Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 56-1 (2023) 216–221

On the Optimality Condition for Optimal Control of Caputo Fractional Differential Equations with State Constraints ⋆

Jun Moon *

* *Department of Electrical Engineering, Hanyang University, Seoul 04763, South Korea (email: junmoon@hanyang.ac.kr).*

Abstract: We consider the fractional optimal control problem with state constraints. The fractional calculus of derivatives and integrals can be viewed as generalizations of their classical ones to any arbitrary real order. In our problem setup, the dynamic constraint is captured by the Caputo fractional differential equation with order $\alpha \in (0, 1)$, and the objective functional is formulated by the left Riemann-Liouville fractional integral with order $\beta \geq 1$. In addition, there are terminal and running state constraints; while the former is described by initial and final states within a convex set, the latter is given by an explicit instantaneous inequality state constraint. We obtain the maximum principle, the first-order necessary optimality condition, for the problem of this paper. Due to the inherent complex nature of the fractional control problem, the presence of the terminal and running state constraints, and the generalized standing assumptions, the maximum principle of this paper is new in the optimal control problem context, and its proof requires to develop new variational and duality analysis using fractional calculus and functional analysis, together with the Ekeland variational principle and the spike variation.

Keywords: Fractional calculus, fractional differential equations, maximum principle, variational analysis, Ekeland variational principle.

1. INTRODUCTION

Fractional calculus of derivatives and integrals can be viewed as generalizations of their classical notions to any real arbitrary order. One important application of fractional calculus is a class of *fractional differential equations*, which enables us to describe more general and extraordinary phenomena observed in real world. Various types of fractional differential equations and their applications have been studied in applied mathematics, science, engineering, and economics; see (Kilbas et al., 2006; Diethelm, 2010; Malinowska et al., 2015) and the references therein.

Along with fractional differential equations, fractional optimal control problems in the sense of Riemann-Liouville (RL) and/or Caputo, and their numerical methods and applications have been studied extensively in the literature under various different formulations. The reader is referred to see (Gomoyunov, 2022; Liu et al., 2022; Gong et al., 2021; Rahimkhani and Ordokhani, 2021; Gomoyunov, 2020) and the references therein.

The maximum principle, i.e., the first-order necessary optimality condition, for the fractional optimal control problems was studied in several different directions; see (Jelicic and Petrovacki, 2008; Kamocki, 2014; Pooseh et al., 2014; Ali et al., 2016; Bergounioux and Bourdin, 2020; Almeida et al., 2021; Yusubov and Mahmudov, 2021) and the references therein. Specifically, in (Kamocki, 2014), a simple convex variation was applied to obtain the maximum principle for the RL fractional optimal control problem. The Caputo fractional optimal control problem without state constraints was studied in (Pooseh et al., 2014), where the sufficient condition was also obtained under the convexity assumption. In (Ali et al., 2016; Almeida et al., 2021), the Caputo and Cucker-Smale multi-agent fractional control problems without state constraints were studied. In addition, (Yusubov and Mahmudov, 2021) studied the singular fractional optimal control problem. Recently, the Caputo fractional optimal control problem with the terminal state constraint only was considered in (Bergounioux and Bourdin, 2020).

In this paper, we consider the fractional optimal control problem with terminal and running state constraints. The precise problem statement is given in Section 3. In our problem setup, the dynamic constraint is captured by the fractional differential equation in the sense of Caputo with order $\alpha \in (0, 1)$, and the objective functional is formulated by the left Riemann-Liouville (RL) fractional integral with order $\beta \geq 1$. In addition, there are terminal and running state constraints; while the former is described by initial and final states within a convex set, the latter is expressed by an explicit instantaneous inequality state constraint.

The main result of this paper is the maximum principle (see Theorem 2). In the proof (see Section 5), we need to

⋆ This work was supported in part by the Technology Innovation Program (20018112) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea), in part by the National Research Foundation of Korea (NRF) Grant funded by the Ministry of Science and ICT, South Korea (NRF-2021R1A2C2094350), and in part by Institute of Information & communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2020-0-01373).

10.1016/j.ifacol.2023.02.037

formulate the penalized unconstrained fractional control problem, for which the Ekeland variational principle is applied. Then to develop the variational and duality analysis via the spike variation, we have to use the fractional calculus and the intrinsic properties of distance functions. Moreover, the proofs for the complementary slackness and the transversality condition are essentially required due to the presence of the terminal and running state constraints.

