Mathematics

Research article

The infinite-dimensional Pontryagin maximum principle for optimal control problems of fractional evolution equations with endpoint state constraints

Yuna Oh[1] **and Jun Moon**[2,*]

[1] Department of Artificial Intelligence, Hanyang University, Seoul 04763, South Korea

[2] Department of Electrical Engineering, Hanyang University, Seoul 04763, South Korea

* **Correspondence:** Email: junmoon@hanyang.ac.kr; Tel: +8222204064.

Abstract: In this paper, we study the infinite-dimensional endpoint state-constrained optimal control problem for fractional evolution equations. The state equation is modeled by the X-valued left Caputo fractional evolution equation with the analytic semigroup, where X is a Banach space. The objective functional is formulated by the Bolza form, expressed in terms of the left Riemann-Liouville (RL) fractional integral running and initial/terminal costs. The endpoint state constraint is described by initial and terminal state values within convex subsets of X. Under this setting, we prove the Pontryagin maximum principle. Unlike the existing literature, we do not assume the strict convexity of X^*, the dual space of X. This assumption is particularly important, as it guarantees the differentiability of the distance function of the endpoint state constraint. In the proof, we relax this assumption via a separation argument and constructing a family of spike variations for the Ekeland variational principle. Subsequently, we prove the maximum principle, including nontriviality, adjoint equation, transversality, and Hamiltonian maximization conditions, by establishing variational and duality analysis under the finite codimensionality of initial- and end-point variational sets. Our variational and duality analysis requires new representation results on left Caputo and right RL linear fractional evolution equations in terms of (left and right RL) fractional state transition operators. Indeed, due to the inherent complex nature of the problem of this paper, our maximum principle and its proof technique are new in the optimal control context. As an illustrative example, we consider the state-constrained fractional diffusion PDE control problem, for which the optimality condition is derived by the maximum principle of this paper.

1. Introduction

Let X be a Banach space. Consider the following X-valued left Caputo fractional evolution equation:

$$\begin{cases} {}^{\mathrm{C}}\mathrm{D}_{0+}^{\alpha}[X](t) + AX(t) = f(t, X(t), u(t)), \ t \in (0, T], \\ X(0) = X_0 \in \mathsf{X}, \end{cases} \tag{1.1}$$

where ${}^{\mathrm{C}}\mathrm{D}_{0+}^{\alpha}$ with $\alpha \in (0, 1)$ denotes the left Caputo derivative, $X : [0, T] \to \mathsf{X}$ is the state with $-A$ being the generator of the analytic semigroup, and $u : [0, T] \to \mathsf{U}$ is the control. The problem of this paper is to minimize the following Bolza-type Riemann-Liouville (RL) fractional integral objective functional over $u(\cdot) \in \mathcal{U}$ (note that $\mathcal{U}$ denotes the set of admissible controls)

$$\textbf{(P)} \ \ J(X_0; u(\cdot)) = \int_0^T \frac{(T - s)^{\beta-1}}{\Gamma(\beta)} l(s, X(s), u(s)) \mathrm{d}s + m(X_0, X(T)), \tag{1.2}$$

subject to (1.1) and *the endpoint state constraint*

$$(X_0, X(T)) \in \mathsf{F}_0 \times \mathsf{F}_T \subset \mathsf{X} \times \mathsf{X} \ \ \text{(endpoint state constraint)}. \tag{1.3}$$

Under (1.1)–(1.3), the precise statement of **(P)** is as follows:

$$\textbf{(P)} \ \ \inf_{u(\cdot) \in \mathcal{U}} J(X_0; u(\cdot)), \ \text{subject to (1.1) and (1.3)}.$$

(P) can be referred to as the infinite-dimensional optimal control problem for fractional evolution equations with the endpoint state constraint. The precise problem statement is given in Section 2. The main goal of this paper is to derive the Pontryagin maximum principle for **(P)**, which constitutes the necessary condition for optimality.

Fractional differential and evolution equations can be viewed as generalizations of classical differential and evolution equations using the fractional calculus framework. They can be applied to study and model various situations in economics, mathematical finance, viscoelasticity, aerodynamics, engineering, and biology, where the fractional phenomena can be observed; see [11, 24, 31, 32, 47] and the references therein. It is worth noting that (1.1) can be viewed as a class of fractional evolution equations, which covers Caputo fractional differential equations when $\mathsf{X} = \mathbb{R}^n$. Indeed, fractional evolution equations can be applied to study fractional partial differential equations (PDEs) and/or delay situations using the generator A and the operator f [8, 31, 47]. Some numerical approaches for solving fractional PDEs can be found in [20, 37] and the references therein.

There are various different formulations and results on optimal control of fractional differential equations, which corresponds to $\mathsf{X} = \mathbb{R}^n$ in **(P)**. In particular, their notable results can be found in [1, 2, 4, 23, 28, 34, 44], where they obtained the Pontryagin maximum principles via variational and duality approaches. On the other hand, we mention that the general infinite-dimensional fractional optimal control problem has not yet been sufficiently studied. Indeed, it is quite surprising that there are no concrete results on the infinite-dimensional maximum principle for optimal control of fractional evolution equations except [12, 33]. Below, we provide a detailed literature review on optimal control for fractional differential and evolution equations.

As mentioned above, to the best of our knowledge, the study on optimal control for fractional evolution equations has not been fully investigated, which we address in this paper through **(P)** (see

Theorem 3.1). There are some results on the controllability and the existence of optimal controls for the fractional control problem similar to **(P)** under additional assumptions [41, 47] (see Remark 3.1). However, the earlier works did not study explicit optimality conditions in terms of the Pontryagin maximum principle. Both [33, Theorem 3] and [12, Theorem 5.3] considered the fractional optimal control problem for fractional evolution equations. However, they did not consider the state constraints in the corresponding optimal control problems and their maximum principles, i.e., $F_0 = F_T = X$ in (1.3) (see Remark 3.1). In addition, although [12] considered the Volterra integral-type singular evolution equation with the analytic semigroup, the overall state equation as well as the adjoint equation in the corresponding maximum principle are $\mathbb{R}$-valued finite dimensional ones, which belong to $L^p([0, T]; \mathbb{R})$ (see [12, Theorems 3.1 and 5.3]). Furthermore, the equivalence between the integral-type singular evolution equation in [12] and the Caputo- and RL-type fractional differential equations holds only in the finite-dimensional case (see [12, Section 4]). The preceding discussion implies that **(P)** is different from the earlier works [12, 33]. Indeed, **(P)** can be viewed as a generalization of [33] to the state-constrained control case.

In the finite-dimensional case (i.e., when $X = \mathbb{R}^n$ in **(P)**), the various versions of maximum principles for fractional optimal control problems were studied in [1,2,4,23,28,34,44]. However, their approaches cannot be applied to solve the infinite-dimensional problem **(P)**, since the infinite-dimensional control problem includes several technical intricacies as discussed in [26, Chapter 4] and [15, Chapter 10]. Moreover, the classical infinite-dimensional maximum principles for nonfractional optimal control problems (e.g., [14, 15, 17, 25, 26, 30, 35]) cannot be used to solve **(P)** directly, since the fractional control problem requires different techniques using fractional calculus and analysis (see Remark 3.1). Therefore, we have to develop new techniques to obtain the maximum principle of **(P)**.

The summary of the preceding discussion is stated as follows. It is not possible to use the earlier works [12, 33], since they considered the infinite-dimensional fractional optimal control problem without the state constraints.* Indeed, they did not need to use the key concepts and techniques established in this paper due to the absence of the state constraints (see the statements (a)–(c) and (i)–(iii) below). Morevoer, we are not able to use the approaches for the earlier works on the finite-dimensional case ($X = \mathbb{R}^n$) with the state constraints (e.g., [1, 2, 4, 23, 28, 34, 44]), as the infinite-dimensional problem requires different technical methods (see [26, Chapter 4] and [15, Chapter 10]). Finally, since the fractional control problem requires fractional calculus and analysis, the approaches for the classical nonfractional infinite-dimensional control problems (e.g., [14, 15, 17, 25, 26, 30, 35]) cannot be used directly to obtain the maximum principle of **(P)**.

In the following, we state the main contributions of this paper:

(a) We prove the maximum principle for **(P)** (see Theorem 3.1 and Section 5). Note that the key concept of proving Theorem 3.1 is to apply the Ekeland variational principle under the family of spike variations, where the intrinsic properties of fractional integral and derivative, analysis of variational equations in Lemmas 5.1 and 5.2, and the representation results of linear fractional evolutions in Lemmas B.3–B.5 are essentially required. Indeed, the proof of Theorem 3.1 is divided into nine steps, which is provided in Section 5.

(b) To prove Theorem 3.1, we obtain the explicit representation formulas of linear left Caputo

---

*As mentioned above, [12] cannot be viewed as the infinite-dimensional problem, since the overall state equation as well as the adjoint equation in the corresponding maximum principle are $\mathbb{R}$-valued finite dimensional ones, which belong to $L^p([0, T]; \mathbb{R})$ (see [12, Theorems 3.1 and 5.3]).

fractional evolution equations with initial conditions and right RL fractional evolution equations with terminal conditions using RL fractional state-transition evolution operators (see Lemmas B.3–B.5 in Appendix 6).

(c) We apply the maximum principle in Theorem 3.1 to the optimal control problem of fractional diffusion PDEs. In fact, the necessary condition for optimality is derived using the maximum principle in Theorem 3.1 (see Section 4).

The detailed statements of the main results of this paper, including the comparisons with the existing literature, are as follows:

(i) As mentioned in (a), we state and prove the maximum principle for **(P)** (see Theorem 3.1). As noted above, our paper extends [33] to the state-constrained case. Indeed, unlike the unconstrained case in [33, Theorem 3] and [12, Theorem 5.3], the maximum principle in Theorem 3.1 requires additional nontriviality and transversality conditions to cope with the state constraint in (1.3). Hence, our proof has to be different from that in [12, 33] (see Remarks 3.1 and 3.2).[†] Also, different from the existing literature for standard nonfractional infinite-dimensional and PDE optimal control problems (e.g., [15, 26, 30, 35, 46]), we do not assume the strict convexity of X^*, the dual space of X. This assumption is particularly important, as it guarantees the differentiability of the distance function of the endpoint state constraint. In the proof (see Section 5), we relax this assumption via a separation argument and constructing a family of spike variations for the Ekeland variational principle, which can be viewed as an extension of [25] to the fractional control problem (see Remarks 3.1 and 3.2). Subsequently, we prove the maximum principle (see Section 5), including nontriviality, adjoint equation, transversality, and Hamiltonian maximization conditions, by establishing variational and duality analysis under the finite codimensionality of initial- and end-point variational sets. Indeed, in our variational analysis, we need to apply fractional calculus to obtain the precise estimates of the variational equations under the family of spike variations (see Lemmas 5.1 and 5.2). Furthermore, in our duality analysis, it is essential to obtain explicit representations of the solutions to linear fractional variational and adjoint equations in terms of the associated RL fractional state-transition evolution operators. These results have not been reported in the existing literature (see also (ii) below and Appendix 6). Hence, the proof in Section 5 has to be different from that for the maximum principles studied in the existing literature.

(ii) As for (b), we obtain the explicit representation results of linear left Caputo fractional evolution equations with initial conditions and right RL fractional evolution equations with terminal conditions (see Lemmas B.3–B.5 in Appendix 6). In both cases, to obtain the desired representation formulas, a detailed analysis of the fundamental solution of the left and right RL fractional state transition operators has to be carried out (see Lemma B.3 and Remark B.1). Also, the careful use of Fubini's formula as well as properties of fractional derivatives and integrals is needed. Notice that the finite-dimensional version of the representation result for forward linear Caputo evolution equations was presented in [19, Theorem 5.1], and Lemma B.4 can be viewed as a generalization [19, Theorem 5.1] to the infinite-dimensional case. On the other hand, the

---

[†]As mentioned in Section 1, although [12] considered the infinite-dimensional Volterra integral-type singular evolution equation, the overall state equation as well as the adjoint equation in the corresponding maximum principle are $\mathbb{R}$-valued finite dimensional ones.

representation result of linear right RL fractional evolution equations with terminal conditions in Lemma B.5 has not been reported in the existing literature. Note that the results in Appendix 6 are independent from the main result of this paper, and we provide their complete proofs, which are omitted in [33].

(iii) Finally, we apply Theorem 3.1 to the optimal control problem of fractional diffusion PDEs, which can be applied to analyze general diffusion problems appeared in engineering and applied sciences (see Section 4). In particular, we first convert the fractional diffusion PDE into the abstract X-valued left Caputo fractional evolution equation in (1.1). Then the fixed end-points constraint optimal control problem is formulated using (1.2) and (1.3). Under this setting, the corresponding optimality condition is derived by the maximum principle of this paper, where the bang-bang type optimal control is obtained via the Hamiltonian maximization condition (see Proposition 4.1). Notice that the result in Section 4 can be viewed as an extension of that for nonfractional versions (e.g., [13, 15, 21, 26, 30]) to the state-constrained fractional PDE control problem.

Based on the above discussion, we mention that the technique for proving the maximum principle (see Theorem 3.1 and Section 5) is significantly different from the unconstrained case (e.g., [12, 33]), the finite-dimensional fractional control problems (e.g., [1, 4, 23]), and the classical nonfractional infinite-dimensional control problems (e.g., [15, 25, 26]). Indeed, due to the inherent complex nature of **(P)**, the maximum principle in this paper is new in the optimal control problem context and its detailed proof must be different from that of the existing literature.

The organization of the paper is as follows. We provide the precise problem statement of **(P)** in Section 2, where the key notation, definitions, and mathematical framework of fractional integrals and derivatives are also given. The statement of the Pontryagin maximum principle is given in Section 3. The application to the fractional diffusion PDE is studied in Section 4. The proof of the maximum principle is given in Section 5. Appendix 6 provides the proof for the well-posedness of (1.1) (see Theorem 2.1). In Appendix 6, we prove the representation results on linear fractional evolution equations. Appendix 6 provides some technical lemmas used in Section 5.

2. Notation and problem statement

In this section, we first state the mathematical framework of fractional integrals and derivatives, including their key definitions and the notations employed in the analysis. Then we provide the problem statement of this paper.

2.1. Notation

Let $\mathbb{R}^n$ be the n-dimensional Euclidean space and $\mathbb{R} := \mathbb{R}^1$, where the norm in $\mathbb{R}^n$ is defined by $|\cdot| := |\cdot|_{\mathbb{R}^n}$. Let Γ be the Gamma function. Let $[0, T]$ with $T < \infty$ be the (finite) time-horizon. In this paper, $M \geq 0$ is a generic constant, and its exact value varies from line to line.

Let $(X, |\cdot|_X)$ be a Banach space, and X^* the dual space of X, i.e., the space of all bounded (or continuous) linear functionals on X. The norm on X^* is defined by $|f|_{X^*} := \sup_{x \in X, |x|_X \leq 1} |f(x)|$ for $f \in X^*$. Let $\langle \cdot, \cdot \rangle_{X^*, X}$ be the usual duality paring between X and X^*. The set of linear bounded operators from X to another Banach space Y is denoted by $\mathcal{L}(X, Y)$. Let $\mathcal{L}(X) := \mathcal{L}(X, X)$. Let $|A|_{\mathcal{L}(X,Y)}$ be the (operator) norm of $A \in \mathcal{L}(X, Y)$. Let $I \in \mathcal{L}(X)$ be an identity operator. Let $A^* \in \mathcal{L}(Y^*, X^*)$ be the adjoint

operator of $A \in \mathcal{L}(\mathsf{X}, \mathsf{Y})$, i.e., $\langle A^* y^*, x \rangle_{\mathsf{X}^*, \mathsf{X}} = \langle y^*, Ax \rangle_{\mathsf{Y}^*, \mathsf{Y}}$ for $x \in \mathsf{X}$ and $y^* \in \mathsf{X}^*$. Clearly, $A^* \in \mathcal{L}(\mathsf{Y}^*, \mathsf{X}^*)$ is also linear and bounded.

We say that f is a Banach space valued function on $[0, T]$ if $f : [0, T] \to \mathsf{X}$. The integration of Banach spaced valued functions is understood in the Bochner sense [26, page 45]. For $1 \le p \le \infty$, let $L^p([0, T]; \mathsf{X})$ be the usual L^p-space of X-valued functions, where L^∞ is defined with *essential supremum*. Let $C([0, T]; \mathsf{X})$ be the space of X-valued continuous functions on $[0, T]$ with norm $\|\cdot\|_\infty$. Let $AC([0, T]; \mathsf{X})$ be the space of absolutely continuous X-valued functions on $[0, T]$.

We state definitions of fractional integrals and derivatives. More detailed results on fractional calculus can be found in [11, 24].

Definition 1. *(i) For $f(\cdot) \in L^1([0, T]; \mathsf{X})$, the left Riemann-Liouville (RL) fractional integral $I_{0+}^\alpha[f]$ of order $\alpha > 0$ is defined by*

$$I_{0+}^\alpha[f](t) := \int_0^t \frac{(t - s)^{\alpha-1}}{\Gamma(\alpha)} f(s) \mathrm{d}s.$$

(ii) For $f(\cdot) \in L^1([0, T]; \mathsf{X})$, the right RL fractional integral $I_{T-}^\alpha[f]$ of order $\alpha > 0$ is defined by

$$I_{T-}^\alpha[f](t) := \int_t^T \frac{(s - t)^{\alpha-1}}{\Gamma(\alpha)} f(s) \mathrm{d}s.$$

Definition 2. *(i) For $f(\cdot) \in L^1([0, T]; \mathsf{X})$, the left RL fractional derivative ${}^{RL}D_{0+}^\alpha[f]$ of order $\alpha \in (0, 1)$ is defined by*

$${}^{RL}D_{0+}^\alpha[f](t) := \frac{\mathrm{d}}{\mathrm{d}t}\Big[I_{0+}^{1-\alpha}[f]\Big](t),$$

where $I_{0+}^{1-\alpha}[f](\cdot) \in AC([0, T]; \mathsf{X})$.

(ii) For $f(\cdot) \in L^1([0, T]; \mathsf{X})$, the right RL fractional derivative ${}^{RL}D_{T-}^\alpha[f]$ of order $\alpha \in (0, 1)$ is defined by

$${}^{RL}D_{T-}^\alpha[f](t) := -\frac{\mathrm{d}}{\mathrm{d}t}\Big[I_{T-}^{1-\alpha}[f]\Big](t),$$

where $I_{T-}^{1-\alpha}[f](\cdot) \in AC([0, T]; \mathsf{X})$.

