MULTISCALE MODEL. SIMUL.
Vol. 21, No. 4, pp. 1690–1716

ANALYSIS AND SIMULATION OF OPTIMAL CONTROL FOR A TWO-TIME-SCALE FRACTIONAL ADVECTION-DIFFUSION-REACTION EQUATION WITH SPACE-TIME-DEPENDENT ORDER AND COEFFICIENTS*

YIQUN LI†, HONG WANG†, AND XIANGCHENG ZHENG‡

Abstract. We investigate an optimal control model with pointwise constraints governed by a two-time-scale time-fractional advection-diffusion-reaction equation with space-time-dependent fractional order and coefficients, which describes, e.g., the contaminant in groundwater under various transport scales or miscible displacement of hydrocarbon by injected fluid through heterogeneous porous media. To accommodate for the effects of complex fractional order and coefficients, an auxiliary equation method is proposed, along with the Fredholm alternative for compact operators, to analyze the well-posedness of the state equation. Additionally, a bootstrapping argument is utilized to progressively improve the solution regularity through a carefully designed pathway, leading to the maximal regularity estimates. Subsequently, we analyze the adjoint equation derived from the first-order optimality condition, which requires more subtle treatments due to the presence of hidden-memory variable-order fractional operators. Based on these findings, we ultimately analyze the well-posedness, first-order optimality conditions and maximal regularity estimates for the optimal control problem, and we conduct numerical experiments to investigate its behavior in potential applications.

Key words. two-time-scale time-fractional, optimal control, variable order, space-time dependent, well-posedness, regularity

MSC codes. 35R11, 65M15, 65M60

DOI. 10.1137/23M1573537

1. Introduction.

1.1. Problem formulation.

Optimal control problems are widely used in various fields and, as a result, have been the subject of extensive research in the literature [2, 3, 22, 25, 28, 30, 49]. In particular, optimal control problems governed by time-fractional partial differential equations are attracting increasing attention since they have been shown to provide competitive descriptions of challenging physical phenomena such as anomalously diffusive transport [2, 14, 26, 33, 45, 46, 48, 50, 67]. Let u be the concentration of the component of interest in the fluid mixture. In the context of enhanced oil recovery, u represents the concentration of the injected fluid, which consists of surfactant or other chemicals that are mixed with water and is injected into the oil reservoir to formulate a fully miscible fluid phase aiming at displacing the resident oil out of the oil reservoir [4, 20]. In the context of contaminant transport and remediation, u represents the concentration of the contaminant solute that is present in groundwater [8, 6, 45], and we aim at optimizing the pollution sources so that the concentration of the contaminant solute reaches an ideal value at minimum

*Received by the editors May 18, 2023; accepted for publication (in revised form) September 8, 2023; published electronically December 6, 2023.
https://doi.org/10.1137/23M1573537
Funding: This work was partially supported by the Taishan Scholars Program of Shandong Province No. tsqn202306083, the National Science Foundation under grant DMS-2012291, and the National Natural Science Foundation of China under grant 12301555.
†Department of Mathematics, University of South Carolina, Columbia, SC 29208 USA (yiqunl@email.sc.edu, hwang@math.sc.edu).
‡Corresponding author. School of Mathematics, Shandong University, Jinan 250100, China (xzheng@sdu.edu.cn).

cost. Let q denote the control variable that adjusts the injection and extraction rate of the injecting fluid from the admissible set

$$U_{ad} := \left\{ q \in L^2(0,T;L^2(\Omega)) : q_* \leq q(\boldsymbol{x},t) \leq q^* \text{ a.e. in } \Omega \times [0,T] \right\} \tag{1.1}$$

for some q_*, $q^* \in \mathbb{R}$ so that we can formulate the following optimal control problem:

$$\min_{q \in U_{ad}, u} J(u,q) = \frac{1}{2}\|u - u_d\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\gamma}{2}\|q\|^2_{L^2(0,T;L^2(\Omega))}, \tag{1.2}$$

where u_d is the prescribed target function and $\gamma > 0$ is a fixed penalty parameter. The optimization problem (1.2) is constrained by a two-time-scale time-fractional advection-diffusion-reaction equation with space-time-dependent fractional order and coefficients

$$\begin{gathered} \partial_t u + k(\boldsymbol{x},t)\partial_t^{\alpha(\boldsymbol{x},t)} u + \mathcal{L}(\boldsymbol{x},t)u = f(\boldsymbol{x},t) + q(\boldsymbol{x},t), \quad (\boldsymbol{x},t) \in \Omega \times (0,T], \\ u(x,0) = 0, \quad \boldsymbol{x} \in \Omega; \quad u(\boldsymbol{x},t) = 0, \quad (\boldsymbol{x},t) \in \partial\Omega \times [0,T]. \end{gathered} \tag{1.3}$$

Here $\Omega \subset \mathbb{R}^d$ $(d = 1,2,3)$ is a convex polygonal domain, $k(\boldsymbol{x},t) \geq 0$ is the partition coefficient, and $\mathcal{L}(\boldsymbol{x},t)$ denotes a second-order elliptic partial differential operator

$$\mathcal{L}(\boldsymbol{x},t)u := -\nabla \cdot \big(\boldsymbol{A}(\boldsymbol{x},t)\nabla u\big) + \boldsymbol{b}(\boldsymbol{x},t) \cdot \nabla u + c(\boldsymbol{x},t)u, \tag{1.4}$$

where $\nabla := (\partial/\partial x_1, \ldots, \partial/\partial x_d)^\top$, $\boldsymbol{A}(\boldsymbol{x},t) := (a_{ij}(\boldsymbol{x},t))_{i,j=1}^d$ is the symmetric diffusivity tensor, $\boldsymbol{b}(\boldsymbol{x},t) := (b_i(\boldsymbol{x},t))_{i=1}^d$ is the fluid velocity field, $c(\boldsymbol{x},t)$ is the reaction coefficient, f is the source and sink, and $\partial_t^{\alpha(\boldsymbol{x},t)}$, with $0 \leq \alpha(\boldsymbol{x},t) \leq \alpha^* < 1$, is the variable-order fractional differential operator defined by [41]

$$\partial_t^{\alpha(\boldsymbol{x},t)} g := I_t^{1-\alpha(\boldsymbol{x},t)} \partial_t g, \quad I_t^{1-\alpha(\boldsymbol{x},t)} g := \int_0^t \frac{g(\boldsymbol{x},s)ds}{\Gamma(1-\alpha(\boldsymbol{x},t))(t-s)^{\alpha(\boldsymbol{x},t)}}. \tag{1.5}$$

1.2. Motivation and contribution. The optimal control model presented in (1.2)–(1.3) faces mathematical challenges that are not typically encountered in traditional fractional optimal control problems. Specifically, the adjoint equation of (1.3) produces a Riemann–Liouville time-fractional equation (6.1) with the so-called hidden-memory variable order $\alpha = \alpha(\boldsymbol{x},s)$ for $s \in [0,t]$ [41, 64]. This equation has more complex properties than (1.5), which makes the analysis challenging. Moreover, due to the space-time-dependent fractional orders and coefficients, existing research methods for constant-order and even variable-order time-fractional problems and their optimal control, such as analytical techniques [15, 50, 51], spectral decomposition methods [57, 65], and solution operator methods [38], are not applicable. Additionally, energy arguments used in, for example, [43, 44] cannot be directly performed, as the variable-order fractional operators lose coercivity.

Motivated by recent works [31, 36, 61] in which the Fredholm alternative is applied to analyze the time-fractional problems with time-dependent coefficients, we adopt this approach to analyze the optimal control problem (1.2), which is constrained by the two-time-scale time-fractional advection-diffusion-reaction transport equation (1.3) with space-time-dependent fractional order, diffusivity coefficient, fluid velocity field, and reaction coefficient. The approach of combining semigroup theory with the Laplace transform in Banach space [38, 66] and the perturbation argument in [34] to analyze problem (1.3), though not impossible, may be extremely complicated. To apply the approach in [31, 36, 61], we need to properly define a compact operator by the compact embedding $H^1(0,T) \to H^{1-\varepsilon}(0,T)$ for $0 < \varepsilon \ll 1$. This in turn forces the domain of the compact operator to be $H^{1-\varepsilon}(0,T)$ in time. However, the variable-order fractional derivative $\partial_t^{\alpha(\boldsymbol{x},t)}u$ in the model (1.3) would require, e.g., $u \in H^1(0,T)$ in time and thus is not well-defined for $u \in H^{1-\varepsilon}(0,T)$. To address this issue, we introduce the auxiliary problem (3.1) with $\partial_t^{\alpha(\boldsymbol{x},t)}u$ in problem (1.3) replaced by a weaker formulation that is well-defined for $u \in H^{1-\varepsilon}(0,T)$ in time and subsequently improve the solution regularity to recover and analyze the original problem (1.3). The adjoint equation of (1.3) arising from the first-order optimality condition is then analyzed with more complicated treatments due to the hidden-memory variable-order operators. Furthermore, we combine maximal regularity estimates for classical parabolic equations [27] with the bootstrapping argument to progressively lift the regularity of the solutions to state and adjoint equations via a carefully designed pathway and finally prove the maximal regularity estimates for the optimal control problem.

The rest of the paper is organized as follows: In section 2, we present multiscale effects of the proposed model to show its novelty. In section 3, we analyze the well-posedness of an auxiliary problem. In section 4, we prove the well-posedness and maximal regularity of the forward problem (1.3) based on that of the auxiliary problem, and we then prove those for the adjoint equation in section 5. In section 6, we prove the well-posedness of the optimal control problem (1.2)–(1.3), derive its first order optimality conditions, and then analyze the maximal regularity of solutions. We then discretize the optimal control model and perform numerical experiments for demonstration in section 7. We address concluding remarks in the last section.

1.3. Spaces and assumptions. Let $C^\mu(\mathcal{I})$ with $0 \le \mu \le 1$ be the space of Hölder continuous functions of index μ on the interval $\mathcal{I}$, and let $L^p(\Omega)$ with $1 \le p \le \infty$ be the Banach space of pth power Lebesgue integrable functions on Ω. For a positive integer m, let $C^m(\Omega)$ and $W^{m,p}(\Omega)$ be the spaces of continuous functions with continuous derivatives up to order m and the Sobolev space of L^p functions with mth weakly derivatives in $L^p(\Omega)$, respectively. Let $H^m(\Omega) = W^{m,2}(\Omega)$, and let $H_0^m(\Omega)$ be the completion of $C_0^\infty(\Omega)$, the space of infinitely differentiable functions with compact support in Ω, in $H^m(\Omega)$. For a noninteger $s \ge 0$, the fractional Sobolev space $H^s(\Omega)$ is defined by interpolation [1]. For a Banach space $\mathcal{X}$, let $W^{m,p}(0,T;\mathcal{X})$ be the space of functions in $W^{m,p}(0,T)$ with respect to $\|\cdot\|_{\mathcal{X}}$. All spaces are equipped with standard norms [1, 19].

For convenience, we may drop the subscript L^2 and the notation Ω in the inner product and Sobolev spaces and norms, and write $W^{m,p}(X)$ for $W^{m,p}(0,T;X)$ when no confusion occurs, e.g., we write $L^2(L^2)$ instead of $L^2(0,T;L^2(\Omega))$ for simplicity. In subsequent sections, we use Q and M to denote positive constants where Q may assume different values at different occurrences.

Throughout the paper, we make the following *assumptions*:

(a) $\alpha \in W^{1,\infty}(L^\infty)$ and $0 \le \alpha(\boldsymbol{x},t) \le \alpha^* < 1$ on $\Omega \times [0,T]$.

(b) $\boldsymbol{A}$ is symmetric and positive definite uniformly on Ω for each $t \in [0,T]$, and $a_{ij} \in W^{1,\infty}(W^{1,\infty}), b_i \in L^\infty(W^{1,\infty}), c \in L^\infty(L^\infty)$ for $1 \le i,j \le d$.

(c) $k \in W^{1,\infty}(L^\infty)$ and $f,\ u_d \in L^2(L^2)$.

(d) $a_{ij}(\boldsymbol{x},\cdot), b_i(\boldsymbol{x},\cdot), c(\boldsymbol{x},\cdot) \in C^\mu[0,T]$ with $\mu > 1/2$ for a.e. $\boldsymbol{x} \in \Omega$ and $1 \le i,j \le d$.

2. Modeling issues. We present novelties of the proposed two-time-scale model (1.3) from the viewpoints of the initial behavior and multiscale effects by means of mean square displacements (MSDs). For simplicity we consider $d = 1$ with $\mathcal{L}(x,t) = A\partial_x^2$, i.e., the pure diffusion case, for model (1.3) throughout this section and compare it with the following single-time-scale time-fractional diffusion equation (tFDE) [5, 54]

$$\partial_t^\alpha u - A\Delta u = 0,\ 0 < \alpha < 1,\ A > 0. \tag{2.1}$$

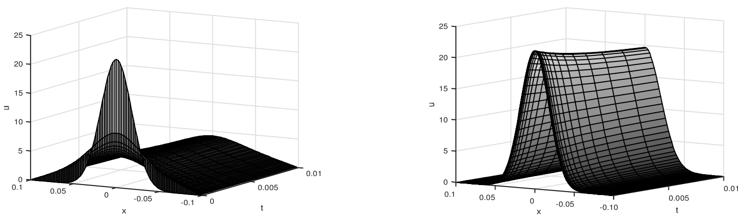


FIG. 2.1. *Solutions to the single-time-scale tFDE* (2.1) *(left) and the two-time-scale tFDE* (1.3) *(right).*

This model has garnered much attention as it admits power-law decaying solutions and is better suited than its integer-order analogue to model physical processes exhibiting power-law memory properties such as anomalously diffusive transport through heterogeneous porous media [7, 9, 12, 13, 16, 17, 24, 32, 39, 45, 46, 59]. However, this tFDE yields solutions with nonphysical initial singularities as it is derived as the diffusion limit of a continuous time random walk when the number of particle jumps tends to infinity and thus only holds for large time [45, 46]. The two-time-scale tFDE (1.3) aims to resolve this issue. In this model, a $k/(k+1)$ portion of total solute mass gets absorbed into the aquifers and undergoes subdiffusive transport modeled by the fractional derivative term, while the remaining portion of the solute mass in the bulk fluid phase forms a mobile phase that undergoes a Brownian motion modeled by the integer-order derivative term [38, 53].

