} \tag{28}$$

with $Q = Q(\alpha^*, Q_0, k, T)$. If $u, u_d \in H^1(L^2) \cap L^2(\check{H}^2)$, then

$$\left\|\partial_t^2 z\right\|_{L^1(L^2)} + \left\|\partial_t z\right\|_{L^1(\check{H}^2)} \le Q\left(\|u - u_d\|_{H^1(L^2)} + \|u - u_d\|_{L^2(\check{H}^2)}\right). \tag{29}$$

Theorem 4. *Suppose that assumptions (a) and (b) hold. If $f, c \in L^2(L^2)$, the state Equation (2) has a unique solution $u \in H^1(L^2) \cap L^2(\check{H}^2)$ and*

$$\|u\|_{H^1(L^2)} + \|u\|_{L^2(\check{H}^2)} \le Q\|f + c\|_{L^2(L^2)}. \tag{30}$$

If $f, u_d \in H^1(L^2) \cap L^2(\check{H}^2)$ and $c \in H^1(L^2)$, then

$$\begin{aligned}\left\|\partial_t^2 u\right\|_{L^1(L^2)} + \left\|\partial_t u\right\|_{L^1(\check{H}^2)} \le Q(&\|f\|_{H^1(L^2)} + \|f\|_{L^2(\check{H}^2)} + \|c\|_{H^1(L^2)} \\ &+ \|u_d\|_{H^1(L^2)} + \|u_d\|_{L^2(\check{H}^2)}).\end{aligned} \tag{31}$$

Proof. By the assumptions estimate (30) holds. However, the regularity estimate of $\partial_t^2 u$ requires $c \in L^2(\check{H}^2)$ that is not true by (16). $z \in H^1(L^2) \cap L^2(\check{H}^2)$ and estimates (28) and (29) hold by Corollary 1. To bound $\partial_t^2 u$, we re-estimate $\partial_t^2 G_1$ in (19) with p replaced by c.

We use the second equation in (11) to get

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

We differentiate the equation with respect to t to find

$$\partial_t^2 \int_0^t e^{-(t-s)\mathcal{B}} c(\boldsymbol{x},s)ds = -e^{-t\mathcal{B}} \mathcal{B} c(\boldsymbol{x},0) + \partial_t c(\boldsymbol{x},t) + \partial_t \int_0^t e^{-(t-s)\mathcal{B}} \partial_s c(\boldsymbol{x},s)ds. \tag{32}$$

Use Young's inequality to bound the last term on the right-hand side

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

We use (5), (12), (16), (29) and the equivalence between $\check{H}^{\frac{1}{2}-\varepsilon}$ and $H^{\frac{1}{2}-\varepsilon}$ to bound the first term on the right-hand side of (32) for $0 < \varepsilon \ll 1$ by

$$\left\| e^{-t\mathcal{B}} \mathcal{B} c(\boldsymbol{x},0) \right\|_{L^2} = \left\| e^{-t\mathcal{B}} c(\boldsymbol{x},0) \right\|_{\check{H}^2} \leq Qt^{-\left(\frac{3}{4}+\frac{\varepsilon}{2}\right)} \|c(\boldsymbol{x},0)\|_{\check{H}^{\frac{1}{2}-\varepsilon}}$$
$$\leq Qt^{-\left(\frac{3}{4}+\frac{\varepsilon}{2}\right)} \|c(\boldsymbol{x},0)\|_{H^{\frac{1}{2}-\varepsilon}} \leq Qt^{-\left(\frac{3}{4}+\frac{\varepsilon}{2}\right)} \|z(\boldsymbol{x},0)\|_{H^{\frac{1}{2}-\varepsilon}}$$
$$\leq Qt^{-\left(\frac{3}{4}+\frac{\varepsilon}{2}\right)} \|\partial_t z\|_{L^1\left(H^{\frac{1}{2}-\varepsilon}\right)} \leq Qt^{-\left(\frac{3}{4}+\frac{\varepsilon}{2}\right)} \left(\|u-u_d\|_{H^1(L^2)} + \|u-u_d\|_{L^2(\check{H}^2)} \right).$$

Consequently we have $\left\| e^{-t\mathcal{B}} \mathcal{B} c(\boldsymbol{x},0) \right\|_{L^1(L^2)} \leq Q\left(\|u-u_d\|_{H^1(L^2)} + \|u-u_d\|_{L^2(\check{H}^2)} \right)$. The remaining analysis can be carried out as in Theorem 3 and is omitted. ■

Theorem 5. *Under assumptions (a)–(c), problem (3) and (2) has a unique solution $u \in H^1(L^2) \cap L^2(\check{H}^2)$ and $c \in H^1(L^2)$. Equation (13) has a unique solution $z \in H^1(L^2) \cap L^2(\check{H}^2)$. There is a positive constant $Q = Q(\alpha^*, Q_0, k, T)$ such that*

$$\|c\|_{H^1(L^2)} \leq Q\left(\|u_d\|_{H^1(L^2)} + \|u_d\|_{L^2(\check{H}^2)} + \|f\|_{L^2(L^2)} + \|c\|_{L^2(L^2)} \right), \tag{33}$$

$$\begin{aligned} &\|u\|_{H^1(L^2)} + \|u\|_{L^2(\check{H}^2)} + \|\partial_t^2 u\|_{L^1(L^2)} + \|\partial_t u\|_{L^1(\check{H}^2)} \\ &\leq Q\left(\|u_d\|_{H^1(L^2)} + \|u_d\|_{L^2(\check{H}^2)} + \|f\|_{H^1(L^2)} + \|f\|_{L^2(\check{H}^2)} + \|c\|_{L^2(L^2)} \right), \end{aligned} \tag{34}$$

$$\begin{aligned} &\|z\|_{H^1(L^2)} + \|z\|_{L^2(\check{H}^2)} + \left\|\partial_t^2 z\right\|_{L^1(L^2)} + \left\|\partial_t z\right\|_{L^1(\check{H}^2)} \\ &\leq Q\left(\|u_d\|_{H^1(L^2)} + \|u_d\|_{L^2(\check{H}^2)} + \|f\|_{L^2(L^2)} + \|c\|_{L^2(L^2)} \right). \end{aligned} \tag{35}$$

Proof. The proof could be performed following that of [58, theorem 5] or [57, theorem 4.3] and is thus omitted. ■

Remark 2. In this paper, we assume the operator $\mathcal{B}$ in (2) is self-adjoint and symmetric, the obtained theoretical results cannot be directly extended to the variably distributed-order time-fractional advection-diffusion-reaction equation where the operator $\mathcal{B}$ is non-symmetric. However, in our recent work, we obtain the well-posedness and maximal regularity estimates for the optimal control of a fractional advection-diffusion-reaction equation with space-time-dependent order and coefficients by utilizing the Fredholm alternative for compact operators and a bootstrapping argument [33]. Consequently, the optimal control of variably distributed-order time-fractional advection-diffusion-reaction equation can also be analyzed similarly.

5 | DISCRETIZATION

In this section we discretize the optimal control model.