Let ${}^{RL}D_{0+}^\alpha[f(\cdot)](\cdot) := {}^{RL}D_{0+}^\alpha[f](\cdot)$ and ${}^{RL}D_{T-}^\alpha[f(\cdot)](\cdot) := {}^{RL}D_{T-}^\alpha[f](\cdot)$, where $(\cdot)$ in the square bracket is used to emphasize the integration of f with respect to the other variable.

Definition 3. *(i) For $f(\cdot) \in C([0, T]; \mathsf{X})$, the left Caputo fractional derivative ${}^C D_{0+}^\alpha[f]$ of order $\alpha \in (0, 1)$ is defined by*

$${}^C D_{0+}^\alpha[f](t) := {}^{RL}D_{0+}^\alpha[f(\cdot) - f(0)](t),$$

where $f(\cdot) - f(0)$ is left RL fractional differentiable in the sense of Definition 2.

(ii) For $f(\cdot) \in C([0, T]; \mathsf{X})$, the right Caputo fractional derivative ${}^C D_{T-}^\alpha[f]$ of order $\alpha \in (0, 1)$ is defined by

$${}^C D_{T-}^\alpha[f](t) := {}^{RL}D_{T-}^\alpha[f(\cdot) - f(T)](t),$$

where $f(\cdot) - f(T)$ is right RL fractional differentiable in the sense of Definition 2.

2.2. Problem statement

Let $(\mathsf{X}, |\cdot|_\mathsf{X})$ be a Banach space, representing the state space. We consider the X-valued left Caputo fractional evolution equation on $[0, T]$ with order $\alpha \in (\frac{1}{2}, 1)$:

$$\begin{cases} {}^{\mathrm{C}}\mathrm{D}_{0+}^{\alpha}[X](t) + AX(t) = f(t, X(t), u(t)), \ t \in (0, T], \\ X(0) = X_0 \in \mathsf{X}, \end{cases} \tag{2.1}$$

where $X_0 \in \mathsf{X}$ is the initial condition, $-A$ is the generator of the analytic semigroup, and $f : [0, T] \times \mathsf{X} \times \mathsf{U} \to \mathsf{X}$ is the driver of (2.1) with U being the control space. Let $\mathcal{U} := \mathcal{U}[0, T] = \{u : [0, T] \to \mathsf{U} \mid u(\cdot) \text{ is measurable}\}$ be a set of admissible controls for (2.1). We sometimes write $X(\cdot; X_0, u) := X(\cdot)$ to emphasize that (2.1) is dependent on $(X_0, u(\cdot)) \in \mathsf{X} \times \mathcal{U}$.

Assumption 1. *(i) $A : D(A) \subset \mathsf{X} \to \mathsf{X}$, where $D(A)$ is the domain of A, i.e., the subset of X such that A exists, is a linear operator such that $-A$ is the generator of the compact analytic semigroup $(\mathcal{T}(t))_{t \geq 0}$ of uniformly bounded linear operators with $\mathcal{T} : [0, T] \to \mathcal{L}(\mathsf{X})$ [36].*

(ii) (U, ρ) is a separable metric space.

(iii) $t \mapsto f$ are measurable for any fixed $(X, u) \in \mathsf{X} \times \mathsf{U}$ and $f(\cdot, X, u) \in L^\infty([0, T]; \mathsf{X})$. $(X, u) \mapsto f$ is continuous in both variables. There is a constant $M \geq 0$ such that $|f(t, 0, u)|_\mathsf{X} \leq M$ and $|f(t, X, u) - f(t, X', u')|_\mathsf{X} \leq M(|X - X'|_\mathsf{X} + \rho(u, u'))$ for any $X, X' \in \mathsf{X}$ and $u, u' \in \mathsf{U}$.

(iv) $X \mapsto f$ is continuously Fréchet differentiable, denoted by $\partial_X f(t, X, u)$, for any fixed $(t, u) \in [0, T] \times \mathsf{U}$, where $t \mapsto \partial_X f$ is continuous. Note that $\partial_X f(t, \cdot, u) : [0, T] \times \mathsf{U} \to \mathcal{L}(\mathsf{X})$. It also holds that there is a constant $M \geq 0$ such that $|\partial_X f(t, X, u)|_\mathsf{X} \leq M$.

(v) $\frac{MT^\alpha \sup_{t \in [0,T]} |\mathcal{T}(t)|_\mathsf{X}}{\Gamma(1+\alpha)} < 1$.

We state the well-posedness result (equivalently, the existence and uniqueness of the solution) of (2.1). The proof is given in Appendix 6.

Theorem 2.1. *Let Assumption 1 hold. Then (2.1) admits a unique (mild) solution. Moreover, the solution of (2.1) can be expressed by the left RL fractional integral form:*

$$X(t) = X_0 - I_{0+}^\alpha[AX(\cdot)](t) + I_{0+}^\alpha[f(\cdot, X(\cdot), u(\cdot))](t), \ t \in [0, T]. \tag{2.2}$$

Finally, there is a constant $M \geq 0$ such that for any $X_0, X_0' \in \mathsf{X}$ (with $X(t) := X(t; X_0, u)$ and $X'(t) := X(t; X_0', u)$),

$$\sup_{t \in [0,T]} |X(t) - X'(t)|_\mathsf{X} \leq M|X_0 - X_0'|_\mathsf{X} + \int_0^T \sum_{k=1}^\infty \frac{(M\Gamma(\alpha))^k}{\Gamma(k\alpha)}(T - s)^{k\alpha - 1}|X_0 - X_0'|_\mathsf{X} \mathrm{d}s,$$

$$\sup_{t \in [0,T]} |X(t)|_\mathsf{X} \leq (|X_0|_\mathsf{X} + M) + \int_0^T \sum_{k=1}^\infty \frac{(M\Gamma(\alpha))^k}{\Gamma(k\alpha)}(T - s)^{k\alpha - 1}(|X_0|_\mathsf{X} + M)\mathrm{d}s.$$

The objective functional to be minimized over $u(\cdot) \in \mathcal{U}$ is given by the following left RL fractional integral with order $\beta \geq 1$:

$$J(X_0; u(\cdot)) = I_{0+}^\beta[l(\cdot, X(\cdot), u(\cdot))](T) + m(X_0, X(T)), \tag{2.3}$$

where by Definition 1,

$$I_{0+}^{\beta}[l(\cdot, X(\cdot), u(\cdot)](T) = \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} l(s, X(s), u(s)) \mathrm{d}s.$$

In (2.3), $l : [0, T] \times \mathsf{X} \times \mathsf{U} \to \mathbb{R}$ is the running cost, while $m : \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ is the terminal cost. We introduce the associated *endpoint state constraint*:

$$(X_0, X(T)) \in \mathsf{F}_0 \times \mathsf{F}_T \subset \mathsf{X} \times \mathsf{X} \quad \text{(endpoint state constraint)}. \tag{2.4}$$

Based on the preceding setting, the problem of this paper is stated as follows:

$$\textbf{(P)} \quad \inf_{u(\cdot) \in \mathcal{U}} J(X_0; u(\cdot)), \text{ subject to (2.1) and (2.4).}$$

Indeed, **(P)** can be regarded as *the optimal control problem for fractional evolution equations with endpoint state constraint.* The main goal is to derive the Pontryagin maximum principle for **(P)**.

Assumption 2. *(i) $t \mapsto l$ is measurable for any fixed $(X, u) \in \mathsf{X} \times \mathsf{U}$ and $l(\cdot, X, u) \in L^\infty([0, T]; \mathbb{R})$. $(X, u) \mapsto l$ and $(X, X') \mapsto m$ are continuous in both variables. There is a constant $M \geq 0$ such that $|l(t, 0, u, w)| + |m(0, 0)| \leq M$.*

(ii) $X \mapsto l$ is continuously Fréchet differentiable, denoted by $\partial_X l(t, X, u)$, for any fixed $(t, u) \in [0, T] \times \mathsf{U}$, where $t \mapsto \partial_X l$ is continuous. Note that $\partial_X l(t, \cdot, u) : [0, T] \times \mathsf{U} \to \mathcal{L}(\mathsf{X}, \mathbb{R}) = \mathsf{X}^*$. $(X, X') \mapsto m$ is continuously Fréchet differentiable, denoted by $\partial_X m(X, X')$ and $\partial_{X'} m(X, X')$, respectively, where $\partial_X m(\cdot, X'), \partial_{X'} m(X, \cdot) \in \mathcal{L}(\mathsf{X}, \mathbb{R}) = \mathsf{X}^*$. There is a constant $M \geq 0$ such that $|\partial_X l(t, X, u)| + |\partial_X m(X, X')| + |\partial_{X'} m(X, X')| \leq M$.

(iii) F_0 and F_T are closed convex subsets of X.

Remark 2.1. *(i) Assumptions 1 and 2 are standard for fractional evolution equations in Theorem 3.1 and their optimal control problems. Indeed, these assumptions have been used in various types of (finite- and infinite-dimensional) optimal control problems and their maximum principles; see [1, 4, 12, 15, 23, 25, 26, 28, 34, 41] and the references therein.*

*(ii) As (2.1) is an infinite-dimensional evolution equation driven by a (possibly) unbounded linear operator $-A$ (the generator for the analytic semigroup), **(P)** is different from the finite-dimensional fractional optimal control problems in [1, 4, 23] (when $\mathsf{X} = \mathbb{R}^n$).*

*(iii) We note that the analyticity of $-A$ in Assumption 1 is used for the well-posedness of (2.1) in Theorem 2.1 (see [48, Theorem 3.3] and [41, Theorem 3.1]). In fact, (v) of Assumption 1 is due to [48, Theorem 3.3]. Then Theorem 2.1 is used to prove the well-posedness of linear Caputo fractional evolution equations in Lemma B.4. Note that we can replace the analyticity of $-A$ with $-A$ being a generator of an (equicontinuous) C_0-semigroup. In this case, we need additional assumptions on f for the well-posedness of (2.1); [9, Section 3] and [38, Sections 3 and 4]. See also [47, Chapter 2], [18, Chapters 2 and 3] and [31] for the well-posedness of (2.1) under different assumptions. In this paper, since the main goal is to derive the maximum principle of **(P)**, we do not focus on the generality of the assumptions.*

3. Statement of main result: Pontryagin maximum principle of (P)

Assume that the pair $(\overline{X}(\cdot), \overline{u}(\cdot)) \in C([0, T]; \mathsf{X}) \times \mathcal{U}$ is the optimal solution of **(P)**. We observe that $\overline{X}(\cdot) \in C([0, T]; \mathsf{X})$ is the corresponding optimal state trajectory generated by the optimal control $\overline{u}(\cdot) \in \mathcal{U}$ satisfying $(\overline{X}_0, \overline{X}(T)) = (\overline{X}_0, \overline{X}(T; \overline{X}_0, \overline{u})) \in \mathsf{F}_0 \times \mathsf{F}_T$. We introduce the notation:

$$\overline{f}(\cdot) := f(\cdot, \overline{X}(\cdot), \overline{u}(\cdot)), \ \partial_X \overline{f}(\cdot) := \partial_X f(\cdot, \overline{X}(\cdot), \overline{u}(\cdot)), \ \overline{l}(\cdot) := l(\cdot, \overline{X}(\cdot), \overline{u}(\cdot)), \ \partial_X \overline{l}(\cdot) := \partial_X l(\cdot, \overline{X}(\cdot), \overline{u}(\cdot)),$$

$$\overline{m} := m(\overline{X}_0, \overline{X}(T)), \ \partial_{X_0} \overline{m} := \partial_{X_0} m(\overline{X}_0, \overline{X}(T)), \ \partial_X \overline{m} := \partial_X m(\overline{X}_0, \overline{X}(T)).$$

Let $N_{\mathsf{F}_0}(X_0)$ be the normal cone to F_0 at $X_0 \in \mathsf{F}_0$ defined by

$$N_{\mathsf{F}_0}(X_0) := \{X' \in \mathsf{X}^* \mid \langle X', y - X_0 \rangle_{\mathsf{X} \times \mathsf{X}^*} \leq 0, \ \forall y \in \mathsf{F}_0\}.$$

The normal cone $N_{\mathsf{F}_T}(X)$ to F_T at $X \in \mathsf{F}_T$ can be defined analogously. Let $\Omega := \cup_{k=1}^{\infty} \Omega_k$, where for a positive integer k,

$$\Omega_k := \left\{ \left(\{u_j(\cdot)\}_{j=1}^k, \{\theta_j\}_{j=1}^k \right) \mid u_j(\cdot) \in \mathcal{U}, \ \theta_j \geq 0, \ \forall j = 1, \ldots, k, \ \sum_{j=1}^k \theta_j = 1 \right\}.$$

For $(\{u_j(\cdot)\}_{j=1}^k, \{\theta_j\}_{j=1}^k) \in \Omega$, we define

$$\mathcal{R} := \left\{ \xi(T) \in \mathsf{X} \mid \xi(t) = \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \left(-A + \partial_X \overline{f}(s) \right) \xi(s) \mathrm{d}s + \sum_{j=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \right. \tag{3.1}$$

$$\left. \times \left(f(t, \overline{X}(s), u_j(s)) - f(t, \overline{X}(s), \overline{u}(s)) \right) \mathrm{d}s, \ t \in [0, T], \ \left(\{u_j(\cdot)\}_{j=1}^k, \{\theta_j\}_{j=1}^k \right) \in \Omega \right\},$$

$$Q := \left\{ \hat{X} - \zeta(T) \in \mathsf{X} \mid \zeta(t) = \hat{X}_0 + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \left(-A + \partial_X \overline{f}(s) \right) \zeta(s) \mathrm{d}s, \right. \tag{3.2}$$

$$\left. t \in [0, T], \ (\hat{X}_0, \hat{X}) \in \mathsf{F}_0 \times \mathsf{F}_T \right\}.$$

Assumption 3. $\mathcal{R} - Q$ *is finite codimensional in* X.

We state the main result of this paper, i.e., the Pontryagin maximum principle of **(P)**, in Theorem 3.1 below. Indeed, Theorem 3.1 constitutes the necessary condition for optimality of **(P)**. The proof of Theorem 3.1 is relegated to Section 5.

Theorem 3.1. *Let Assumptions 1–3 hold. Assume that* $(\overline{X}(\cdot), \overline{u}(\cdot)) \in C([0, T]; \mathsf{X}) \times \mathcal{U}$ *is the optimal solution of* **(P)**. *Then there exist* $(\mathsf{p}^0, \varphi_1, \varphi_2) \in \mathbb{R} \times \mathsf{X}^* \times \mathsf{X}^*$ *such that the following conditions hold:*

(i) **Nontriviality condition:** $(\mathsf{p}^0, \varphi_1, \varphi_2) \neq 0$, *where* $\mathsf{p}^0 \leq 0$ *and* $(\varphi_1, \varphi_2) \in N_{\mathsf{F}_0}(\overline{X}_0) \times N_{\mathsf{F}_T}(\overline{X}(T))$.

(ii) **Adjoint equation:** $P(\cdot) \in L^1([0, T]; \mathsf{X}^*)$ *is a unique solution to the following right RL fractional evolution equation:*

$$^{RL}D_{T-}^{\alpha}[P](t) = -A^* P(t) + \partial_X \overline{f}(t)^* P(t) - \mathsf{p}^0 \frac{(T-t)^{\beta-1}}{\Gamma(\beta)} \partial_X \overline{l}(t), \ t \in [0, T). \tag{3.3}$$

(iii) ***Transversality condition:*** *For any $(X'', X') \in \mathsf{F}_0 \times \mathsf{F}_T$,*

$$\langle I_{T_-}^{1-\alpha}[P](0) - \mathsf{p}^0 \partial_{X_0} \overline{m}, (X'' - \overline{X}_0) \rangle_{\mathsf{X} \times \mathsf{X}^*} - \langle I_{T_-}^{1-\alpha}[P](T) - \mathsf{p}^0 \partial_X \overline{m}, (X' - \overline{X}(T)) \rangle_{\mathsf{X} \times \mathsf{X}^*} \leq 0,$$

and the adjoint equation holds the following boundary conditions:

$$\begin{cases} I_{T_-}^{1-\alpha}[P](T) = -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, \\ I_{T_-}^{1-\alpha}[P](0) = -(-\varphi_1 + \mathsf{p}^0 \partial_{X_0} \overline{m}). \end{cases}$$

(iv) ***Hamiltonian maximization condition:*** *$\overline{u}$ (pointwisely) maximizes the Hamiltonian, i.e.,*

$$\max_{u \in \mathsf{U}} H(t, \overline{X}(t), P(t), \mathsf{p}^0, u) = H(t, \overline{X}(t), P(t), \mathsf{p}^0, \overline{u}(t)), \ \ a.e. \ t \in [0, T],$$

where $H : [0, T] \times \mathsf{X} \times \mathsf{X}^* \times \mathbb{R} \times \mathsf{U} \to \mathbb{R}$ is the Hamiltonian defined by

$$H(t, X, P, \mathsf{p}^0, u) := \langle P, f(t, X, u) \rangle_{\mathsf{X} \times \mathsf{X}^*} + \mathsf{p}^0 \frac{(T - t)^{\beta - 1}}{\Gamma(\beta)} l(t, X, u).$$

Some important remarks of Theorem 3.1 are given below.

Remark 3.1. *(i) We have assumed the existence of the optimal solution in Theorem 3.1. Indeed, the problem for existence of optimal controls requires different technical analysis under additional restrictive conditions (e.g., convexity and/or linearity) of (2.1) and (2.3). Some results on these additional conditions as well as the techniques for proving the existence (including the finite-dimensional case) can be found in [4, 22, 41, 42, 47] and the references therein.*

*(ii) When there is no endpoint state constraint, i.e., $\mathsf{F}_0 = \mathsf{F}_T = \mathsf{X}$, Theorem 3.1 is reduced to the unconstrained case studied in [33, Theorem 3]. Moreover, Theorem 3.1 is different from the classical infinite-dimensional nonfractional maximum principles in [15, 25, 26], where Theorem 3.1 involves the RL fractional adjoint equation and the fractional Hamiltonian. If $\mathsf{X} = \mathbb{R}^n$ and $\mathsf{U} \subset \mathbb{R}^p$ in (**P**), then Theorem 3.1 is reduced to the maximum principle for the finite-dimensional fractional control problem with the state constraint studied in [1, 4, 23].*

(iii) As mentioned in Section 1, the proof of Theorem 3.1 is different from that for the unconstrained case in [33], the finite-dimensional fractional control problem in [1, 4, 23], and the infinite-dimensional nonfractional control problem in [15, 25, 26]. Specifically, in this paper, we need to establish new variational and duality analysis via several technical lemmas for fractional evolution equations, which are new in the context of infinite-dimensional optimal control problems.