To demonstrate the improvement of the two-time-scale model on the initial behavior, the solutions to the single-time-scale tFDE (2.1) and the two-time-scale tFDE (1.3) with $\alpha = 0.1$, $A = 0.01$, $k = 100$ (i.e., the fractional derivative is dominant), $u(x,0) = e^{-\frac{x^2}{2\sigma^2}}/(\sqrt{2\pi}\sigma)$ with $\sigma = 0.018$ on a domain $[-0.1, 0.1]$ over a short time period $[0, 0.01]$ are presented in Figure 2.1, which indicate the initial singularity and the Fickian diffusion behavior of the solutions to (2.1) and (1.3), respectively. Later, we will further elucidate this phenomenon from the viewpoint of MSDs.

Then we turn to exhibit the novelty of the two-time-scale model from its multiscale effects. In the literature, [42] presents that the two-time-scale model bridges between Fickian fluxes at early times and non-Gaussian behavior at late times. In this work we demonstrate the multiscale effects by means of the MSDs (denoted by $\langle X(t)^2\rangle$), where in general $\langle X(t)^2\rangle \sim t^\beta$ for some β representing the scale of diffusion. For instance, $\langle X(t)^2\rangle \sim t$ for the Fickian diffusion described by the integer-order diffusion equation and $\langle X(t)^2\rangle \sim t^\alpha$ for the subdiffusion modeled by the single-time-scale tFDE (2.1) [45].

Let $\Omega = [-10, 10]$, $T = 100$, and we consider the integer-order and constant-fractional-order cases of models (2.1) and (1.3) with $k = 0.01$, 1, and 100 in the left plot of Figure 2.2 (the other data are the same as before). We have the following observations: (i) The MSD of two-time-scale tFDE has almost the same behavior as the integer-order diffusion equation for small k, while it has the same slope 0.1 as that of the single-time-scale tFDE for large k, which indicates that the two-time-scale model provides a framework containing both integer-order and single-time-scale models. (ii) For large k, i.e., the fractional derivative is dominant, the magnitude of the MSD for the single-time-scale tFDE is much larger than that for the two-time-scale tFDE, indicating the rapid initial spread, and hence the initial singularity of the solutions to the single-time-scale tFDE due to mass conservation. This is

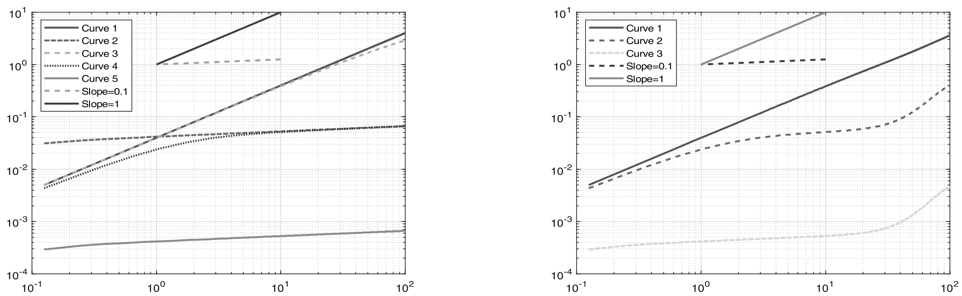


FIG. 2.2. *Log-log plots of MSDs. Left: the integer-order diffusion equation (curve* 1*), the single-time-scale tFDE with* $\alpha = 0.1$ *(curve* 2*), and the two-time-scale tFDE with* $\alpha = 0.1$ *and* $k = 0.01$ *(curve* 3*),* 1 *(curve* 4*),* 100 *(curve* 5*). Right: the two-time-scale variable-order tFDE with* $\alpha(0) = 0.1$*,* $\alpha(T) = 0.6$ *and* $k = 0.01$ *(curve* 1*),* 1 *(curve* 2*),* 100 *(curve* 3*).*

also consistent with the observations in Figure 2.1. (iii) The MSD for two-time-scale tFDE with $k = 1$ switches smoothly from the initial Fickian diffusion behavior to the long-term subdiffusive behavior (of order $\alpha = 0.1$), which implies that the two-time-scale tFDE provides a proper approximation for the single-time-scale tFDE for t away from the initial time. Thus, in comparison with the single-time-scale tFDE, one novelty of the two-time-scale tFDE lies in that it captures the long-term subdiffusion behavior typical of the single-time-scale tFDE while eliminating its nonphysical initial singularity, and thus provides a physically relevant extension of single-time-scale tFDE to the entire time interval, including the initial time $t = 0$.

Furthermore, in practical applications such as unconventional hydrocarbon or gas recovery [21], the structure of porous materials may change over time while staying heterogeneous in space, which results in a modification of the fractal dimension of the porous material via the Hurst index [18, 45], leading to tFDEs with space-time-dependent variable fractional order [31, 35, 57, 62, 63]. In addition, it is pointed out in [55] that the variable order could efficiently quantify the transitions between different diffusive states at various transport scales, which further demonstrates the advantages of variable-order fractional models in characterizing multiscale behaviors.

To substantiate the advantages of variable order, in the right plot of Figure 2.2 we present MSDs for the two-time-scale tFDE with variable order

$$\alpha(t) = \alpha(T) + (\alpha(0) - \alpha(T))\Big(1 - \frac{t}{T} - \frac{1}{2\pi}\sin\Big(2\pi\Big(1 - \frac{t}{T}\Big)\Big)\Big),$$

a smooth and monotonic function on $[0, T]$ with end values $\alpha(0)$ and $\alpha(T)$. The other data follows as above. In addition to similar observations as before, we find that the MSDs to the two-time-scale variable-order tFDE exhibit much richer structures and behaviors, e.g., the change of slope and the convexity. This justifies the novelty of the two-time-scale variable-order tFDE in modeling complex multiscale physical phenomena.

3. Analysis of auxiliary problem.

We analyze the following auxiliary problem:

$$\begin{aligned} &\partial_t u + k(\boldsymbol{x},t)\, {}^R D_t^{\alpha(\boldsymbol{x},t)} u + \mathcal{L}(\boldsymbol{x},t) u = f(\boldsymbol{x},t) + q(\boldsymbol{x},t), \quad (\boldsymbol{x},t) \in \Omega \times (0,T], \\ &u(x,0) = 0, \quad \boldsymbol{x} \in \Omega; \quad u(\boldsymbol{x},t) = 0, \quad (\boldsymbol{x},t) \in \partial\Omega \times [0,T], \end{aligned} \tag{3.1}$$

where we still denote the unknown solution by u for the sake of simplicity. Here the variable-order operator ${}^R D_t^{\alpha(\boldsymbol{x},t)}$ in (3.1) is defined by [68]

$$ {}^R D_t^{\alpha(\boldsymbol{x},t)} g := \frac{1}{\Gamma(1-\alpha(\boldsymbol{x},t))}\left[\frac{d}{d\xi}\int_0^\xi \frac{g(\boldsymbol{x},s)}{(\xi-s)^{\alpha(\boldsymbol{x},t)}}ds\right]_{\xi=t}. \tag{3.2} $$

We demonstrate the motivation of introducing this auxiliary problem by the following two lemmas, which indicate that the fractional derivative defined in (3.2) requires less temporal regularity for g (i.e., $g \in H^{1-\varepsilon}(L^2)$) than that in (1.5) (i.e., $g \in H^1(L^2)$), which barely accounts for the low regularity of the solutions before it is improved, and thus facilitate the analysis.

LEMMA 3.1 ([63, 68]). *Let $g \in H^1(L^2)$ with $g(\boldsymbol{x},0)=0$. Then*

$$ \partial_t^{\alpha(\boldsymbol{x},t)} g = {}^R D_t^{\alpha(\boldsymbol{x},t)} g. \tag{3.3} $$

LEMMA 3.2. *Let $0<\varepsilon<1-\alpha^*$. For any $g \in H^{1-\varepsilon}(L^2)$ with $g(\boldsymbol{x},0)=0$*

$$ \left\|{}^R D_t^{\alpha(\boldsymbol{x},t)} g\right\|_{L^2(L^2)} \le Q(T,\alpha^*,\varepsilon)\|g\|_{H^{1-\varepsilon}(L^2)}. $$

Proof. We follow [23, 60, 61] to conclude that there exists an $h \in L^2(L^2)$ such that $g = I_t^{1-\varepsilon}h$. We incorporate this with (3.2) to arrive at

$$ \begin{aligned} {}^R D_t^{\alpha(\boldsymbol{x},t)} g &= \frac{1}{\Gamma(1-\alpha(\boldsymbol{x},t))}\left[\frac{d}{d\xi}\int_0^\xi \frac{I_s^{1-\varepsilon}h}{(\xi-s)^{\alpha(\boldsymbol{x},t)}}ds\right]_{\xi=t} \\ &= \frac{1}{\Gamma(1-\varepsilon)\Gamma(1-\alpha(\boldsymbol{x},t))}\left[\frac{d}{d\xi}\int_0^\xi\int_0^s \frac{(s-y)^{-\varepsilon}h(\boldsymbol{x},y)dy}{(\xi-s)^{\alpha(\boldsymbol{x},t)}}ds\right]_{\xi=t} \\ &= \frac{1}{\Gamma(1-\varepsilon)\Gamma(1-\alpha(\boldsymbol{x},t))}\left[\frac{d}{d\xi}\int_0^\xi h(\boldsymbol{x},y)\int_y^\xi (s-y)^{-\varepsilon}(\xi-s)^{-\alpha(\boldsymbol{x},t)}dsdy\right]_{\xi=t} \\ &= \frac{1}{\Gamma(2-\alpha(\boldsymbol{x},t)-\varepsilon)}\left[\frac{d}{d\xi}\int_0^\xi (\xi-y)^{1-\varepsilon-\alpha(\boldsymbol{x},t)}h(\boldsymbol{x},y)dy\right]_{\xi=t} \\ &= \frac{1}{\Gamma(1-\alpha(\boldsymbol{x},t)-\varepsilon)}\int_0^t (t-s)^{-\varepsilon-\alpha(\boldsymbol{x},t)}h(\boldsymbol{x},s)ds. \end{aligned} \tag{3.4} $$

We then utilize the relation

$$ (t-s)^{-\varepsilon-\alpha(\boldsymbol{x},t)} = (t-s)^{-\varepsilon-\alpha^*}(t-s)^{\alpha^*-\alpha(\boldsymbol{x},t)} \le \max\{1,T\}(t-s)^{-\varepsilon-\alpha^*} \tag{3.5} $$

and (3.4) to obtain

$$ \left|{}^R D_t^{\alpha(\boldsymbol{x},t)} g\right| \le Q\int_0^t \left|h(\boldsymbol{x},s)\right|(t-s)^{-\alpha^*-\varepsilon}ds. $$

Use Young's convolutional inequality and $\|h\|_{L^2(L^2)} \sim \|g\|_{H^{1-\varepsilon}(L^2)}$ [23] to obtain $\left\|{}^R D_t^{\alpha(\boldsymbol{x},t)} g\right\|_{L^2(L^2)} \le Q\left\|t^{-\alpha^*-\varepsilon}\right\|_{L^1(0,T)}\|h\|_{L^2(L^2)}$, which concludes the proof. ∎

3.1. A uniqueness result. We prove a uniqueness result for future use.

LEMMA 3.3. *The homogeneous problem*

$$ \begin{aligned} &\partial_t v + k(\boldsymbol{x},t){}^R D_t^{\alpha(\boldsymbol{x},t)} v + \mathcal{L}(\boldsymbol{x},t)v = 0, \quad (\boldsymbol{x},t) \in \Omega \times (0,T], \\ &v(\boldsymbol{x},0)=0, \quad \boldsymbol{x} \in \Omega; \quad v(\boldsymbol{x},t)=0, \quad (\boldsymbol{x},t) \in \partial\Omega \times [0,T] \end{aligned} \tag{3.6} $$

admits a unique trivial solution in $H^1(L^2) \cap L^2(H^2)$.

Proof. We introduce a time-dependent bilinear form

$$B[u,v;t] := \int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla u \cdot \nabla v + \boldsymbol{b}(\boldsymbol{x},t)\cdot \nabla u v + c(\boldsymbol{x},t) uv \, d\boldsymbol{x} \tag{3.7}$$

for $u, v \in H_0^1(\Omega)$ and a.e. $t \in [0,T]$, which is derived according to the operator (1.4). By the *assumption* (b) and Poincaré's inequality, we conclude that there exist some constants $\beta > 0$ and $\gamma \geq 0$ such that

$$\beta \|v\|^2_{H_0^1(\Omega)} \leq B[v,v;t] + \gamma \|v\|^2_{L^2(\Omega)}, \quad \big|B[u,v;t]\big| \leq Q\|u\|_{H_0^1(\Omega)}\|v\|_{H_0^1(\Omega)}. \tag{3.8}$$

We employ the relation (from [19])

$$\int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla v \cdot \partial_t \nabla v \, d\boldsymbol{x} = \frac{1}{2}\partial_t \int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla v\cdot\nabla v\, d\boldsymbol{x} - \frac{1}{2}\int_\Omega \partial_t \boldsymbol{A}(\boldsymbol{x},t)\nabla v \cdot \nabla v \, d\boldsymbol{x}$$

to reformulate $B[v, \partial_t v; t]$ as

$$\begin{aligned} &B[v,\partial_t v;t] = \frac{1}{2}\partial_t \int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla v\cdot\nabla v\, d\boldsymbol{x} + \mathcal{A}, \\ &\qquad \mathcal{A} := -\frac{1}{2}\int_\Omega \partial_t \boldsymbol{A}(\boldsymbol{x},t)\nabla v\cdot\nabla v + \boldsymbol{b}(\boldsymbol{x},t)\cdot\nabla v \partial_t v + c(\boldsymbol{x},t) v\partial_t v\, d\boldsymbol{x}, \end{aligned} \tag{3.9}$$

and we bound $\mathcal{A}$ by the assumption (b) and arithmetic-geometric inequality

$$|\mathcal{A}| \leq \frac{Q}{\varepsilon}\|v(\cdot,t)\|^2_{H_0^1(\Omega)} + \frac{\varepsilon}{2}\|\partial_t v(\cdot,t)\|^2_{L^2(\Omega)} \tag{3.10}$$

for a sufficiently small $\varepsilon > 0$. We now multiply (3.6) by $\partial_t v$ and integrate the resulting equation over Ω to get

$$\big\|\partial_t v(\cdot,t)\big\|^2_{L^2(\Omega)} + B[v,\partial_t v;t] = -\int_\Omega k(\boldsymbol{x},t)\partial_t v(\boldsymbol{x},t)\, {}^R D_t^{\alpha(\boldsymbol{x},t)} v(\boldsymbol{x},t)\, d\boldsymbol{x}. \tag{3.11}$$