5.1 | Time discretization of state Equation (2) and adjoint state Equation (13)

Partition $[0, T]$ by $t_n := n\tau$ for $0 \le n \le N$ and $\tau := T/N$, and $[0, \alpha^*]$ by $\alpha_m := m\sigma$ for $0 \le m \le M$ and $\sigma := \alpha^*/M$. Let $f_n := f(\boldsymbol{x}, t_n)$, $c_n := c(\boldsymbol{x}, t_n)$, $u_n := u(\boldsymbol{x}, t_n)$, and $\omega_n^m := \omega(\alpha_m, t_n)$. We discretize $\partial_t u$ by the implicit Euler method and $\mathbb{D}_t^{\omega} u$ by the composite trapezoidal rule on $[0, \alpha^*]$ and the L1 method for each $\partial_t^{\alpha_m}$

$$
\begin{aligned}
\partial_t u(\boldsymbol{x}, t_n) &= \delta_\tau u_n + E_n := \frac{u_n - u_{n-1}}{\tau} + \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \partial_t^2 u(\boldsymbol{x}, t)(t - t_{n-1})dt, \\
\mathbb{D}_t^{\omega} u(\boldsymbol{x}, t_n) &= \sum_{m=0}^{M} \sigma_m \omega_n^m \delta_\tau^{\alpha_m} u_n + R_n + S_n =: \check{\mathbb{D}}_\tau^{\omega} u_n + R_n + S_n, \\
\delta_\tau^{\alpha_m} u_n &:= \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \frac{\delta_\tau u_k \, ds}{\Gamma(1-\alpha_m)(t_n - s)^{\alpha_m}} = \sum_{k=1}^{n} b_{n,k}^m (u_k - u_{k-1}) \\
&= b^m u_n + \sum_{k=1}^{n-1} (b_{n,k}^m - b_{n,k+1}^m) u_k, \\
b_{n,k}^m &:= \frac{(t_n - t_{k-1})^{1-\alpha_m} - (t_n - t_k)^{1-\alpha_m}}{\Gamma(2-\alpha_m)\tau}, \quad 0 \le m \le M, \quad 1 \le n \le N; \\
\sigma_0 &= \sigma_M = \sigma/2, \quad \sigma_1 = \ldots = \sigma_{M-1} = \sigma.
\end{aligned}
\tag{36}
$$

Here the local truncation errors R_n and S_n are given by

$$
\begin{aligned}
S_n &:= \mathbb{D}_t^{\omega} u(\boldsymbol{x}, t_n) - \sum_{m=0}^{M} \sigma_m \omega_n^m \partial_t^{\alpha_m} u(\boldsymbol{x}, t_n), \\
R_n &:= \sum_{m=0}^{M} \sigma_m \omega_n^m R_n^m := \sum_{m=0}^{M} \sigma_m \omega_n^m \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \frac{\partial_s u(\boldsymbol{x}, s) - \delta_\tau u_k}{\Gamma(1-\alpha_m)(t_n - s)^{\alpha_m}} ds \\
&= \sum_{m=0}^{M} \sigma_m \omega_n^m \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \frac{(t_n - s)^{-\alpha_m}}{\tau\Gamma(1-\alpha_m)} \left[\int_{t_{k-1}}^{t_k} \int_z^s \partial_\theta^2 u(\boldsymbol{x}, \theta) d\theta dz \right] ds.
\end{aligned}
\tag{37}
$$

We plug (36) into (2) and integrate the equation multiplied by $\chi \in H_0^1(\Omega)$ on Ω to obtain the following for Equation (2) for any $\chi \in H_0^1$ and $n = 1, 2, \ldots, N$

$$
(\delta_\tau u_n + k\check{\mathbb{D}}_\tau^{\omega} u_n + \mathscr{B} u_n, \chi) = (f_n + c_n, \chi) - (k(R_n + S_n) + E_n, \chi). \tag{38}
$$

We discretize $-\partial_t z$ and apply ${}^R\hat{\partial}_t^{\alpha} g = \hat{\partial}_t^{\alpha} g$ if $g(T) = 0$ to discretize ${}^R\mathbb{D}_t^{\alpha^*}(\omega z)$ backward in time for $n = N, N-1, \ldots, 1$ as follows

$$
\begin{aligned}
-\partial_t z(\boldsymbol{x}, t_{n-1}) &= -\delta_\tau z_n + \hat{E}_{n-1} := \frac{z_{n-1} - z_n}{\tau} + \frac{1}{\tau}\int_{t_{n-1}}^{t_n} \partial_t^2 z(\boldsymbol{x}, t)(t_n - t)dt, \\
{}^R\mathbb{D}_t^{\alpha^*}(\omega z(\boldsymbol{x}, t_{n-1})) &= \int_0^{\alpha^*} {}^R\hat{\partial}_t^{\alpha}(\omega z)|_{t=t_{n-1}} d\alpha = \int_0^{\alpha^*} \hat{\partial}_t^{\alpha}(\omega z)|_{t=t_{n-1}} d\alpha \\
&= \sum_{m=0}^{M} \sigma_m \hat{\delta}_\tau^{\alpha_m}(\omega_{n-1}^m z_{n-1}) + \hat{R}_{n-1} + \hat{S}_{n-1} \\
&=: {}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m z_{n-1}) + \hat{R}_{n-1} + \hat{S}_{n-1}, \\
\hat{\delta}_\tau^{\alpha_m}(\omega_{n-1}^m z_{n-1}) &:= \sum_{k=n}^{N} \int_{tk-1}^{tk} \frac{-\delta_\tau(\omega_k^m\, z_k)\, ds}{\Gamma(1-\alpha_m)(s - t_{n-1})^{\alpha_m}} \\
&= b_{n,n}\omega_{n-1}^m z_{n-1} + \sum_{k=n+1}^{N} (b_{k,n} - b_{k,n+1})\omega_{k-1}^m z_{k-1}.
\end{aligned}
\tag{39}
$$

Here the local truncation errors $\hat{R}_{n-1}$ and $\hat{S}_{n-1}$ are given by

$$
\begin{aligned}
\hat{S}_{n-1} &:= \int_0^{\alpha^*} \hat{\partial}_t^{\alpha}(\omega z)|_{t=t_{n-1}} d\alpha - \sum_{m=0}^{M} \sigma_m\ \hat{\partial}_t^{\alpha_m}\ (\omega(\alpha_m, t) z(\boldsymbol{x}, t))\ |_{t=t_{n-1}} \\
\hat{R}_{n-1} &:= \sum_{m=0}^{M} \sigma_m \hat{R}_{n-1}^m := \sum_{m=0}^{M} \sigma_m \sum_{k=n}^{N} \int_{t_{k-1}}^{t_k} \frac{\delta_\tau(\omega_k^m z_k) - \partial_s(\omega(\alpha_m, s) z(\boldsymbol{x}, s))}{\Gamma(1-\alpha_m)(s - t_{n-1})^{\alpha_m}} ds \\
&= \sum_{m=0}^{M} \sigma_m \sum_{k=n}^{N} \int_{t_{k-1}}^{t_k} \frac{(s - t_{n-1})^{-\alpha_m}}{\tau\Gamma(1-\alpha_m)} \left[\int_{t_{k-1}}^{t_k}\int_s^z \partial_\theta^2(\omega(\alpha_m, \theta) z(\boldsymbol{x}, \theta)) d\theta dz\right] ds.
\end{aligned}
\tag{40}
$$

We plug (39) into (13) to get the following equation for adjoint state Equation (13): For $n = N, N-1, \ldots, 1$, find $z(\boldsymbol{x}, t_{n-1})$ such that for $\chi \in H_0^1$

$$
\begin{aligned}
&\left(-\delta_\tau z_n + k\,{}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m z_{n-1}) + \mathcal{B} z_{n-1}, \chi\right) \\
&\quad = (u_{n-1} - u_d(\cdot, t_{n-1}), \chi) - \left(k(\hat{R}_{n-1} + \hat{S}_{n-1}) + \hat{E}_{n-1}, \chi\right).
\end{aligned}
\tag{41}
$$

5.2 | A discretization of the optimal control model

Define a quasi-uniform partition of Ω with the mesh diameter h. Let S_h be the space of continuous piecewise linear functions on Ω with respect to the partition, and $\mathcal{B}_h : S_h \to S_h$ be defined by $(\mathcal{B}_h\zeta, \chi) = (\boldsymbol{K}\nabla\zeta, \nabla\chi)$ for $\zeta, \chi \in S_h$. Let $\Lambda^\tau(l, r)$ be a time discretization of $\Lambda(l, r)$ in (4) such that $c(\boldsymbol{x}, t) = c(\boldsymbol{x}, t_{n-1})$ on each $[t_{n-1}, t_n) \subset [0, T]$. Namely,

$$
\Lambda^\tau(l, r) := \{\boldsymbol{C} := \{C_{n-1}(\boldsymbol{x})\}_{n=1}^N :\quad l \le C_{n-1} \le r, \quad 1 \le n \le N\}.
$$

We drop the last term on the right-hand side of (38) to get a discrete control model

$$
\min_{\boldsymbol{C}\in\Lambda^\tau(l,r)} J^\tau(\boldsymbol{U}, \boldsymbol{C}) := \frac{\tau}{2}\sum_{n=1}^{N} \left(\|U_n - u_d(\cdot, t_n)\|^2 + \gamma\|C_{n-1}\|^2\right), \tag{42}
$$

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

$$\left(\delta_\tau U_n + k\check{\mathbb{D}}_\tau^\omega U_n + \mathcal{B}_h U_n, \chi\right) = (f_n + C_{n-1}, \chi), \quad \chi \in S_h, \quad 1 \le n \le N. \tag{43}$$

By a similar proof as Theorem 1, we reach the following theorem.