Remark 3.2. *(i) Assumption 3 states that the algebraic difference between initial- and end-point variational sets $\mathcal{R}$ and $\mathcal{Q}$ is finite codimensional in X (see [26, Definition 1.5, Chapter 4]). Some conditions and approaches to check the finite codimensionality can be found in [26, Chapter 1.6 and 4], [30, Theorems 3.1 and 3.2], and [14, 29, 39]. Note that [26, Proposition 3.2 and Corollary 3.3, Chapter 4] provide the general conditions for the finite codimensionality, which can be applied to the case of fractional evolution equations. Assumption 3 is very common in various infinite-dimensional maximum principles and optimal control for partial differential*

equations [5, 14, 15, 21, 26, 27, 29, 30, 35, 39, 40, 45, 46]. In fact, the finite codimensionality guarantees that the variables $(\mathsf{p}^0, \varphi_1, \varphi_2) \in \mathbb{R} \times \mathsf{X}^* \times \mathsf{X}^*$ in Theorem 3.1 are nontrivial, i.e., $(\mathsf{p}^0, \varphi_1, \varphi_2) \neq 0$. On the other hand, in the finite-dimensional case (when $\mathsf{X} = \mathbb{R}^n$), the associated nontriviality can be obtained without Assumption 3.

(ii) Unlike [26], we do not need the strict convexity of X^*, the dual space of X. This assumption is satisfied when X is separable or reflexive. The strict convexity of X^* has been used in various earlier infinite-dimensional maximum principles, which is crucial to obtain the explicit derivative of the distance function in variational analysis [5, 15, 21, 26, 27, 35, 40, 46]. We relax this assumption in Theorem 3.1 by establishing the separation argument and constructing a family of spike variations using Ω, which can be viewed as an extension of [25] to the fractional control problem. Specifically, in Section 5.4, our proof shows that the directional derivative of the distance function is expressed by the duality paring between X and X^*, where the element of X^* is regarded as the generalized gradient of the distance function. This alternative duality paring form is used in the rest of the proof to derive the optimality conditions in Theorem 3.1.

4. An example: fractional diffusion PDE

In this section, we apply **(P)** for optimal control of fractional diffusion PDEs, which can be applied to analyze general diffusion aspects appeared in engineering and applied sciences [30, 41, 48]. The purpose of this section is to show the conversion of the fractional PDE control problem into the abstract infinite-dimensional problem **(P)** and to demonstrate the application of Theorem 3.1.

Let $O \subset \mathbb{R}^3$ be open with smooth boundary $\partial O \subset \mathbb{R}^3$. Let $L^p(O)$ be the set of real-valued pth-integrable functions on O (when $p = \infty$, it is the space of real-valued essential bounded functions). We denote $H^k(O)$ by the usual Sobolev space of real-valued functions on O, whose distributional derivatives, up to the order k, are square-integrable, and $H_0^k(O)$ by the closure of $C_c^\infty(O)$ in $H^k(O)$, where $C_c^\infty(O)$ is the set of continuously differentiable real-valued functions on O having compact support.

Consider the following *fractional diffusion PDE*:

$$\begin{cases} {}^{\mathrm{C}}\mathrm{D}_{0+}^\alpha[x(\cdot, w)](t) - \Delta x(t, w) = f_1(t, x(t, w)) + f_2(t)v(t, w), \ (t, w) \in (0, T) \times O, \\ x(t, w) = 0, \ (t, w) \in [0, T] \times \partial O, \end{cases} \tag{4.1}$$

where $x : [0, T] \times O \to \mathbb{R}$ is the state trajectory describing the evolution of the diffusion process with initial and boundary conditions, Δ is the Laplacian operator, and $v : [0, T] \times O \to [-1, 1] \subset \mathbb{R}$ is the control input (the external force acting at every point of O). The objective functional is given by

$$J(x_0; v) = \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \int_O \Big(l_1(s, x(s, w)) + l_2(s)v(s, w)\Big) \mathrm{d}w \mathrm{d}s. \tag{4.2}$$

The minimization problem of (4.2) subject to (4.1) can be regarded as optimal control of fractional diffusion equations, where its nonfractional standard version was considered in various settings in the existing literature; see [13, 26, 30] and the references therein. Notice that (4.2) is the objective functional with RL fractional integral, which can be regarded as the minimization of the general time fractional and/or nonlocal diffusion situations of fractional diffusion PDEs.

Let $\mathsf{X} = L^2(O)$ and $X(t) := x(t, \cdot) \in \mathsf{X}$. Let $Ax := -\Delta x$, where $x \in D(A) = H^2(O) \cap H_0^1(O)$. By [36, Theorem 2.7, Chapter 7] and the embedding theorem of Sobolev spaces, $-A$ is the generator

of the compact analytic semigroup of uniformly bounded linear operators, and $-A$ is self adjoint. Let $u(t) := v(t, \cdot)$ and $\mathcal{U} = L^\infty([0, T]; \mathsf{U})$, where $\mathsf{U} = L^2(O; [-1, 1])$, the space of square-integrable functions on O taking values in $[-1, 1]$. It is easy to see that $\mathcal{U}$ is a nonempty closed and bounded convex subset of $L^\infty([0, T]; \mathsf{X})$, which has at least one interior point.

Under the above setup, (4.1) can be written as the following X-valued Caputo fractional evolution equation (with the Nemytskii operator $f_1(t, X(t))(w) := f_1(t, x(t, w))$):

$$^\mathrm{C}\mathrm{D}_{0+}^\alpha[X](t) + AX(t) = f_1(t, X(t)) + f_2(t)u(t), \ t \in (0, T]. \tag{4.3}$$

Similar to [45, Section 2.4, page 1324], we consider the following endpoint state constraint for (4.1) and (4.2) (note that $X(0) = x(0, \cdot)$ and $X(T) = x(T, \cdot)$):

$$(X(0), X(T)) \in \mathsf{F}_0 \times \mathsf{F}_T, \tag{4.4}$$

where for $a_i, b_i \in \mathbb{R}$ and $h_i(\cdot), g_i(\cdot) \in \mathsf{X}, \ i = 1, \dots, m,$

$$\mathsf{F}_0 = \Big\{ y''(\cdot) \in \mathsf{X} \mid \int_O y''(w)h_i(w)\mathrm{d}w = a_i, \ i = 1, \dots, m \Big\},$$

$$\mathsf{F}_T = \Big\{ y'(\cdot) \in \mathsf{X} \mid \int_O y'(w)g_i(w)\mathrm{d}w = b_i, \ i = 1, \dots, m \Big\}.$$

We assume that F_0 and F_T are nonempty. It is easy to see that F_0 and F_T are closed convex subsets of X. Let $L_1(s, X(s)) := \int_O l_1(s, x(s, w))\mathrm{d}w$ and $L_2(s, u(s)) := \int_O l_2(s)v(s, w)\mathrm{d}w$. Then (4.2) becomes

$$J(X_0; u(\cdot)) = \int_0^T \frac{(T - s)^{\beta-1}}{\Gamma(\beta)}\Big(L_1(s, X(s)) + L_2(s, u(s))\Big)\mathrm{d}s.$$

Let us assume that $f_1(\cdot, x), f_2(\cdot), l_1(\cdot, x), l_2(\cdot) \in C([0, T]; \mathbb{R})$ hold Assumptions 1 and 2.

Based on the above discussion, **(P)** can be applied to formulate the optimal control problem of (4.3) with the end-points state constraint in (4.4):

$$\inf_{u(\cdot)\in\mathcal{U}} J(X_0; u(\cdot)), \ \text{subject to (4.3) and (4.4)}. \tag{4.5}$$

Suppose that $(\overline{X}(\cdot), \overline{u}(\cdot)) = (\overline{x}, \overline{v}) \in C([0, T]; \mathsf{X}) \times \mathcal{U}$ is the optimal solution of (4.5). In view of (5.28) and (5.29) with Π from (B.1), $\mathcal{R}$ and Q can be written as follows (see also [26, page 160, Chapter 4]):

$$\mathcal{R} = \Big\{ \xi(T) \in \mathsf{X} \mid \xi(t) = \sum_{j=1}^k \theta_j \int_0^t \Pi(t, s)f_2(s)(u_j(s) - \overline{u}(s))\mathrm{d}s, \ t \in [0, T], \ \big(\{u_j(\cdot)\}_{j=1}^k, \{\theta_j\}_{j=1}^k\big) \in \Omega \Big\},$$

$$Q = \mathsf{F}_T - \Big(\mathrm{I} + \int_0^T \Pi(T, s)\big(-A + \partial_X f_1(t, \overline{X}(s)\big)\mathrm{d}s\Big)\mathsf{F}_0.$$

For Assumption 3, the finite codimensionality of F_T in X is crucial. Indeed, by [26, Proposition 3.4 and page 160, Chapter 4]‡, the finite codimensionality of F_T in X implies the finite codimensionality of Q in X, which (by [26, Proposition 3.4, Chapter 4] again) further implies the finite codimensionality

---

‡Note that [26, Proposition 3.4, Chapter 4] does not need the fractional calculus in the proof.

of $\mathcal{R} - Q$ in X in Assumption 3. Then it follows from [45, Section 2.4, page 1324]§ that F_T is finite codimensional in X. Hence, Assumption 3 holds.

Since $X^* = X$, we have (note that $\overline{X}(0) = \overline{x}(0, \cdot)$ and $\overline{X}(T) = \overline{x}(T, \cdot)$)

$$N_{\mathsf{F}_0}(\overline{x}(0, \cdot)) = \left\{ z_1(\cdot) \in X \mid \int_O z_1(w)(x''(w) - \overline{x}(0, w))dw \le 0, \ \forall x''(\cdot) \in \mathsf{F}_0 \right\}, \tag{4.6}$$

$$N_{\mathsf{F}_T}(\overline{x}(T, \cdot)) = \left\{ z_2(\cdot) \in X \mid \int_O z_2(w)(x'(w) - \overline{x}(T, w))dw \le 0, \ \forall x'(\cdot) \in \mathsf{F}_T \right\}. \tag{4.7}$$

Based on the preceding analysis and Theorem 3.1, we state the following result:

Proposition 4.1. *Assume that* $(\overline{X}(\cdot), \overline{u}(\cdot)) = (\overline{x}, \overline{v}) \in C([0, T]; X) \times \mathcal{U}$ *is the optimal solution of (4.5). Then there exist* $(\mathsf{p}^0, \varphi_1, \varphi_2) \in \mathbb{R} \times X \times X$ *such that the following conditions hold:*

(i) **Nontriviality condition:** $(\mathsf{p}^0, \varphi_1, \varphi_2) \ne 0$, *where* $\mathsf{p}^0 \le 0$ *and* $(\varphi_1, \varphi_2) \in N_{\mathsf{F}_0}(\overline{x}(0, \cdot)) \times N_{\mathsf{F}_T}(\overline{x}(T, \cdot))$.

(ii) **Adjoint equation and boundary condition:** $p \in L^1([0, T]; X)$ *is the unique solution to the following RL fractional evolution equation with the boundary condition:*

$$\begin{cases} {}^{RL}D_{T-}^\alpha[p(\cdot, w)](t) - \Delta p(t, w) \\ \qquad = \partial_x f_1(t, \overline{x}(t, w))p(t, w) - \mathsf{p}^0 \dfrac{(T-t)^{\beta-1}}{\Gamma(\beta)} \partial_x l_1(t, \overline{x}(t, w)), \ (t, w) \in (0, T) \times O, \\ p(t, w) = 0, \ (t, w) \in [0, T] \times \partial O, \\ I_{T-}^{1-\alpha}[p(\cdot, w)](T) = -\varphi_2(w), \ w \in O, \\ I_{T-}^{1-\alpha}[p(\cdot, w)](0) = \varphi_1(w), \ w \in O. \end{cases}$$

(iii) **Maximum condition:** *The optimal control* $\overline{v}$ *satisfies the following maximum condition:*

$$\int_O \left(p(t, w)f_2(t) + \mathsf{p}^0 \frac{(T-t)^{\beta-1}}{\Gamma(\beta)} l_2(t) \right) \overline{v}(t, w)dw$$

$$\ge \int_O \left(p(t, w)f_2(t) + \mathsf{p}^0 \frac{(T-t)^{\beta-1}}{\Gamma(\beta)} l_2(t) \right) v(t, w)dw, \ \forall v(t, \cdot) \in \mathsf{U}, \ a.e. \ t \in [0, T],$$

which is equivalent to the bang-bang control:

$$\overline{v}(t, w) = \begin{cases} 1, & \text{if } p(t, w)f_2(t) + \mathsf{p}^0 \dfrac{(T-t)^{\beta-1}}{\Gamma(\beta)} l_2(t) > 0, \\ -1, & \text{if } p(t, w)f_2(t) + \mathsf{p}^0 \dfrac{(T-t)^{\beta-1}}{\Gamma(\beta)} l_2(t) < 0. \end{cases}$$

Proof. The entire proof is an application of Theorem 3.1. We observe that (i) follows from the nontriviality condition of Theorem 3.1. Then (ii) can be deduced from the adjoint equation and its boundary conditions in Theorem 3.1, where $P(\cdot) \in L^1([0, T]; X)$ with $P(t) := p(t, \cdot)$. Note that the

---

§Note that [45, Section 2.4, page 1324] does not need the fractional calculus in the corresponding result.

inequality in the transversality condition of Theorem 3.1 holds immediately, since by (4.6) and (4.7), for any $(x''(\cdot), x'(\cdot)) \in \mathsf{F}_0 \times \mathsf{F}_T$, it follows that

$$
\int_O \varphi_1(w)(x''(w) - \overline{x}(0, w)) \mathrm{d}w + \int_O \varphi_2(w)(x'(w) - \overline{x}(T, w)) \mathrm{d}w \leq 0
$$
$$
\Leftrightarrow \langle \mathrm{I}_{T^-}^{1-\alpha}[P](0), (X'' - \overline{X}_0) \rangle_{\mathsf{X} \times \mathsf{X}^*} - \langle \mathrm{I}_{T^-}^{1-\alpha}[P](T), (X' - \overline{X}(T)) \rangle_{\mathsf{X} \times \mathsf{X}^*} \leq 0.
$$

The maximum condition and its equivalent bang-bang control in (iii) are obtained from the Hamiltonian maximization condition in Theorem 3.1. We complete the proof. $\qquad\square$

5. Proof of Theorem 3.1

This section is devoted to proving Theorem 3.1. We use $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_{\mathsf{X} \times \mathsf{X}^*}$ when there is no confusion.

The key concept of proving the maximum principle in Theorem 3.1 is to apply the Ekeland variational principle under the family of spike variations, where the key techniques are applications of intrinsic properties of fractional integral and derivative, analysis of variational equations in Lemmas 5.1 and 5.2, and the representation results of linear fractional evolutions in Lemmas B.3–B.5. Specifically, the proof is divided into nine steps, where the outline of each step is given below:

Step 1. Ekeland variational principle: We formulate the unconstrained fractional optimal control problem of (**P**) with the penalized objective functional. Then the Ekeland variational principle is applied to obtain the ϵ-optimal solution to the corresponding unconstrained control problem as well as the associated inequalities (see (5.2)).

Step 2. Family of spike variations: By Lemma C.6 and (5.3), the family of spike variations is defined on $\{E_j^\delta \subset [0, T], \ j = 1 \ldots, k\}$, where $\cup_{j=1}^k \mathrm{meas}|E_j^\delta| = \delta T$ (see (5.4)). Then the estimate of l and f under the spike variations is obtained with respect to δ (see (5.6)).

Step 3. Variational analysis of δ: By Step 2, variational analysis of the state trajectories and the objective functionals controlled by the family of spike variations and the ϵ-optimal solution is obtained with respect to δ (see Lemma 5.1). Based on Lemma 5.1, by taking $\delta \downarrow 0$ of the inequalities in Step 1, we obtain the variational inequality as well as the multiplier condition (see (5.12) and (5.13)).

Step 4. Separation: By the separation argument and Lemma C.7, we obtain the alternative forms of (5.12) and (5.13) as well as the transversality-like conditions for the endpoint state constraints (see (5.19)–(5.22)).

Step 5. Variational analysis of ϵ: Similar to Step 3, variational analysis of the state trajectories and the objective functionals controlled by the ϵ-optimal solution and the optimal solution is obtained with respect to ϵ (see Lemma 5.2). Based on Lemma 5.2 and the multiplier condition (5.20) in Step 4, by taking $\epsilon \downarrow 0$ of the inequality in Step 4, we obtain the variational inequality (see (5.26)).

Step 6. Proof of (i), Nontriviality condition: By the finite codimensionality in Assumption 3 and [26, Lemma 3.6, Chapter 4], together with the variational inequality in Step 5, the multiplier condition (5.20) in Step 4, and Lemma B.4 (see Appendix 6), we prove $(\mathsf{p}^0, \varphi_1, \varphi_2) \neq 0$ with

$\mathsf{p}^0 \leq 0.$[¶] Then we show $(\varphi_1, \varphi_2) \in N_{\mathsf{F}_0}(\overline{X}_0) \times N_{\mathsf{F}_T}(\overline{X}(T))$ by the definition of subdifferential of convex functions. From the nontriviality condition, we obtain the limiting form of the variational inequality and multiplier condition in Step 5 with respect to $\epsilon \downarrow 0$ (see (5.30)).

Step 7. Proof of (ii), Duality and adjoint equation: By Lemma B.5, we prove (ii) of Theorem 3.1. Then using Lemmas B.3–B.5, we obtain the duality form of the variational inequality in Step 6 (see (5.33)).

Step 8. Proof of (iii), Transversality condition: By the definitions of (RL and Caputo) fractional integrals, we prove (iii) of Theorem 3.1 when $u = \overline{u}$ of the (duality form) variational inequality in Step 7.

Step 9. Proof of (iv), Hamiltonian maximization condition: Using the (duality form) variational inequality in Step 7 and [6, Theorem 5.6.2], we prove (iv) of Theorem 3.1.

5.1. Ekeland variational principle

Recall that $(\overline{X}(\cdot), \overline{u}(\cdot)) \in C([0, T]; \mathsf{X}) \times \mathcal{U}$ is the optimal solution of **(P)**. Let $d_{\mathsf{F}_0} : \mathsf{X} \to \mathbb{R}_+$ be the distance function of F_0 defined by $d_{\mathsf{F}_0}(X) := \inf_{y \in \mathsf{F}_0} |y - X|$ for $X \in \mathsf{X}$. Let $d_{\mathsf{F}_T} : \mathsf{X} \to \mathbb{R}_+$ be the distance function of F_T defined similarly. Then the endpoint state constraint $(X_0, X(T)) \in \mathsf{F}_0 \times \mathsf{F}_T$ is equivalent to $d_{\mathsf{F}_0}(X_0) = 0$ and $d_{\mathsf{F}_T}(X(T)) = 0$.