To bound the right-hand side term of (3.11), we employ the relation (3.3) by $v \in H^1(L^2)$ and $v(\boldsymbol{x},0) = 0$, together with Cauchy's inequality, Fubini's theorem, and an estimate similar to (3.5) to obtain, for $0 < \varepsilon \ll 1$,

$$\begin{aligned} &\int_\Omega \big|k(\boldsymbol{x},t)\partial_t v(\boldsymbol{x},t)\, {}^R D_t^{\alpha(\boldsymbol{x},t)} v(\boldsymbol{x},t)\big| d\boldsymbol{x} \\ &\quad = \int_\Omega \big|k(\boldsymbol{x},t)\partial_t v(\boldsymbol{x},t)\partial_t^{\alpha(\boldsymbol{x},t)} v(\boldsymbol{x},t)\big| d\boldsymbol{x} \\ &\quad \leq Q \int_\Omega \big|\partial_t v(\boldsymbol{x},t)\big| \int_0^t (t-s)^{-\alpha^*}\big|\partial_s v(\boldsymbol{x},s)\big| ds d\boldsymbol{x} \\ &\quad = Q\int_0^t (t-s)^{-\alpha^*}\int_\Omega \big|\partial_t v(\boldsymbol{x},t)\big|\big|\partial_s v(\boldsymbol{x},s)\big| d\boldsymbol{x} ds \\ &\quad \leq \frac{\varepsilon}{2}\big\|\partial_t v(\cdot,t)\big\|^2_{L^2(\Omega)} + \frac{Q}{\varepsilon}\int_0^t (t-s)^{-\alpha^*}\big\|\partial_s v(\cdot,s)\big\|^2_{L^2(\Omega)} ds. \end{aligned} \tag{3.12}$$

We invoke the aforementioned relations in (3.11) to obtain

$$\begin{aligned} &\big\|\partial_t v(\cdot,t)\big\|^2_{L^2(\Omega)} + \frac{1}{2}\partial_t \int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla v\cdot\nabla v\, d\boldsymbol{x} \\ &\quad \leq \frac{Q}{\varepsilon}\|v(\cdot,t)\|^2_{H_0^1(\Omega)} + \varepsilon\|\partial_t v(\cdot,t)\|^2_{L^2(\Omega)} + \frac{Q}{\varepsilon}\int_0^t (t-s)^{-\alpha^*}\big\|\partial_s v(\cdot,s)\big\|^2_{L^2(\Omega)} ds. \end{aligned} \tag{3.13}$$

Choose ε sufficiently small in (3.13) to cancel the like terms on both sides of the inequality, integrate the resulting inequality from 0 to t, and then incorporate Poincaré's inequality to obtain

$$\begin{aligned}&\int_0^t \left\|\partial_\tau v(\cdot,\tau)\right\|_{L^2(\Omega)}^2 d\tau + \|v(\cdot,t)\|_{H_0^1(\Omega)}^2\\ &\quad\le Q\int_0^t \|v(\cdot,\tau)\|_{H_0^1(\Omega)}^2 d\tau + Q\int_0^t\int_0^\tau (\tau-y)^{-\alpha^*}\left\|\partial_y v(\cdot,y)\right\|_{L^2(\Omega)}^2 dy d\tau.\end{aligned} \tag{3.14}$$

Apply Gronwall's inequality to (3.14) and use integration by parts to obtain

$$\begin{aligned}&\|v(\cdot,t)\|_{H_0^1(\Omega)}^2 + \int_0^t \left\|\partial_\tau v(\cdot,\tau)\right\|_{L^2(\Omega)}^2 d\tau\\ &\quad\le Q\int_0^t\int_0^\tau (\tau-y)^{-\alpha^*}\left\|\partial_y v(\cdot,y)\right\|_{L^2(\Omega)}^2 dy d\tau\\ &\quad= Q\int_0^t \left\|\partial_y v(\cdot,y)\right\|_{L^2(\Omega)}^2 \int_y^t (\tau-y)^{-\alpha^*} d\tau dy\\ &\quad= \frac{Q}{1-\alpha^*}\int_0^t (t-\tau)^{1-\alpha^*}\left\|\partial_\tau v(\cdot,\tau)\right\|_{L^2(\Omega)}^2 d\tau\\ &\quad= \frac{Q}{1-\alpha^*}\int_0^t (t-\tau)^{1-\alpha^*}\frac{d}{d\tau}\left[\int_0^\tau \left\|\partial_y v(\cdot,y)\right\|_{L^2(\Omega)}^2 dy\right] d\tau\\ &\quad= Q\int_0^t (t-\tau)^{-\alpha^*}\int_0^\tau \left\|\partial_y v(\cdot,y)\right\|_{L^2(\Omega)}^2 dy d\tau,\end{aligned} \tag{3.15}$$

and we conclude from the weak singular Gronwall inequality (cf. [58]) that

$$\int_0^t \left\|\partial_\tau v(\cdot,\tau)\right\|_{L^2(\Omega)}^2 d\tau = 0, \quad t\in[0,T], \tag{3.16}$$

which, together with $v(\boldsymbol{x},0)=0$ and $v(\boldsymbol{x},t)=\int_0^t \partial_\tau v(\boldsymbol{x},\tau)d\tau$, yields $\|v\|_{L^2(L^2)}=0$. Consequently, $v=0$ for a.e. $(\boldsymbol{x},t)\in\Omega\times[0,T]$ and thus $\|v\|_{H^1(L^2)}=\|v\|_{L^2(H^2)}=0$, which completes the proof of the lemma. ∎

3.2. Well-posedness of (3.1). We prove the well-posedness of the auxiliary problem (3.1).

THEOREM 3.4. *If assumptions* (a)–(c) *hold, the auxiliary problem* (3.1) *admits a unique solution* $u\in H^1(L^2)\cap L^2(H^2)$ *such that*

$$\|u\|_{H^1(L^2)} + \|u\|_{L^2(H^2)} \le Q\|f+q\|_{L^2(L^2)}, \tag{3.17}$$

where Q *is independent of* u, f, *or* q.

Proof. We prove this theorem in three steps.

Step 1*: Analysis of a data-to-solution map* K_α. Define an operator K_α that maps v to w determined by

$$\partial_t w + \mathcal{L}(\boldsymbol{x},t)w = -k(\boldsymbol{x},t)\,{}^R D_t^{\alpha(\boldsymbol{x},t)} v \tag{3.18}$$

over $\Omega\times(0,T]$, equipped with homogeneous initial and boundary conditions. By Lemma 3.2 and $k\in L^\infty(L^\infty)$, the right-hand-side term of (3.18) belongs to $L^2(L^2)$ for $v\in H^{1-\varepsilon}(L^2)$. Then the regularity results for classical parabolic equations [40, section 4.7.1, p. 243] ensure that (3.18) admits a unique solution $w = K_\alpha v \in H^1(L^2) \cap L^2(H^2)$ such that, according to Lemma 3.2,

$$\begin{aligned}\|w\|_{H^1(L^2)} + \|w\|_{L^2(H^2)} &\le Q\big\|k(\boldsymbol{x},t)\, {}^R D_t^{\alpha(\boldsymbol{x},t)} v\big\|_{L^2(L^2)} \\ &\le Q\big\|{}^R D_t^{\alpha(\boldsymbol{x},t)} v\big\|_{L^2(L^2)} \le Q\|v\|_{H^{1-\varepsilon}(L^2)},\end{aligned} \tag{3.19}$$

which indicates that $K_\alpha : H^{1-\varepsilon}(L^2) \to H^1(L^2) \cap L^2(H^2)$ is a bounded linear operator.

We then show that $K_\alpha : H^{1-\varepsilon}(L^2) \to H^{1-\varepsilon}(L^2)$ is a compact operator. For a bounded sequence $\{u_n\}_{n=1}^\infty \subset H^{1-\varepsilon}(L^2)$, the estimate (3.19) yields

$$\|K_\alpha u_n\|_{H^1(L^2)} + \|K_\alpha u_n\|_{L^2(H^2)} \le Q\|u_n\|_{H^{1-\varepsilon}(L^2)}, \tag{3.20}$$

which implies that $\{K_\alpha u_n\}_{n=1}^\infty$ is a bounded sequence in $H^1(L^2) \cap L^2(H^2)$. By [40, Theorem 16.2, Chapter 1], $H^1(L^2) \cap L^2(H^2)$ is compactly embedded in $H^{1-\varepsilon}(L^2)$, which implies there exists a subsequence $\{K_\alpha u_{n_j}\}_{j=1}^\infty$ converging in $H^{1-\varepsilon}(L^2)$. Thus we conclude that $K_\alpha : H^{1-\varepsilon}(L^2) \to H^{1-\varepsilon}(L^2)$ is a compact operator.

Step 2: Well-posedness of an abstract equation $u := u_f + K_\alpha u$. We intend to prove that $u := u_f + K_\alpha u$ admits a unique solution in $H^{1-\varepsilon}(L^2)$, where u_f is determined by the problem

$$\partial_t u_f + \mathcal{L}(\boldsymbol{x},t) u_f = f(\boldsymbol{x},t) + q(\boldsymbol{x},t) \tag{3.21}$$

over $\Omega \times (0,T]$ with homogeneous initial and boundary conditions. Similar to the above, the assumption (c) and regularity results for classical parabolic equations [40, section 4.7.1, p. 243] give the uniqueness of the solution $u_f \in H^1(L^2) \cap L^2(H^2)$ and

$$\|u_f\|_{H^1(L^2)} + \|u_f\|_{L^2(H^2)} \le Q\|f+q\|_{L^2(L^2)}. \tag{3.22}$$

By the Fredholm alternative of compact operators, $u = u_f + K_\alpha u$ admits a unique solution in $H^{1-\varepsilon}(L^2)$ as long as $N(I - K_\alpha) = \{0\}$, where $N(I - K_\alpha)$ represents the null set of $I - K_\alpha$ [19]. Thus we remain to prove that $(I - K_\alpha)v = 0$ has only the trivial solution in $H^{1-\varepsilon}(L^2)$.

We note from (3.18) that $v = K_\alpha v$ is exactly (3.6), and $k \in L^\infty(L^\infty)$ and Lemma 3.2 indicate that the fractional term of (3.6) belongs to $L^2(L^2)$, provided that $v \in H^{1-\varepsilon}(L^2)$, and thus the solution $v \in H^1(L^2) \cap L^2(H^2)$ could be proved in an analogous manner to (3.19). Then we obtain from the uniqueness result in Lemma 3.3 that $\|v\|_{H^1(L^2)} = 0$. By the Sobolev embedding $\|v\|_{H^{1-\varepsilon}(L^2)} \le Q\|v\|_{H^1(L^2)} = 0$, we conclude that $(I - K_\alpha)v = 0$ has only the trivial solution in $H^{1-\varepsilon}(L^2)$, which, together with the Fredholm alternative, indicates that there exists a unique solution of $u = K_\alpha u + u_f$ in $H^{1-\varepsilon}(L^2)$. In addition, $N(I - K_\alpha) = \{0\}$ implies that the operator $I - K_\alpha$ is one-to-one onto $H^{1-\varepsilon}(L^2)$ by the Fredholm alternative [19], and we further conclude that the operator $(I - K_\alpha)^{-1} : H^{1-\varepsilon}(L^2) \to H^{1-\varepsilon}(L^2)$ is a bounded linear operator [52, Theorem 3.8]. Consequently, $u = (I - K_\alpha)^{-1} u_f$ could be estimated based on (3.22),

$$\begin{aligned}\|u\|_{H^{1-\varepsilon}(L^2)} &\le \big\|(I - K_\alpha)^{-1}\big\|_{H^{1-\varepsilon}(L^2) \to H^{1-\varepsilon}(L^2)} \|u_f\|_{H^{1-\varepsilon}(L^2)} \\ &\le Q\|u_f\|_{H^1(L^2)} \le Q\|f+q\|_{L^2(L^2)},\end{aligned} \tag{3.23}$$

which, together with (3.19), leads to the following regularity improvement:

$$\begin{aligned} \|u\|_{H^1(L^2)} + \|u\|_{L^2(H^2)} &\le \|K_\alpha u\|_{H^1(L^2)} + \|K_\alpha u\|_{L^2(H^2)} + \|u_f\|_{H^1(L^2)} + \|u_f\|_{L^2(H^2)} \\ &\le Q\|u\|_{H^{1-\varepsilon}(L^2)} + Q\|f+q\|_{L^2(L^2)} \le Q\|f+q\|_{L^2(L^2)}. \end{aligned} \tag{3.24}$$

Step 3*: Analysis of model* (3.1). By the definitions of u_f and $K_\alpha u$, we have

$$\partial_t u_f + \mathcal{L}(\boldsymbol{x},t)u_f = f(\boldsymbol{x},t) + q(\boldsymbol{x},t),\ \partial_t K_\alpha u + k(\boldsymbol{x},t)\, {}^R D_t^{\alpha(\boldsymbol{x},t)} u + \mathcal{L}(\boldsymbol{x},t) K_\alpha u = 0$$

over $\Omega \times (0,T]$, equipped with homogeneous initial and boundary conditions. We then add these two equations and apply $u = u_f + K_\alpha u$ to find that the resulting equation is exactly model (3.1), that is, model (3.1) has a solution $u \in H^2(L^2) \cap L^2(H^2)$ with the estimate (3.24). The uniqueness of the solutions to model (3.1) follows from that for $u = u_f + K_\alpha u$, which completes the proof of the theorem. ∎

4. Analysis of forward problem. We prove the well-posedness of the forward problem (1.3) based on that of the auxiliary problem (3.1).

THEOREM 4.1. *If assumptions* (a)–(c) *hold, the forward problem* (1.3) *admits a unique solution* $u \in H^1(L^2) \cap L^2(H^2)$ *such that*

$$\|u\|_{H^1(L^2)} + \|u\|_{L^2(H^2)} \le Q\|f+q\|_{L^2(L^2)}, \tag{4.1}$$

where Q *is independent of* u, f, *or* q.