Theorem 6. *The discrete optimal control system* (42) *and* (43) *admits a unique solution* $(\boldsymbol{U}, \boldsymbol{C})$, *and a adjoint state* $\boldsymbol{Z} = \{Z_n\}_{n=0}^{N-1} \subset S_h$ *with* $Z_N = 0$ *such that*

$$\left(-\delta_\tau Z_n + k\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m Z_{n-1}) + \mathcal{B}_h Z_{n-1}, \chi\right) = (U_n - u_d(\cdot, t_n), \chi), \quad \forall \chi \in S_h \tag{44}$$

for $n = N, N-1, \dots, 1$ *and* $(\gamma C_{n-1} + Z_{n-1}, v - C_{n-1}) \ge 0$ *for any* $v \in L^2$ with $l \le v \le r$. *Furthermore*,

$$C_{n-1}(\boldsymbol{x}) = \mathcal{P}\left(-Z_{n-1}(\boldsymbol{x})/\gamma\right), \quad 1 \le n \le N. \tag{45}$$

6 | STABILITY AND ERROR ESTIMATES OF THE DISCRETIZATION

We start with the stability analysis of the time-discretized Equations (38) and (41).

6.1 | Stability of the time-discretized equations

Theorem 7. *The solutions* $(\boldsymbol{U}, \boldsymbol{Z})$ *of schemes* (43) *and* (44) *are stable*

$$\|\boldsymbol{U}\|_{\hat{L}^\infty(L^2)} \le Q\left(\|\boldsymbol{f}\|_{\hat{L}^1(L^2)} + \|\boldsymbol{C}\|_{\hat{L}^1(L^2)}\right), \quad \|\boldsymbol{Z}\|_{\hat{L}^\infty(L^2)} \le Q\|\boldsymbol{U} - \boldsymbol{u}_d\|_{\hat{L}^1(L^2)}. \tag{46}$$

Here the discrete norms are defined by

$$\|\boldsymbol{V}\|_{\hat{L}^\infty(L^2)} := \max_{1 \le n \le N} \|V_n\|, \quad \|\boldsymbol{V}\|_{\hat{L}^p(L^2)} := \left[\tau \sum_{n=1}^N \left\|V_n\right\|^p\right]^{\frac{1}{p}}, \ p = 1, 2 \tag{47}$$

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

Proof. We set $\chi = Z_{n-1}$ in (44) and use expression (40) for $\hat{\delta}_\tau^{\alpha_m}$ to get

$$\begin{aligned}&\left(1 + k\tau\sum_{m=0}^M \sigma_m b_{n,n}^m \omega_{n-1}^m\right)\|Z_{n-1}\|^2 + \tau\left(\mathcal{B}_h Z_{n-1}, Z_{n-1}\right)\\&= (Z_n, Z_{n-1}) + k\tau\sum_{m=0}^M \sigma_m \sum_{k=n+1}^N (b_{k,n+1}^m - b_{k,n}^m)(\omega_{k-1}^m Z_{k-1}, Z_{n-1})\\&\quad + \tau(U_n - u_d(\cdot, t_n), Z_{n-1}).\end{aligned}$$

We use $b_{k,n+1}^m > b_{k,n}^m$ for $n+1 \le k \le N$ and $0 \le m \le M$ to cancel $\|Z_{n-1}\|$ on both sides to obtain

$$\begin{aligned}&\left(1 + k\tau\sum_{m=0}^M \sigma_m b_{n,n}^m \omega_{n-1}^m\right)\|Z_{n-1}\|\\&\le \|Z_n\| + k\tau\sum_{m=0}^M \sigma_m \sum_{k=n+1}^N \left(b_{k,n+1}^m - b_{k,n}^m\right)\omega_{k-1}^m \|Z_{k-1}\| + \tau\|U_n - u_d(\cdot, t_n)\|.\end{aligned} \tag{48}$$

We use assumption (a) and the expression $b_{k,n}^m$ in (37) to get the estimate for $0 \le m \le M$ and $1 \le n \le N$

$$\sum_{k=n}^{N} b_{k,n}^m |\omega_k^m - \omega_{k-1}^m| \le \frac{Q_0}{\Gamma(1-\alpha_m)} \sum_{k=n}^{N} \int_{t_{k-1}}^{t_k} \frac{1}{(s-t_{n-1})^{\alpha_m}} ds$$
$$= \frac{Q_0 (T-t_{n-1})^{1-\alpha_m}}{\Gamma(2-\alpha_m)} \le \frac{Q_0 \max\{1,T\}}{\Gamma(\gamma_0)} =: Q_2$$

where $\gamma_0 \approx 1.46$ is the minimizer of $\Gamma(\cdot)$ on $(0,\infty)$. Set $n = N$ in (48) to get

$$\|Z_{N-1}\| \le \tau \left\| U_N - u_d(\cdot, t_N) \right\| \le \tau (1 + Q_2 k\tau) \left\| U_N - u_d(\cdot, t_N) \right\|.$$

Suppose that the following estimate holds for $n+1 \le m \le N$

$$\|Z_{m-1}\| \le A_m \tau \sum_{j=m}^{N} \left\| U_j - u_d(\cdot, t_j) \right\|, \quad A_m := (1 + Q_2 k\tau)^{N-m+1}. \tag{49}$$

We use (48) and (49), $A_1 > A_2 > \cdots > A_N > 1$ and

$$\sum_{k=n+1}^{N} \left(b_{k,n+1}^m - b_{k,n}^m\right) \omega_{k-1}^m = \sum_{k=n+1}^{N} \left(b_{k-1,n}^m \omega_{k-1}^m - b_{k,n}^m \omega_k^m\right) + \left(b_{k,n}^m \omega_k^m - b_{k,n}^m \omega_{k-1}^m\right)$$
$$\le b_{n,n}^m \omega_n^m + \sum_{k=n+1}^{N} b_{k,n}^m \left(\omega_k^m - \omega_{k-1}^m\right)$$
$$= b_{n,n}^m \omega_{n-1}^m + \sum_{k=n}^{N} b_{k,n}^m \left(\omega_k^m - \omega_{k-1}^m\right) \le b_{n,n}^m \omega_{n-1}^m + Q_2$$

to obtain

$$\left(1 + k\tau \sum_{m=0}^{M} \sigma_m b_{n,n}^m \omega_{n-1}^m\right) \|Z_{n-1}\| \le A_{n+1} \tau \sum_{j=n}^{N} \left\| U_j - u_d(\cdot, t_j) \right\|$$
$$+ k\tau \sum_{m=0}^{M} \sigma_m \left(b_{n,n}^m \omega_{n-1}^m + Q_2\right) \left[A_{n+1} \tau \sum_{j=n+1}^{N} \left\| U_j - u_d(\cdot, t_j) \right\|\right]$$
$$\le \left(1 + k\tau \sum_{m=0}^{M} \sigma_m b_{n,n}^m \omega_{n-1}^m + \alpha^* Q_2 k\tau\right) A_{n+1} \tau \sum_{j=n}^{N} \left\| U_j - u_d(\cdot, t_j) \right\|.$$