We introduce the penalized objective functional

$$J_\epsilon(X_0; u(\cdot)) = \left(([J(X_0; u(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+)^2 + d_{\mathsf{F}_0}(X_0)^2 + d_{\mathsf{F}_T}(X(T))^2 \right)^{\frac{1}{2}}, \qquad (5.1)$$

where $[\cdot]^+ := \max\{\cdot, 0\}$. Define the Ekeland metric by

$$\tilde{d}(u(\cdot), \widetilde{u}(\cdot)) := \mathrm{meas}|\{t \in [0, T] \mid u(t) \neq \widetilde{u}(t)\}|,$$

where $\mathrm{meas}|\cdot|$ is understood as the Lebesgue measure of the corresponding set. Let us define the metric

$$\widehat{d}((X_0, X_0'), (u(\cdot), \widetilde{u}(\cdot))) := |X_0 - X_0'|_{\mathsf{X}} + \tilde{d}(u(\cdot), \widetilde{u}(\cdot)).$$

We observe that $(\mathsf{X} \times \mathcal{U}, \widehat{d})$ is a complete metric space [26, Proposition 3.10, Chapter 4]. Furthermore, by our assumptions and Theorem 2.1, J_ϵ is a continuous functional on $(\mathsf{X} \times \mathcal{U}, \widehat{d})$.

We can observe that

$$J_\epsilon(X_0; u(\cdot)) > 0, \ \forall (X_0, u(\cdot)) \in \mathsf{X} \times \mathcal{U},$$

$$J_\epsilon(\overline{X}_0; \overline{u}(\cdot)) = \epsilon \leq \inf_{(X_0, u(\cdot)) \in \mathsf{X} \times \mathcal{U}} J_\epsilon(X_0; u(\cdot)) + \epsilon.$$

By the Ekeland variational principle [26, Chapter 4.2], there exist a pair $(X_0^\epsilon, u^\epsilon(\cdot)) \in \mathsf{X} \times \mathcal{U}$ such that

$$\begin{cases} \widehat{d}((X_0^\epsilon, \overline{X}_0), (u^\epsilon(\cdot), \overline{u}(\cdot))) \leq \sqrt{\epsilon}, \\ J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot)) \leq J_\epsilon(\overline{X}_0; \overline{u}(\cdot)) = \epsilon, \\ J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot)) \leq J_\epsilon(X_0; u(\cdot)) + \sqrt{\epsilon}\widehat{d}((X_0^\epsilon, X_0), (u^\epsilon(\cdot), u(\cdot))), \ \forall u(\cdot) \in \mathcal{U}. \end{cases} \qquad (5.2)$$

Let $X^\epsilon(\cdot) := X(\cdot; X_0^\epsilon, u^\epsilon) \in C([0, T]; \mathsf{X})$ be the state trajectory of (2.1) controlled by $u^\epsilon(\cdot) \in \mathcal{U}$ with $X^\epsilon(0) = X_0^\epsilon$. It follows from Theorem 2.1 that $\sup_{t \in [0, T]} |X^\epsilon(t)| \leq M$.

[¶] In this proof, we first show $\mathsf{p} \geq 0$, and then $\mathsf{p}^0 := -\mathsf{p}$.

5.2. Family of spike variations

Recall that $\omega := (\{u_j(\cdot)\}_{j=1}^k, \{\theta_j\}_{j=1}^k) \in \Omega$, where $\sum_{j=1}^k \theta_j = 1$ and $\theta_j \geq 0$. Let us define

$$h_j^\epsilon(s) := \begin{bmatrix} l(s, X^\epsilon(s), u_j(s)) - l(s, X^\epsilon(s), u^\epsilon(s)) \\ f(s, X^\epsilon(s), u_j(s)) - f(s, X^\epsilon(s), u^\epsilon(s)) \end{bmatrix} \in \mathbb{R} \times \mathsf{X}.$$

Clearly, $h_j^\epsilon(\cdot) \in L^1([0, T]; \mathbb{R} \times \mathsf{X})$, $j = 1, \ldots, k$ by Assumptions 1 and 2. Hence, by invoking Lemma C.6 in Appendix 6, for any $\delta \in (0, 1)$, there are $\{E_j^\delta \subset [0, T], \ j = 1 \ldots, k\}$ such that for $\omega \in \Omega$ uniformly,

$$\begin{cases} E_j^\delta \cap E_i^\delta = \emptyset, \ j \neq i, \\ \cup_{j=1}^k \mathrm{meas}|E_j^\delta| = \delta T, \\ \sup_{t \in [0,T]} \left| \delta \sum_{j=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_j^\epsilon(s) \mathrm{d}s - \sum_{j=1}^k \int_{E_j^\delta \cap [0,t]} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_j^\epsilon(s) \mathrm{d}s \right|_{\mathbb{R} \times \mathsf{X}} = o(\delta). \end{cases} \tag{5.3}$$

We define the family of spike variations by

$$u_\omega^{\epsilon,\delta}(t) := \begin{cases} u_j(t), & \text{for } t \in E_j^\delta, \ j = 1, \ldots, k, \\ u^\epsilon(t), & \text{for } t \in [0, T] \setminus \cup_{j=1}^k E_j^\delta. \end{cases} \tag{5.4}$$

Clearly $u_\omega^{\epsilon,\delta}(\cdot) \in \mathcal{U}$ and $\tilde{d}(u_\omega^{\epsilon,\delta}(\cdot), u^\epsilon(\cdot)) \leq \mathrm{meas}|E_\delta| = \delta T$, where $E^\delta := \cup_{j=1}^k E_j^\delta$. Note that

$$\sum_{j=1}^k \int_{E_j^\delta \cap [0,t]} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_j^\epsilon(s) \mathrm{d}s = \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_\omega^{\epsilon,\delta}(s) \mathrm{d}s, \tag{5.5}$$

where $h_\omega^{\epsilon,\delta}(\cdot) \in \mathbb{R} \times \mathsf{X}$ is defined by

$$h_\omega^{\epsilon,\delta}(s) := \begin{bmatrix} l(s, X^\epsilon(s), u_\omega^{\epsilon,\delta}(s)) - l(s, X^\epsilon(s), u^\epsilon(s)) \\ f(s, X^\epsilon(s), u_\omega^{\epsilon,\delta}(s)) - f(s, X^\epsilon(s), u^\epsilon(s)) \end{bmatrix}.$$

Then from (5.3) and (5.5), it follows that

$$\sup_{(t,\omega) \in [0,T] \times \Omega} \left| \delta \sum_{k=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_j^\epsilon(s) \mathrm{d}s - \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_\omega^{\epsilon,\delta}(s) \mathrm{d}s \right|_{\mathbb{R} \times \mathsf{X}} = o(\delta). \tag{5.6}$$

Notice that (5.6) will be used to prove the variational analysis in Lemma 5.1.

We consider the state variation $X_\omega^{\epsilon,\delta}(\cdot) := X(\cdot; X_0^\epsilon + \delta a, u_\omega^{\epsilon,\delta}) \in C([0, T]; \mathsf{X})$, where $a \in \mathsf{X}$. Note that $X_\omega^{\epsilon,\delta}$ is the variational equation under (5.4) and the perturbed initial condition $a \in \mathsf{X}$. Let us define (with $\widetilde{X}_\omega^{\epsilon,\delta}(\cdot) := X_\omega^{\epsilon,\delta}(\cdot) - X^\epsilon(\cdot)$)

$$f^\epsilon(s) := f(s, X^\epsilon(s), u^\epsilon(s)), \ \partial_X f^\epsilon(s) := \partial_X f(s, X^\epsilon(s), u^\epsilon(s)),$$

$$\widehat{f_j^\epsilon}(s) := f(s, X^\epsilon(s), u_j(s)) - f(s, X^\epsilon(s), u^\epsilon(s)), \ f_X^{\epsilon,\delta}(s) := \int_0^1 \partial_X f(s, X^\epsilon(s) + r\widetilde{X}_\omega^{\epsilon,\delta}(s), u_\omega^{\epsilon,\delta}(s)) \mathrm{d}r,$$

$$l^\epsilon(s) := l(s, X^\epsilon(s), u^\epsilon(s)), \ \partial_X l^\epsilon(s) := \partial_X l(s, X^\epsilon(s), u^\epsilon(s)),$$

$$\widehat{l_j^\epsilon}(s) := l(s, X^\epsilon(s), u_j(s)) - l(s, X^\epsilon(s), u^\epsilon(s)), \quad l_X^{\epsilon,\delta}(s) := \int_0^1 \partial_X l(s, X^\epsilon(s) + r\widetilde{X}_\omega^{\epsilon,\delta}(s), u_\omega^{\epsilon,\delta}(s)) dr,$$

$$m^\epsilon := m(X_0^\epsilon, X^\epsilon(t_f)), \quad \partial_{X_0} m^\epsilon := \partial_{X_0} m(X_0^\epsilon, X^\epsilon(t_f)), \quad \partial_X m^\epsilon := \partial_X m(X_0^\epsilon, X^\epsilon(t_f)),$$

$$m_{X_0}^{\epsilon,\delta}(t_f) := \int_0^1 \partial_{X_0} m(X_0^\epsilon + r\delta a, X^\epsilon(t_f) + r\widetilde{X}_\omega^{\epsilon,\delta}(r)) dr, \quad m_X^{\epsilon,\delta}(t_f) := \int_0^1 \partial_X m(X_0^\epsilon + r\delta a, X^\epsilon(t_f) + r\widetilde{X}_\omega^{\epsilon,\delta}(r)) dr,$$

$$\widehat{f_j}(s) := f(t, \overline{X}(s), u_j(s)) - f(t, \overline{X}(s), \overline{u}(s)), \quad \widehat{l_j}(s) := l(s, \overline{X}(s), u_j(s)) - l(t, \overline{X}(s), \overline{u}(s)).$$

5.3. Variational analysis of δ

We state the variational analysis with respect to δ.

Lemma 5.1. *The following estimates hold:*

$$\sup_{\omega \in \Omega} \sup_{t \in [0,T]} \left| \frac{X_\omega^{\epsilon,\delta}(t) - X^\epsilon(t)}{\delta} - Z_\omega^\epsilon(t) \right|_X = o(1),$$

$$\sup_{\omega \in \Omega} \left| \frac{J(X_0^\epsilon + \delta a; u^{\epsilon,\delta}(\cdot)) - J(X_0^\epsilon; u^\epsilon(\cdot))}{\delta} - \widehat{Z}_\omega^\epsilon \right| = o(1),$$

where Z_ω^ϵ is the X-valued variational equation given by

$$\begin{cases} {}^C D_{0+}^\alpha [Z_\omega^\epsilon](t) + A Z_\omega^\epsilon(t) = \partial_X f^\epsilon(t) Z_\omega^\epsilon(t) + \sum_{j=1}^k \theta_j \widehat{f_j^\epsilon}(t), \ t \in (0, T], \\ Z_\omega^\epsilon(0) = a \in X, \end{cases} \tag{5.7}$$

and $\widehat{Z}_\omega^\epsilon$ is the $\mathbb{R}$-valued variational equation:

$$\widehat{Z}_\omega^\epsilon = \sum_{j=1}^k \theta_j \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \left(\partial_X l^\epsilon(s) Z_\omega^\epsilon(s) + \widehat{l_j^\epsilon}(s) \right) ds + \partial_{X_0} m^\epsilon a + \partial_X m^\epsilon Z_\omega^\epsilon(T). \tag{5.8}$$

Proof. Let $Z_\omega^{\epsilon,\delta}(\cdot) := \frac{X_\omega^{\epsilon,\delta}(\cdot) - X^\epsilon(\cdot)}{\delta}$. Then $Z_\omega^{\epsilon,\delta}$ satisfies

$$\begin{aligned} Z_\omega^{\epsilon,\delta}(t) = &\frac{1}{\delta} \Big(X_0 + \delta a + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \big(-A X_\omega^{\epsilon,\delta}(s) + f(s, X_\omega^{\epsilon,\delta}(s), u_\omega^{\epsilon,\delta}(s)) \big) ds \\ &- X_0 - \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \big(-A X^\epsilon(s) + f(s, X^\epsilon(s), u^\epsilon(s)) \big) ds \Big) \\ = &- \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} A Z_\omega^{\epsilon,\delta}(s) ds \\ &+ \frac{1}{\delta} \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \big(f(s, X_\omega^{\epsilon,\delta}(s), u_\omega^{\epsilon,\delta}(s)) - f(s, X^\epsilon(s), u_\omega^{\epsilon,\delta}(s)) \\ &+ f(s, X^\epsilon(s), u_\omega^{\epsilon,\delta}(s)) - f(s, X^\epsilon(s), u^\epsilon(s)) \big) ds. \end{aligned}$$

By (5.6), we get

$$Z_\omega^{\epsilon,\delta}(t) = \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \big(-A + f_X^{\epsilon,\delta}(s) \big) Z_\omega^{\epsilon,\delta}(s) ds + \sum_{j=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \widehat{f_j^\epsilon}(s) ds + \frac{o(\delta)}{\delta}.$$

We then have

$$Z_\omega^{\epsilon,\delta}(t) - Z_\omega^\epsilon(t) = -\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} A(Z_\omega^{\epsilon,\delta}(s) - Z_\omega^\epsilon(s))\mathrm{d}s + o(1)$$

$$+ \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} f_X^{\epsilon,\delta}(s)(Z_\omega^{\epsilon,\delta}(s) - Z_\omega^\epsilon(s))\mathrm{d}s$$

$$+ \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \Big(f_X^{\epsilon,\delta}(s) - \partial_X f^\epsilon(s)\Big)Z_\omega^\epsilon(s)\mathrm{d}s.$$

Since $Z^{\epsilon;\omega}$ is a linear Caputo fractional evolution equation, by Theorem 2.1, it follows that

$$\sup_{(\omega,t)\in\Omega\times[0,T]} |Z_\omega^\epsilon(t)|_X \le M. \tag{5.9}$$

Similarly, by (5.4) and Theorem 2.1, it follows that

$$\sup_{(\omega,t)\in\Omega\times[0,T]} |X_\omega^{\epsilon,\delta}(s) - X^\epsilon(s)|_X \le M\widehat{d}(u_\omega^{\epsilon,\delta}(\cdot), u^\epsilon(\cdot)) \le M\delta. \tag{5.10}$$

Recall $E^\delta := \cup_{j=1}^k E_j^\delta$ and $\{E_j^\delta \subset [0,T],\ j = 1\ldots,k\}$ with $\cup_{j=1}^k \mathrm{meas}|E_j^\delta| = \delta T$. We need to obtain the estimate of the following expression:

$$\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \Big(f_X^{\epsilon,\delta}(s) - \partial_X f^\epsilon(s)\Big)Z_\omega^\epsilon(s)\mathrm{d}s$$

$$= \int_{[0,t]\cap E^\delta} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \Big(f_X^{\epsilon,\delta}(s) - \partial_X f^\epsilon(s)\Big)Z_\omega^\epsilon(s)\mathrm{d}s$$

$$+ \int_{[0,t]\setminus E^\delta} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \Big(f_X^{\epsilon,\delta}(s) - \partial_X f^\epsilon(s)\Big)Z_\omega^\epsilon(s)\mathrm{d}s = o(1).$$

By (5.9) and the fact that $\mathrm{meas}|E^\delta| = \delta T$, we have

$$\sup_{(\omega,t)\in\Omega\times[0,T]} \Big|\int_{[0,t]\cap E^\delta} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \Big(f_X^{\epsilon,\delta}(s) - \partial_X f^\epsilon(s)\Big)Z_\omega^\epsilon(s)\mathrm{d}s\Big|_X = o(1).$$

Also, by (5.9) and (5.10), we also have

$$\sup_{(\omega,t)\in\Omega\times[0,T]} \Big|\int_{[0,t]\setminus E^\delta} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \Big(f_X^{\epsilon,\delta}(s) - \partial_X f^\epsilon(s)\Big)Z^\epsilon(s)\mathrm{d}s\Big|_X$$

$$\le M \sup_{(\omega,t)\in\Omega\times[0,T]} \Big|\int_{[0,t]\setminus E^\delta} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} |X_\omega^{\epsilon,\delta}(s) - X^\epsilon(s)|_X |Z_\omega^\epsilon(s)|_X \mathrm{d}s\Big|_X \le M\delta = o(1).$$

Based on the preceding estimates, it follows that

$$Z_\omega^{\epsilon,\delta}(t) - Z_\omega^\epsilon(t) \le \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \Big(-A + M\Big)(Z_\omega^{\epsilon,\delta}(s) - Z_\omega^\epsilon(s))\mathrm{d}s + o(1).$$

By Lemma A.2,

$$\sup_{(\omega,t)\in\Omega\times[0,T]} |Z_\omega^{\epsilon,\delta}(t) - Z_\omega^\epsilon(t)|_X \le o(1)E_\alpha(M\Gamma(\alpha)T^\alpha).$$

As $\delta \downarrow 0$, the first estimate follows.

To prove the second estimate, we note that

$$
\frac{J(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) - J(X_0^\epsilon; u^\epsilon(\cdot))}{\delta} - \widehat{Z}_\omega^\epsilon
$$

$$
= \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} l_X^{\epsilon,\delta}(s)(Z_\omega^{\epsilon,\delta}(s) - Z_\omega^\epsilon(s)) \mathrm{d}s
$$

$$
+ \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \Big(l_X^{\epsilon,\delta}(s) - \partial_X l^\epsilon(s) \Big) Z_\omega^\epsilon(s) \mathrm{d}s + (m_{X_0}^{\epsilon,\delta}(t_f) - \partial_{X_0} m^\epsilon) a
$$

$$
+ m_X^{\epsilon,\delta}(T)(Z_\omega^{\epsilon,\delta}(T) - Z_\omega^\epsilon(T)) + (m_X^{\epsilon,\delta}(T) - \partial_X m^\epsilon) Z_\omega^\epsilon(T).
$$

Then the rest of the proof is analogous to that for the first estimate. This completes the proof. $\qquad\square$

Now, consider the situation of $\delta \downarrow 0$. For any $\omega = (\{u_j(\cdot)\}_{j=1}^k, \{\theta_j\}_{j=1}^k) \in \Omega$, in view of (5.2), it follows that

$$
\frac{J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) - J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))}{\delta} \geq -\frac{1}{\delta} \widehat{d}((X_0^\epsilon + \delta X_0, X_0^\epsilon)(u_\omega^{\epsilon,\delta}(\cdot), u^\epsilon(\cdot))) \geq -\sqrt{\epsilon}(T + |a|_\mathsf{X}).
$$

We then have

$$
-\sqrt{\epsilon}(T + |a|_\mathsf{X}) \leq \frac{1}{\delta} \frac{1}{J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) + J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} \tag{5.11}
$$

$$
\times \bigg(\Big(([J(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+)^2 + d_{\mathsf{F}_0}(X_0^\epsilon + \delta a)^2 + d_{\mathsf{F}_T}(X_\omega^{\epsilon,\delta}(T))^2 \Big)
$$

$$
- \Big(([J(X_0^\epsilon; u^\epsilon(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+)^2 + d_{\mathsf{F}_0}(X_0^\epsilon)^2 + d_{\mathsf{F}_T}(X^\epsilon(T))^2 \Big) \bigg).
$$

Let us consider the case when $\delta \downarrow 0$ in (5.11).