The paper is organized as follows. Preliminaries on fractional calculus are given in Section 2. We formulate the problem in Section 3. The maximum principle of this paper is stated in Section 4, and its proof sketch is provided in Section 5. We conclude our paper in Section 6.

2. PRELIMINARIES ON FRACTIONAL CALCULUS

In this section, we provide some preliminary results on fractional calculus. More detailed results on fractional calculus can be found in (Kilbas et al., 2006; Diethelm, 2010).

Let $\mathbb{R}^n$ be the n-dimensional Euclidean space. For $A \in \mathbb{R}^{m\times n}$, $A^\top$ denotes the transpose of A. Let $\langle x, y\rangle := x^\top y$ and $|x| := \langle x, x\rangle^{1/2}$ be the norm. Let I_n be an $n\times n$ identity matrix. Define $\mathbb{1}_A(\cdot)$ by the indicator function of any set A. Let Γ be the Gamma function.

Define the following spaces for $t_0, t_f \in [0, T]$ with $t_0 < t_f$:

- $L^p([t_0, t_f]; \mathbb{R}^n)$, $p \geq 1$: the space of functions $\psi : [t_0, t_f] \to \mathbb{R}^n$ such that ψ is measurable and $\|\psi(\cdot)\|_{L^{p,n}} := (\int_{t_0}^{t_f} |\psi(t)|_{\mathbb{R}^n}^p \mathrm{d}t)^{\frac{1}{p}} < \infty$;
- $L^\infty([t_0, t_f]; \mathbb{R}^n)$: the space of functions $\psi : [t_0, t_f] \to \mathbb{R}^n$ such that ψ is measurable and $\|\psi(\cdot)\|_{L^{\infty,n}} := \operatorname{ess\,sup}_{t\in[t_0,t_f]} |\psi(t)| < \infty$;
- $C([t_0, t_f]; \mathbb{R}^n)$: the space of functions $\psi : [t_0, t_f] \to \mathbb{R}^n$ such that ψ is continuous and $\|\psi(\cdot)\|_\infty := \sup_{t\in[t_0,t_f]} |\psi(t)| < \infty$;
- $\mathrm{AC}([t_0, t_f]; \mathbb{R}^n)$: the space of functions $\psi : [t_0, t_f] \to \mathbb{R}^n$ such that ψ is absolutely continuous;
- $\mathrm{BV}([t_0, t_f]; \mathbb{R}^n)$: the space of functions $\psi : [t_0, t_f] \to \mathbb{R}^n$ such that ψ is of bounded variation on $[t_0, t_f]$.

The norm on $\mathrm{BV}([t_0, t_f]; \mathbb{R}^n)$ is defined by $\|\psi(\cdot)\|_{\mathrm{BV}^n} := \psi(t_0) + \mathrm{TV}(\psi)$, where $\mathrm{TV}(\psi) := \sup_{(t_k)_k} \{\sum_k |\psi(t_{k+1}) - \psi(t_k)|\} < \infty$ with the supremum being taken by all partitions of $[t_0, t_f]$. Let $\mathrm{NBV}([t_0, t_f]; \mathbb{R}^n)$ be the space of functions $\psi(\cdot) \in \mathrm{BV}([t_0, t_f]; \mathbb{R}^n)$ such that ψ is normalized, i.e., $\psi(t_0) = 0$ and ψ is left continuous. The norm on $\mathrm{NBV}([t_0, t_f]; \mathbb{R}^n)$ is defined by $\|\psi(\cdot)\|_{\mathrm{NBV}^n} := \mathrm{TV}(\psi)$. When $\psi(\cdot) \in \mathrm{NBV}([t_0, t_f]; \mathbb{R})$ is monotonically nondecreasing, $\|\psi(\cdot)\|_{\mathrm{NBV}} := \|\psi(\cdot)\|_{\mathrm{NBV}^1} = \psi(t_f)$.