We divide both sides by $(1 + k\tau \sum_{m=0}^{M} \sigma_m b_{n,n}^m \omega_{n-1}^m)$ to get that for $1 \le n \le N$

$$\|Z_{n-1}\| \le \left(1 + \frac{Q_2 k\tau}{1 + k\tau \sum_{m=0}^{M} \sigma_m b_{n,n}^m \omega_{n-1}^m}\right) A_{n+1} \tau \sum_{j=n}^{N} \left\| U_j - u_d(\cdot, t_j) \right\|$$
$$\le (1 + kQ_2\tau) A_{n+1} \tau \sum_{j=n}^{N} \left\| U_j - u_d(\cdot, t_j) \right\| = A_n \tau \sum_{j=n}^{N} \left\| U_j - u_d(\cdot, t_j) \right\|.$$

Thus, (49) holds for $m = n$ and for all $1 \le m \le N$ by induction. We have proved the second estimate in (46). The first one can be proved in a similar manner. ■

6.2 | An optimal-order error estimate of the discretization

Let $\Pi_h : H_0^1(\Omega) \to S_h$ be the Ritz projection operator: for any $v \in H_0^1(\Omega)$

$$(\boldsymbol{K}\nabla(v - \Pi_h v), \nabla \chi) = 0, \quad \forall \chi \in S_h.$$

It is well known that the following approximation property holds [52]

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

Theorem 8. *Suppose that assumptions (a)–(c) hold. The optimal-order error estimate holds for the discrete optimal control system*

$$\|u - U\|_{\hat{L}^\infty(L^2)} + \|z - Z\|_{\hat{L}^\infty(L^2)} + \|c - C\|_{\hat{L}^\infty(L^2)} \le Q Q^* \left(\tau + h^2 + \sigma^2\right) \tag{51}$$

with $Q^* := \|u_d\|_{H^1(L^2)} + \|u_d\|_{L^2(\check{H}^2)} + \|f\|_{H^1(L^2)} + \|f\|_{L^2(\check{H}^2)} + \|c\|_{L^2(L^2)}$.

Proof. We subtract Equation (38) from scheme (43) to obtain the error equation

$$\begin{aligned} &\left(\delta_\tau(U_n - u_n) + k\check{\mathbb{D}}_\tau^\omega(U_n - u_n), \chi\right) + (\boldsymbol{K}\nabla(U_n - u_n), \nabla\chi) \\ &\quad = (C_{n-1} - c_n, \chi) + (k(R_n + S_n) + E_n, \chi), \quad \forall \chi \in S_h. \end{aligned} \tag{52}$$

We decompose $U_n - u_n = \xi_n + \eta_n$, where $\xi_n = U_n - \Pi_h u_n \in S_h$ and $\eta(\boldsymbol{x}, t) := \Pi_h u(\boldsymbol{x}, t) - u(\boldsymbol{x}, t)$. We rewrite Equation (52) in terms of ξ and η as follows

$$\begin{aligned} &\left(\delta_\tau \xi_n + k\check{\mathbb{D}}_\tau^\omega \xi_n, \chi\right) + (\boldsymbol{K}\nabla\xi_n, \nabla\chi) \\ &\quad = -\left(\delta_\tau \eta_n + k\check{\mathbb{D}}_\tau^\omega \eta_n, \chi\right) + (C_{n-1} - c_n, \chi) + (k(R_n + S_n) + E_n, \chi). \end{aligned} \tag{53}$$

We apply Theorem 7 to Equation (53) to obtain the following estimate

$$\begin{aligned} \|\xi\|_{\hat{L}^\infty(L^2)} \le{}& Q\Big(\|\delta_\tau \eta\|_{\hat{L}^1(L^2)} + k\|\check{\mathbb{D}}_\tau^\omega \eta\|_{\hat{L}^1(L^2)} + k\|S\|_{\hat{L}^1(L^2)} + \|E\|_{\hat{L}^1(L^2)} \\ &+ \|R\|_{\hat{L}^1(L^2)} + \|C - c\|_{\hat{L}^1(L^2)}\Big) + Q\tau \sum_{n=1}^N \|c_n - c_{n-1}\|. \end{aligned} \tag{54}$$

We use estimate (34) and (50) to get

$$\|\eta\|_{L^\infty(L^2)} \le Q h^2 \|u\|_{L^\infty(H^2)} \le Q h^2 \|u\|_{W^{1,1}(H^2)} \le Q Q^* h^2. \tag{55}$$

We apply Lemmas 1,2 and 3 in the appendix to obtain

$$\begin{aligned} &\left\|\check{\mathbb{D}}_\tau^\omega \eta\right\|_{\hat{L}^1(L^2)} + \|E\|_{\hat{L}^1(L^2)} + \|R\|_{\hat{L}^1(L^2)} + \|S\|_{\hat{L}^1(L^2)} + \|C - c\|_{\hat{L}^1(L^2)} \\ &\quad \le Q Q^* \left(\tau + h^2 + \sigma^2\right). \end{aligned} \tag{56}$$

We use estimate (33) to directly bound

$$\tau \sum_{n=1}^N \|c_n - c_{n-1}\| \le Q\tau \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\partial_t c\| dt \le Q\tau \|c\|_{H^1(L^2)} \le Q Q^* \tau. \tag{57}$$

We use estimates (34) and (50) to obtain

$$\|\delta_\tau \eta\|_{\hat{L}^1(L^2)} = \sum_{n=1}^N \left\| \int_{t_{n-1}}^{t_n} \partial_t \eta dt \right\| \le Q \|\Pi_h \partial_t u - \partial_t u\|_{L^1(L^2)} \le Q Q^* h^2. \tag{58}$$

We incorporate estimates (56)–(58) and Lemma 3 into (54) and combine the resulting estimate with (55) to prove the estimate (51) for $U - u$.

We subtract Equation (41) from scheme (44) to conclude that for any $\chi \in S_h$

$$\begin{aligned}&\left(-\delta_\tau(Z_n - z_n) + k^R\hat{\mathbb{D}}_\tau^{\alpha^*}\left(\omega_{n-1}^m(Z_{n-1} - z_{n-1})\right), \chi\right)\\&\quad + (\boldsymbol{K}\nabla(Z_{n-1} - z_{n-1}), \nabla\chi)\\&= (U_n - u_d(\cdot, t_n) - u_{n-1} + u_d(\cdot, t_{n-1}), \chi) + \left(k(\hat{R}_{n-1} + \hat{S}_{n-1}) + \hat{E}_{n-1}, \chi\right).\end{aligned} \tag{59}$$

We similarly decompose $Z_n - z_n = \hat{\xi}_n + \hat{\eta}_n$ for $n = 0, 1, \ldots, N-1$, where $\hat{\xi}_n = Z_n - \Pi_h z_n \in S_h$ and $\hat{\eta}(\boldsymbol{x}, t) := \Pi_h z(\boldsymbol{x}, t) - z(\boldsymbol{x}, t)$. We rewrite Equation (59) in terms of $\hat{\xi}$ and $\hat{\eta}$ for $n = N, N-1, \ldots, 1$ and $\chi \in S_h$ as follows

$$\begin{aligned}&\left(-\delta_\tau\hat{\xi}_n + k^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m\hat{\xi}_{n-1}), \chi\right) + \left(\boldsymbol{K}\nabla\hat{\xi}_{n-1}, \nabla\chi\right)\\&\quad = \left(\delta_\tau\hat{\eta}_n - k^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m\hat{\eta}_{n-1}), \chi\right) + (U_n - u_n, \chi) + (u_n - u_{n-1}, \chi)\\&\qquad + (u_d(\cdot, t_{n-1}) - u_d(\cdot, t_n), \chi) + \left(k(\hat{R}_{n-1} + \hat{S}_{n-1}) + \hat{E}_{n-1}, \chi\right).\end{aligned} \tag{60}$$