(F.1) Note that $\lim_{\delta \downarrow 0} J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) = J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot)) + o(1)$ by Lemma 5.1. Hence,

$$
\lim_{\delta \downarrow 0} \frac{1}{\delta} \frac{1}{J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) + J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))}
$$

$$
\times \bigg(([J(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+)^2 - ([J((X_0^\epsilon; u^\epsilon(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+)^2 \bigg)
$$

$$
= \lim_{\delta \downarrow 0} \frac{1}{J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) + J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))}
$$

$$
\times \lim_{\delta \downarrow 0} \bigg(([J(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+) + ([J((X_0^\epsilon; u^\epsilon(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+) \bigg)
$$

$$
\times \lim_{\delta \downarrow 0} \frac{1}{\delta} \bigg(([J(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+) - ([J((X_0^\epsilon; u^\epsilon(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+) \bigg)
$$

$$
= \frac{[J(X_0^\epsilon; u^\epsilon(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} \widehat{Z}_\omega^\epsilon =: \lambda_\epsilon \widehat{Z}_\omega^\epsilon,
$$

where by definition, $\lambda_\epsilon \geq 0$.

(F.2) By Lemma 5.1, it follows that

$$\lim_{\delta\downarrow 0} \frac{1}{\delta} \frac{1}{J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) + J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} \Big(d_{\mathsf{F}_T}(X_\omega^{\epsilon,\delta}(T))^2 - d_{\mathsf{F}_T}(X^\epsilon(T))^2\Big)$$

$$= \lim_{\delta\downarrow 0} \frac{1}{J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) + J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))}$$

$$\times \lim_{\delta\downarrow 0}\big(d_{\mathsf{F}_T}(X_\omega^{\epsilon,\delta}(T)) + d_{\mathsf{F}_T}(X^\epsilon(T))\big) \lim_{\delta\downarrow 0} \frac{1}{\delta}\big(d_{\mathsf{F}_T}(X_\omega^{\epsilon,\delta}(T)) - d_{\mathsf{F}_T}(X^\epsilon(T))\big)$$

$$= \frac{1}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} d_{\mathsf{F}_T}(X^\epsilon(T)) \lim_{\delta\downarrow 0} \frac{1}{\delta}\Big(d_{\mathsf{F}_T}(X^\epsilon(T) + \delta Z_\omega^\epsilon(T)) - d_{\mathsf{F}_T}(X^\epsilon(T))\Big)$$

$$+ \frac{1}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} d_{\mathsf{F}_T}(X^\epsilon(T)) \lim_{\delta\downarrow 0} \frac{1}{\delta}\Big(d_{\mathsf{F}_T}(X^\epsilon(T) + \delta Z_\omega^\epsilon(T) + o(\delta)) - d_{\mathsf{F}_T}(X^\epsilon(T) + \delta Z_\omega^\epsilon(T))\Big)$$

$$= \frac{1}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} d_{\mathsf{F}_T}(X^\epsilon(T)) \nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z_\omega^\epsilon(T)) =: \nu_\epsilon \nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z_\omega^\epsilon(T)),$$

where $\nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z_\omega^\epsilon(T))$ is the directional (or Gâteaux) derivative of d_{F_T} at $X^\epsilon(T)$ and $\nu_\epsilon \geq 0$.

(F.3) Similarly, we can prove that

$$\lim_{\delta\downarrow 0} \frac{1}{\delta} \frac{1}{J_\epsilon(X_0^\epsilon + \delta a; u_\omega^{\epsilon,\delta}(\cdot)) + J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))}\Big(d_{\mathsf{F}_0}(X_0^\epsilon + \delta a)^2 - d_{\mathsf{F}_0}(X_0^\epsilon)^2\Big)$$

$$= \frac{1}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} d_{\mathsf{F}_0}(X_0^\epsilon) \nabla d_{\mathsf{F}_0}(X_0^\epsilon; a) =: \eta_\epsilon \nabla d_{\mathsf{F}_0}(X_0^\epsilon; a),$$

where $\nabla d_{\mathsf{F}_0}(X_0^\epsilon; a)$ is the directional (or Gâteaux) derivative of d_{F_0} at X_0^ϵ and $\eta_\epsilon \geq 0$.

Based on Lemma 5.1 and (F.1)–(F.3) above, as $\delta \downarrow 0$, (5.11) becomes

$$-\sqrt{\epsilon}(T + |a|_\mathsf{X}) \leq \frac{[J(X_0^\epsilon; u^\epsilon(\cdot)) - J(\overline{X}_0; \overline{u}(\cdot)) + \epsilon]^+}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} \widehat{Z}_\omega^\epsilon \tag{5.12}$$

$$+ \frac{d_{\mathsf{F}_0}(X_0^\epsilon)}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} \nabla d_{\mathsf{F}_0}(X_0^\epsilon; a) + \frac{d_{\mathsf{F}_T}(X^\epsilon(T))}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} \nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z_\omega^\epsilon(T))$$

$$= \lambda_\epsilon \widehat{Z}_\omega^\epsilon + \eta_\epsilon \nabla d_{\mathsf{F}_0}(X_0^\epsilon; a) + \nu_\epsilon \nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z_\omega^\epsilon(T)) =: \Psi_\omega^\epsilon(\lambda_\epsilon, \eta_\epsilon, \nu_\epsilon),$$

where $\lambda_\epsilon, \eta_\epsilon, \nu_\epsilon \geq 0$ and by definition of J_ϵ, the multiplier condition is given by

$$\lambda_\epsilon^2 + \eta_\epsilon^2 + \nu_\epsilon^2 = 1. \tag{5.13}$$

5.4. Separation

This subsection is not needed when X^*, the dual space of X, is strictly convex (see Remark 3.2). Note that the strict convexity of X^* is crucial in various results on infinite-dimensional maximum principles to obtain the explicit derivative of the distance function. In this subsection, we obtain an alternative expression of (5.12) and (5.13) via the separation argument.

Let us define

$$\mathcal{A}^\epsilon := \{(a, Z, \widehat{Z}, \gamma) \in \mathsf{X} \times \mathsf{X} \times \mathbb{R} \times \mathbb{R} \mid \lambda_\epsilon \widehat{Z} + \eta_\epsilon \nabla d_{\mathsf{F}_0}(X_0^\epsilon; a) + \nu_\epsilon \nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z) \leq \gamma\},$$

$$\mathcal{B}^{\epsilon} := \{(\bar{a}, \bar{Z}, \widehat{\bar{Z}}, \bar{\gamma}) \in \mathsf{X} \times \mathsf{X} \times \mathbb{R} \times \mathbb{R} \mid (\bar{a}, Z, \widehat{Z}, \gamma) = (\bar{a}, Z_{\omega}^{\epsilon}(T), \widehat{Z}_{\omega}^{\epsilon}, -\sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}}))\}.$$

By Lemma C.7 in Appendix 6, $\mathcal{B}^{\epsilon}$ is a convex subset of $\mathsf{X} \times \mathsf{X} \times \mathbb{R} \times \mathbb{R}$. In addition, $\mathcal{A}^{\epsilon}$ is convex and nonempty, where the former follows from the fact that the Gâteaux derivative is (Lipschitz) continuous and convex its direction, while the latter is due to the finiteness of the Gâteaux derivatives [10, Chapter 2]. Note also that $\mathcal{B}^{\epsilon}$ contains no interior points of $\mathcal{A}^{\epsilon}$.

By the separation theorem (also known as the geometric version of Hahn-Banach theorem) for convex sets, there exist $(\varphi_1^{\epsilon}, \varphi_2^{\epsilon}, \mathsf{p}^{\epsilon}, \mathsf{q}^{\epsilon}) \in \mathsf{X}^* \times \mathsf{X}^* \times \mathbb{R} \times \mathbb{R}$ with $(\varphi_1^{\epsilon}, \varphi_2^{\epsilon}, \mathsf{p}^{\epsilon}, \mathsf{q}^{\epsilon}) \neq 0$ such that for any $(a, Z, \widehat{Z}, \gamma) \in \mathcal{A}^{\epsilon}$ and $(\bar{a}, \bar{Z}, \widehat{\bar{Z}}, -\sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}})) \in \mathcal{B}^{\epsilon}$,

$$\langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z \rangle + \mathsf{p}^{\epsilon} \widehat{Z} + \mathsf{q}^{\epsilon} \gamma \leq \langle \varphi_1^{\epsilon}, \bar{a} \rangle + \langle \varphi_2^{\epsilon}, \bar{Z} \rangle + \mathsf{p}^{\epsilon} \widehat{\bar{Z}} - \mathsf{q}^{\epsilon} \sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}}). \tag{5.14}$$

Notice that $(0, 0, 0, -\sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}})) \in \mathcal{B}^{\epsilon}$, which corresponds to the case of no variations.

Let us assume that the left-hand side of (5.14) holds for some $(a, Z, \widehat{Z}, \gamma) \in \mathcal{A}^{\epsilon}$,

$$\langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z \rangle + \mathsf{p}^{\epsilon} \widehat{Z} + \mathsf{q}^{\epsilon} \gamma > 0.$$

Since $(ka, kZ, k\widehat{Z}, k\gamma) \in \mathcal{A}^{\epsilon}$ for $k \geq 1$, it follows that $\langle \varphi_1^{\epsilon}, ka \rangle + \langle \varphi_2^{\epsilon}, kZ \rangle + k\mathsf{p}^{\epsilon} \widehat{Z} + k\mathsf{q}^{\epsilon} \to \infty$ as $k \to \infty$, which contradicts (5.14). Therefore, for any $(a, Z, \widehat{Z}, \gamma) \in \mathcal{A}^{\epsilon}$ and $(\bar{a}, \bar{Z}, \widehat{\bar{Z}}, -\sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}})) \in \mathcal{B}^{\epsilon}$, (5.14) can be rewritten as

$$\langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z \rangle + \mathsf{p}^{\epsilon} \widehat{Z} + \mathsf{q}^{\epsilon} \gamma \leq 0 \leq \langle \varphi_1^{\epsilon}, \bar{a} \rangle + \langle \varphi_2^{\epsilon}, \bar{Z} \rangle + \mathsf{p}^{\epsilon} \widehat{\bar{Z}} - \mathsf{q}^{\epsilon} \sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}}). \tag{5.15}$$

By (5.15) and (5.12), for a fixed $\omega \in \Omega$, it follows that

$$\langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z_{\omega}^{\epsilon}(T) \rangle + \mathsf{p}^{\epsilon} \widehat{Z}_{\omega}^{\epsilon} + \mathsf{q}^{\epsilon} \Psi_{\omega}^{\epsilon}(\lambda_{\epsilon}, \eta_{\epsilon}, \nu_{\epsilon}) \leq \langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z_{\omega}^{\epsilon}(T) \rangle + \mathsf{p}^{\epsilon} \widehat{Z}_{\omega}^{\epsilon} - \mathsf{q}^{\epsilon}(\sqrt{\epsilon}(T + |a|_{\mathsf{X}})),$$

which leads to $\mathsf{q}^{\epsilon} \Psi_{\omega}^{\epsilon}(\lambda_{\epsilon}, \eta_{\epsilon}, \nu_{\epsilon}) \leq -\mathsf{q}^{\epsilon}(\sqrt{\epsilon}(T + |a|_{\mathsf{X}}))$. By (5.12), this implies $\mathsf{q}^{\epsilon} = -1$. Note that when $\mathsf{q}^{\epsilon} = 0$, from (5.15), we have

$$\langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z \rangle + \mathsf{p}^{\epsilon} \widehat{Z} \leq 0, \ \forall (a, Z, \widehat{Z}) \in \mathsf{X}^* \times \mathsf{X}^* \times \mathbb{R}.$$

This is possible only if $\varphi_1^{\epsilon} = \varphi_2^{\epsilon} = 0 \in \mathsf{X}^*$ and $\mathsf{p}^{\epsilon} = 0$, which is impossible. Hence, only $\mathsf{q}^{\epsilon} = -1$ is possible in (5.15).

From (5.15), it follows that for any $\omega \in \Omega$ and $(\bar{a}, \bar{Z}, \widehat{\bar{Z}}, -\sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}})) \in \mathcal{B}^{\epsilon}$,

$$0 \leq \langle \varphi_1^{\epsilon}, \bar{a} \rangle + \langle \varphi_2^{\epsilon}, Z \rangle + \mathsf{p}^{\epsilon} \widehat{Z} + \sqrt{\epsilon}(T + |\bar{a}|_{\mathsf{X}}),$$

which implies

$$0 \leq \langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z_{\omega}^{\epsilon}(T) \rangle + \mathsf{p}^{\epsilon} \widehat{Z}_{\omega}^{\epsilon} + \sqrt{\epsilon}(T + |a|_{\mathsf{X}}). \tag{5.16}$$

Also, from (5.15) and the definition of $\mathcal{A}^{\epsilon}$, for any $(a, Z, \widehat{Z}) \in \mathsf{X} \times \mathsf{X} \times \mathbb{R}$,

$$\langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z \rangle + \mathsf{p}^{\epsilon} \widehat{Z} \leq \lambda_{\epsilon} \widehat{Z} + \eta_{\epsilon} \nabla d_{\mathsf{F}_0}(X_0^{\epsilon}; a) + \nu_{\epsilon} \nabla d_{\mathsf{F}_T}(X^{\epsilon}(T); Z). \tag{5.17}$$

Thus, with $a = Z = 0 \in \mathsf{X}$ in (5.17), we have $\mathsf{p}^{\epsilon} \widehat{Z} \leq \lambda_{\epsilon} \widehat{Z}$, which leads to $\mathsf{p}^{\epsilon} = \lambda_{\epsilon} \geq 0$. Therefore, it follows that for $(a, Z) \in \mathsf{X} \times \mathsf{X}$, (5.17) becomes

$$\langle \varphi_1^{\epsilon}, a \rangle + \langle \varphi_2^{\epsilon}, Z \rangle \leq \eta_{\epsilon} \nabla d_{\mathsf{F}_0}(X_0^{\epsilon}; a) + \nu_{\epsilon} \nabla d_{\mathsf{F}_T}(X^{\epsilon}(T); Z). \tag{5.18}$$

We now consider three different cases of (5.18) to obtain alternative form of (5.13):

- When $\eta_\epsilon = \nu_\epsilon = 0$, (5.18) holds for any $(a, Z) \in \mathsf{X} \times \mathsf{X}$. Hence, $\varphi_1^\epsilon = \varphi_2^\epsilon = 0 \in \mathsf{X}^*$, by which $\lambda_\epsilon^2 = |\mathsf{p}^\epsilon|^2 = 1$ with $\lambda_\epsilon = \mathsf{p}^\epsilon \geq 0$.
- When $\eta_\epsilon = 0$ and $\nu_\epsilon \neq 0$, we have $\varphi_1^\epsilon = 0 \in \mathsf{X}^*$ and (5.18) becomes

$$\Big\langle \frac{\varphi_2^\epsilon}{\nu_\epsilon}, Z \Big\rangle \leq \nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z).$$

Then by [10, Proposition 2.1.2], $\frac{\varphi_2^\epsilon}{\nu_\epsilon} \in \partial d_{\mathsf{F}_T}(X^\epsilon(T))$, i.e., $\frac{\varphi_2^\epsilon}{\nu_\epsilon} \in \mathsf{X}^*$ is a generalized gradient of $d_{\mathsf{F}_T}(X^\epsilon(T))$, and $|\frac{\varphi_2^\epsilon}{\nu_\epsilon}|_{\mathsf{X}^*} = 1$, i.e., $|\varphi_2^\epsilon|_{\mathsf{X}^*} = \nu_\epsilon$. Similarly, when $\eta_\epsilon \neq 0$ and $\nu_\epsilon = 0$, we have $\varphi_2^\epsilon = 0 \in \mathsf{X}^*$ and $\varphi_1^\epsilon \in \partial d_{\mathsf{F}_0}(X_0^\epsilon)$, implying $|\frac{\varphi_1^\epsilon}{\eta_\epsilon}|_{\mathsf{X}^*} = 1$, i.e., $|\varphi_1^\epsilon|_{\mathsf{X}^*} = \eta_\epsilon$.
- When $\eta_\epsilon \neq 0$ and $\nu_\epsilon \neq 0$, by (5.18), we have

$$\eta_\epsilon \Big\langle \frac{\varphi_1^\epsilon}{\eta_\epsilon}, a \Big\rangle + \nu_\epsilon \Big\langle \frac{\varphi_2^\epsilon}{\nu_\epsilon}, Z \Big\rangle \leq \eta_\epsilon \nabla d_{\mathsf{F}_0}(X_0^\epsilon; a) + \nu_\epsilon \nabla d_{\mathsf{F}_T}(X^\epsilon(T); Z).$$

Analogously, $\frac{\varphi_1^\epsilon}{\eta_\epsilon} \in \partial d_{\mathsf{F}_0}(X_0^\epsilon)$ and $\frac{\varphi_2^\epsilon}{\nu_\epsilon} \in \partial d_{\mathsf{F}_T}(X^\epsilon(T))$, which imply $|\varphi_1^\epsilon|_{\mathsf{X}^*} = \eta_\epsilon$ and $|\varphi_2^\epsilon|_{\mathsf{X}^*} = \nu_\epsilon$.