Definition 1. For $f(\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$ and $t \in [t_0, t_f]$, the left Riemann-Liouville (RL) fractional integral $\mathrm{I}^\alpha_{t_0+}[f]$ of order $\alpha > 0$ is defined by

$$\mathrm{I}^\alpha_{t_0+}[f](t) := \int_{t_0}^{t} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} f(s)\mathrm{d}s.$$

For $f(\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$ and $t \in [t_0, t_f]$, the right RL fractional integral $\mathrm{I}^\alpha_{t_f-}[f]$ of order $\alpha > 0$ is defined by

$$\mathrm{I}^\alpha_{t_f-}[f](t) := \int_{t}^{t_f} \frac{(s-t)^{\alpha-1}}{\Gamma(\alpha)} f(s)\mathrm{d}s.$$

For $\alpha = 0$, we set $\mathrm{I}^0_{t_0+}[f](\cdot) := \mathrm{I}^0_{t_f-}[f](\cdot) := f(\cdot)$. □

Definition 2. (i) For $f(\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$, the left RL fractional derivative $\mathrm{D}^\alpha_{t_0+}[f]$ of order $\alpha \in (0, 1)$ is defined by

$$\mathrm{D}^\alpha_{t_0+}[f](t) := \frac{\mathrm{d}}{\mathrm{d}t}\Big[\mathrm{I}^{1-\alpha}_{t_0+}[f]\Big](t),$$

provided that $\mathrm{I}^{1-\alpha}_{t_0+}[f](\cdot) \in \mathrm{AC}([t_0, t_f]; \mathbb{R}^n)$. In this case, $\mathrm{D}^\alpha_{t_0+}[f](\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$. Let ${}^{\mathrm{RL}}_{\ n}\mathcal{D}^\alpha_{t_0+}$ be the set of functions $f(\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$ such that f admits the left RL fractional derivative of order $\alpha \in (0, 1)$;

(ii) For $f(\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$, the right RL fractional derivative $\mathrm{D}^\alpha_{t_f-}[f]$ of order $\alpha \in (0, 1)$ is defined by

$$\mathrm{D}^\alpha_{t_f-}[f](t) := -\frac{\mathrm{d}}{\mathrm{d}t}\Big[\mathrm{I}^{1-\alpha}_{t_f-}[f]\Big](t),$$

provided that $\mathrm{I}^{1-\alpha}_{t_f-}[f](\cdot) \in \mathrm{AC}([t_0, t_f]; \mathbb{R}^n)$. In this case, $\mathrm{D}^\alpha_{t_f-}[f](\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$. Let ${}^{\mathrm{RL}}_{\ n}\mathcal{D}^\alpha_{t_f-}$ be the set of functions $f(\cdot) \in L^1([t_0, t_f]; \mathbb{R}^n)$ such that f admits the right RL fractional derivative of order $\alpha \in (0, 1)$. □

Definition 3. (i) For $f(\cdot) \in C([t_0, t_f]; \mathbb{R}^n)$, the left Caputo fractional derivative ${}^{\mathrm{C}}\mathrm{D}^\alpha_{t_0+}[f]$ of order $\alpha \in (0, 1)$ is defined by

$${}^{\mathrm{C}}\mathrm{D}^\alpha_{t_0+}[f](t) := \mathrm{D}^\alpha_{t_0+}[f(\cdot) - f(t_0)](t),$$

where $f(\cdot) - f(t_0) \in {}^{\mathrm{RL}}_{\ n}\mathcal{D}^\alpha_{t_0+}$. In this case, ${}^{\mathrm{C}}\mathrm{D}^\alpha_{t_0+}[f] \in L^1([t_0, t_f]; \mathbb{R}^n)$. Let ${}^{\mathrm{C}}_{n}\mathcal{D}^\alpha_{t_0+}$ be the set of functions $f \in C([t_0, t_f]; \mathbb{R}^n)$ such that f admits the left Caputo fractional derivative of order $\alpha \in (0, 1)$.