We apply Theorem 7 to Equation (60) to obtain the following estimate

$$\begin{aligned}\|\hat{\xi}\|_{\hat{L}^\infty(L^2)} \leq{}& Q(\|\delta_\tau\hat{\eta}\|_{\hat{L}^1(L^2)} + \left\|{}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega\hat{\eta})\right\|_{\hat{L}^1(L^2)} + \|\hat{E}\|_{\hat{L}^1(L^2)}\\&+ \|\hat{R}\|_{\hat{L}^1(L^2)} + \|\hat{S}\|_{\hat{L}^1(L^2)} + \|U - u\|_{\hat{L}^1(L^2)})\\&+ Q\tau\sum_{n=1}^N(\|u_n - u_{n-1}\| + \|u_d(\cdot, t_n) - u_d(\cdot, t_{n-1})\|).\end{aligned}$$

$\|\hat{E}\|_{\hat{L}^1(L^2)}$, $\|\hat{R}\|_{\hat{L}^1(L^2)}$, $\|\hat{S}\|_{\hat{L}^1(L^2)}$, and $\|{}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega\hat{\eta})\|_{\hat{L}^1(L^2)}$ are bounded in Lemmas 1 and 2. $\|\delta_\tau\hat{\eta}\|_{\hat{L}^1(L^2)}$ was bounded in (58), $\|U - u\|_{\hat{L}^1(L^2)}$ already estimated. The last term can be estimated as (57). We put these estimates into the preceding inequality and bound $\|\hat{\eta}\|_{L^\infty(L^2)}$ as in (55) to prove the estimate for $Z - z$ in (51).

To bound $C - c\|_{\hat{L}^\infty(L^2)}$, we use (16) and (45) to find that for $1 \leq n \leq N$

$$\left|c_{n-1}(\boldsymbol{x}) - C_{n-1}(\boldsymbol{x})\right| = \left|\mathcal{P}(-z_{n-1}(\boldsymbol{x})/\gamma) - \mathcal{P}(-Z_{n-1}(\boldsymbol{x})/\gamma)\right|. \tag{61}$$

If both $-z_{n-1}(\boldsymbol{x})/\gamma, -Z_{n-1}(\boldsymbol{x})/\gamma \in [l, r]$, then

$$\left|c_{n-1}(\boldsymbol{x}) - C_{n-1}(\boldsymbol{x})\right| = \left|z_{n-1}(\boldsymbol{x}) - Z_{n-1}(\boldsymbol{x})\right|/\gamma.$$

Otherwise, say $-z_{n-1}/\gamma \leq l$ and $-Z_{n-1}/\gamma \geq r$, we have from (61)

$$\left|c_{n-1}(\boldsymbol{x}) - C_{n-1}(\boldsymbol{x})\right| = r - l \leq \left|z_{n-1}(\boldsymbol{x}) - Z_{n-1}(\boldsymbol{x})\right|/\gamma.$$

We similarly bound $|c_{n-1} - C_{n-1}|$ by $|z_{n-1} - Z_{n-1}|/\gamma$ for other cases, and obtain

$$\|C - c\|_{\hat{L}^\infty(L^2)} \leq Q\|Z - z\|_{\hat{L}^\infty(L^2)} \leq QQ^*(\tau + h^2 + \sigma^2).$$

We thus complete the proof of (51). ∎

7 | NUMERICAL EXPERIMENTS

We carry out numerical experiments to investigate the performance of the discretization of the optimal control model by measuring its convergence rate υ with respect to the time step size τ, the convergence rate ι with respect to the spatial mesh size h, and the convergence rate κ with respect to the quadrature mesh size σ in discretizing the distribute-order differential operator. A uniform spatial partition is used in all the experiments.

7.1 | The approximation to the optimal control (3) and (2) in one space dimension

In the numerical experiments the data are as follows: $\Omega = (0,1)$, $[0,T] = [0,1]$, $k = 1$, $K = 0.01$, $l = 0.2$, $r = 0.3$, $\gamma = 1$, $f = 1$, $u_d = 1$ and $\omega(\alpha,t) = (t+4)\alpha^{t+3}/0.8^{t+4}$ on supp $\omega = [0,0.8]\times[0,1]$. As the closed-form analytical solution to the problem is not available, we use the numerical solution computed with $(\tau_f, h_f, \sigma_f) = (1/720, 1/120, 1/120)$ as the reference solution to test the temporal convergence rates υ, $(\tau_f, h_f, \sigma_f) = (1/720, 1/360, 1/120)$ for the spatial convergence rates ι, and $(\tau_f, h_f, \sigma_f) = (1/720, 1/120, 1/360)$ for the convergence rates κ of the quadrature error. When measuring υ, we adopt the same mesh sizes for h and σ as used for the reference solution. We similarly measure ι and κ. We present the numerical results in Table 1 and observe second-order accuracy on h and σ and first-order convergence on τ as proved in Theorem 8.

TABLE 1 Accuracy of the discretization of the one-dimensional optimal control in Section 7.1.

τ	1/8	1/16	1/24	1/32	υ
$\|c-C\|_{\hat{L}^\infty(L^2)}$	1.92E−02	9.57E−03	6.26E−03	4.29E−03	1.00
$\|u-U\|_{\hat{L}^\infty(L^2)}$	1.38E−02	7.21E−03	4.87E−03	3.24E−03	0.96
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	2.37E−02	1.21E−02	8.05E−03	5.33E−03	0.99
h	1/60	1/72	1/90	1/120	ι
$\|c-C\|_{\hat{L}^\infty(L^2)}$	4.66E−04	3.73E−04	2.10E−04	1.16E−04	2.07
$\|u-U\|_{\hat{L}^\infty(L^2)}$	1.28E−03	8.98E−04	5.73E−04	3.12E−04	2.04
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	9.37E−04	6.58E−04	4.22E−04	2.31E−04	2.02
σ	1/30	1/60	1/90	1/120	κ
$\|c-C\|_{\hat{L}^\infty(L^2)}$	8.90E−05	2.18E−05	9.33E−06	4.97E−06	2.08
$\|u-U\|_{\hat{L}^\infty(L^2)}$	7.09E−04	1.74E−04	7.45E−05	3.97E−05	2.07
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	1.23E−04	3.00E−05	1.29E−05	6.85E−06	2.08

TABLE 2 Accuracy of the discretization of the two-dimensional optimal control in Section 7.2.

τ	1/16	1/32	1/64	1/128	υ
$\|c-C\|_{\hat{L}^\infty(L^2)}$	1.18E−02	6.31E−03	3.21E−03	1.63E−03	0.95
$\|u-U\|_{\hat{L}^\infty(L^2)}$	1.59E−02	8.05E−03	4.04E−03	2.02E−03	0.99
$\|z-Z\hat{L}^\infty(L^2)$	2.20E−02	1.10E−02	5.52E−03	2.76E−03	1.00
h	1/8	1/16	1/24	1/32	ι
$\|c-C\|_{\hat{L}^\infty(L^2)}$	6.10E−03	1.63E−03	7.16E−04	4.08E−04	1.95
$\|u-U\|_{\hat{L}^\infty(L^2)}$	7.74E−03	1.95E−03	8.69E−04	4.89E−04	1.99
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	1.07E−02	2.68E−03	1.19E−03	6.70E−04	2.00
σ	1/10	1/20	1/30	1/40	κ
$\|c-C\|_{\hat{L}^\infty(L^2)}$	4.08E−03	1.04E−03	4.61E−04	2.58E−04	1.99
$\|u-U\|_{\hat{L}^\infty(L^2)}$	5.15E−03	1.28E−03	5.57E−04	3.04E−04	2.04
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	7.05E−03	1.75E−03	7.63E−04	4.19E−04	2.03

7.2 | The approximation to the optimal control (3) and (2) in two space dimensions

In this set of numerical experiments, we let $\Omega = (0,1)^2$, $[0,T] = [0,1]$, $k = 1$, $\boldsymbol{K} = \text{diag}(0.01, 0.01)$, $\omega(\alpha,t) = (1+2\alpha t)/(0.3+0.3^2 t)$ on supp $\omega = [0,0.3]\times[0,1]$. The solutions are chosen to be $u(\boldsymbol{x},t) = t^{2-\alpha(0)}\sin(\pi x_1)\sin(\pi x_2)$, $z(\boldsymbol{x},t) = (T-t)^{2-\alpha(T)}\sin(\pi x_1)\sin(\pi x_2)$, $c(\boldsymbol{x},t) = \max\{l, \min\{-z(\boldsymbol{x},t)/\gamma, r\}\}$ with $l = -0.2$ and $r = -0.1$, $\gamma = 1$, and f and u_d calculated accordingly. Mesh sizes of $h = 1/64$ and $\sigma = 1/100$ are used to measure the temporal convergence rate υ, while $\tau = 2h^2$ and $\sigma = 1/100$ are used to measure the spatial convergence rate ι. To measure the convergence rates κ of the quadrature error, we apply $\tau = 2\sigma^2$ and $h = 1/30$. We present the numerical results in Table 2, which again substantiate the second-order accuracy of ι and κ and first-order convergence of υ as proved in Theorem 8.