Hence, by (5.16), for any $\omega \in \Omega$, we have

$$-\sqrt{\epsilon}(T + |a|_{\mathsf{X}}) \leq \langle \varphi_1^\epsilon, a \rangle_{\mathsf{X} \times \mathsf{X}^*} + \langle \varphi_2^\epsilon, Z_\omega^\epsilon(T) \rangle_{\mathsf{X} \times \mathsf{X}^*} + \mathsf{p}^\epsilon \widehat{Z_\omega^\epsilon}, \tag{5.19}$$

where in view of all three cases discussed above and (5.13), it follows that the multiplier condition is equivalent to (note that $\mathsf{p}^\epsilon \geq 0$)

$$|\mathsf{p}^\epsilon|^2 + |\varphi_1^\epsilon|_{\mathsf{X}^*}^2 + |\varphi_2^\epsilon|_{\mathsf{X}^*}^2 = 1. \tag{5.20}$$

To obtain the transversality condition, by the fact that F_0 and F_T are convex, the definition of subdifferential for convex functions (see [10, page 36]) implies for any $(X'', X') \in \mathsf{X} \times \mathsf{X}$,

$$\Big\langle \frac{\varphi_1^\epsilon}{\eta_\epsilon}, X'' - X_0^\epsilon \Big\rangle \leq d_{\mathsf{F}_0}(X'') - d_{\mathsf{F}_0}(X_0^\epsilon),$$

$$\Big\langle \frac{\varphi_2^\epsilon}{\nu_\epsilon}, X' - X^\epsilon(T) \Big\rangle \leq d_{\mathsf{F}_T}(X') - d_{\mathsf{F}_T}(X^\epsilon(T)).$$

Hence, for $(X'', X') \in \mathsf{F}_0 \times \mathsf{F}_T$, we have (see (F.2) and (F.3))

$$\frac{d_{\mathsf{F}_0}(X_0^\epsilon)^2}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} + \langle \varphi_1^\epsilon, X'' - X_0^\epsilon \rangle \leq 0, \tag{5.21}$$

$$\frac{d_{\mathsf{F}_T}(X^\epsilon(T))^2}{J_\epsilon(X_0^\epsilon; u^\epsilon(\cdot))} + \langle \varphi_2^\epsilon, X' - X^\epsilon(T) \rangle \leq 0. \tag{5.22}$$

5.5. Variational analysis of ϵ

The following lemma shows the variational analysis with respect to ϵ. The proof is analogous to that of Lemma 5.1.

Lemma 5.2. *The following estimates hold*

$$\sup_{\omega \in \Omega} \sup_{t \in [0,T]} |Z_\omega^\epsilon(t) - Z_\omega(t)|_X = o(1), \qquad \sup_{\omega \in \Omega} |\widehat{Z}_\omega^\epsilon - \widehat{Z}_\omega| = o(1),$$

where Z_ω is the X-valued variational equation given by

$$\begin{cases} {}^C D_{0+}^\alpha [Z_\omega](t) + AZ_\omega(t) = \partial_X \overline{f}(t) Z_\omega(t) + \sum_{j=1}^k \theta_j \widehat{f}_j(t), \ t \in (0, T], \\ Z_\omega(0) = a \in X, \end{cases} \tag{5.23}$$

and $\widehat{Z}_\omega$ is the $\mathbb{R}$-valued variational equation:

$$\widehat{Z}_\omega = \sum_{j=1}^k \theta_j \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \left(\langle \partial_X \overline{l}(s), Z_\omega(s) \rangle + \widehat{l}_j(s) \right) ds + \langle \partial_{X_0} \overline{m}, a \rangle + \langle \partial_X \overline{m}, Z_\omega(T) \rangle. \tag{5.24}$$

From (5.21), (5.22) and the fact that $|\varphi_1^\epsilon|_{X^*} \le 1$ and $|\varphi_2^\epsilon|_{X^*} \le 1$, for $(X'', X') \in F_0 \times F_T$, as $\epsilon \downarrow 0$,

$$\begin{cases} \langle \varphi_1^\epsilon, X'' - \overline{X}_0 \rangle = \langle \varphi_1^\epsilon, X'' - X_0^\epsilon \rangle + \langle \varphi_1^\epsilon, X_0^\epsilon - \overline{X}_0 \rangle \le |\varphi_1^\epsilon|_{X^*} |X_0^\epsilon - \overline{X}_0|_X \to 0, \\ \langle \varphi_2^\epsilon, X' - \overline{X}(T) \rangle \le |\varphi_2^\epsilon|_{X^*} |X^\epsilon(T) - \overline{X}(T)|_X \to 0. \end{cases} \tag{5.25}$$

By Lemma 5.2, for any $\omega \in \Omega$, (5.19) becomes

$$\begin{aligned} &\langle \varphi_1^\epsilon, a \rangle + \langle \varphi_2^\epsilon, Z_\omega(T) \rangle + \mathsf{p}^\epsilon \widehat{Z}_\omega \\ &= \langle \varphi_1^\epsilon, a \rangle + \langle \varphi_2^\epsilon, Z_\omega^\epsilon(T) \rangle + \mathsf{p}^\epsilon \widehat{Z}_\omega^\epsilon - \langle \varphi_2^\epsilon, Z_\omega^\epsilon(T) - Z_\omega(T) \rangle - \mathsf{p}^\epsilon(\widehat{Z}_\omega^\epsilon - \widehat{Z}_\omega) \\ &\ge -\sqrt{\epsilon}(T + |a|_X) - |\varphi_2^\epsilon|_{X^*} |Z_\omega^\epsilon(T) - Z_\omega(T)|_X - |\mathsf{p}^\epsilon| |\widehat{Z}_\omega^\epsilon - \widehat{Z}_\omega| \\ &\ge -\sqrt{\epsilon}(T + |a|_X) - |Z_\omega^\epsilon(T) - Z_\omega(T)|_X - |\widehat{Z}_\omega^\epsilon - \widehat{Z}_\omega| =: c_\epsilon. \end{aligned}$$

Note that by Lemma 5.2, for any $\omega \in \Omega$, we have $\lim_{\epsilon \downarrow 0} c_\epsilon = 0$. Then using (5.25), it follows that for any $\omega \in \Omega$ and $(X'', X') \in F_0 \times F_T$

$$\langle \varphi_1^\epsilon, a - (X'' - \overline{X}_0) \rangle_{X \times X^*} + \langle \varphi_2^\epsilon, Z_\omega(T) - (X' - \overline{X}(T)) \rangle_{X \times X^*} + \mathsf{p}^\epsilon \widehat{Z}_\omega \ge c_\epsilon \to 0 \text{ as } \epsilon \downarrow 0. \tag{5.26}$$

5.6. Proof of (i): nontriviality condition

Recall from Lemma B.4 in Appendix 6 that the variational equation in (5.23) can be written as

$$Z_\omega(t) = a + \int_0^t \Pi(t, s) \left(-A + \partial_X \overline{f}(s) \right) a \, ds + \sum_{j=1}^k \theta_j \int_0^t \Pi(t, s) \widehat{f}_j(s) ds, \ t \in [0, T]. \tag{5.27}$$

In view of (5.27) and Lemma B.4, $\mathcal{R}$ and $\mathcal{Q}$ in (3.1) and (3.2) can be rewritten as

$$\mathcal{R} = \left\{ \xi(T) \in X \mid \xi(t) = \sum_{j=1}^k \theta_j \int_0^t \Pi(t, s) \widehat{f}_j(s) ds, \ t \in [0, T], \ \left(\{u_j(\cdot)\}_{j=1}^k, \{\theta_j\}_{j=1}^k \right) \in \Omega \right\}, \tag{5.28}$$

$$Q = \left\{ \hat{X} - \hat{X}_0 - \int_0^T \Pi(T, s)\left(-A + \partial_X \overline{f}(s)\right)\hat{X}_0 \mathrm{d}s \in \mathsf{X} \mid (\hat{X}_0, \hat{X}) \in \mathsf{F}_0 \times \mathsf{F}_T \right\}. \tag{5.29}$$

Let us define

$$\widehat{\mathcal{R}} := \left\{ (\hat{X}_0, \hat{X}) \in \mathsf{X} \times \mathsf{X} \mid \hat{X} = \hat{X}_0 + \int_0^T \Pi(T, s)\left(-A + \partial_X \overline{f}(s)\right)\hat{X}_0 \mathrm{d}s + \xi, \ \xi \in \mathcal{R}, \ \hat{X}_0 \in \mathsf{X} \right\}.$$

Since $\mathcal{R} - Q$ is finite codimensional in X, by [26, Proposition 3.5, Chapter 4], it follows that $\widehat{\mathcal{R}} - \mathsf{F}$ (note that $\mathsf{F} = \mathsf{F}_0 \times \mathsf{F}_T$) is finite codimensional in $\mathsf{X} \times \mathsf{X}$. Then by [26, Proposition 3.4, Chapter 4], $\widehat{\mathcal{R}} - \mathsf{F} + \begin{bmatrix} \overline{X}_0 \\ \overline{X}(T) \end{bmatrix}$ is also finite codimensional in $\mathsf{X} \times \mathsf{X}$. Hence, (5.26) can be written as for any $\omega \in \Omega$,

$$\left\langle \begin{bmatrix} \varphi_1^{\epsilon} \\ \varphi_2^{\epsilon} \end{bmatrix}, \begin{bmatrix} X'' \\ X' \end{bmatrix} \right\rangle + \mathsf{p}^{\epsilon} \widehat{Z}_{\omega} \geq c_{\epsilon}, \ \forall \begin{bmatrix} X'' \\ X' \end{bmatrix} \in \widehat{\mathcal{R}} - \mathsf{F} + \begin{bmatrix} \overline{X}_0 \\ \overline{X}(T) \end{bmatrix}.$$

Let $(\mathsf{p}^{\epsilon_k}, \varphi_1^{\epsilon_k}, \varphi_2^{\epsilon_k}) \in \mathbb{R} \times \mathsf{X}^* \times \mathsf{X}^*$ be the sequence satisfying (5.20) for $k \geq 1$. Consider $\epsilon_k \downarrow 0$ as $k \to \infty$. Then by the Banach-Alaoglu theorem and (5.20), we can extract the subsequence, still denoted by $(\mathsf{p}^{\epsilon_k}, \varphi_1^{\epsilon_k}, \varphi_2^{\epsilon_k}) \in \mathbb{R} \times \mathsf{X}^* \times \mathsf{X}^*$, such that $(\mathsf{p}^{\epsilon_k}, \varphi_1^{\epsilon_k}, \varphi_2^{\epsilon_k}) \to (\mathsf{p}, \varphi_1, \varphi_2) \neq 0$. Notice that $\mathsf{p} \geq 0$ and the convergence of $(\varphi_1^{\epsilon_k}, \varphi_2^{\epsilon_k})$ to (φ_1, φ_2) as $k \to \infty$ is understood in the weak$-^*$ sense. Here, the nontriviality $\mathbb{R} \times \mathsf{X}^* \times \mathsf{X}^* \ni (\mathsf{p}, \varphi_1, \varphi_2) \neq 0$ follows from [26, Lemma 3.6, Chapter 4] and by the fact that $\widehat{\mathcal{R}} - \mathsf{F}$ is finite codimensional in $\mathsf{X} \times \mathsf{X}$. In particular, when $\mathsf{p} \neq 0$ ($\mathsf{p} > 0$), we are done. Otherwise, if $\mathsf{p}^{\epsilon_k} \to 0$ as $k \to \infty$, then $|\varphi_1^{\epsilon_k}|_{\mathsf{X}^*} + |\varphi_2^{\epsilon_k}|_{\mathsf{X}^*} > 0$ for sufficiently large k, which implies the nontriviality $\mathsf{X}^* \times \mathsf{X}^* \ni (\varphi_1, \varphi_2) \neq 0$ [26, Lemma 3.6, Chapter 4].

Hence, as $\epsilon \downarrow 0$, (5.26) can be written as for any $\omega \in \Omega$ and $(X'', X') \in \mathsf{F}_0 \times \mathsf{F}_T$,

$$\langle \varphi_1, a - (X'' - \overline{X}_0) \rangle + \langle \varphi_2, Z_{\omega}(T) - (X' - \overline{X}(T)) \rangle + \mathsf{p}\widehat{Z}_{\omega} \geq 0, \tag{5.30}$$

where $(\mathsf{p}, \varphi_1, \varphi_2) \in \mathbb{R} \times \mathsf{X}^* \times \mathsf{X}^*$ with $(\mathsf{p}, \varphi_1, \varphi_2) \neq 0$, $\mathsf{p} \geq 0$, and $|\mathsf{p}|^2 + |\varphi_1|_{\mathsf{X}^*}^2 + |\varphi_2|_{\mathsf{X}^*}^2 \leq 1$ from (5.20). Notice that by (5.25), it follows that for any $(X'', X') \in \mathsf{F}_0 \times \mathsf{F}_T$,

$$\langle \varphi_1, X'' - \overline{X}_0 \rangle \leq 0, \ \langle \varphi_2, X' - \overline{X}(T) \rangle \leq 0.$$

Then by definition of subdifferential of convex functions, it follows that

$$(\varphi_1, \varphi_2) \in N_{\mathsf{F}_0}(\overline{X}_0) \times N_{\mathsf{F}_T}(\overline{X}(T)),$$

see [10, Proposition 2.4.4]. This proves the nontriviality condition of Theorem 3.1.

5.7. Proof of (ii): duality and adjoint equation

Let $\mathsf{p}^0 := -\mathsf{p} \leq 0$. The proof of (ii) is given in Lemma B.5 in Appendix 6, by which we recall that the adjoint equation in (3.3) can be written as

$$P(t) = \Pi(T, t)^* \left(-\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}\right) + \mathsf{p}^0 \int_t^T \frac{(T - s)^{\beta - 1}}{\Gamma(\beta)} \Pi(s, t)^* \partial_X \overline{l}(s) \mathrm{d}s. \tag{5.31}$$

Then for any $\omega \in \Omega$ and $(X'', X') \in \mathsf{F}_0 \times \mathsf{F}_T$, (5.30) becomes

$$\langle \varphi_1, (X'' - \overline{X}_0) \rangle + \langle \varphi_1, -a \rangle - \langle -\varphi_2, (X' - \overline{X}(T)) \rangle - \langle -\varphi_2, -Z_\omega(T) \rangle + \mathsf{p}^0 \widehat{Z}_\omega \leq 0. \qquad (5.32)$$

By (5.27) and $\widehat{Z}_\omega$ in (5.24), we have

$$\langle \varphi_1, -a \rangle - \langle -\varphi_2, -Z_\omega(T) \rangle + \mathsf{p}^0 \widehat{Z}_\omega$$

$$= \langle -\varphi_1 + \mathsf{p}^0 \partial_{X_0} \overline{m}, a \rangle + \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \sum_{j=1}^k \theta_j \widehat{l}_j(s) \mathrm{d}s$$

$$+ \langle -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, a \rangle + \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \langle \partial_X \overline{l}(s), a \rangle \mathrm{d}s$$

$$+ \left\langle -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, \int_0^T \Pi(T, s) \left(-A + \partial_X \overline{f}(s) \right) a \mathrm{d}s \right\rangle$$

$$+ \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \left\langle \partial_X \overline{l}(s), \int_0^s \Pi(s, \tau) \left(-A + \partial_X \overline{f}(\tau) \right) a \mathrm{d}\tau \right\rangle \mathrm{d}s$$

$$+ \left\langle -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, \sum_{j=1}^k \theta_j \int_0^T \Pi(T, s) \widehat{f}_j(s) \mathrm{d}s \right\rangle$$

$$+ \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \left\langle \partial_X \overline{l}(s), \int_0^s \Pi(s, \tau) \sum_{j=1}^k \theta_j \widehat{f}_j(\tau) \mathrm{d}\tau \right\rangle \mathrm{d}s.$$

By Fubini's formula ([6, Theorem 3.4.4]),

$$\int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \left\langle \partial_X \overline{l}(s), \int_0^s \Pi(s, \tau) \sum_{j=1}^k \theta_j \widehat{f}_j(\tau) \mathrm{d}\tau \right\rangle \mathrm{d}s$$

$$= \int_0^T \left\langle \int_s^T \frac{(T-\tau)^{\beta-1}}{\Gamma(\beta)} \Pi(\tau, s)^* \partial_X \overline{l}(\tau) \mathrm{d}\tau, \sum_{j=1}^k \theta_j \widehat{f}_j(s) \right\rangle \mathrm{d}s.$$

Then using (5.31) and Lemma B.3 in Appendix 6, we have

$$\left\langle -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, \sum_{j=1}^k \theta_j \int_0^T \Pi(T, s) \widehat{f}_j(s) \mathrm{d}s \right\rangle$$

$$+ \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \left\langle \partial_X \overline{l}(s), \int_0^s \Pi(s, \tau) \sum_{j=1}^k \theta_j \widehat{f}_j(\tau) \mathrm{d}\tau \right\rangle \mathrm{d}s$$

$$= \int_0^T \left\langle \Pi(T, s)^* \left(-\varphi_2 + \mathsf{p}^0 \partial_X \overline{m} \right) \right.$$

$$\left. + \mathsf{p}^0 \int_s^T \frac{(T-\tau)^{\beta-1}}{\Gamma(\beta)} \Pi(\tau, s)^* \partial_X \overline{l}(\tau) \mathrm{d}\tau, \sum_{k=1}^j \theta_j \widehat{f}_j(s) \right\rangle \mathrm{d}s$$

$$= \int_0^T \left\langle P(s), \sum_{j=1}^k \theta_j \widehat{f}_j(s) \right\rangle_{\mathsf{X} \times \mathsf{X}^*} \mathrm{d}s,$$

and similarly

$$\left\langle -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, \int_0^T \Pi(T, s)\left(-A + \partial_X \overline{f}(s)\right)a\,\mathrm{d}s \right\rangle$$

$$+ \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \left\langle \partial_X \bar{l}(s), \int_0^s \Pi(s, \tau)\left(-A + \partial_X \overline{f}(\tau)\right)a\,\mathrm{d}\tau \right\rangle \mathrm{d}s$$

$$= \int_0^T \left\langle \Pi(T, s)^*\left(-\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}\right) + \mathsf{p}^0 \int_s^T \frac{(T-\tau)^{\beta-1}}{\Gamma(\beta)} \Pi(\tau, s)^* \partial_X \bar{l}(\tau)\mathrm{d}\tau, \left(-A + \partial_X \overline{f}(s)\right)a \right\rangle \mathrm{d}s$$

$$= \int_0^T \left\langle P(s), \left(-A + \partial_X \overline{f}(s)\right)a \right\rangle_{\mathsf{X} \times \mathsf{X}^*} \mathrm{d}s.$$

Using the preceding analysis, (5.32) can be written as the following duality form:

$$\langle \varphi_1, (X'' - \overline{X}_0) \rangle + \langle \varphi_1, -a \rangle - \langle -\varphi_2, (X' - \overline{X}(T)) \rangle - \langle -\varphi_2, -Z_\omega(T) \rangle + \mathsf{p}^0 \widehat{Z}_\omega \qquad (5.33)$$

$$= \langle \varphi_1, (X'' - \overline{X}_0) \rangle - \langle -\varphi_2, (X' - \overline{X}(T)) \rangle + \langle -\varphi_1 + \mathsf{p}^0 \partial_{X_0} \overline{m}, -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m} a \rangle$$

$$+ \int_0^T \left\langle P(s), \left(-A + \partial_X \overline{f}(s)\right)a \right\rangle \mathrm{d}s + \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \langle \partial_X \bar{l}(s), a \rangle \mathrm{d}s$$

$$+ \int_0^T \left\langle P(s), \sum_{j=1}^k \theta_j \widehat{f}_j(s) \right\rangle \mathrm{d}s + \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \sum_{j=1}^k \theta_j \widehat{l}_j(s)\mathrm{d}s \leq 0.$$

We use (5.33) to derive the transversality condition as well as the Hamiltonian maximization condition in Theorem 3.1.