(ii) For $f(\cdot) \in C([t_0, t_f]; \mathbb{R}^n)$, the right Caputo fractional derivative ${}^{\mathrm{C}}\mathrm{D}^\alpha_{t_f-}[f]$ of order $\alpha \in (0, 1)$ is defined by

$${}^{\mathrm{C}}\mathrm{D}^\alpha_{t_f-}[f](t) := \mathrm{D}^\alpha_{t_f-}[f(\cdot) - f(t_f)](t),$$

where $f(\cdot) - f(t_f) \in {}^{\mathrm{RL}}_{\ n}\mathcal{D}^\alpha_{t_f-}$. In this case, ${}^{\mathrm{C}}\mathrm{D}^\alpha_{t_f-}[f] \in L^1([t_0, t_f]; \mathbb{R}^n)$. Let ${}^{\mathrm{C}}_{n}\mathcal{D}^\alpha_{t_f-}$ be the set of functions $f \in C([t_0, t_f]; \mathbb{R}^n)$ such that f admits the right Caputo fractional derivative of order $\alpha \in (0, 1)$. □

3. FRACTIONAL OPTIMAL CONTROL PROBLEM

Consider the following $\mathbb{R}^n$-valued left Caputo fractional differential equation with order $\alpha \in (0, 1)$:

$$\begin{cases} {}^{\mathrm{C}}\mathrm{D}^\alpha_{t_0+}[X](t) = f(t, X(t), u(t)),\ t \in (t_0, t_f], \\ X(t_0) = X_0 \in \mathbb{R}^n, \end{cases} \tag{1}$$

where $X(\cdot) \in \mathbb{R}^n$ is the state, $u : [t_0, t_f] \to U \subset \mathbb{R}^r$ is the control input with U being the control space, and $f : [t_0, t_f] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ is the driver of (1). Let $\mathcal{U} := \{u : [t_0, t_f] \to U \mid u \text{ is measurable } t \in [t_0, t_f]\}$ be the set of admissible controls for (1).

Assumption 1. (i) $U \subset \mathbb{R}^r$ is a separable metric space;

(ii) f holds that $t \mapsto f(t, X, u)$ is continuous, $f(\cdot, X, u) \in L^\infty([t_0, t_f]; \mathbb{R}^n)$, and $(X, u) \mapsto f(t, X, u)$ is Lipschitz continuous, and $|f(t, 0, u)| \leq L(1 + |X|)$;

(iii) $X \mapsto f(t, X, u)$ is continuously differentiable with $(t, X, u) \mapsto \partial_X f(t, X, u)$ being bounded and $(X, u) \mapsto \partial_X f(t, X, u)$ being Lipschitz continuous. □

Theorem 1. Suppose that Assumption 1 holds. Then for any $(X_0, u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$, (1) has a unique solution of $X(\cdot) \in$

${}^{\mathrm{C}}_{n}\mathcal{D}^{\alpha}_{t_0+}$. In addition, for any $(X_0, u(\cdot)), (X_0', u'(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$ (with $X(\cdot; X_0, u) := X(\cdot)$), we have

$$\sup_{t\in[t_0,t_f]} |X(t;X_0,u) - X'(t;X_0',u')| \leq b(t_f) + \int_{t_0}^{t_f} \sum_{k=1}^{\infty} \frac{(L\Gamma(\alpha))^k}{\Gamma(k\alpha)} (t_f - s)^{k\alpha-1} b(s)\mathrm{d}s$$

$$\sup_{t\in[t_0,t_f]} |X(t;X_0,u)| \leq b'(t_f) + \int_{t_0}^{t_f} \sum_{k=1}^{\infty} \frac{(L\Gamma(\alpha))^k}{\Gamma(k\alpha)} (t_f - s)^{k\alpha-1} b'(s)\mathrm{d}s,$$

where $b(t) := |X_0 - X_0'| + L\int_{t_0}^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} |(u(s) - u'(s)|\mathrm{d}s$ and $b'(t) := |X_0| + L\int_{t_0}^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}\mathrm{d}s$.

Proof. The existence and uniqueness of the solution can be shown by using the contraction mapping theorem under the Bielecki norm on $C([t_0,t_f];\mathbb{R}^n)$. The estimates follow from Assumption 1 and the fractional Gronwall's inequality. Note that the detailed proof is omitted due to space restriction. We complete the proof. □

The objective functional is given by the following left RL fractional integral with order $\beta \geq 1$:

$$J(X_0; u(\cdot)) = I^{\beta}_{t_0+}[l(\cdot, X(\cdot), u(\cdot)](t_f) + m(X_0, X(t_f)).$$