8 | CONCLUDING REMARKS

In this paper, we analyzed the well-posedness, smoothing properties and numerical discretization of the optimal control problem governed by a time two-scale variably distributed-order tFDE in heterogeneous porous media, where the integer-order term refers to the normal diffusion, and the variably distributed-order fractional derivative is employed to describe the anomalous diffusion caused by, for example, the absorption to heterogeneous porous matrix. Therefore, the considered model enjoys more powerful modeling capacities than the single-term variably distributed-order fractional model or the conventional integer-order model. However, in the current analysis framework, the first-order time derivative is crucial to derive the well-posedness and error estimates of the time two-scale variably distributed-order fractional optimal control problem, due to the limited smooth property of the pure variably distributed-order fractional derivative. We will deeply investigate the mathematical analysis of the variably distributed-order fractional optimal control problem without the first-order time derivative in our future work.

ACKNOWLEDGMENTS

This research was supported by the Taishan Scholars Program of Shandong Province of China (Nos. tsqn201909044, tsqn202306083), the National Natural Science Foundation of China under Grants (Nos. 12001337, 12301555, 12271303), the National Science Foundation under Grants DMS-2012291 and DMS-2245097, the Major Fundamental Research Project of Shandong Province of China (No. ZR2023ZD33), and the Fundamental Research Funds for the Central Universities (No. 2022110089). The authors thank the reviewers for their valuable comments and suggestions, which greatly improved the quality of this paper.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interests.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Hong Wang https://orcid.org/0000-0002-1668-7332

REFERENCES

[1] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, Elsevier, San Diego, 2003.
[2] G. Akrivis, B. Li, and C. Lubich, *Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations*, Math. Comput. 86 (2017), 1527–1552.
[3] A. Alikhanov, *Numerical methods of solutions of boundary value problems for the multi-term variable-distributed order diffusion equation*, Appl. Math. Comput. 268 (2015), 12–22.
[4] H. Antil, E. Otárola, and A. Salgado, *A space-time fractional optimal control problem: Analysis and discretization*, SIAM J. Control. Optim. 54 (2016), 1295–1328.
[5] H. Antil, E. Otárola, and A. Salgado, *Some applications of weighted norm inequalities to the error analysis of PDE-constrained optimization problems*, IMA J. Numer. Anal. 38 (2018), 852–883.
[6] M. Athans and P. L. Falb, *Optimal control: An introduction to the theory and its applications*, Dover Publications, Mineola, NY, 2006.
[7] P. Baveye, P. Vandevivere, B. L. Hoyle, P. C. DeLeo, and D. S. de Lozada, *Environmental impact and mechanisms of the biological clogging of saturated soils and aquifer materials*, Crit. Rev. Env. Sci. Tec. 28 (2006), 123–191.
[8] J. Bear, *Dynamics of fluids in porous media*, Elsevier, New York, 1972.
[9] D. Benson, R. Schumer, M. M. Meerschaert, and S. W. Wheatcraft, *Fractional dispersion, Lévy motions, and the MADE tracer tests*, Transp Porous Med. 42 (2001), 211–240.
[10] L. Caffarelli and L. Silvestre, *An extension problem related to the fractional Laplacian*, Comm. Partial Differ. Equ. 32 (2007), 1245–1260.
[11] A. Chechkin, R. Goreno, I. Sokolov, and V. Gonchar, *Distributed order time fractional diffusion equation*, Fract. Calc. Appl. Anal. 6 (2003), 259280.
[12] P. Chen, A. Quarteroni, and G. Rozza, *Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by stokes equations*, Numer. Math. 133 (2015), 67–102.
[13] Y. Chen, *Superconvergence of quadratic optimal control problems by triangular mixed finite element methods*, Int. J. Numer. Methods Eng. 75 (2008), 881–898.
[14] M. D'Elia, C. Glusa, and E. Otarola, *A priori error estimates for the optimal control of the integral fractional Laplacian*, SIAM J. Control. Optim. 57 (2019), 2775–2798.
[15] M. D'Elia and M. Gunzburger, *Optimal distributed control of nonlocal steady diffusion problems*, SIAM J. Control. Optim. 52 (2014), no. 1, 243–273.
[16] K. Diethelm and N. Ford, *Numerical analysis for distributed-order differential equations*, J. Comput. Appl. Math. 225 (2009), 96–104.
[17] V. J. Ervin, *Regularity of the solution to fractional diffusion, advection, reaction equations in weighted Sobolev spaces*, J. Diff. Equ. 278 (2021), 294–325.
[18] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, V 19, American Mathematical Society, Rhode Island, 1998.
[19] W. Fan, X. Hu, and S. Zhu, *Numerical reconstruction of a discontinuous diffusive coefficient in variable-order time-fractional subdiffusion*, J. Sci. Comput. 96 (2023), no. 1, 13.
[20] L. Gandossi and U. Von Estorff, *An overview of hydraulic fracturing and other formation stimulation technologies for shale gas production, scientific and technical research reports, Joint Research Centre of the European Commission*, Publications Office of the European Union, Luxembourg, 2015. https://doi.org/10.2790/379646.
[21] G. Gao, A. A. Alikhanov, and Z. Sun, *The temporal second order difference schemes based on the interpolation approximation for solving the time multi-term and distributed-order fractional sub-diffusion equations*, J. Sci. Comput. 73 (2017), 93–121.
[22] R. Garrappa, A. Giusti, and F. Mainardi, *Variable-order fractional calculus: A change of perspective*, Commun. Nonlinear Sci. Numer. Simul. 102 (2021), 105904.
[23] W. Gong, M. Hinze, and Z. Zhou, *Space-time finite element approximation of parabolic optimal control problems*, J. Numer. Math. 20 (2012), 111–145.
[24] R. Gorenflo, Y. Luchko, and M. Stojanović, *Fundamental solution of a distributed order time-fractional diffusion-wave equation as probability density*, Fract. Calc. Appl. Anal. 16 (2013), 297–316.
[25] M. Gunzburger and A. Kunoth, *Space-time adaptive wavelet methods for control problems constrained by parabolic evolution equations*, SIAM J. Contr. Optim. 49 (2011), 1150–1170.
[26] M. Gunzburger and J. Wang, *Error analysis of fully discrete finite element approximations to an optimal control problem governed by a time-fractional PDE*, SIAM J. Contr. Optim. 57 (2019), 241–263.
[27] K. Ito and K. Kunisch, *Semismooth Newton methods for time-optimal control for a class of ODEs*, SIAM J. Control. Optim. 48 (2010), 3997–4013.

[28] B. Jin, B. Li, and Z. Zhou, *Correction of high-order BDF convolution quadrature for fractional evolution equations*, SIAM J. Sci. Comput. 39 (2017), A3129–A3152.

[29] B. Jin, B. Li, and Z. Zhou, *Pointwise-in-time error estimates for an optimal control problem with subdiffusion constraint*, IMA J. Numer. Anal. 40 (2020), 377–404.

[30] E. Kharazmi, M. Zayernouri, and G. Karniadakis, *Petrov-Galerkin and spectral collocation methods for distributed order differential equations*, SIAM J. Sci. Comput. 39 (2017), A1003–A1037.