Proof. Theorem 3.4 proves that the solution u to the auxiliary problem (3.1) belongs to $H^1(L^2) \cap L^2(H^2)$, which together with $u(\boldsymbol{x},0) = 0$ and Lemma 3.1 implies that $\partial_t^{\alpha(\boldsymbol{x},t)} u = {}^R D_t^{\alpha(\boldsymbol{x},t)} u$. Thus the auxiliary problem (3.1) is exactly the forward problem (1.3) and the solution $u \in H^1(L^2) \cap L^2(H^2)$ to the auxiliary problem (3.1) also solves (1.3) with the regularity estimate (4.1). The uniqueness of the solutions to the forward problem (1.3) in $H^1(L^2) \cap L^2(H^2)$ follows from that of the auxiliary problem (3.1) by Lemma 3.1 and Theorem 3.4, which completes the proof. ∎

We prove maximal regularity results for (1.3) based on Theorem 4.1.

THEOREM 4.2. *Suppose that assumptions* (a)–(d) *hold and* $f, q \in L^p(L^2)$ *for* $2 < p < \infty$. *Then the solution to* (1.3) *has the following regularity estimates:*

$$\|u\|_{W^{1,p}(L^2)} + \big\|\mathcal{L}u\big\|_{L^p(L^2)} \le Q\|f+q\|_{L^p(L^2)}, \tag{4.2}$$

where Q *is independent of* u, f, *or* q.

Proof. Based on Theorem 4.1, we reformulate problem (1.3) with $u \in H^1(L^2) \cap L^2(H^2)$ as

$$\partial_t u + \mathcal{L}(\boldsymbol{x},t)u = \hat{f} := f(\boldsymbol{x},t) + q(\boldsymbol{x},t) - k(\boldsymbol{x},t)\partial_t^{\alpha(\boldsymbol{x},t)} u. \tag{4.3}$$

Case 1: $0 \le \alpha^* < 1/2$. By (1.5) and Young's convolution inequality [1], we find $1 < s < 2$ such that $1 + 1/p = 1/s + 1/2$ and

$$
\begin{aligned}
\big\|k(\boldsymbol{x},t)\partial_t^{\alpha(\boldsymbol{x},t)}u\big\|_{L^p(L^2)} &\le Q\bigg\|\int_0^t (t-s)^{-\alpha^*}|\partial_s u(\boldsymbol{x},s)|ds\bigg\|_{L^p(L^2)}\\
&= Q\big\|t^{-\alpha^*} * |\partial_t u|\big\|_{L^p(L^2)}\\
&\le Q\big\|t^{-\alpha^*}\big\|_{L^s(0,T)}\|\partial_t u\|_{L^2(L^2)}\\
&\le Q\|\partial_t u\|_{L^2(L^2)}\\
&\le Q\|f+q\|_{L^2(L^2)},
\end{aligned} \tag{4.4}
$$

where $*$ represents the symbol of convolution and we used the fact that $\alpha^* s < 2\alpha^* < 1$. The estimate (4.4) implies that $\hat{f} \in L^p(L^2)$ such that we combine (3.8), assumption (d), and (4.4) with the maximal regularity estimates for the parabolic equation $\partial_t v + \mathcal{L}(\boldsymbol{x},t)v = g$ over $\Omega \times (0,T]$ with homogeneous initial and boundary conditions [27, Theorem 1.2]

$$
\|v\|_{W^{1,r}(L^2)} + \|\mathcal{L}v\|_{L^r(L^2)} \le Q\|g\|_{L^r(L^2)}, \quad r \in (1,\infty), \tag{4.5}
$$

for (4.3) to conclude that

$$
\|u\|_{W^{1,p}(L^2)} + \big\|\mathcal{L}u\big\|_{L^p(L^2)} \le Q\big\|\hat{f}\big\|_{L^p(L^2)} \le Q\|f+q\|_{L^p(L^2)}. \tag{4.6}
$$

Case 2: $1/2 \le \alpha^* < 1$. Let $1 < s_0 < 1/\alpha^* \le 2$ and $m \in \mathbb{N}^+$ satisfy

$$
m \le \frac{\frac{1}{2} - \frac{1}{p}}{1 - \frac{1}{s_0}} < m+1,
$$

and we assume $m > 1$ without loss of generality. Then we accordingly define an equidistant sequence $\{1/r_i\}_{i=0}^m$ by

$$
\frac{1}{2} = \frac{1}{r_0} > \frac{1}{r_1} > \cdots > \frac{1}{r_m} \ge \frac{1}{p} \text{ with the equidistance } \frac{1}{r_{i-1}} - \frac{1}{r_i} = 1 - \frac{1}{s_0}.
$$

We intend to combine the maximal regularity estimate (4.5) with the bootstrapping argument and Young's convolution inequality to progressively lift the temporal regularity of the solution u to problem (4.3) via the pathway

$$
H^1(L^2) = W^{1,r_0}(L^2) \to W^{1,r_1}(L^2) \to \cdots \to W^{1,r_m}(L^2) \to W^{1,p}(L^2).
$$

By Young's convolution inequality we have

$$
\begin{aligned}
\big\|k(\boldsymbol{x},t)\partial_t^{\alpha(\boldsymbol{x},t)}u\big\|_{L^{r_1}(L^2)} &\le \bigg\|\int_0^t (t-s)^{-\alpha^*}|\partial_s u(\boldsymbol{x},s)|ds\bigg\|_{L^{r_1}(L^2)}\\
&\le Q\big\|t^{-\alpha^*}\big\|_{L^{s_0}(0,T)}\|\partial_t u\|_{L^{r_0}(L^2)}\\
&\le Q\|\partial_t u\|_{L^{r_0}(L^2)}\\
&\le Q\|f+q\|_{L^{r_0}(L^2)},
\end{aligned} \tag{4.7}
$$

where we used the estimate (4.1) and $s_0\alpha^* < 1$. Thus we obtain $\hat{f} \in L^{r_1}(L^2)$ and apply the maximal regularity estimate (4.5) to find that

$$
\|u\|_{W^{1,r_1}(L^2)} + \big\|\mathcal{L}u\big\|_{L^{r_1}(L^2)} \le Q\big\|\hat{f}\big\|_{L^{r_1}(L^2)} \le Q\|f+q\|_{L^{r_1}(L^2)}. \tag{4.8}
$$

We repeat the above procedure $m-1$ times with r_{i-1} replaced by r_i and r_i replaced by r_{i+1} in (4.7)–(4.8) for $1 \le i \le m-1$ to derive that

$$\|u\|_{W^{1,r_m}(L^2)} + \big\|\mathcal{L}u\big\|_{L^{r_m}(L^2)} \le Q\|f+q\|_{L^{r_m}(L^2)}. \tag{4.9}$$

Finally, we select s_1 such that $1+1/p = 1/s_1 + 1/r_m$. As $r_m \le p$, we have $s_1 \ge 1$ and apply

$$\frac{1}{2} - \frac{1}{p} < (m+1)\Big(1 - \frac{1}{s_0}\Big), \ \frac{1}{r_m} = \frac{1}{r_0} - m\Big(1 - \frac{1}{s_0}\Big)$$

to find

$$\begin{aligned} \frac{1}{s_1} - \frac{1}{s_0} &= 1 + \frac{1}{p} - \frac{1}{r_m} - \frac{1}{s_0} \\ &= 1 + \frac{1}{p} - \left(\frac{1}{r_0} - m\Big(1 - \frac{1}{s_0}\Big)\right) - \frac{1}{s_0} \\ &= (m+1)\Big(1 - \frac{1}{s_0}\Big) - \frac{1}{2} + \frac{1}{p} > 0, \end{aligned} \tag{4.10}$$

that is, $s_1 < s_0$ and thus $\alpha^* s_1 < \alpha^* s_0 < 1$. We then apply Young's convolution inequality with r_1, s_0, and r_0 in (4.7) replaced by p, s_1, and r_m, respectively, to obtain $\|k(\boldsymbol{x},t)\partial_t^{\alpha(\boldsymbol{x},t)} u\|_{L^p(L^2)} \le Q\|f+q\|_{L^{r_m}(L^2)}$. We then apply the maximal L^p regularity result (4.5) for problem (4.3) to obtain

$$\|u\|_{W^{1,p}(L^2)} + \big\|\mathcal{L}u\big\|_{L^p(L^2)} \le Q\big\|\hat{f}\big\|_{L^p(L^2)} \le Q\|f+q\|_{L^p(L^2)}, \tag{4.11}$$

which completes the proof for this case. $\square$

5. Analysis of adjoint equation. We prove the well-posedness of the following variable-order time-fractional PDE, which is indeed the forward-in-time analogue of the adjoint equation (6.1):

$$\begin{aligned} &\partial_t z(\boldsymbol{x},t) + {}^R\bar{\partial}_t^{\alpha(\boldsymbol{x},t)}(kz) + \hat{\mathcal{L}}(\boldsymbol{x},t) z(\boldsymbol{x},t) = F(\boldsymbol{x},t), \quad (\boldsymbol{x},t) \in \Omega \times (0,T]; \\ &z(\boldsymbol{x},0) = 0, \ \boldsymbol{x} \in \Omega; \quad z(\boldsymbol{x},t) = 0, \quad (\boldsymbol{x},t) \in \partial\Omega \times [0,T]. \end{aligned} \tag{5.1}$$

Here ${}^R\bar{\partial}_t^{\alpha(t)}$ is the forward-in-time analogue of (6.3),

$${}^R\bar{\partial}_t^{\alpha(\boldsymbol{x},t)} g := \partial_t \, {}_0\bar{I}_t^{1-\alpha(\boldsymbol{x},t)} g, \quad {}_0\bar{I}_t^{1-\alpha(t)} g := \int_0^t \frac{g(\boldsymbol{x},s)ds}{\Gamma(1-\alpha(\boldsymbol{x},s))(t-s)^{\alpha(\boldsymbol{x},s)}}, \tag{5.2}$$

and the adjoint operator $\hat{\mathcal{L}}(\boldsymbol{x},t)$ of $\mathcal{L}(\boldsymbol{x},t)$ is defined as

$$\hat{\mathcal{L}}(\boldsymbol{x},t)u := -\nabla \cdot \big(\boldsymbol{A}(\boldsymbol{x},t)\nabla u\big) - \nabla \cdot (\boldsymbol{b}(\boldsymbol{x},t)u) + c(\boldsymbol{x},t)u.$$

LEMMA 5.1 ([66, Lemma 1]). *Under assumption* (a)*, the following estimate holds for $g \in H^1(0,T)$ with $g(0) = 0$,*

$$\big|{}^R\bar{\partial}_t^{\alpha(t)} g\big| \le Q \int_0^t |\partial_s g(s)|(t-s)^{-\alpha^*} ds \tag{5.3}$$

for $0 \le t \le T$. Here $Q = Q(\|\alpha\|_{W^{1,\infty}}, T, \alpha^*)$.

LEMMA 5.2. *Under assumptions* (a) *and* (c), $g \in H^1(L^2)$ *with* $g(\boldsymbol{x},0)=0$ *has the estimate*

$$\left|{}^R\bar{\partial}_t^{\alpha(\boldsymbol{x},t)}(kg)\right| \le Q\int_0^t |\partial_s g(\boldsymbol{x},s)|(t-s)^{-\alpha^*}ds \tag{5.4}$$

for $0 \le t \le T$ *and a.e.* $\boldsymbol{x} \in \Omega$. *Here* $Q = Q(\|\alpha\|_{W^{1,\infty}(L^\infty)}, \|k\|_{W^{1,\infty}(L^\infty)}, T, \alpha^*)$.

Proof. Since $g \in H^1(L^2)$ and $k \in W^{1,\infty}(L^\infty)$, we have $kg \in H^1(L^2)$ with $(kg)(\boldsymbol{x},0)=0$. Then (5.3) yields

$$\begin{aligned}\left|{}^R\bar{\partial}_t^{\alpha(\boldsymbol{x},t)}(kg)\right| &\le Q\int_0^t |\partial_s(kg)(\boldsymbol{x},s)|(t-s)^{-\alpha^*}ds\\ &\le Q\int_0^t |k(\boldsymbol{x},s)||\partial_s g(\boldsymbol{x},s)|(t-s)^{-\alpha^*}ds\\ &\quad + Q\int_0^t |\partial_s k(\boldsymbol{x},s)||g(\boldsymbol{x},s)|(t-s)^{-\alpha^*}ds\\ &\le Q\int_0^t |\partial_s g(\boldsymbol{x},s)|(t-s)^{-\alpha^*}ds\\ &\quad + Q\int_0^t\int_0^s |\partial_\tau g(\boldsymbol{x},\tau)|d\tau(t-s)^{-\alpha^*}ds,\end{aligned} \tag{5.5}$$

where the last term on the right-hand side of (5.5) could be further bounded by

$$\begin{aligned}\int_0^t\int_0^s |\partial_\tau g(\boldsymbol{x},\tau)|(t-s)^{-\alpha^*}d\tau ds &= \int_0^t |\partial_\tau g(\boldsymbol{x},\tau)|\int_\tau^t (t-s)^{-\alpha^*}dsd\tau\\ &\le Q\int_0^t |\partial_s g(\boldsymbol{x},s)|(t-s)^{1-\alpha^*}ds\\ &\le Q\int_0^t |\partial_s g(\boldsymbol{x},s)|(t-s)^{-\alpha^*}ds,\end{aligned} \tag{5.6}$$

which together with the estimate (5.5) yields (5.4). ∎

LEMMA 5.3. *Let* $\varepsilon > 0$ *be sufficiently small such that* $1-\alpha^*-\varepsilon > 0$. *Then there exists a positive constant* $Q = Q(\|\alpha\|_{W^{1,\infty}(L^\infty)}, \|k\|_{W^{1,\infty}(L^\infty)}, T, \alpha^*, \varepsilon)$ *such that the estimate*

$$\left\|{}^R\bar{\partial}_t^{\alpha(\boldsymbol{x},t)}(kg)\right\|_{L^2(L^2)} \le Q\|g\|_{H^{1-\varepsilon}(L^2)} \tag{5.7}$$

holds for any $g \in H^{1-\varepsilon}(L^2)$ *with* $g(\boldsymbol{x},0)=0$.