The fractional optimal control problem of this paper is

$$\textbf{(P)} \quad \inf_{u(\cdot)\in\mathcal{U}} J(X_0; u(\cdot))$$

subject to the terminal and running state constraints given by

$$\begin{cases} (X_0, X(t_f)) \in F \subset \mathbb{R}^{2n}, \\ G_i(t, X(t)) \leq 0, \ \forall t \in [t_0, t_f], \ i = 1, \ldots, q. \end{cases} \tag{2}$$

Assumption 2. (i) $l : [t_0,t_f]\times\mathbb{R}^n\times U \to \mathbb{R}$ is the running cost, where $t \mapsto l(t,X,u)$ is continuous, $l(\cdot,X,u) \in L^{\infty}([t_0,t_f];\mathbb{R})$, and $(X,u) \mapsto l(t,X,u)$ is Lipschitz continuous. $m : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is the terminal cost, where $(X,X') \mapsto m(X,X')$ is Lipschitz continuous;
(ii) $X \mapsto l(t,X,u)$ is continuously differentiable, where $(t,X,u) \mapsto \partial_X l(t,X,u)$ is bounded and $(X,u) \mapsto \partial_X l(t,X,u)$ is Lipschitz continuous. $(X,X') \mapsto m(X,X')$ is continuously differentiable, where $(X,X') \mapsto \partial_X m(X,X')$ and $(X,X') \mapsto \partial_{X'} m(X,X')$ are bounded and Lipschitz continuous;
(iii) F is a nonempty closed convex subset of $\mathbb{R}^{2n}$;
(iv) $(t,X) \mapsto G(t,X) := [G_1(t,X) \cdots G_q(t,X)]^{\top}$ with $G_i : [t_0,t_f] \times \mathbb{R}^n \to \mathbb{R}$, $i = 1,\ldots,q$, is continuous, where $X \mapsto G(t,X)$ is continuously differentiable with $(t,X) \mapsto \partial_X G(t,x)$ being bounded. □

4. MAIN RESULT

Assume that $(\overline{u}(\cdot), \overline{X}(\cdot)) \in \mathcal{U} \times {}^{\mathrm{C}}_{n}\mathcal{D}^{\alpha}_{t_0+}$ is the optimal solution of **(P)**, i.e., $\overline{u}(\cdot) \in \mathcal{U}$ is the optimal solution of **(P)** and $\overline{X}(\cdot) := \overline{X}(\cdot; \overline{X}_0, \overline{u}) \in {}^{\mathrm{C}}_{n}\mathcal{D}^{\alpha}_{t_0+}$ is the corresponding optimal state trajectory of (1) controlled by $\overline{u}(\cdot) \in \mathcal{U}$. $\overline{X}(\cdot) \in {}^{\mathrm{C}}_{n}\mathcal{D}^{\alpha}_{t_0+}$ holds the state constraints in (2). We let

$$\begin{aligned} \overline{f}(t) &:= f(t, \overline{X}(t), \overline{u}(t)), \ \partial_X \overline{f}(t) := \partial_X f(t, \overline{X}(t), \overline{u}(t)) \\ \overline{l}(t) &:= l(\cdot, \overline{X}(t), \overline{u}(t)), \ \partial_X \overline{l}(t) := \partial_X l(t, \overline{X}(t), \overline{u}(t)) \\ \partial_{X_0}\overline{m} &:= \partial_{X_0} m(\overline{X}_0, \overline{X}(t_f)), \ \partial_X \overline{m} := \partial_X m(\overline{X}_0, \overline{X}(t_f)) \\ \overline{G}_i(t) &:= G_i(t, \overline{X}(t)), \ \partial_X \overline{G}_i(t) := \partial_X G_i(t, \overline{X}(t)). \end{aligned}$$

We state the main result of this paper.

Theorem 2. Suppose that Assumptions 1-2 hold. Assume that $(\overline{u}(\cdot), \overline{X}(\cdot)) \in \mathcal{U} \times {}^{\mathrm{C}}_{n}\mathcal{D}^{\alpha}_{t_0+}$ is the optimal solution of **(P)**. Then there exists a tuple (λ, ξ, θ), where $\lambda \in \mathbb{R}$, $\xi = (\xi_1, \xi_2) \in \mathbb{R}^{2n}$ with $\xi_1, \xi_2 \in \mathbb{R}^n$, and $\theta(\cdot) = (\theta_1(\cdot), \ldots, \theta_q(\cdot)) \in \mathrm{NBV}([t_0,t_f];\mathbb{R}^q)$ with $\theta_i(\cdot) \in \mathrm{NBV}([t_0,t_f];\mathbb{R})$, $i = 1,\ldots,q$, such that the following conditions are satisfied:

(i) **Nontriviality condition:** $(\lambda, \xi, \theta) \neq 0$ with $\xi = (\xi_1, \xi_2) \in N_F(\overline{X}_0, \overline{X}(t_f))$ and $\|\theta_i\|_{\mathrm{NBV}} = \theta_i(t_f) \geq 0$ for $i = 1\ldots,q$, where $N_F(\overline{X}_0, \overline{X}(t_f))$ is the normal cone to the convex set F at $(\overline{X}_0, \overline{X}(t_f)) \in F$ defined in (4), and θ_i is finite, nonnegative and monotonically nondecreasing on $[t_0, t_f]$;
(ii) **Nonnegativity condition:** $\lambda \geq 0$ and $\mathrm{d}\theta_i(t) \geq 0$ for $t \in [t_0,t_f]$ and $i = 1,\ldots,q$, where $\mathrm{d}\theta_i$ denotes the Lebesgue-Stieltjes measure on $[t_0,t_f]$ corresponding to θ_i, $i = 1,\ldots,q$;
(iii) **Adjoint equation:** there exists a nontrivial $p(\cdot) \in {}^{\mathrm{RL}}_{n}\mathcal{D}^{\alpha}_{t_f-}$ such that p is the unique solution of the following right RL fractional differential equation:

$$\mathrm{d}\big[\mathrm{I}^{1-\alpha}_{t_f-}[p]\big](t) = -\Big[\partial_X \overline{f}(t)^{\top} p(t) + \lambda \frac{(t_f - t)^{\beta-1}}{\Gamma(\beta)} \partial_X \overline{l}(t)^{\top}\Big]\mathrm{d}t - \sum_{i=1}^{q} \partial_X \overline{G}_i(t)^{\top} \mathrm{d}\theta_i(t), \tag{3}$$

(iv) **Transversality condition:**

$$\mathrm{I}^{1-\alpha}_{t_f-}[p](t_0) = -(\xi_1 + \lambda \partial_{X_0}\overline{m}^{\top}),$$
$$\mathrm{I}^{1-\alpha}_{t_f-}[p](t_f) = \xi_2 + \lambda \partial_X \overline{m}^{\top};$$

(v) **Complementary slackness condition:**

$$\int_{t_0}^{t_f} \overline{G}_i(s)\mathrm{d}\theta_i(s) = 0, \ i = 1,\ldots,q;$$

(vi) **Hamiltonian minimization condition:**

$$\min_{u\in U} \mathcal{H}(\overline{X}(t), p(t); u) = \mathcal{H}(\overline{X}(t), p(t); \overline{u}(t)),$$

a.e. $t \in [t_0,t_f]$, where $\mathcal{H}$ is the Hamiltonian:

$$\mathcal{H}(X, p; u) := \langle p, f(t,X,u)\rangle + \lambda \frac{(t_f - t)^{\beta-1}}{\Gamma(\beta)} l(t,X,u).$$

Remark 1. Without (2), Theorem 2 is reduced to (Ali et al., 2016, Theorem 3.1). In this case, only (iii), (iv) and (vi) of Theorem 2 are needed with $\lambda = 1$, $\xi = 0$, and $\theta = 0$. Note also that without the running state constraint, Theorem 2 is specialized to (Bergounioux and Bourdin, 2020, Theorem 3.12) with $\theta = 0$. □

5. SKETCH OF THE PROOF FOR THEOREM 2

In this section, the sketch of the proof for Theorem 2 is presented. Note that the full proof cannot be included in this paper due to space restriction.