[31] K. Kunisch and B. Vexler, *Constrained Dirichlet boundary control in L^2 for a class of evolution equations*, SIAM J. Control. Optim. 46 (2007), 1726–1753.

[32] I. Lasiecka, *Ritz-Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary conditions*, SIAM J. Control. Optim. 22 (1984), 477–500.

[33] Y. Li, H. Wang, and X. Zheng, *Analysis and simulation of optimal control for a two-time- scale fractional advection-diffusion-reaction equation with space-time-dependent order and coefficients*, Multiscale Model. Sim. 21 (2023), no. 4, 1690–1716.

[34] Z. Li, Y. Luchko, and M. Yamamoto, *Analyticity of solutions to a distributed order time-fractional diffusion equation and its application to an inverse problem*, Comput. Math. Appl. 73 (2017), 1041–1052.

[35] H. Liao, D. Li, and J. Zhang, *Sharp error estimate of the nonuniform L1 formula for linear reaction-subdiffusion equations*, SIAM J. Numer. Anal. 56 (2018), no. 2, 1112–1133.

[36] H. Liao, W. McLean, and J. Zhang, *A discrete Gronwall inequality with applications to numerical schemes for subdiffusion problems*, SIAM J. Numer. Anal. 57 (2019), no. 1, 218–237.

[37] Y. Lin and C. Xu, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys. 225 (2007), 1533–1552.

[38] C. F. Lorenzo and T. T. Hartley, *Variable order and distributed order fractional operators*, Nonlinear Dyn. 29 (2002), 57–98.

[39] C. Lubich, T. Sloan, and V. Thomée, *Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term*, Math. Comput. 65 (1996), 1–17.

[40] Y. Luchko, *Boundary value problems for the generalized time-fractional diffusion equation of distributed order*, Fract. Calc. Appl. Anal. 12 (2009), 409–422.

[41] F. Mainardi, G. Pagnini, and R. Gorenflo, *Some aspects of fractional diffusion equations of single and distributed order*, Appl. Math. Comput. 187 (2007), 295–305.

[42] M. M. Meerschaert and A. Sikorskii, *Stochastic models for fractional calculus*, De Gruyter, Berlin, 2011.

[43] D. Meidner and B. Vexler, *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 (2008), 1301–1329.

[44] R. Metzler and J. Klafter, *The random walk's guide to anomalous diffusion: A fractional dynamics approach*, Phys. Rep. 339 (2000), 1–77.

[45] G. Mophou, *Optimal control of fractional diffusion equation*, Comput. Math. Appl. 61 (2011), 68–78.

[46] P. Neittaanmaki, D. Tiba, and M. Dekker, *Optimal control of nonlinear parabolic systems: Theory, Algorithms and Applications*, Marcel Dekker, New York, 1994.

[47] R. Nochetto, E. Otárola, and A. Salgado, *A PDE approach to space-time fractional parabolic problems*, SIAM J. Numer. Anal. 54 (2016), 848–873.

[48] I. Podlubny, *Fractional differential equations*, Academic Press, New York, 1999.

[49] K. Sakamoto and M. Yamamoto, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems*, J. Math. Anal. Appl. 382 (2011), 426–447.

[50] R. Schumer, D. A. Benson, M. M. Meerschaert, and B. Baeumer, *Fractal mobile/immobile solute transport*, Water Resour. Res. 39 (2003), 1–12.

[51] M. Stynes, E. O'Riordan, and J. L. Gracia, *Error analysis of a finite difference method on graded mesh for a time-fractional diffusion equation*, SIAM J. Numer. Anal. 55 (2017), 1057–1079.

[52] V. Thomée, *Galerkin finite element methods for parabolic problems*, Lecture Notes in Mathematics, Vol 1054, Springer-Verlag, New York, 1984.

[53] Z. Yang, X. Zheng, and H. Wang, *A variably distributed-order time-fractional diffusion equation: Analysis and approximation*, Comput. Meth. Appl. Mech. Engrg. 367 (2020), 113118.

[54] M. Zaky and J. Machado, *On the formulation and numerical simulation of distributed-order fractional optimal control problems*, Commun. Nonlinear Sci. Numer. Simul. 52 (2017), 177–189.

[55] M. Zayernouri, M. Ainsworth, and G. Karniadakis, *A unified Petrov-Galerkin spectral method for fractional PDEs*, Comput. Meth. Appl. Mech. Eng. 283 (2015a), 1545–1569.

[56] M. Zayernouri, M. Ainsworth, and G. Karniadakis, *Tempered fractional Sturm-Liouville eigenproblems*, SIAM J. Sci. Comput. 37 (2015b), A1777–A1800.

[57] X. Zheng and H. Wang, *A hidden-memory variable-order fractional optimal control model: Analysis and approximation*, SIAM J. Control. Optim. 59 (2021), 1851–1880.

[58] X. Zheng and H. Wang, *Discretization and analysis of an optimal control of a variable-order fractional diffusion equation with pointwise constraints*, J. Sci. Comput. 91 (2022), 56.

[59] Z. Zhou and W. Gong, *Finite element approximation of optimal control problems governed by time fractional diffusion equation*, Comput. Math. Appl. 71 (2016), 301–318.

How to cite this article: X. Zheng, H. Liu, H. Wang, and X. Guo, *Optimal control of variably distributed-order time-fractional diffusion equation: Analysis and computation*, Numer. Methods Partial Differ. Eq. (2024), e23134. https://doi.org/10.1002/num.23134

APPENDIX A

We prove Lemmas 1,2 and 3 used earlier.

A.1 | Truncation and quadrature errors

Lemma 1. *Under assumptions (a)–(c), E_n, R_n, S_n, $\hat{E}_n$, $\hat{R}_n$ and $\hat{S}_n$ in (37) and (40) are bounded by*

$$\|E\|_{\hat{L}^1(L^2)} + \|R\|_{\hat{L}^1(L^2)} + \|S\|_{L^1(L^2)} + \|\hat{E}\|_{\hat{L}^1(L^2)} + \|\hat{R}\|_{\hat{L}^1(L^2)} + \|\hat{S}\|_{\hat{L}^1(L^2)} \leq QQ^*(\tau+\sigma^2). \tag{A1}$$

Here Q^* is given in (51) and Q is independent of N, M, and h.

Proof. We use (35) and (47) to bound $\hat{E}_{k-1}$ in (39) by

$$\|\hat{E}\|_{\hat{L}^1(L^2)} \leq \tau\|\partial_t^2 z\|_{L^1(L^2)} \leq QQ^*\tau.$$

We use (35), assumption (a), the boundedness of $\partial_\alpha(1/\Gamma(1-\alpha))$ on $[0,\alpha^*]$, and the Sobolev embedding $W(0,T)^{2,1} \hookrightarrow W^{1,\infty}(0,T)$ to bound $\|\hat{S}_{n-1}\|$ by

$$\begin{aligned}
\|\hat{S}_{n-1}\| &\leq Q\sigma^2\left\|\sup_{\alpha\in[0,\alpha^*]}\left|\partial_\alpha^2\left({}_tI_T^{1-\alpha}\partial_t(\omega z)|_{t=t_{n-1}}\right)\right|\right\| \\
&\leq Q\sigma^2\left\|\sup_{\alpha\in[0,\alpha^*]}\int_{t_{n-1}}^T \frac{|\ln(s-t_{n-1})|^2|\partial_s z(\boldsymbol{x},s)|}{(s-t_{n-1})^\alpha}d\alpha\right\| \\
&\leq Q\sigma^2\left\|\int_{t_{n-1}}^T \frac{|\partial_s z(\boldsymbol{x},s)|}{(s-t_{n-1})^{\alpha^*+\varepsilon}}d\alpha\right\| \leq Q\sigma^2\|\partial_t z\|_{L^\infty(L^2)} \leq QM\sigma^2
\end{aligned}$$

for some $0<\varepsilon<1-\alpha^*$, which immediately yields the estimate of $\|\hat{S}\|_{\hat{L}^1(L^2)}$.