5.8. Proof of (iii): transversality condition

When $u = \overline{u}$, (5.33) becomes

$$\langle \varphi_1, (X'' - \overline{X}_0) \rangle - \langle -\varphi_2, (X' - \overline{X}(T)) \rangle + \langle -\varphi_1 + \mathsf{p}^0 \partial_{X_0} \overline{m} - \varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, a \rangle \qquad (5.34)$$

$$+ \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \langle \partial_X \bar{l}(s), a \rangle \mathrm{d}s + \int_0^T \left\langle P(s), \left(-A + \partial_X \overline{f}(s)\right)a \right\rangle \mathrm{d}s \leq 0.$$

Let us take

$$\begin{cases} \mathrm{I}_{T-}^{1-\alpha}[P](T) = -\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}, \\ \mathrm{I}_{T-}^{1-\alpha}[P](0) = -(-\varphi_1 + \mathsf{p}^0 \partial_{X_0} \overline{m}). \end{cases} \qquad (5.35)$$

Using the adjoint equation in (3.3),

$$\mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \langle \partial_X \bar{l}(s), a \rangle \mathrm{d}s + \int_0^T \left\langle P(s), \left(-A + \partial_X \overline{f}(s)\right)a \right\rangle \mathrm{d}s$$

$$= \int_0^T \left\langle {}^{\mathrm{RL}}\mathrm{D}_{T-}^{\alpha}[P](s), a \right\rangle \mathrm{d}s = \int_0^T \left\langle -\frac{\mathrm{d}}{\mathrm{d}s}\left[\mathrm{I}_{T-}^{1-\alpha}[P]\right](s), a \right\rangle \mathrm{d}s$$

$$= \left\langle \mathrm{I}_{T-}^{1-\alpha}[P](0) - \mathrm{I}_{T-}^{1-\alpha}[P](T), a \right\rangle = \left\langle \left(-(-\varphi_1 + \mathsf{p}^0 \partial_{X_0} \overline{m})\right) - \left(-\varphi_2 + \mathsf{p}^0 \partial_X \overline{m}\right), a \right\rangle.$$

Then (5.34) becomes for $(X'', X') \in \mathsf{F}_0 \times \mathsf{F}_T$,

$$\langle \mathrm{I}_{T-}^{1-\alpha}[P](0) - \mathsf{p}^0 \partial_{X_0} \overline{m}, (X'' - \overline{X}_0) \rangle - \langle \mathrm{I}_{T-}^{1-\alpha}[P](T) - \mathsf{p}^0 \partial_X \overline{m}, (X' - \overline{X}(T)) \rangle \leq 0. \qquad (5.36)$$

Therefore, (5.35) and (5.36) prove the transversality condition in Theorem 3.1.

5.9. Proof of (iv): Hamiltonian maximization condition

We prove the Hamiltonian maximization condition in Theorem 3.1. Notice that when $(X'', X') = (\overline{X}_0, \overline{X}(T))$ and $a = 0$, (5.33) becomes

$$\int_0^T \Big\langle P(s), \sum_{j=1}^k \theta_j \widehat{f}_j(s) \Big\rangle \mathrm{d}s + \mathsf{p}^0 \int_0^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \sum_{j=1}^k \theta_j \widehat{l}_j(s) \mathrm{d}s \le 0.$$

Take $k = 1$. The definition of the Hamiltonian in Theorem 3.1 implies

$$\int_0^T H(s, \overline{X}(s), P(s), \mathsf{p}^0, u(s)) \mathrm{d}s \le \int_0^T H(s, \overline{X}(s), P(s), \mathsf{p}^0, \overline{u}(s)).$$

Since U is separable, there exists a countable dense set $\mathsf{U}_0 = \{u_i, \ i \ge 1\} \subset \mathsf{U}$. Moreover, there exists a measurable set $S_i \subset [0, T]$ such that $|S_i| = T$ and any $t \in S_i$ is the Lebesgue point of $H(t, \overline{X}(t), P(t), \mathsf{p}^0, u(t))$ due to the fact that $H(\cdot, X, P, \mathsf{p}^0, u) \in L^1([0, T]; \mathbb{R})$, i.e., we have $\lim_{\tau \downarrow 0} \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} H(s, \overline{X}(s), P(s), \mathsf{p}^0, u(s)) \mathrm{d}s = H(t, \overline{X}(t), P(t), \mathsf{p}^0, u(t))$ [6, Theorem 5.6.2]. We fix $u_i \in \mathsf{U}_0$. For any $t \in S_i$, define

$$u(s) := \begin{cases} \overline{u}(s), & s \in [0, T] \setminus (t - \tau, t + \tau), \\ u_i, & s \in (t - \tau, t + \tau). \end{cases}$$

It then follows that

$$\lim_{\tau \downarrow 0} \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} H(s, \overline{X}(s), P(s), \mathsf{p}^0, u(s)) \mathrm{d}s$$
$$\le \lim_{\tau \downarrow 0} \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} H(s, \overline{X}(s), P(s), \mathsf{p}^0, \overline{u}(s)) \mathrm{d}s$$
$$\Rightarrow \ H(t, \overline{X}(t), P(t), \mathsf{p}^0, u(t)) \le H(t, \overline{X}(t), P(t), \mathsf{p}^0, \overline{u}(t)).$$

Since $\cap_{i \ge 1} S_i = [0, T]$, H is continuous in $u \in \mathsf{U}$, and U is separable, it follows that

$$\max_{u \in \mathsf{U}} H(t, \overline{X}(t), P(t), \mathsf{p}^0, u) = H(t, \overline{X}(t), P(t), \mathsf{p}^0, \overline{u}(t)), \text{ a.e. } t \in [0, T],$$

which shows the Hamiltonian maximization condition. This is the end of the proof of Theorem 3.1.

6. Conclusions

We state the following interesting potential future research problems of this paper:

- One important future research problem is to study the finite codimensionality in Assumption 3 for fractional evolution equations. This should require different techniques, compared with the nonfractional case studied in the earlier works (see Remark 3.2 and [5, 14, 15, 21, 26, 27, 29, 30, 35, 39, 40, 45, 46]). Indeed, it is important to study the equivalence between the finite codimensionality and the finite codimensional exactly controllability for fractional evolution equations as in the nonfractional case in [30, Theorem 3.2] and [29, Corollary 4.1].

- We may consider the applicability of the derived maximum principle to specific fractional evolution equations arising in different scientific and engineering applications, including various classes of fractional PDEs and differential equations with delay.
- It would be interesting to consider more general endpoint state constraints, such as time-dependent or nonlinear constraints, and to investigate the stability and sensitivity analysis of the fractional optimal control problem in the presence of uncertainties or disturbances.
- Finally, we plan to study the control problem in a distributed control framework and analyze the optimal control of fractional evolution equations on networks or graphs.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors would like to thank four anonymous reviewers for their careful reading and suggestions, which have helped to improve the paper. The authors particularly express thanks to the reviewer who pointed out the conditions in Theorem 2.1 and the important issue of the finite codimensionality in Section 4 of the earlier version of the manuscript.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. H. M. Ali, F. L. Pereira, S. M. A. Gamma, A new approach to the Pontryagin maximum principle for nonlinear fractional optimal control problems, *Math. Methods Appl. Sci.*, **39** (2016), 3640–3649. https://doi.org/10.1002/mma.3811

2. R. Almeida, R. Kamocki, A. B. Malinowska, T. Odzijewicz, On the necessary optimality conditions for the fractional Cucker-Smale optimal control problem, *Commun. Nonlinear Sci. Numer. Simul.*, **96** (2021), 105678. https://doi.org/10.1016/j.cnsns.2020.105678

3. E. G. Bajlekova, *Fractional evolution equations in Banach spaces*, Ph.D. Thesis, Technische Universiteit Eindhoven, 2001.

4. M. Bergounioux, L. Bourdin, Pontryagin maximum principle for general Caputo fractional optimal control problems with Bolza cost and terminal constraints, *ESAIM Control Optim. Calc. Var.*, **26** (2020), 35. https://doi.org/10.1051/cocv/2019021

5. M. Bergounioux, H. Zidani, Pontryagin maximum principle for optimal control of variational inequalities, *SIAM J. Control Optim.*, **37** (1999), 1273–1290. https://doi.org/10.1137/S0363012997328087

6. V. I. Bogachev, *Measure theory*, Springer, 2000. https://doi.org/10.1007/978-3-540-34514-5

7. L. Bourdin, Cauchy-Lipschitz theory for fractional multi-order dynamics-state-transition matrices, Duhamel formulas and duality theorems, *Differ. Integral Equ.*, **31** (2017), 559–594. https://doi.org/10.57262/die/1526004031

8. R. Chaudhary, S. Reich, Existence and controllability results for Hilfer fractional evolution equations via integral contractors, *Fract. Calc. Appl. Anal.*, **25** (2022), 2400–2419. https://doi.org/10.1007/s13540-022-00099-z

9. P. Chen, Y. Li, Q. Chen, B. Feng, On the initial value problem of fractional evolution equations with noncompact semigroup, *Comput. Math. Appl.*, **67** (2014), 1108–1115. https://doi.org/10.1016/j.camwa.2014.01.002

10. F. H. Clarke, *Optimization and nonsmooth analysis*, SIAM, 1990. https://doi.org/10.1137/1.9781611971309

11. K. Diethelm, *The analysis of fractional differential equations*, An Application-Oriented Exposition Using Differential Operators of Caputo Type, Springer, 2010. https://doi.org/10.1007/978-3-642-14574-2

12. X. L. Ding, I. Area, J. Nieto, Controlled singular evolution equations and Pontryagin type maximum principle with applications, *Evol. Equ. Control The.*, **11** (2022), 1655–1679. https://doi.org/10.3934/eect.2021059

13. G. Fabbri, F. Gozzi, A. Święch, *Stochastic optimal control in infinite dimension*, Dynamic Programming and HJB Equations, Springer, 2017. https://doi.org/10.1007/978-3-319-53067-3

14. H. O. Fattorini, A unified theory of necessary conditions for nonlinear nonconvex control systems, *Appl. Math. Optim.*, **15** (1987), 141–185. https://doi.org/10.1007/BF01442651

15. H. O. Fattorini, *Infinite dimensional optimization and control theory*, Cambridge University Press, 1999. https://doi.org/10.1017/CBO9780511574795

16. T. M. Flett, *Differential analysis*, Cambridge University Press, 1980. https://doi.org/10.1017/CBO9780511897191

17. H. Frankowska, E. M. Marchini, M. Mazzola, Necessary optimality conditions for infinite dimensional state constrained control problems, *J. Differ. Equations*, **264** (2018), 7294–7327. https://doi.org/10.1016/j.jde.2018.02.012

18. C. G. Gal, M. Warma, *Fractional-in-time semilinear parabolic equations and applications*, Springer, 2020. https://doi.org/10.1007/978-3-030-45043-4

19. M. Gomoyunov, On representation formulas for solutions of linear differential equations with Caputo fractional derivatives, *Fract. Calc. Appl. Anal.*, **23** (2020), 1141–1160. https://doi.org/10.1515/fca-2020-0058

20. H. Hassani, A. Avazzadeh, P. Agarwal, M. Javad Ebadi, A. Bayati Eshkaftaki, Generalized Bernoulli-Laguerre polynomials: applications in coupled nonlinear system of variable-order fractional PDEs, *J. Optim. Theory Appl.*, **200** (2023), 371–393. https://doi.org/10.1007/s10957-023-02346-6

21. B. Hu, J. Yong, Pontryagin maximum principle for semilinear and quasilinear parabolic equations with pointwise state constraints, *SIAM J. Control Optim.*, **33** (1995), 1857–1880. https://doi.org/10.1007/s10957-023-02346-6

22. R. Kamocki, On the existence of optimal solutions to fractional optimal control problems, *Appl. Math. Comput.*, **235** (2014), 94–104. https://doi.org/10.1016/j.amc.2014.02.086

23. R. Kamocki, Pontryagin maximum principle for fractional ordinary optimal control problems, *Math. Methods Appl. Sci.*, **37** (2014), 1668–1686. https://doi.org/10.1002/mma.2928

24. A. A. Kilbas, H. M. Srivastava, J. J. Trujillo, *Theory and applications of fractional differential equations*, Elsevier, 2006.

25. M. I. Krastanov, N. K. Ribarska, T. Y. Tsachev, A Pontryagin maximum principle for infinite dimensional problems, *SIAM J. Control Optim.*, **49** (2011), 2155–2182. https://doi.org/10.1137/100799009

26. X. Li, J. Yong, *Optimal control theory for infinite dimensional systems*, Birkhauser, 1995. https://doi.org/10.1007/978-1-4612-4260-4

27. X. Li, S. Chow, Maximum principle of optimal control for functional differential systems, *J. Optim. Theory Appl.*, **54** (1987), 335–360. https://doi.org/10.1007/BF00939438

28. P. Lin, J. Yong, Controlled singular Volterra integral equations and Pontryagin maximum principle, *SIAM J. Control Optim.*, **58** (2020), 136–164. https://doi.org/10.1137/19M124602X

29. X. Liu, Q. Lü, H. Zhang, X. Zhang, Finite codimensionality method for infinite-dimensional optimization problems, *arXiv*, 2022. https://arxiv.org/abs/2102.00652

30. X. Liu, Q. Lü, X. Zhang, Finite codimensional controllability and optimal control problems with endpoint state constraints, *J. Math. Pures Appl.*, **138** (2020), 164–203. https://doi.org/10.1016/j.matpur.2020.03.004

31. C. Lizama, Abstract nonlinear fractional evolution equations, In: A. Kochubei, Y. Luchko, *Handbook of fractional calculus with applications: volume 2 fractional differential equations*, Boston: De Gruyter, 2019, 499–514. https://doi.org/10.1515/9783110571660-022

32. F. Mainardi, *Fractional calculus and waves in linear viscoelasticity*, An Introduction to Mathematical Models, Imperial College Press, 2010.

33. J. Moon, A necessary optimality condition for optimal control of caputo fractional evolution equations, *IFAC-PapersOnLine*, **56** (2023), 7480–7485. https://doi.org/10.1016/j.ifacol.2023.10.1299

34. J. Moon, A Pontryagin maximum principle for terminal state-constrained optimal control problems of Volterra integral equations with singular kernels, *AIMS Math.*, **8** (2023), 22924–22943. https://doi.org/10.3934/math.20231166

35. M. McAsey, L. Mou, A proof of a general maximum principle for optimal controls via a multiplier rule on metric space, *J. Math. Anal. Appl.*, **337** (2008), 1072–1088. https://doi.org/10.1016/j.jmaa.2007.04.029

36. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, New York: Springer, 1983. https://doi.org/10.1007/978-1-4612-5561-1

37. M. Radmanesh, M. J. Ebadi, A local mesh-less collocation method for solving a class of time-dependent fractional integral equations: 2D fractional evolution equation, *Eng. Anal. Bound. Elem.*, **113** (2020), 372–381. https://doi.org/10.1016/j.enganabound.2020.01.017

38. C. S. Sin, H. C. In, K. C. Kim, Existence and uniqueness of mild solutions to initial value problems for fractional evolution equations, *Adv. Differ. Equ.*, **2018** (2018), 61. https://doi.org/10.1186/s13662-018-1519-9

39. C. Wang, S. Chen, Maximum principle for optimal control of some parabolic systems with two point boundary conditions, *Numer. Func. Anal. Optim.*, **20** (1999), 163–174. https://doi.org/10.1080/01630569908816886

40. G. Wang, L. Wang, State-constrained optimal control governed by non-well-posed parabolic differential equations, *SIAM J. Control Optim.*, **40** (2002), 1517–1539. https://doi.org/10.1137/S0363012900377006

41. J. Wang, Y. Zhou, A class of fractional evolution equations and optimal controls, *Nonlinear Anal.: Real World Appl.*, **12** (2011), 262–272. https://doi.org/10.1016/j.nonrwa.2010.06.013

42. J. Wang, Y Zhou, W. Wei, Optimal feedback control for semilinear fractional evolution equations in Banach spaces, *Syst. Control Lett.*, **61** (2012), 472–476. https://doi.org/10.1016/j.sysconle.2011.12.009

43. H. Ye, J. Gas, Y. Ding, A generalized Gronwall inequality and its application to a fractional differential equation, *J. Math. Anal. Appl.*, **328** (2007), 1075–1081. https://doi.org/10.1016/j.jmaa.2006.05.061

44. T. A. Yıldız, A. Jajarmi, B. Yıldız, D. Baleanu, New aspects of time fractional optimal control problems within operators with nonsingular kernel, *Discrete Cont. Dyn. Syst. (Ser. S)*, **13** (2020), 407–428. https://doi.org/10.3934/dcdss.2020023

45. J. Yong, Pontryagin maximum principle for semilinear second order elliptic partial differential equations and variational inequalities with state constraints, *Differ. Integral Equ.*, **5** (1992), 1307–1334. https://doi.org/10.57262/die/1370875549

46. X. Zhang, H. Li, C. Liu, Optimal control problem for the Cahn-Hilliard/Allen-Cahn equation with state constraint, *Appl. Math. Optim.*, **82** (2020), 721–754. https://doi.org/10.1007/s00245-018-9546-1

47. Y. Zhou, *Fractional evolution equations and inclusions: analysis and control*, Academic Press, 2016. https://doi.org/10.1016/C2015-0-00813-9

48. Y. Zhou, F. Jiao, Existence of mild solutions for fractional neutral evolution equations, *Comput. Math. Appl.*, **59** (2010), 1063–1077. https://doi.org/10.1016/j.camwa.2009.06.026

Appendix

Appendix A. Proof of Theorem 2.1

Lemma A.1. *[24, Lemma 2.3] and [3, page 10] For any $f(\cdot) \in L^1([0,T];\mathsf{X})$ and $\alpha, \beta > 0$, it holds that*

$$I_{0+}^{\alpha}[I_{0+}^{\beta}[f]](\cdot) = I_{0+}^{\alpha+\beta}[f](\cdot) = I_{0+}^{\beta+\alpha}[f](\cdot) = I_{0+}^{\beta}[I_{0+}^{\alpha}[f]](\cdot)$$

and

$$I_{T-}^{\alpha}[I_{T-}^{\beta}[f]](\cdot) = I_{T-}^{\alpha+\beta}[f](\cdot) = I_{T-}^{\beta+\alpha}[f](\cdot) = I_{T-}^{\beta}[I_{T-}^{\alpha}[f]](\cdot).$$

Lemma A.2. *[43, Theorem 1] Let $z(\cdot) \in L^1([0,T];\mathbb{R})$ and $b(\cdot) \in L^1([0,T];\mathbb{R})$ be nonnegative functions on $[0,T]$, and $h(\cdot) \in C([0,T];\mathbb{R})$ be nonnegative and nondecreasing. Assume that*

$$z(t) \le b(t) + h(t) \int_0^t (t-s)^{\alpha-1} z(s) \mathrm{d}s \ \ for \ \ t \in [0,T].$$

Then it holds that $z(t) \le b(t) + \int_0^t \sum_{k=1}^{\infty} \frac{(h(t)\Gamma(\alpha))^k}{\Gamma(k\alpha)} (t-s)^{k\alpha-1} b(s) \mathrm{d}s$ for $t \in [0,T]$. In addition, when b is nondecreasing, we have $z(t) \le b(t) E_{\alpha}[h(t)\Gamma(\alpha)t^{\alpha}]$, where $E_{\alpha}(t) := \sum_{k=0}^{\infty} \frac{t^k}{\Gamma(\alpha k+1)}$ is the Mittag-Leffler function.