5.1 Preliminaries on Distance Functions

Let $d_F : \mathbb{R}^{2n} \to \mathbb{R}_+$ be the standard Euclidean distance function to F defined by $d_F(x) := \inf_{y\in F} |x-y|$ for $x \in \mathbb{R}^{2n}$. By the projection theorem (Ruszczynski, 2006, Theorem 2.10), there is a unique $P_F(x) \in F$ with $P_F(x) : \mathbb{R}^{2n} \to F \subset \mathbb{R}^{2n}$, the projection of $x \in \mathbb{R}^{2n}$ onto F, such that $d_F(x) = \inf_{y\in F} |x-y| = |x - P_F(x)|$. In view of (Ruszczynski, 2006, Definition 2.37), we have $x - P_F(x) \in N_F(P_F(x))$ for $x \in \mathbb{R}^{2n}$, where $N_F(x)$ is the normal cone to the convex set F at a point $x \in F$ defined by

$$N_F(x) := \{y \in \mathbb{R}^{2n} \mid \langle y, y'-x\rangle \leq 0,\ \forall y' \in F\}. \quad (4)$$

Lemma 3. (Clarke, 1990, Proposition 2.5.4) The function $d_F(x)^2$ is Fréchet differentiable on $\mathbb{R}^{2n}$ with the Fréchet differentiation of $d_F(x)^2$ at $x \in \mathbb{R}^{2n}$ given by $Dd_F(x)^2(h) = 2\langle x - P_F(x), h\rangle$ for $h \in \mathbb{R}^{2n}$. □

Define $\psi : C([t_0,t_f];\mathbb{R}^n) \to C([t_0,t_f];\mathbb{R}^q)$ by $\psi(X(\cdot)) := G(\cdot, X(\cdot)) = [G_1(\cdot,X(\cdot)) \cdots G_q(\cdot,X(\cdot))]^\top$. Let $S \subset C([t_0,t_f];\mathbb{R}^q)$ be a nonempty closed convex cone of $C([t_0,t_f];\mathbb{R}^q)$ defined by $S := C([t_0,t_f];\mathbb{R}^q_-)$, where $\mathbb{R}^q_- := \mathbb{R}_- \times \cdots \times \mathbb{R}_-$. Note that S has a nonempty interior.

The normal cone to S at $x \in S$ is defined by

$$N_S(x) := \{\kappa \in C([t_0,t_f];\mathbb{R}^q)^* \mid \langle \kappa, \kappa' - x\rangle_{C_q^* \times C_q} \leq 0,\ \forall \kappa' \in S\}, \quad (5)$$

where $\langle \cdot,\cdot\rangle_{C_q^*\times C_q} := \langle\cdot,\cdot\rangle_{C([t_0,t_f];\mathbb{R}^q)^*\times C([t_0,t_f];\mathbb{R})^q}$ stands for the duality paring of $C([t_0,t_f];\mathbb{R}^q)$ and $C([t_0,t_f];\mathbb{R}^q)^*$ with $C([t_0,t_f];\mathbb{R}^q)^*$ being the dual space of $C([t_0,t_f];\mathbb{R}^q)$.

Let us define the distance function to S by $d_S(x) := \inf_{y\in S} \|x-y\|_{C([t_0,t_f];\mathbb{R}^q)}$ for $x \in C([t_0,t_f];\mathbb{R}^q)$.

Lemma 4. (Clarke, 1990, Proposition 2.4.1 and page 53) d_S is nonexpansive, continuous, and convex. □

Lemma 5. (Mordukhovich, 2006, Theorem 3.54) $d_S(x)^2$ is strictly Hadamard differentiable on $C([t_0,t_f];\mathbb{R}^q) \setminus S$ with the Hadamard differential given by $Dd_S(x)^2 = 2d_S(x)Dd_S(x)$ for $x \in C([t_0,t_f];\mathbb{R}^q)\setminus S$. Moreover, $d_S(x)^2$ is Fréchet differentiable on S with the Fréchet differential being $Dd_S(x)^2 = 0 \in C([t_0,t_f];\mathbb{R}^q)^*$ for all $x \in S$. □

5.2 Ekeland Variation Principle

For $\epsilon \geq 0$, define the penalized objective functional

$$J_\epsilon(X_0;u(\cdot)) = \Big(([J(X_0;u(\cdot)) - J(\overline{X}_0;\overline{u}(\cdot)) + \epsilon]^+)^2 + d_F(X_0,X(t_f))^2 + d_S(\psi(X(\cdot)))^2\Big)^{\frac{1}{2}}. \quad (6)$$

Define the Ekeland metric by

$$\widehat{d}((X_0,u(\cdot)),(\tilde{X}_0,\tilde{u}(\cdot))) = |X_0 - \tilde{X}_0| + \tilde{d}(u(\cdot),\tilde{u}(\cdot)), \quad (7)$$

where $\tilde{d}(u(\cdot),\tilde{u}(\cdot)) := |\{t \in [t_0,t_f] \mid u(t) \neq \tilde{u}(t)\}|$. $(\mathbb{R}^n \times \mathcal{U}, \widehat{d})$ is a complete metric space, and J_ϵ is continuous on $(\mathbb{R}^n \times \mathcal{U}, \widehat{d})$ by Assumptions 1-2 and Lemmas 3-4.