Proof. By the assumptions on g, we follow [23] to conclude that there exists an element $h \in L^2(L^2)$ such that $g = I_t^{1-\varepsilon}h$ and $\|h\|_{L^2(L^2)} \le Q\|g\|_{H^{1-\varepsilon}(L^2)}$. We employ (5.2) and then interchange the order of integration to arrive at

$$\begin{aligned}{}^R\bar{\partial}_t^{\alpha(\boldsymbol{x},t)}(kg) &= \partial_t\left[\int_0^t \frac{k(\boldsymbol{x},s)I_s^{1-\varepsilon}hds}{\Gamma(1-\alpha(\boldsymbol{x},s))(t-s)^{\alpha(\boldsymbol{x},s)}}\right]\\ &= \frac{1}{\Gamma(1-\varepsilon)}\partial_t\left[\int_0^t\int_0^s \frac{k(\boldsymbol{x},s)(s-y)^{-\varepsilon}h(\boldsymbol{x},y)dyds}{\Gamma(1-\alpha(\boldsymbol{x},s))(t-s)^{\alpha(\boldsymbol{x},s)}}\right]\end{aligned}$$

$$
\begin{aligned}
(5.8)\qquad &= \frac{1}{\Gamma(1-\varepsilon)}\partial_t\Big[\int_0^t h(\boldsymbol{x},y)\int_y^t \frac{k(\boldsymbol{x},s)(s-y)^{-\varepsilon}(t-s)^{-\alpha(\boldsymbol{x},s)}}{\Gamma(1-\alpha(\boldsymbol{x},s))}dsdy\Big]\\
&= \frac{1}{\Gamma(1-\varepsilon)}\int_0^t h(\boldsymbol{x},y)\partial_t\int_y^t \frac{k(\boldsymbol{x},s)(s-y)^{-\varepsilon}(t-s)^{-\alpha(\boldsymbol{x},s)}}{\Gamma(1-\alpha(\boldsymbol{x},s))}dsdy,
\end{aligned}
$$

where we have used the fact that

$$
\begin{aligned}
&\Big|\int_y^t \frac{k(\boldsymbol{x},s)(s-y)^{-\varepsilon}(t-s)^{-\alpha(\boldsymbol{x},t)}}{\Gamma(1-\alpha(\boldsymbol{x},s))}ds\Big|\\
&\quad \le Q\int_y^t (s-y)^{-\varepsilon}(t-s)^{-\alpha^*}ds = Q(t-y)^{1-\varepsilon-\alpha^*}\to 0,\quad \text{as } y\to t^-.
\end{aligned}
$$

We use the variable substitution $z=(s-y)/(t-y)$, which implies

$$
(5.9)\quad t-s=(t-y)(1-z),\quad s-y=(t-y)z,\quad ds=(t-y)dz,\quad s=zt+y(1-z),
$$

to reformulate the inner integral of (5.8) as

$$
\begin{aligned}
&\int_y^t \frac{k(\boldsymbol{x},s)(s-y)^{-\varepsilon}(t-s)^{-\alpha(\boldsymbol{x},s)}}{\Gamma(1-\alpha(\boldsymbol{x},s))}ds\\
&= \int_0^1 \frac{z^{-\varepsilon}(1-z)^{-\alpha(\boldsymbol{x},zt+y(1-z))}}{\Gamma(1-\alpha(\boldsymbol{x},zt+y(1-z)))}(t-y)^{1-\alpha(\boldsymbol{x},zt+y(1-z))-\varepsilon}k(\boldsymbol{x},zt+y(1-z))dz.
\end{aligned}
$$

Differentiate the above equation to get

$$
\begin{aligned}
&\partial_t\int_y^t \frac{k(\boldsymbol{x},s)(s-y)^{-\varepsilon}(t-s)^{-\alpha(\boldsymbol{x},s)}}{\Gamma(1-\alpha(\boldsymbol{x},s))}ds\\
&\quad = \int_0^1 \partial_t\Big[\frac{(1-z)^{-\alpha(\boldsymbol{x},zt+y(1-z))}z^{-\varepsilon}}{\Gamma(1-\alpha(\boldsymbol{x},zt+y(1-z)))}(t-y)^{1-\alpha(\boldsymbol{x},zt+y(1-z))-\varepsilon}\\
(5.10)\qquad &\qquad\qquad \times k(\boldsymbol{x},zt+y(1-z))\Big]dz\\
&\quad = \int_0^1 \frac{(1-z)^{-\alpha(\boldsymbol{x},zt+y(1-z))}z^{-\varepsilon}}{\Gamma(1-\alpha(\boldsymbol{x},zt+y(1-z)))}(t-y)^{1-\alpha(\boldsymbol{x},zt+y(1-z))-\varepsilon}\mathcal{K}(\boldsymbol{x},z,y,t)dz
\end{aligned}
$$

with

(5.11)

$$
\begin{aligned}
\mathcal{K}(\boldsymbol{x},z,y,t) &= k(\boldsymbol{x},zt+y(1-z))\big[-\partial_t\alpha(\boldsymbol{x},zt+y(1-z))\ln(1-z)\\
&\quad + \frac{1-\alpha(\boldsymbol{x},zt+y(1-z))-\varepsilon}{t-y} - \partial_t\alpha(\boldsymbol{x},zt+y(1-z))\ln(t-y)\\
&\quad + \frac{\Gamma'(1-\alpha(\boldsymbol{x},zt+y(1-z)))}{\Gamma(1-\alpha(\boldsymbol{x},zt+y(1-z)))}\partial_t\alpha(\boldsymbol{x},zt+y(1-z))\big] + \partial_t k(\boldsymbol{x},zt+y(1-z)).
\end{aligned}
$$

We combine assumptions (a) and (c) to bound the leading terms in (5.11) as

$$
\begin{aligned}
(5.12)\qquad \big|\mathcal{K}(\boldsymbol{x},z,y,t)\big| &\le Q|\ln(1-z)| + Q(t-y)^{-1} + Q|\ln(t-y)|\\
&\le Q|\ln(1-z)| + Q(t-y)^{-1}.
\end{aligned}
$$

We then employ this with the estimate (3.5) as well as the definition of the Beta function $B(\cdot,\cdot)$ [50] to bound (5.10),

$$
\begin{aligned}
&\left|\partial_t \int_y^t \frac{k(\boldsymbol{x},s)(s-y)^{-\varepsilon}}{\Gamma(1-\alpha(\boldsymbol{x},s))(t-s)^{\alpha(\boldsymbol{x},s)}} ds dy\right| \\
&\quad \le Q\int_0^1 (1-z)^{-\alpha^*} z^{-\varepsilon}(t-y)^{1-\alpha^*-\varepsilon} Q\big(|\ln(1-z)|+(t-y)^{-1}\big)dz \\
&\quad \le Q(t-y)^{-\alpha^*-\varepsilon}\int_0^1 \frac{|\ln(1-z)|}{(1-z)^{\alpha^*}z^{\varepsilon}}dz \\
&\quad \le Q(t-y)^{-\alpha^*-\varepsilon}\int_0^1 (1-z)^{-\alpha^*-\varepsilon}z^{-\varepsilon}dz \\
&\quad = Q(t-y)^{-\alpha^*-\varepsilon}B(1-\alpha^*-\varepsilon,1-\varepsilon) \le Q(t-y)^{-\alpha^*-\varepsilon},
\end{aligned}
\tag{5.13}
$$

and then incorporate the above estimate into (5.8) to obtain

$$
\left|{}^R\bar\partial_t^{\alpha(\boldsymbol{x},t)}(kg)\right| \le Q\int_0^t \big|h(\boldsymbol{x},y)\big|(t-y)^{-\alpha^*-\varepsilon}dy. \tag{5.14}
$$

The estimate (5.7) could thus be derived by taking the $L^2(L^2)$ norm on both sides of this equation and applying Young's convolution inequality and the estimate $\|h\|_{L^2(L^2)} \le Q\|g\|_{H^{1-\varepsilon}(L^2)}$. ∎

We then prove a uniqueness result.

LEMMA 5.4. *The homogeneous problem*

$$
\begin{aligned}
&\partial_t v + \hat{\mathcal{L}}(\boldsymbol{x},t)v = -{}^R\bar\partial_t^{\alpha(\boldsymbol{x},t)}(kv), \quad (\boldsymbol{x},t)\in\Omega\times(0,T], \\
&v(\boldsymbol{x},0)=0, \quad \boldsymbol{x}\in\Omega; \quad v(\boldsymbol{x},t)=0, \quad (\boldsymbol{x},t)\in\partial\Omega\times[0,T]
\end{aligned}
\tag{5.15}
$$

has only a trivial solution in $H^1(L^2)\cap L^2(H^2)$.

Proof. We introduce a time-dependent bilinear form $\hat B[u,v;t]$ derived from the operator (6.3):

$$
\begin{aligned}
\hat B[u,v;t] &:= \int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla u\cdot\nabla v - \nabla\cdot(\boldsymbol{b}(\boldsymbol{x},t)u)v + c(\boldsymbol{x},t)uv\, d\boldsymbol{x} \\
&= \int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla u\cdot\nabla v + \boldsymbol{b}(\boldsymbol{x},t)\cdot\nabla v u + c(\boldsymbol{x},t)uv\, d\boldsymbol{x}
\end{aligned}
\tag{5.16}
$$

for $u,v\in H_0^1(\Omega)$ and a.e. $t\in[0,T]$. As $\hat B[v,v;t]=B[v,v;t]$, (3.8) yields

$$
\beta\|v\|^2_{H_0^1(\Omega)} \le \hat B[v,v;t]+\gamma\|v\|^2_{L^2(\Omega)}, \quad \big|\hat B[u,v;t]\big| \le Q\|u\|_{H_0^1(\Omega)}\|v\|_{H_0^1(\Omega)}. \tag{5.17}
$$

Similar to the derivations of (3.9) we have

$$
\begin{aligned}
&\hat B[v,\partial_t v;t] = \frac12\partial_t\int_\Omega \boldsymbol{A}(\boldsymbol{x},t)\nabla v\cdot\nabla v\, d\boldsymbol{x} + \hat{\mathcal{A}}, \\
&\hat{\mathcal{A}} := -\frac12\int_\Omega \partial_t\boldsymbol{A}(\boldsymbol{x},t)\nabla v\cdot\nabla v - \boldsymbol{b}(\boldsymbol{x},t)\cdot\nabla v\partial_t v\, d\boldsymbol{x} \\
&\qquad + \int_\Omega -\nabla\cdot\boldsymbol{b}(\boldsymbol{x},t)v\partial_t v + c(\boldsymbol{x},t)v\partial_t v\, d\boldsymbol{x}
\end{aligned}
\tag{5.18}
$$

with the estimate $|\hat{\mathcal{A}}| \leq \frac{Q}{\varepsilon}\|v\|^2_{H^1_0(\Omega)} + \frac{\varepsilon}{2}\|\partial_t v\|^2_{L^2(\Omega)}$ for a sufficiently small $\varepsilon > 0$. We also follow an estimate similar to that of (3.12) to apply (5.4), obtaining for $0<\varepsilon\ll 1$

$$\int_\Omega \left|\partial_t v^R\bar{\partial}_t^{\alpha(\boldsymbol{x},t)}(kv)\right|dx \leq \frac{\varepsilon}{2}\left\|\partial_t v(\cdot,t)\right\|^2_{L^2(\Omega)} + \frac{Q}{\varepsilon}\int_0^t (t-s)^{-\alpha^*}\left\|\partial_s v(\cdot,s)\right\|^2_{L^2(\Omega)}ds.$$

We employ the above estimates and follow the proof of Lemma 3.3 to arrive at the uniqueness result, the details of which are omitted due to similarity. ∎

We then follow an argument similar to that in the proofs of Theorems 4.1 and 4.2 to obtain the well-posedness and regularity estimates of problem (5.1).

THEOREM 5.5. *Suppose assumptions* (a)–(c) *hold and* $F \in L^2(L^2)$*; then* (5.1) *admits a unique solution* $z\in H^1(L^2)\cap L^2(H^2)$ *such that*

$$\|z\|_{H^1(L^2)} + \|z\|_{L^2(H^2)} \leq Q\|F\|_{L^2(L^2)}, \tag{5.19}$$

where Q *is independent of* z *or* F.

Proof. Based on Lemmas 5.3 and 5.4, the proof of this theorem follows from that of Theorem 4.1 and is thus omitted. ∎

THEOREM 5.6. *Suppose that assumptions* (a)–(d) *hold and* $F \in L^p(L^2)$ *for* $2 < p<\infty$*; then the solution to* (5.1) *has the following regularity estimate:*

$$\|z\|_{W^{1,p}(L^2)} + \left\|\hat{\mathcal{L}}z\right\|_{L^p(L^2)} \leq Q\|F\|_{L^p(L^2)}. \tag{5.20}$$

Here the positive constant Q *does not depend on* z *or* F.

Proof. The proof could be carried out following that of Theorem 4.2 with the assistance of Lemma 5.2 and (5.17), and is thus omitted. ∎

The following corollary is a direct consequence of the above two theorems.

COROLLARY 5.7. *Suppose assumptions* (a)–(d) *hold. If* u, $u_d \in L^2(L^2)$, *the adjoint problem* (6.1) *has a unique solution* $z\in H^1(L^2)\cap L^2(H^2)$ *and*

$$\|z\|_{H^1(L^2)} + \|z\|_{L^2(H^2)} \leq Q\|u-u_d\|_{L^2(L^2)}. \tag{5.21}$$

If further u, $u_d \in L^p(L^2)$ *for* $2<p<\infty$, *we have*

$$\|z\|_{W^{1,p}(L^2)} + \left\|\hat{\mathcal{L}}z\right\|_{L^p(L^2)} \leq Q\|u-u_d\|_{L^p(L^2)}. \tag{5.22}$$

Here the positive constant Q *is independent of* u, z, *or* u_d.

6. Analysis of optimal control problem.

We prove the well-posedness of the optimal control problem (1.2)–(1.3) and derive its first-order optimality conditions, as well as analyze the maximal regularity of solutions.

THEOREM 6.1. *Suppose assumptions* (a)–(c) *hold. The optimal control problem* (1.2)–(1.3) *admits a unique solution* (u,q) *such that* $u\in H^1(L^2)\cap L^2(H^2)$.