Proof of Theorem 2.1. The existence and uniqueness of the (mild) solution to (2.1) under (v) of Assumption 1 follows from [48, Theorem 3.3] (see also [41, Theorem 3.1]), in which the analyticity of $-A$, the Laplace transformation, and the contraction mapping theorem are key ingredients in the proof. The integral expression in (2.2) can be obtained from Definition 1 and the semigroup property of fractional integral in Lemma A.1. The last two estimates can be shown using Assumption 1 and Lemma A.2. In particular, by Assumption 1 (including the compactness of $(\mathcal{T}(t))_{t\ge 0}$), we can show that

$$|X(t) - X'(t)|_{\mathsf{X}} \le M|X_0 - X'_0|_{\mathsf{X}} + M \int_0^t (t-s)^{\alpha-1} |X(s) - X'(s)|_{\mathsf{X}} \mathrm{d}s.$$

Then we apply Lemma A.2 to get the first one. A similar approach can be applied to show the second estimate. This completes the proof. □

Appendix B. Representation of linear fractional evolution equations

This appendix provides the representation results on linear fractional evolution equations. We provide the complete proofs, which are omitted in [33].

Lemma B.3. *Let $\Pi : [0,T] \times [0,T] \to \mathcal{L}(\mathsf{X})$ be the left RL fractional state-transition evolution operator defined by*

$$\Pi(t,s)x = \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} Ix + \int_s^t \frac{(t-r)^{\alpha-1}}{\Gamma(\alpha)} \Big(-A + \partial_X \overline{f}(r)\Big) \Pi(r,s)x \mathrm{d}r. \tag{B.1}$$

Then (B.1) can be written as the following right RL fractional state-transition evolution operator form:

$$\Pi(t,s)x = \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} Ix + \int_s^t \frac{(r-s)^{\alpha-1}}{\Gamma(\alpha)} \Pi(t,r)\Big(-A + \partial_X \overline{f}(r)\Big)x \mathrm{d}r. \tag{B.2}$$

Proof. By the right-hand-side of (B.2), let

$$\widehat{\Pi}(t, \tau) := \frac{(t - \tau)^{\alpha-1}}{\Gamma(\alpha)} I + \int_{\tau}^{t} \frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)} \Pi(t, r)\big(-A + \partial_X \overline{f}(r)\big) dr. \tag{B.3}$$

Then we need to show the equivalence of (B.1) and (B.3). Hence, we prove the following result:

$$\int_{\tau}^{t} \frac{(t - r)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big)\widehat{\Pi}(r, \tau) dr = \int_{\tau}^{t} \frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)} \Pi(t, r)\big(-A + \partial_X \overline{f}(r)\big) dr. \tag{B.4}$$

Clearly, (B.4) holds when $t = \tau$. Then (B.4) follows by Fubini's formula (see [6, Theorem 3.4.4]), since we can show that

$$\int_{\tau}^{t} \frac{(t - r)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big)\widehat{\Pi}(r, \tau) dr$$

$$= \int_{\tau}^{t} \frac{(t - r)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big)\frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)} I_n dr$$

$$+ \int_{\tau}^{t} \frac{(t - r)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big) \int_{\tau}^{r} \frac{(v - \tau)^{\alpha-1}}{\Gamma(\alpha)} \Pi(r, v) F(v) dv dr$$

$$= \int_{\tau}^{t} \frac{(t - r)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big)\frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)} I dr$$

$$+ \int_{\tau}^{t} \int_{r}^{t} \frac{(t - v)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(v)\big)\Pi(v, r) dv \frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big) dr$$

$$= \int_{\tau}^{t} \frac{(t - r)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big)\frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)} I dr$$

$$+ \int_{\tau}^{t} \Big(\Pi(t, r) - \frac{(t - r)^{\alpha-1}}{\Gamma(\alpha)} I\Big)\frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)}\big(-A + \partial_X \overline{f}(r)\big) dr$$

$$= \int_{\tau}^{t} \frac{(r - \tau)^{\alpha-1}}{\Gamma(\alpha)} \Pi(t, r)\big(-A + \partial_X \overline{f}(r)\big) dr.$$

This completes the proof. □

Remark B.1. *Notice that by Definition 1, $I_{s+}^{1-\alpha}[\Pi](s, s) = I$ and $I_{t-}^{1-\alpha}[\Pi](t, t) = I$ for $s, t \in [0, T]$ with $s \leq t$. Hence, we can see that (B.1) is the left RL fractional state-transition evolution operator with the initial condition, whereas (B.2) is the right RL fractional state-transition evolution operator with the terminal condition.*

The next result states the representation result for linear forward Caputo fractional evolution equations. The finite-dimensional case was reported in [7, 19].

Lemma B.4. *The (mild) solution to the variational equation in (5.23) is as follows:*

$$Z_\omega(t) = a + \int_{0}^{t} \Pi(t, s)\big(-A + \partial_X \overline{f}(s)\big) a ds + \sum_{j=1}^{k} \theta_j \int_{0}^{t} \Pi(t, s)\widehat{f_j}(s) ds, \ t \in [0, T], \tag{B.5}$$

where Π is the X-valued left RL fractional state-transition evolution operator in (B.1).

Proof. Since (5.23) is a linear fractional evolution equation, the existence and uniqueness of the solution follows from Theorem 2.1 under (v) of Assumption 1.

We now prove (B.5). By Theorem 2.1, it follows that

$$Z_\omega(t) = a + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\left(-A + \partial_X\overline{f}(s)\right)Z_\omega(s)\mathrm{d}s + \sum_{j=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\widehat{f}_j(s)\mathrm{d}s.$$

Hence, to prove (B.5), we have to show the following equivalence:

$$\int_0^t \Pi(t,s)\left(-A + \partial_X\overline{f}(s)\right)a\mathrm{d}s + \sum_{j=1}^k \theta_j \int_0^t \Pi(t,s)\widehat{f}_j(s)\mathrm{d}s \tag{B.6}$$

$$= \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\left(-A + \partial_X\overline{f}(s)\right)Z_\omega(s)\mathrm{d}s + \sum_{j=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\widehat{f}_j(s)\mathrm{d}s.$$

Clearly, it holds when $t = 0$. Then by Fubini's formula,

$$\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\left[\left(-A + \partial_X\overline{f}(s)\right)Z_\omega(s) + \sum_{j=1}^k \theta_j\widehat{f}_j(s)\right]\mathrm{d}s$$

$$= \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\left(-A + \partial_X\overline{f}(s)\right)a\mathrm{d}s + \sum_{j=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\widehat{f}_j(s)\mathrm{d}s$$

$$+ \int_0^t \int_s^t \frac{(t-\tau)^{\alpha-1}}{\Gamma(\alpha)}\left(-A + \partial_X\overline{f}(\tau)\right)\Pi(\tau,s)\mathrm{d}\tau\left(-A + \partial_X\overline{f}(s)\right)a\mathrm{d}s$$

$$+ \int_0^t \int_s^t \frac{(t-\tau)^{\alpha-1}}{\Gamma(\alpha)}\left(-A + \partial_X\overline{f}(\tau)\right)\Pi(\tau,s)\mathrm{d}\tau \sum_{k=1}^j \theta_j\widehat{f}_j(s)\mathrm{d}s.$$

Rearranging the right-hand side of the above expression yields

$$\int_0^t \left[\frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\mathrm{I} + \int_s^t \frac{(t-\tau)^{\alpha-1}}{\Gamma(\alpha)}\left(-A + \partial_X\overline{f}(\tau)\right)\Pi(\tau,s)\mathrm{d}\tau\right]\left(-A + \partial_X\overline{f}(s)\right)a\mathrm{d}s$$

$$+ \int_0^t \left[\frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\mathrm{I} + \int_s^t \frac{(t-\tau)^{\alpha-1}}{\Gamma(\alpha)}\left(-A + \partial_X\overline{f}(\tau)\right)\Pi(\tau,s)\mathrm{d}\tau\right]\sum_{j=1}^k \theta_j\widehat{f}_j(s)\mathrm{d}s$$

$$= \int_0^t \Pi(t,s)\left(-A + \partial_X\overline{f}(s)\right)a\mathrm{d}s + \int_0^t \Pi(t,s)\sum_{j=1}^k \theta_j\widehat{f}_j(s)\mathrm{d}s.$$

Clearly, (B.6) holds and the conclusion follows. We complete the proof. $\qquad\square$

The following result shows the backward representation of the adjoint equation, which has not been reported in the existing literature.

Lemma B.5. *The (mild) solution to the adjoint equation in (3.3) can be written as follows:*

$$\begin{cases} P(t) = \Pi(T,t)^*P(T) + \mathsf{p}^0 \int_t^T \frac{(T-s)^{\beta-1}}{\Gamma(\beta)}\Pi(s,t)^*\partial_X\overline{l}(s)\mathrm{d}s, \ t \in [0,T), \\ I_{T-}^{1-\alpha}[P](T) = P(T), \end{cases} \tag{B.7}$$

where Π is the X-valued right RL fractional state-transition evolution operator in (B.2).

Proof. Notice that by definition of Fréchet differentiation, we have $\partial_X \bar{l}(\cdot) \in X^* = \mathcal{L}(X, \mathbb{R})$ [16, page 166]. Then by definition of Π in (B.2), together with Assumptions 1 and 2, (B.7) is a well-defined RL fractional integral equation. We can easily observe that P is the X^*-valued function and $P(\cdot) \in L^1([0, T]; X^*)$ [24, Theorem 3.3, Chapter 3]. Hence, if (B.7) is the integral representation of (3.3), then we prove (ii) of Theorem 3.1.

To show (B.7), we obtain (3.3) from (B.7). Using (B.2) and with $W(s) := \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \partial_X \bar{l}(s)$,

$$
\begin{aligned}
P(t) = {} & \frac{(T-t)^{\alpha-1}}{\Gamma(\alpha)} P(T) + \mathsf{p}^0 \int_t^T \frac{(s-t)^{\alpha-1}}{\Gamma(\alpha)} W(s) \mathrm{d}s \\
& + \int_t^T \frac{(r-t)^{\alpha-1}}{\Gamma(\alpha)} \left(-A^* + \partial_X \bar{f}(r)^* \right) \Pi(T, r)^* P(T) \mathrm{d}r \\
& + \mathsf{p}^0 \int_t^T \int_t^s \frac{(r-t)^{\alpha-1}}{\Gamma(\alpha)} \left(-A^* + \partial_X \bar{f}(r)^* \right) \Pi(s, r)^* \mathrm{d}r W(s) \mathrm{d}s.
\end{aligned}
$$

By Fubini's formula,

$$
\begin{aligned}
& \int_t^T \int_t^s \frac{(r-t)^{\alpha-1}}{\Gamma(\alpha)} \left(-A^* + \partial_X \bar{f}(r)^* \right) \Pi(s, r)^* \mathrm{d}r W(s) \mathrm{d}s \\
& = \int_t^T \frac{(s-t)^{\alpha-1}}{\Gamma(\alpha)} \left(-A^* + \partial_X \bar{f}(r)^* \right) \int_s^T \Pi(r, s)^* W(r) \mathrm{d}r \mathrm{d}s.
\end{aligned}
$$

Hence, from Definitions 1 and 2, it follows that

$$
\begin{aligned}
P(t) = {} & \frac{(T-t)^{\alpha-1}}{\Gamma(\alpha)} P(T) + \mathsf{p}^0 \int_t^T \frac{(s-t)^{\alpha-1}}{\Gamma(\alpha)} W(s) \mathrm{d}s + \int_t^T \frac{(s-t)^{\alpha-1}}{\Gamma(\alpha)} (-A^* + \partial_X \bar{f}(r)^*) P(s) \mathrm{d}s \\
= {} & {}^{\mathrm{RL}}\mathrm{D}_{T-}^{1-\alpha}[P(T)](t) + \mathrm{I}_{T-}^\alpha [-A^* P(\cdot) + \partial_X \bar{f}(\cdot)^* P(\cdot) + \mathsf{p}^0 W(\cdot)](t).
\end{aligned}
$$

Definitions 1 and 2 and Lemma A.1 lead to

$$
\begin{aligned}
-\mathrm{I}_{T-}^{1-\alpha}[P](t) = {} & -\mathrm{I}_{T-}^{1-\alpha}\Big[-\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{I}_{T-}^\alpha[P(T)]](\cdot)\Big](t) - \mathrm{I}_{T-}^1 [-A^* P(\cdot) + \partial_X \bar{f}(\cdot)^* P(\cdot) + W(\cdot)](t) \\
= {} & -P(T) - \mathrm{I}_{T-}^1 [-A^* P(\cdot) + \partial_X \bar{f}(\cdot)^* P(\cdot) + \mathsf{p}^0 W(\cdot)](t).
\end{aligned}
$$

Note that $\mathrm{I}_{T-}^{1-\alpha}[P](T) = P(T)$. This, together with Definition 2 and the fact that $\mathrm{I}_{T-}^{1-\alpha}[P](\cdot) \in AC([0, T]; X^*)$ by our assumptions, implies (with $W(s) := \frac{(T-s)^{\beta-1}}{\Gamma(\beta)} \partial_X \bar{l}(s)$)

$$
\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}\Big[\mathrm{I}_{T-}^{1-\alpha}[P](\cdot)\Big](t) = {}^{\mathrm{RL}}\mathrm{D}_{T-}^{1-\alpha}[P](t) = {} & -\frac{\mathrm{d}}{\mathrm{d}t}\Big[\mathrm{I}_{T-}^1 [-A^* P(\cdot) + \partial_X \bar{f}(\cdot)^* P(\cdot) + \mathsf{p}^0 W(\cdot)]\Big](t) \\
= {} & -A^* P(t) + \partial_X \bar{f}(t)^* P(t) + \mathsf{p}^0 W(t).
\end{aligned}
$$

This shows that (B.7) is the solution to the adjoint equation in (3.3). This completes the proof. $\quad\square$

Appendix C. Technical lemmas for Section 5

In this appendix, we provide two technical lemmas used in duality and variational analysis.

Lemma C.6. *Let $h_j(\cdot) \in L^1([0,T];\mathsf{X})$, $j = 1,\ldots,j$, and $\sum_{j=1}^k \theta_j = 1$ with $\theta_j \geq 0$. Then for any $\delta \in (0,1)$, there exist measurable $E_j^\delta \subset [0,T]$, $j = 1,\ldots,k$, such that*

$$
\begin{cases}
E_j^\delta \cap E_i^\delta = \emptyset, \ j \neq i, \\
\cup_{j=1}^k meas|E_j^\delta| = \delta T, \\
\displaystyle\sup_{t\in[0,T]} \Big| \delta \sum_{k=1}^k \theta_j \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_j(s)\mathrm{d}s - \sum_{j=1}^k \int_{E_j^\delta \cap [0,t]} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h_j(s)\mathrm{d}s \Big|_{\mathsf{X}} = o(\delta).
\end{cases} \tag{C.1}
$$

Proof. Notice that for each $t \in [0,T]$, $\frac{(t-\cdot)^{\alpha-1}}{\Gamma(\alpha)} h_j(\cdot)$ is Bochner integrable. Then the result follows from [27, Lemma 5.5]; completing the proof. □

Lemma C.7. *For any $\epsilon > 0$, $\{(a, Z_\omega^\epsilon(T), \widehat{Z}_\omega^\epsilon) \mid \omega \in \Omega\} \subset \mathsf{X} \times \mathsf{X} \times \mathbb{R}$ is convex.*

Proof. Notice that by Theorem 2.1, $(Z_\omega^\epsilon(\cdot), \widehat{Z}_\omega^\epsilon)$ admit unique solutions in $C([0,T];\mathsf{X}) \times \mathbb{R}$ for any $(t,a,\omega) \in [0,T] \times \mathsf{X} \times \Omega$. For $i = 1,\ldots,l$, consider $\omega_k^{(i)} := (\{u_j^{(i)}(\cdot)\}_{j=1}^k, \{\theta_j^{(i)}\}_{j=1}^k) \in \Omega$. Let $(Z_\omega^{\epsilon,(i)}(\cdot), \widehat{Z}_\omega^{\epsilon,(i)})$ be the solutions of (5.7) and (5.8) under $\omega_k^{(i)}$. Let $a^{(i)} \in \mathsf{X}$. With $\gamma^{(i)} \in [0,1]$ and $\sum_{i=1}^l \gamma^{(i)} = 1$, we define

$$
a = \sum_{i=1}^l \gamma^{(i)} a^{(i)}, \ Z_\omega^\epsilon(T) = \sum_{i=1}^l \gamma^{(i)} Z_\omega^{\epsilon,(i)}(T), \ \widehat{Z}_\omega^\epsilon = \sum_{i=1}^l \gamma^{(i)} \widehat{Z}_\omega^{\epsilon,(i)}. \tag{C.2}
$$

By defining

$$
\omega_k := (\{\gamma^{(i)} \omega_k^{(i)}\}_{i=1}^l) = (\{\{u_j^{(i)}(\cdot)\}_{j=1}^k, \{\gamma^{(i)}\theta_j^{(i)}\}_{j=1}^k\}_{i=1}^l) \in \Omega,
$$

we observe that (C.2) are well-defined solutions and $a \in \mathsf{X}$. This completes the proof. □