We bound $\hat{R}_{n-1}^m$ by using its expression in (40)

$$\begin{aligned}
\left|\hat{R}_{n-1}^m\right| &\leq \sum_{k=n}^N \int_{t_{k-1}}^{t_k} \frac{(s-t_{n-1})^{-\alpha_m}}{\Gamma(1-\alpha_m)}\left[\int_{t_{k-1}}^{t_k}|\partial_\theta^2(\omega(\alpha_m,\theta)z(\boldsymbol{x},\theta))|d\theta\right]ds \\
&= \sum_{k=n}^N \int_{t_{k-1}}^{t_k}|\partial_\theta^2(\omega(\alpha_m,\theta)z(\boldsymbol{x},\theta))|d\theta\left(\frac{(t_k-t_{n-1})^{1-\alpha_m}-(t_k-t_n)^{1-\alpha_m}}{\Gamma(2-\alpha_m)}\right).
\end{aligned}$$

We then use (40) and (47) to bound

$$\begin{aligned}
\|\hat{R}\|_{\hat{L}^1(L^2)} &\leq \tau\sum_{n=1}^{N}\sum_{m=0}^{M}\sigma_m\|\hat{R}_{n-1}^m\| \\
&\leq \tau\sum_{m=0}^{M}\sigma_m\sum_{n=1}^{N}\sum_{k=n}^{N}\int_{t_{k-1}}^{t_k}\|\partial_\theta^2(\omega z)\|d\theta\left(\frac{(t_k-t_{n-1})^{1-\alpha_m}-(t_k-t_n)^{1-\alpha_m}}{\Gamma(2-\alpha_m)}\right) \\
&= \tau\sum_{m=0}^{M}\sigma_m\sum_{k=1}^{N}\int_{t_{k-1}}^{t_k}\|\partial_\theta^2(\omega z)\|d\theta\sum_{n=1}^{k}\left(\frac{(t_k-t_{n-1})^{1-\alpha_m}-(t_k-t_n)^{1-\alpha_m}}{\Gamma(2-\alpha_m)}\right) \\
&\leq Q\tau\|z\|_{W^{2,1}(L^2)} \leq QQ^*\tau.
\end{aligned}$$

All remaining terms in (A1) can be bounded similarly and the proofs are omitted. ■

Lemma 2. *Under assumptions (a)–(c), the following estimate holds*

$$\tau\sum_{n=1}^{N}\left(\left\|\check{\mathbb{D}}_\tau^{\omega}\eta_n\right\|+\left\|{}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m\hat{\eta}_{n-1})\right\|\right)\leq QQ^*h^2.$$

Here Q is independent of M, N, and h.

Proof. We use the expression of ${}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m\hat{\eta}_{n-1})$ in (40) and $\hat{\eta}_N=0$ to obtain

$$\begin{aligned}
&\tau\sum_{n=1}^{N}\left\|{}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega_{n-1}^m\hat{\eta}_{n-1})\right\| \\
&\quad\leq Q\sum_{m=0}^{M}\sigma_m\sum_{n=1}^{N}\sum_{k=n}^{N}\int_{t_{k-1}}^{t_k}\|\partial_t(\omega^m\hat{\eta})\|dt\int_{t_{k-1}}^{t_k}\frac{ds}{\Gamma(1-\alpha_m)(s-t_{n-1})^{\alpha_m}} \\
&\quad= Q\sum_{m=0}^{M}\sigma_m\sum_{k=1}^{N}\int_{t_{k-1}}^{t_k}\|\partial_t(\omega^m\hat{\eta})\|dt\sum_{n=1}^{k}\frac{(t_k-t_{n-1})^{1-\alpha_m}-(t_k-t_n)^{1-\alpha_m}}{\Gamma(2-\alpha_m)} \\
&\quad\leq Qh^2\|\partial_t z\|_{L^1(H^2)}\leq QQ^*h^2.
\end{aligned}$$

The first term can be estimated similarly, and the estimate is omitted. ■

A.2 | Approximation to the control variable

Lemma 3. *Under assumptions (a)–(c), the following estimate holds*

$$\|C-c\|_{\hat{L}^2(L^2)}\leq QQ^*(\tau+h^2+\sigma^2).$$

Proof. Let $U(c)$ satisfy $U_0(c)=0$ and for $1\leq n\leq N$,

$$\left(\delta_\tau U_n(c)+k\check{\mathbb{D}}_\tau^{\omega}U_n(c)+\mathcal{B}U_n(c),\chi\right)=(f_n+c_{n-1},\chi),\quad\forall\chi\in S_h, \tag{A2}$$

and $Z(U(c))$ satisfy $Z_N(U(c))=0$ and for $1\leq n\leq N$ and $\chi\in S_h$

$$\begin{aligned}
&\left(-\delta_\tau Z_n(U(c))+k{}^R\hat{\mathbb{D}}_\tau^{\alpha^*}\left(\omega_{n-1}^m Z_{n-1}(U(c))\right)+\mathcal{B}Z_{n-1}(U(c)),\chi\right) \\
&\quad=(U_n(c)-u_d(\cdot,t_n),\chi).
\end{aligned}$$

Following the standard procedure as [57, lemma 9.4], we reach

$$\|c - C\|_{\hat{L}^2(L^2)} \leq \frac{1}{\gamma}\|Z(U(c)) - z\|_{\hat{L}^2(L^2)} \leq \frac{1}{\gamma}\left(\|z - Z(u)\| + \|Z(u) - Z(U(c))\|\right). \tag{A3}$$

Here $Z(u)$ satisfies $Z_N(u) = 0$ and for $1 \leq n \leq N$ and $\chi \in S_h$

$$\left(-\delta_\tau Z_n(u) + k^R\hat{\mathbb{D}}_\tau^{\alpha^*}\left(\omega_{n-1}^m Z_{n-1}(u)\right) + \mathcal{B}Z_{n-1}(u), \chi\right) = (u_n - u_d(\cdot, t_n), \chi). \tag{A4}$$

To bound the first term on the right side of (A3), we split $Z_n(u) - z_n = \check{\xi}_n + \hat{\eta}_n$ with $\check{\xi}_n = Z_n(u) - \Pi_h z_n \in S_h$. We subtract (41) from (A4) and express the equation in terms of $\check{\xi}$ and $\hat{\eta}$ as

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

We apply Theorem 7 to the equation to arrive at the following estimate

$$\begin{aligned}\|\check{\xi}\|_{\hat{L}^\infty(L^2)} \leq Q\Big(&\|\delta_\tau\hat{\eta}\|_{\hat{L}^1(L^2)} + \|{}^R\hat{\mathbb{D}}_\tau^{\alpha^*}(\omega^m\hat{\eta})\|_{\hat{L}^1(L^2)} + \|\hat{E}\|_{\hat{L}^1(L^2)} + \|\hat{R}\|_{\hat{L}^1(L^2)}\\ &+ \|\hat{S}\|_{\hat{L}^1(L^2)}\Big) + Q\tau\sum_{n=1}^{N}(\|u_n - u_{n-1}\| + \|u_d(\cdot, t_n) - u_d(\cdot, t_{n-1})\|).\end{aligned}$$

The terms in the summation can be bounded as in (57). The remaining terms on the right side were bounded in (58) and Lemmas 1 and 2. We combine these estimates with that of $\|\hat{\eta}\|_{L^\infty(L^2)}$ to get

$$\|z - Z(u)\|_{\hat{L}^\infty(L^2)} \leq QQ^*(\tau + h^2 + \sigma^2). \tag{A5}$$

The second term on the right side of (A3) can be estimated by subtracting (A4) from (9.2) and employing the stability estimate in Theorem 7

$$\|Z(u) - Z(U(c))\|_{\hat{L}^\infty(L^2)} \leq Q\|u - U(c)\|_{\hat{L}^1(L^2)}.$$

The right side can be bounded as in (A5). ∎