Proof. Since $\hat{J}(q) := J(u(q),q) \geq 0$, the infimum $J_{inf} := \inf_{q\in U_{ad}}\hat{J}(q)$ exists. Thus there exists a sequence $\{q^{(l)}\}_{l=1}^\infty \subset U_{ad}$ with $J_{inf} \leq \hat{J}(q^{(l+1)}) \leq \hat{J}(q^{(l)})$ such that $\lim_{l\to\infty}\hat{J}(q^{(l)}) = J_{inf}$ [67, Theorem 3.3]. Let $u^{(l)}$ be the solution to (1.3) with q being replaced by $q^{(l)}$. By $\lim_{l\to\infty}\hat{J}(q^{(l)}) = J_{inf}$ we have $\|q^{(l)}\|_{L^2(L^2)} \leq Q_0$ for some

$Q_0 > 0$ and $l \geq 1$. Then we apply this and the estimate (4.1) (that is, $\|u^{(l)}\|_{H^1(L^2)} + \|u^{(l)}\|_{L^2(H^2)} \leq Q(\|f\|_{L^2(L^2)} + \|q^{(l)}\|_{L^2(L^2)}))$ to obtain

$$\big\|u^{(l)}\big\|_{H^1(L^2)} + \big\|u^{(l)}\big\|_{L^2(H^2)} \leq Q\big(\|f\|_{L^2(L^2)} + Q_0\big),$$

which further implies $\big\|\partial_t^{\alpha(t)} u^{(l)}\big\|_{L^2(L^2)} \leq Q\big(\|f\|_{L^2(L^2)} + Q_0\big)$ by Lemmas 3.1 and 3.2. As a consequence, there exist weakly convergent subsequences $\{u^{(l_j)}\}_{j=1}^\infty \subset \{u^l\}_{l=1}^\infty$ and $\{q^{(l_j)}\}_{j=1}^\infty \subset \{q^{(l)}\}_{l=1}^\infty$ such that $\{u^{(l_j)}\}_{j=1}^\infty \to u_*$ weakly in $H^1(L^2) \cap L^2(H^2)$, $\{\partial_t^{\alpha(\boldsymbol{x},t)} u^{(l_j)}\}_{j=1}^\infty \to \partial_t^{\alpha(\boldsymbol{x},t)} u_*$ weakly in $L^2(L^2)$, and $\{q^{(l_j)}\}_{j=1}^\infty \to q_*$ weakly in $L^2(L^2)$, respectively. Since U_{ad} is a closed convex set, U_{ad} is weakly closed and thus $q^* \in U_{ad}$ [56].

To prove that u_* solves the state equation (1.3) associated with q^*, we multiply (1.3) associated with q_* by any $\phi \in C^\infty(\Omega \times [0,T])$ with $\phi(\boldsymbol{x},T) = 0$ for $\boldsymbol{x} \in \Omega$ and $\phi(\boldsymbol{x},t) = 0$ for any $(\boldsymbol{x},t) \in \partial\Omega \times [0,T]$, and we integrate the resulting equation by parts to obtain

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

which implies that u_* solves the state equation (1.3) associated with q_*. Here we use the fact that ${}^R\hat{\partial}_t^{\alpha(\boldsymbol{x},t)}$ defined by (6.3) is the adjoint operator of $\partial_t^{\alpha(\boldsymbol{x},t)}$ [65]. By the weakly lower semicontinuity of $\hat{J}(q)$, we have $J_{inf} \geq \liminf_{j\to\infty} \hat{J}(q^{(k)}) \geq \hat{J}(q_*) \geq J_{inf}$. Thus, (u_*, q_*) is a solution to the optimal control problem (1.2)–(1.3). The uniqueness of the solution follows from the strict convexity of J. ∎

THEOREM 6.2. *Suppose assumptions* (a)–(c) *hold; there exists an adjoint state z such that (u,q,z) satisfies the state equation* (1.3) *and the adjoint equation*

$$\begin{aligned}
-\partial_t z + {}^R\hat{\partial}_t^{\alpha(\boldsymbol{x},t)}(kz) + \hat{\mathcal{L}}(\boldsymbol{x},t)z = u(\boldsymbol{x},t;q) - u_d(\boldsymbol{x},t),\ (\boldsymbol{x},t) \in \Omega \times [0,T); \\
z(\boldsymbol{x},T) = 0,\ \boldsymbol{x} \in \Omega; \quad z(\boldsymbol{x},t) = 0,\ (\boldsymbol{x},t) \in \partial\Omega \times [0,T]
\end{aligned} \tag{6.1}$$

with the variational inequality

$$\int_0^T \int_\Omega (\gamma q + z)(v - q)d\boldsymbol{x}dt \geq 0 \qquad \forall v \in U_{ad}. \tag{6.2}$$

Here the backward variable-order Riemann–Liouville differential operator ${}^R\hat{\partial}_t^{\alpha(\boldsymbol{x},t)}$ is the adjoint operator of $\partial_t^{\alpha(\boldsymbol{x},t)}$ defined as

(6.3)

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

In addition, suppose assumptions (a)–(d) *hold and* $f, q, u_d \in L^p(L^2)$ *for* $2 \le p < \infty$*; then the optimal control problem* (1.2)–(1.3) *has a unique solution* $u \in H^1(L^2) \cap L^2(H^2)$ *and* $q \in H^1(L^2)$*, and the adjoint equation* (6.1) *has a unique solution* $z \in H^1(L^2) \cap L^2(H^2)$ *such that*

(6.4)
$$
\begin{aligned}
\|u\|_{L^2(H^2)} + \|u\|_{W^{1,p}(L^2)} + \big\|\mathcal{L}u\big\|_{L^p(L^2)} &\le Q\big(\|f\|_{L^p(L^2)} + \|q\|_{L^p(L^2)}\big),\\
\|z\|_{L^2(H^2)} + \|z\|_{W^{1,p}(L^2)} + \big\|\hat{\mathcal{L}}z\big\|_{L^p(L^2)} &\le Q\big(\|u_d\|_{L^p(L^2)} + \|f\|_{L^p(L^2)} + \|q\|_{L^p(L^2)}\big),\\
\|q\|_{H^1(L^2)} &\le Q\big(\|u_d\|_{L^2(L^2)} + \|f\|_{L^2(L^2)} + \|q\|_{L^2(L^2)}\big),
\end{aligned}
$$

where the positive constant Q *is independent of* f, q , u, u_d, *or* z.

Proof. The proof of (6.1)–(6.2) could be carried out following a procedure similar to that of [65, Theorem 2.1], and thus we will provide only a brief outline of the proof here. For any $q \in U_{ad}$ and $0 < \varepsilon \ll 1$, let $\delta q := v - q$, and we have $q + \varepsilon\delta q \in U_{ad}$. We note from (1.3) that $\delta_\varepsilon u(q) := \big(u(q + \varepsilon\delta q) - u(q)\big)/\varepsilon$ satisfies the homogeneous initial-boundary value problem of

$$
\partial_t \delta_\varepsilon u + k(\boldsymbol{x},t)\, \partial_t^{\alpha(\boldsymbol{x},t)} \delta_\varepsilon u + \mathcal{L}(\boldsymbol{x},t)\delta_\varepsilon u = \delta q(\boldsymbol{x},t), \;\; (\boldsymbol{x},t) \in \Omega \times (0,T], \tag{6.5}
$$

which has ε-independent coefficients and right-hand-side term. Hence, the solution $\delta_\varepsilon u = \delta u$ is independent of ε by the uniqueness of the solution to problem (6.5) proved in Theorem 4.1.

We write $\hat{J}(q) = J(u(q), q)$, using $u(q+\varepsilon\delta q) - u(q) = \varepsilon\delta u$ and the adjoint properties of $\partial_t^{\alpha(\boldsymbol{x},t)}$ and ${}^R\hat{\partial}_t^{\alpha(\boldsymbol{x},t)}$, $\mathcal{L}(\boldsymbol{x},t)$ and $\hat{\mathcal{L}}(\boldsymbol{x},t)$ to obtain

$$
\begin{aligned}
0 \le \partial_q \hat{J}(q)\delta q &= \lim_{\varepsilon\to 0^+} \varepsilon^{-1}\big(\hat{J}(q+\varepsilon\delta q) - \hat{J}(q)\big)\\
&= \int_0^T \int_\Omega (u(q) - u_d)\delta u \, d\boldsymbol{x}dt + \int_0^T \int_\Omega \gamma q \delta q \, d\boldsymbol{x}dt\\
&= \int_0^T \int_\Omega \big(- \partial_t z + {}^R\hat{\partial}_t^{\alpha(\boldsymbol{x},t)}(kz) + \hat{\mathcal{L}}(\boldsymbol{x},t)z\big)\delta u \, d\boldsymbol{x}dt + \int_0^T \int_\Omega \gamma q\delta q \, d\boldsymbol{x}dt\\
&= \int_0^T \int_\Omega z\big(\partial_t \delta u + k(\boldsymbol{x},t)\, \partial_t^{\alpha(\boldsymbol{x},t)} \delta u + \mathcal{L}(\boldsymbol{x},t)\delta u\big) d\boldsymbol{x}dt + \int_0^T \int_\Omega \gamma q \delta q \, d\boldsymbol{x}dt\\
&= \int_0^T \int_\Omega \big(\gamma q + z\big)\delta q \, d\boldsymbol{x}dt\\
&= \int_0^T \int_\Omega \big(\gamma q + z\big)(v - q) \, d\boldsymbol{x}dt,
\end{aligned} \tag{6.6}
$$

which proves the first statement of this theorem.

Let $P_{U_{ad}}(g) := \max\{q_*, \min\{g, q^*\}\} \in H^1(L^2)$ be the pointwise projection onto the admissible set U_{ad} with the following estimate [37]:

$$
\|P_{U_{ad}}(g)\|_{H^s} \le \|g\|_{H^s} \quad \forall g \in H^s,\ 0 \le s \le 1. \tag{6.7}
$$

Then inequality (6.2) implies that [26, 67]

$$
q(\boldsymbol{x},t) = P_{U_{ad}}\big(- z(\boldsymbol{x},t)/\gamma\big). \tag{6.8}
$$

For regularity estimates, we use assumptions (a)–(c) and the fact $q \in L^2(L^2)$ and apply Theorem 4.1 to deduce $u \in H^1(L^2) \cap L^2(H^2)$. Then we apply Corollary 5.7 to conclude that the adjoint equation (6.1) has a unique solution $z \in H^1(L^2) \cap L^2(H^2)$. We combine (5.21)–(5.22) with (4.1) as well as the Sobolev embedding $H^1(L^2) \hookrightarrow L^p(L^2)$ for $2 \le p < \infty$ and $\|\hat{\mathcal{L}} z\|_{L^2(L^2)} \le Q\|z\|_{L^2(H^2)}$ to prove the second estimate in (6.4) for z. We use estimates (4.1)–(4.2) and $\|\mathcal{L} u\|_{L^2(L^2)} \le Q\|u\|_{L^2(H^2)}$ to get the estimate of u in (6.4). With the above estimates, we finally combine (6.8) with (6.7) and the estimate of z in (6.4) with $p=2$ to conclude that

$$\|q\|_{H^1(L^2)} \le Q\|z\|_{H^1(L^2)} \le Q\big(\|u_d\|_{L^2(L^2)} + \|f\|_{L^2(L^2)} + \|q\|_{L^2(L^2)}\big), \tag{6.9}$$

which completes the proof of the theorem. ∎

7. Numerical simulation.

We develop the numerical scheme for the state equation (1.3) and the adjoint state equation (6.1) and then perform numerical experiments in one and two space dimensions to study the optimal control problem (1.2)–(1.3) and its potential applications in, e.g., contaminant remediation. All the numerical experiments are implemented using MATLAB R2018a on a Dell XPS 15 laptop with Intel Core i7-8750H CPU Duo of 2.20G CPU and 16.0 GB RAM.

7.1. Discretization scheme.

Partition $[0,T]$ by $t_n := n\tau$ for $\tau := T/N$ and $0 \le n \le N$. Define a quasi-uniform mesh on Ω and let $S(\Omega)$ be the continuous and piecewise-linear finite element space on Ω with respect to the partition. Let $k_n := k(\boldsymbol{x}, t_n)$, $\boldsymbol{A}_n := \boldsymbol{A}(\boldsymbol{x}, t_n)$, $\boldsymbol{b}_n := \boldsymbol{b}(\boldsymbol{x}, t_n)$, $c_n := c(\boldsymbol{x}, t_n)$, $f_n := f(\boldsymbol{x}, t_n)$, $q_n := q(\boldsymbol{x}, t_n)$, $u_n := u(\boldsymbol{x}, t_n)$, and $z_n := z(\boldsymbol{x}, t_n)$. Then we follow [65, 66] to discretize $\partial_t u$, $\partial_t^{\alpha(\boldsymbol{x},t)} u$, $-\partial_t z$ and $\hat{\partial}_t^{\alpha(\boldsymbol{x},t)}(kz)$ at t_n for $1 \le n \le N$ by

$$\begin{aligned} &\partial_t u(\boldsymbol{x}, t_n) \approx \delta_\tau u_n := \frac{u_n - u_{n-1}}{\tau},\ \partial_t^{\alpha(\boldsymbol{x},t_n)} u(\boldsymbol{x}, t_n) \approx \delta_\tau^{\alpha(\boldsymbol{x},t_n)} u_n, \\ &-\partial_t z(\boldsymbol{x}, t_{n-1}) \approx -\delta_\tau z_n,\ \hat{\partial}_t^{\alpha(\boldsymbol{x},t)}(kz)(\boldsymbol{x}, t_{n-1}) \approx \hat{\delta}_\tau^{\alpha(\boldsymbol{x},t_{n-1})}(kz)_{n-1}, \end{aligned} \tag{7.1}$$

where

$$\delta_\tau^{\alpha(\boldsymbol{x},t_n)} u_n := \sum_{k=1}^{n} b_{n,k}(u_k - u_{k-1}),\ b_{n,k} := \frac{(t_n - t_{k-1})^{1-\alpha(\boldsymbol{x},t_n)} - (t_n - t_k)^{1-\alpha(\boldsymbol{x},t_n)}}{\Gamma(2 - \alpha(\boldsymbol{x}, t_n))\tau},$$

$$\hat{\delta}_\tau^{\alpha(\boldsymbol{x},t_{n-1})}(kz)_{n-1} = b_{n,n}(kz)_{n-1} + \sum_{k=n+1}^{N} (b_{k,n} - b_{k,n+1})(kz)_{k-1}.$$

We employ them to get a fully discrete finite element scheme for optimal control following [65, 66]: find $U = \{U_n\}_{n=1}^N \subset S(\Omega)$ with $U_0 = 0$, $Z = \{Z_n\}_{n=0}^{N-1} \subset S(\Omega)$ with $Z_N = 0$ and $Q = \{Q_n\}_{n=0}^{N-1}$ such that for $n = 1, 2, \ldots, N$ and $\chi \in S(\Omega)$

$$\begin{aligned} &(\delta_\tau U_n + k_n \delta_\tau^{\alpha(\boldsymbol{x},t_n)} U_n, \chi) + (\boldsymbol{A}_n \nabla U_n, \nabla\chi) + (\boldsymbol{b}_n \cdot \nabla U_n, \chi) \\ &\qquad + (c_n U_n, \chi) = (f_n + Q_{n-1}, \chi), \end{aligned} \tag{7.2a}$$

$$\begin{aligned} &\big(-\delta_\tau Z_n + \hat{\delta}_\tau^{\alpha(\boldsymbol{x},t_{n-1})}(kZ)_{n-1}, \chi\big) + (\boldsymbol{A}_n \nabla Z_n + \boldsymbol{b}_n Z_n, \nabla\chi) \\ &\qquad + (c_n Z_n, \chi) = \big(U_n - u_d(\cdot, t_n), \chi\big), \end{aligned} \tag{7.2b}$$

$$Q_{n-1}(\boldsymbol{x}) = \mathcal{P}_{U_{ad}}\big(-Z_{n-1}(\boldsymbol{x})/\gamma\big). \tag{7.2c}$$

In practical computations, we iteratively solve the discretized first-order optimality condition starting from an initial guess of Q, solving (7.2a), (7.2b), and (7.2c) in order and then updating Q for the next round computation until the difference of two contiguous Q under a certain norm is less than the tolerance. In this work we adopt the discrete-in-time norms defined by $\|v\|_{\hat{L}^\infty(L^2)} := \max_{1\le j\le N}\|v_j\|$ and $\|v\|_{\hat{L}^2(L^2)} := \big(\tau\sum_{j=1}^N\|v_j\|^2\big)^{\frac{1}{2}}$ for $v=\{v_j\}_{j=1}^N$.

7.2. One-dimensional example.

7.2.1. Convergence test. The data are as follows: $\Omega=(0,1)$, $[0,T]=[0,1]$, $k=1$, $A=0.01$, $b=0.1$, $c=0.01$, $q_*=0.2$, $q^*=0.3$, $\gamma=1$, $f=1$, $u_d(x,t)=1-4(x-\frac{1}{2})^2$, and $\alpha(x,t)=\eta(t)\phi(x)$ with

$$\eta(t)=\alpha(T)+(\alpha(0)-\alpha(T))\Big(1-\frac{t}{T}-\frac{1}{2\pi}\sin\Big(2\pi\Big(1-\frac{t}{T}\Big)\Big)\Big),\ \ \phi(x)=1+\frac{\sin(0.5\pi x)}{10}. \tag{7.3}$$

As the exact solution is not available, we use the numerical solution computed with $\tau_f=1/720$ and $h_f=1/360$ as the reference solution.

We investigate the spatial convergence behavior of the numerical algorithm proposed at the beginning of section 7 under the fine time step size τ_f, and similarly we investigate the temporal convergence behavior under the fine spatial mesh size h_f. We present numerical results in Tables 7.1 and 7.2 and observe that the errors shrink as either the temporal or spatial mesh becomes finer, which indicates the reliability and efficiency of the numerical algorithm to the optimal control problem (1.2)–(1.3) under different variable orders.

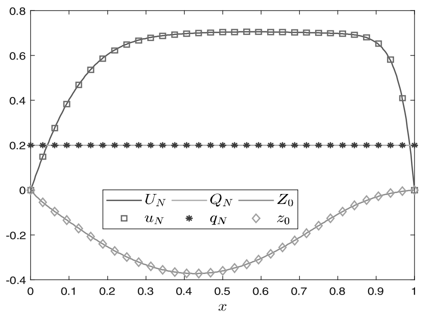
We then present solution curves at the terminal time step in Figure 7.1 by choosing the same data as before, except for $\tau=1/720$ and $h=1/32$, $\alpha(0)=0.3$, and $\alpha(T)=0.8$. Numerical results show that the numerical solutions computed under the coarse grid $h=1/32$ provide an accurate approximation, which again demonstrates

TABLE 7.1
Accuracy of numerical method in section 7.2.1 with $\alpha(0)=0.9$ and $\alpha(T)=0.2$.

τ	1/8	1/16	1/32	1/64
$\|q-Q\|_{\hat{L}^\infty(L^2)}$	3.40E-02	1.81E-02	1.21E-02	8.02E-03
$\|u-U\|_{\hat{L}^\infty(L^2)}$	1.68E-02	8.83E-03	5.97E-03	3.97E-03
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	3.43E-02	1.83E-02	1.22E-02	8.09E-03
h	1/60	1/72	1/90	1/120
$\|q-Q\|_{\hat{L}^\infty(L^2)}$	9.01E-04	5.96E-04	3.83E-04	2.04E-04
$\|u-U\|_{\hat{L}^\infty(L^2)}$	1.49E-03	1.00E-03	6.34E-04	3.38E-04
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	9.58E-04	6.59E-04	4.13E-04	2.21E-04

TABLE 7.2
Accuracy of numerical method in section 7.2.1 with $\alpha(0)=0.3$ and $\alpha(T)=0.8$.

τ	1/8	1/16	1/32	1/64
$\|q-Q\|_{\hat{L}^\infty(L^2)}$	2.58E-02	1.26E-02	8.37E-03	5.49E-03
$\|u-U\|_{\hat{L}^\infty(L^2)}$	2.89E-02	1.50E-02	1.02E-02	6.82E-03
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	2.96E-02	1.51E-02	1.01E-02	6.71E-03
h	1/60	1/72	1/90	1/120
$\|q-Q\|_{\hat{L}^\infty(L^2)}$	1.27E-03	8.68E-04	5.44E-04	2.89E-04
$\|u-U\|_{\hat{L}^\infty(L^2)}$	1.64E-03	1.10E-03	6.98E-04	3.72E-04
$\|z-Z\|_{\hat{L}^\infty(L^2)}$	1.27E-03	8.69E-04	5.44E-04	2.89E-04



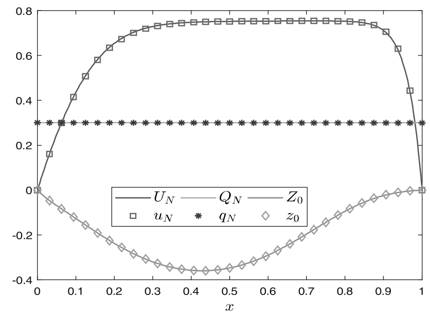


FIG. 7.1. *Plots of* U_N, Z_0, *and* Q_N *in section* 7.2.1 *with* $(q_*, q^*) = (0.2, 0.3)$ *(left) and* $(q_*, q^*) = (0.3, 0.4)$ *(right).*

TABLE 7.3
Performance of the optimal control problem (1.2)–(1.3) *in subsection* 7.2.2.

k	$\Vert U - u_d\Vert_{\hat{L}^2(L^2)}$	$\Vert Q\Vert_{\hat{L}^2(L^2)}$	$J(U,Q)$
0	3.83E-01	2.23E-01	2.03E-01
0.5	3.86E-01	2.45E-01	2.05E-01
5	5.77E-01	2.68E-01	3.02E-01
25	6.91E-01	2.40E-01	3.57E-01

the effectiveness of the numerical method to the optimal control problem (1.2)–(1.3) under different constraints.

7.2.2. Performance under different transport scales. We investigate the performance of the optimal control problem (1.2)–(1.3) under different transport scales, which is inherently characterized by the partition coefficient k as discussed in section 2. Specifically, as k tends to 0, the model (1.2)–(1.3) describes the optimal control of, e.g., contaminant transport mainly under Fickian diffusion, while for large k the transport process is mainly governed by the subdiffusion.

We present optimal control results under different k in Table 7.3 with the same data as those for Figure 7.1, except for $(q_*, q^*) = (0.2, 0.3)$ and $\gamma = 0.1$. We observe that the control variable has similar values under different k, while the difference $U - u_d$ becomes larger from Fickian diffusion to subdiffusion, which indicates that the contaminant remediation in heterogeneous media is more difficult than that in homogeneous media. This phenomenon is probably caused by the fact that as k grows, a larger amount of contaminant solute gets absorbed into the aquifers that undergoes the subdiffusion, which is more difficult to be remediated than those in the bulk fluid phase that undergoes a Brownian motion and may thus require more expensive controls to reach the ideal distribution u_d.

7.2.3. Performance under multiscale coefficients. We consider the optimal control problem (1.2)–(1.3) with multiscale coefficients, which arises from, e.g., the local effective dispersion tensor used to account for the relatively large particle motions; cf. [11, section 1] and [29, 47]. We set $\Omega \times [0,T] = [0,1] \times [0,0.5]$ and

$$A\Big(x, \frac{x}{\epsilon}\Big) = \frac{0.01}{1 + 0.2\sin\left(2\pi x + \frac{2\pi x}{\epsilon}\right)}, \quad b\Big(x, \frac{x}{\epsilon}\Big) = 0.5 + 0.2\sin\left(2\pi x + \frac{2\pi x}{\epsilon}\right),$$
$$(7.4)$$
$$c\Big(x, \frac{x}{\epsilon}\Big) = -\frac{0.4\pi(\epsilon+1)}{\epsilon}\cos\left(2\pi x + \frac{2\pi x}{\epsilon}\right)$$

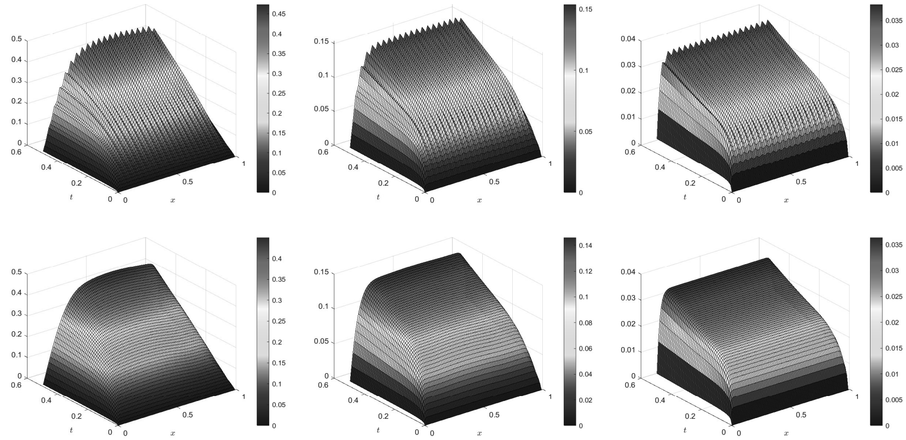


FIG. 7.2. *Surfaces of the state variable U with $k=0.5$ (left), $k=5$ (middle), and $k=25$ (right). First row: multiscale coefficients. Second row: averaged coefficients.*

TABLE 7.4
Performance of the optimal control problem (1.2)–(1.3) *with multiscale coefficients in section* 7.2.3.

k	$\|U-u_d\|_{\hat{L}^2(L^2)}$	$\|Q\|_{\hat{L}^2(L^2)}$	$J(U,Q)$
0.5	3.77E-01	9.47E-02	2.36E-01
5	4.57E-01	7.36E-02	2.65E-01
25	5.00E-01	7.07E-02	2.85E-01

with $\epsilon = 1/63$, which have averaged values $\overline{A} \approx 0.01$, $\overline{b} = 0.5$, and $\overline{c} = 0$ for comparison as performed in [11]. We follow the same data as those for Figure 7.1 except for $q_* = 0.1$, $q^* = 0.5$, $\tau = 1/2048$, $h = 1/256$, and $\alpha = \eta(t)$ with $\eta(t)$ defined as in (7.3). Numerical results are presented in Figure 7.2 and Table 7.4, from which we observe the following: (i) The state variables under multiscale coefficients have similar amplitudes to those under averaged coefficients, while they vary rapidly along the spatial direction that coincides with the multiscale nature of the solutions. (ii) The state variable approaches the ideal distribution u_d more closely as the partition coefficient k gets smaller, which indicates that the heterogeneity of the medium makes the contaminant remediation more difficult, as we have found in section 7.2.2. These observations indicate that the proposed model provides adequate descriptions for contaminant transport through heterogeneous porous media with multiscale coefficients.

7.3. Two-dimensional example. Let $\Omega \times [0,T] = (0,1)^2 \times [0,1]$, $k = 1$, $\boldsymbol{A} = \mathrm{diag}(0.01, 0.01)$, $\boldsymbol{b} = (0.1, 0.1)$, $c = 0.01$, $\gamma = 1$, $u_d = \sin(\pi x)\sin(\pi y)$, and

$$\alpha(x,y,t) = \frac{3+2t}{10}(1-0.2x^2)(1+0.1\sin(\pi y)).$$

The source term f is a mollified point source of the form

$$f = \bar{u} \times \left[\mathrm{erf}\left(\frac{x-z_1}{\sigma}\right) - \mathrm{erf}\left(\frac{x-z_0}{\sigma}\right)\right]\left[\mathrm{erf}\left(\frac{y-z_1}{\sigma}\right) - \mathrm{erf}\left(\frac{y-z_0}{\sigma}\right)\right]$$

with $\bar{u} = 0.5$, $\sigma = 0.001$, $z_0 = 0.4$, and $z_1 = 0.6$. In Figures 7.3 and 7.4, we plot the state variable U and the control variable Q at $t = T/2$ and $t = T$, respectively.

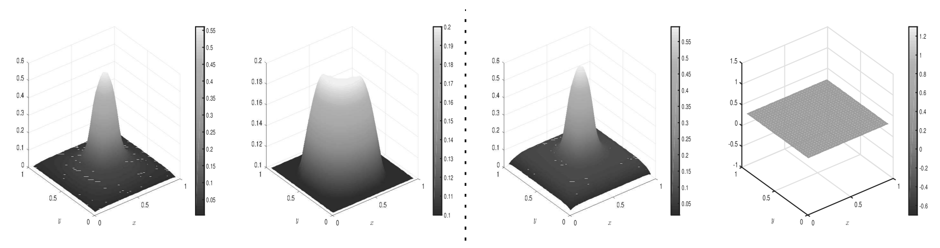


FIG. 7.3. *Optimal control results at* $t = T/2$ *in section* 7.3. *Left-hand panel: plots of* U *(left) and* Q *(right) with* $(q_*, q^*) = (0.1, 0.2)$. *Right-hand panel: corresponding plots with* $(q_*, q^*) = (0.3, 0.4)$.

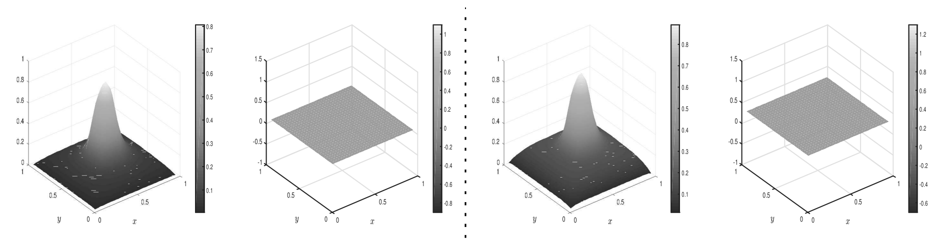


FIG. 7.4. *Optimal control results at* $t = T$ *in section* 7.3. *Left-hand panel: plots of* U *(left) and* Q *(right) with* $(q_*, q^*) = (0.1, 0.2)$. *Right-hand panel: corresponding plots with* $(q_*, q^*) = (0.3, 0.4)$.

We observe that as time evolves, the state variable U gets closer to the ideal distribution u_d (e.g., the hump in U becomes higher from $t = T/2$ to $t = T$ and approaches u_d), which indicates the applicability and effectiveness of the proposed model in modeling challenging phenomena such as contaminant remediation in heterogeneous media.

8. Concluding remarks. This work investigates an optimal control model governed by a two-time-scale fractional advection-diffusion-reaction equation with space-time-dependent fractional order and coefficients. The well-posedness, first-order optimality conditions, and maximal regularity of the optimal control problem are proved, and numerical experiments are performed to investigate its potential applications.

A potential extension of the current work is to investigate the more complicated optimal control of the miscible displacement system, in which the state equations contain additional pressure and velocity equations. The optimal control governed by the integer-order miscible displacement system was investigated in [10], where the diffusion coefficient depends on the concentration. In general, one could combine the techniques in [10] with the methods in the current work to analyze the two-time-scale fractional miscible displacement system, while further studies need to be conducted on how the coupling of the system may affect the analysis of the optimal control.

Another interesting topic is to prove high-order regularity estimates to support the numerical analysis. In principle the developed compact mapping method could be applied, while, due to the effects of the weakly singular kernel in (1.5), $\partial_t^2 u$ could be unbounded near the initial time $t = 0$ [63] such that we should construct the compact mapping between Sobolev spaces $W^{2,p}(0,T)$ based on the derivative properties of variable-order fractional operators. We will investigate this topic in the near future.

As discussed in section 2, the MSD of the two-time-scale model with $k=1$ approximates that of the single-time-scale model away from the initial time, and for large k the MSD of the two-time-scale model exhibits similar behavior to that of the single-time-scale model on the entire time interval. Since the analysis of this work is valid for any bounded k, the derived results hold true for the fractional-order dominant case that approximates the single-time-scale model. Nevertheless, there remains a gap, namely, analyzing the single-time-scale variable-order fractional advection-diffusion-reaction model

$$\partial_t^{\alpha(\boldsymbol{x},t)} u + \mathcal{L}(\boldsymbol{x},t)u = f(\boldsymbol{x},t).$$

Due to the impacts of the variable order, existing methods for its constant-order counterpart such as analytical techniques [15, 50, 51] and spectral expansions [51] are not applicable, while energy arguments in, e.g., [43, 44] cannot be directly performed as the variable-order fractional operators lose the favorable coercivity. It is unclear how the current analysis techniques could be applied to the above model and its optimal control problem.

Acknowledgment. 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.

REFERENCES

[1] R. Adams and J. Fournier, *Sobolev Spaces*, Elsevier, San Diego, 2003.

[2] H. Antil, E. Otárola, and A. J. Salgado, *A space-time fractional optimal control problem: Analysis and discretization*, SIAM J. Control Optim., 54 (2016), pp. 1295–1328, https://doi.org/10.1137/15M1014991.

[3] H. Antil, E. Otárola, and A. Salgado, *Optimization with respect to order in a fractional diffusion model: Analysis, approximation and algorithmic aspects*, J. Sci. Comput., 77 (2018), pp. 204–224.

[4] M. Balhoff, *An Introduction to Multiphase, Multicomponent Reservoir Simulation*, Develop. Petroleum Sci. 75, Elsevier, 2022.

[5] L. Banjai and C. Makridakis, *A posteriori error analysis for approximations of time-fractional subdiffusion problems*, Math. Comp., 91 (2022), pp. 1711–1737.

[6] J. Bear, *Dynamics of Fluids in Porous Media*, Elsevier, New York, 1972.

[7] D. Benson, R. Schumer, M. Meerschaert, and S. Wheatcraft, *Fractional dispersion, Lévy motions, and the MADE tracer tests*, Transp. Porous Media, 42 (2001), pp. 211–240.

[8] D. Benson, S. Wheatcraft, and M. Meerschaeert, *The fractional order governing equations of Lévy motion*, Water Resour. Res., 36 (2000), pp. 1413–1423.

[9] D. L. Brown, J. Gedicke, and D. Peterseim, *Numerical homogenization of heterogeneous fractional Laplacians*, Multiscale Model. Simul., 16 (2018), pp. 1305–1332, https://doi.org/10.1137/17M1147305.

[10] Y. Chang, W. Cao, D. Yang, T. Sun, and W. Liu, *Finite element approximation of optimal control governed by immiscible displacement in porous media*, Int. J. Numer. Anal. Model., 11 (2014), pp. 688–714.

[11] A. Cheng, K. Wang, and H. Wang, *A preliminary study on multiscale ELLAM schemes for transient advection-diffusion equations*, Numer. Methods Partial Differential Equations, 26 (2010), pp. 1405–1419.

[12] E. Chung, Y. Efendiev, W. Leung, M. Vasilyeva, and Y. Wang, *Non-local multi-continua upscaling for flows in heterogeneous fractured media*, J. Comput. Phys., 372 (2018), pp. 22–34.

[13] E. Chung, Y. Efendiev, W. Leung, and M. Vasilyeva, *Nonlocal multicontinua with representative volume elements. Bridging separable and non-separable scales*, Comput. Methods Appl. Mech. Engrg., 377 (2021), 113687.

[14] M. D'Elia, C. Glusa, and E. Otárola, *A priori error estimates for the optimal control of the integral fractional Laplacian*, SIAM J. Control Optim., 57 (2019), pp. 2775–2798, https://doi.org/10.1137/18M1219989.

[15] K. Diethelm and N. Ford, *Analysis of fractional differential equations*, J. Math. Anal. Appl., 265 (2002), pp. 229–248.
[16] W. Deng, B. Li, W. Tian, and P. Zhang, *Boundary problems for the fractional and tempered fractional operators*, Multiscale Model. Simul., 16 (2018), pp. 125–149, https://doi.org/10.1137/17M1116222.
[17] W. Deng, X. Wang, and P. Zhang, *Anisotropic nonlocal diffusion operators for normal and anomalous dynamics*, Multiscale Model. Simul., 18 (2020), pp. 415–443, https://doi.org/10.1137/18M1184990.
[18] P. Embrechts and M. Maejima, *Selfsimilar Processes*, Princeton Ser. Appl. Math., Princeton University Press, Princeton, NJ, 2002.
[19] L. Evans, *Partial Differential Equations*, Grad. Stud. Math. 19, American Mathematical Society, 1998.
[20] R. Ewing, *The Mathematics of Reservoir Simulation*, Frontiers Appl. Math. 1, SIAM, Philadelphia, 1984, https://doi.org/10.1137/1.9781611971071.
[21] 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, 2015, https://doi.org/10.2790/379646.
[22] W. Gong, M. Hinze, and Z. Zhou, *Finite element method and a priori error estimates for Dirichlet boundary control problems governed by parabolic PDEs*, J. Sci. Comput., 66 (2016), pp. 941–967.
[23] R. Gorenflo, Y. Luchko, and M. Yamamoto, *Time-fractional diffusion equation in the fractional Sobolev spaces*, Fract. Calc. Appl. Anal., 18 (2015), pp. 799–820.
[24] R. Garrappa, A. Giusti, and F. Mainardi, *Variable-order fractional calculus: A change of perspective*, Commun. Nonlinear Sci. Numer. Simul., 102 (2021), 105904.
[25] M. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, Philadelphia, 2002, https://doi.org/10.1137/1.9780898718720.
[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. Control Optim., 57 (2019), pp. 241–263, https://doi.org/10.1137/17M1155636.
[27] B. Haak and E. Ouhabaz, *Maximal regularity for non-autonomous evolution equations*, Math. Ann., 363 (2015), pp. 1117–1145.
[28] R. Herzog, G. Stadler, and G. Wachsmuth, *Directional sparsity in optimal control of partial differential equations*, SIAM J. Control Optim., 50 (2012), pp. 943–963, https://doi.org/10.1137/100815037.
[29] T. Hou and X. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
[30] K. Ito and K. Kunisch, *Semismooth Newton methods for time-optimal control for a class of ODEs*, SIAM J. Control Optim., 48 (2010), pp. 3997–4013, https://doi.org/10.1137/090753905.
[31] D. Jiang and Z. Li, *Coefficient inverse problem for variable order time-fractional diffusion equations from distributed data*, Calcolo, 59 (2022), 34.
[32] L. Jiang and N. Ou, *Bayesian inference using intermediate distribution based on coarse multiscale model for time fractional diffusion equations*, Multiscale Model. Simul., 16 (2018), pp. 327–355, https://doi.org/10.1137/17M1110535.
[33] 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), pp. 377–404.
[34] B. Jin, B. Li, and Z. Zhou, *Subdiffusion with a time-dependent coefficient: Analysis and numerical solution*, Math. Comp., 88 (2019), pp. 2157–2186.
[35] Y. Kian, E. Soccorsi, and M. Yamamoto, *On time-fractional diffusion equations with space-dependent variable order*, Ann. Henri Poincaré, 19 (2018), pp. 3855–3881.
[36] A. Kubica, K. Ryszewska, and M. Yamamoto, *Time-Fractional Differential Equations*, Springer Briefs in Math., Springer Nature, Singapore, 2020.
[37] K. Kunisch and B. Vexler, *Constrained Dirichlet boundary control in L^2 for a class of evolution equations*, SIAM J. Control Optim., 46 (2007), pp. 1726–1753, https://doi.org/10.1137/060670110.
[38] B. Li, H. Wang, and J. Wang, *Well-posedness and numerical approximation of a fractional diffusion equation with a nonlinear variable order*, ESAIM Math. Model. Numer. Anal., 55 (2021), pp. 171–207.
[39] G. Lin and A. Tartakovsky, *An efficient, high-order probabilistic collocation method on sparse grids for three-dimensional flow and solute transport in randomly heterogeneous porous media*, Adv. Water Resour., 32 (2009), pp. 712–722.

[40] J. Lions and E. Magenes, *Non-homogeneous Boundary Value Problems and Applications*, Vol. 1, Springer, Berlin, 1972.

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

[42] B. Maryshev, M. Joelson, D. Lyubimov, T. Lyubimova, and M. Néel, *Non Fickian flux for advection-dispersion with immobile periods*, J. Phys. A, 42 (2009), 115001.

[43] W. McLean, K. Mustapha, R. Ali, and O. Knio, *Well-posedness of time-fractional advection-diffusion-reaction equations*, Fract. Calc. Appl. Anal., 22 (2019), pp. 918–944.

[44] W. McLean, K. Mustapha, R. Ali, and O. Knio, *Regularity theory for time-fractional advection-diffusion-reaction equations*, Comput. Math. Appl., 79 (2020), pp. 947–961.

[45] M. Meerschaert and A. Sikorskii, *Stochastic Models for Fractional Calculus*, De Gruyter Stud. Math., 2011.

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

[47] P. Ming and X. Yue, *Numerical methods for multiscale elliptic problems*, J. Comput. Phys., 214 (2006), pp. 421–445.

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

[49] P. Neittaanmaki, D. Tiba, and M. Dekker, *Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications*, Marcel Dekker, New York, 1994.

[50] I. Podlubny, *Fractional Differential Equations*, Academic Press, 1999.

[51] 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), pp. 426–447.

[52] M. Schechter, *Principles of Functional Analysis*, American Mathematical Society, 2001.

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

[54] M. Stynes, E. O'Riordan, and J. L. Gracia, *Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation*, SIAM J. Numer. Anal., 55 (2017), pp. 1057–1079, https://doi.org/10.1137/16M1082329.

[55] H. Sun, Y. Zhang, W. Chen, and D. Reeves, *Use of a variable-index fractional-derivative model to capture transient dispersion in heterogeneous media*, J. Contam. Hydrol., 157 (2014), pp. 47–58.

[56] F. Troltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, Grad. Stud. Math., American Mathematical Society, 2010.

[57] H. Wang and X. Zheng, *Wellposedness and regularity of the variable-order time-fractional diffusion equations*, J. Math. Anal. Appl., 475 (2019), pp. 1778–1802.

[58] J. Webb, *Weakly singular Gronwall inequalities and applications to fractional differential equations*, J. Math. Anal. Appl., 471 (2019), pp. 692–711.

[59] L. Xiao, Q. Cai, Z. Li, H. Zhao, and R. Luo, *A multi-scale method for dynamics simulation in continuum solvent models.* I*: Finite-difference algorithm for Navier-Stokes equation*, Chem. Phys. Lett., 616–617 (2014), pp. 67–74.

[60] M. Yamamoto, *Fractional calculus and time-fractional differential equations: Revisit and construction of a theory*, Mathematics, 10 (2022), 698.

[61] M. Yamamoto, *On time fractional derivatives in fractional Sobolev spaces and applications to fractional ordinary differential equations*, in Nonlocal and Fractional Operators, SEMA SIMAI Springer Ser. 26, Springer, Cham, Switzerland, 2021, pp. 287–308.

[62] F. Zeng, Z. Zhang, and G. E. Karniadakis, *A generalized spectral collocation method with tunable accuracy for variable-order fractional differential equations*, SIAM J. Sci. Comput., 37 (2015), pp. A2710–A2732, https://doi.org/10.1137/141001299.

[63] X. Zheng and H. Wang, *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 (2021), pp. 1522–1545.

[64] X. Zheng and H. Wang, *An error estimate of a numerical approximation to a hidden-memory variable-order space-time fractional diffusion equation*, SIAM J. Numer. Anal., 58 (2020), pp. 2492–2514, https://doi.org/10.1137/20M132420X.

[65] X. Zheng and H. Wang, *A hidden-memory variable-order time-fractional optimal control model: Analysis and approximation*, SIAM J. Control Optim., 59 (2021), pp. 1851–1880, https://doi.org/10.1137/20M1344962.

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

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

[68] P. Zhuang, F. Liu, V. Anh, and I. Turner, *Numerical methods for the variable-order fractional advection-diffusion equation with a nonlinear source term*, SIAM J. Numer. Anal., 47 (2009), pp. 1760–1781, https://doi.org/10.1137/080730597.