Note that $J_\epsilon(X_0,u(\cdot)) > 0$ for $(X_0,u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$ and $J_\epsilon(\overline{X}_0;\overline{u}(\cdot)) = \epsilon \leq \inf_{(X_0,u(\cdot))\in\mathbb{R}^n\times\mathcal{U}} J_\epsilon(X_0,u(\cdot)) + \epsilon$. Then by the Ekeland variational principle (Ekeland, 1974), there exists a pair $(X_0^\epsilon, u^\epsilon(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$ such that

$$\widehat{d}((X_0^\epsilon,u^\epsilon(\cdot)),(\overline{X}_0,\overline{u}(\cdot))) \leq \sqrt{\epsilon},$$

and for any $(X_0,u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$,

$$J_\epsilon(X_0^\epsilon,u^\epsilon(\cdot)) \leq J_\epsilon(X_0;u(\cdot)) + \sqrt{\epsilon}\widehat{d}((X_0^\epsilon,u^\epsilon(\cdot)),(X_0;u(\cdot))). \quad (8)$$

5.3 Spike Variation and Notation

For $\delta \in (0,1)$, define $\mathcal{E}_\delta = \{E \subset [t_0,t_f] \mid |E| = \delta t_f\}$, where $|E|$ denotes the Lebesgue measure of E. For any $E_\delta \in \mathcal{E}_\delta$ and $u(\cdot) \in \mathcal{U}$, we introduce *the spike variation*:

$$u^{\epsilon,\delta}(t) := \begin{cases} u^\epsilon(t), & t \in [t_0,t_f]\setminus E_\delta, \\ u(t), & t \in E_\delta. \end{cases} \quad (9)$$

Clearly $u^{\epsilon,\delta}(\cdot) \in \mathcal{U}$ and $\tilde{d}(u^{\epsilon,\delta}(\cdot),u^\epsilon(\cdot)) \leq |E_\delta| = \delta t_f$. Recall and consider the following variation: $X^\epsilon(\cdot) := X(\cdot;X_0^\epsilon,u^\epsilon)$ and $X^{\epsilon,\delta}(\cdot) := X(\cdot;X_0^\epsilon + \delta a, u^{\epsilon,\delta})$, where $a \in \mathbb{R}^n$ and $u^{\epsilon,\delta} \in \mathcal{U}$ is given in in (9). $X^\epsilon(\cdot)$ is the state trajectory of (1) under $(X_0^\epsilon,u^\epsilon(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$.

With $E^{\epsilon,\delta}(\cdot) := X^{\epsilon,\delta}(\cdot) - X^\epsilon(\cdot)$,
$f^\epsilon(s) := f(s,X^\epsilon(s),u^\epsilon(s))$,
$\partial_X f^\epsilon(s) := \partial_X f(s,X^\epsilon(s),u^\epsilon(s))$,
$\widehat{f}^\epsilon(s) := f(s,X^\epsilon(s),u(s)) - f(s,X^\epsilon(s),u^\epsilon(s))$,
$l^\epsilon(s) := l(s,X^\epsilon(s),u^\epsilon(s))$, $m^\epsilon := m(X_0^\epsilon,X^\epsilon(t_f))$, $\partial_X l^\epsilon(s) := \partial_X l(s,X^\epsilon(s),u^\epsilon(s))$,
$\widehat{l}^\epsilon(s) := l(s,X^\epsilon(s),u(s)) - l(s,X^\epsilon(s),u^\epsilon(s))$, $\partial_{X_0} m^\epsilon := \partial_{X_0} m(X_0^\epsilon,X^\epsilon(t_f))$, $\partial_X m^\epsilon := \partial_X m(X_0^\epsilon,X^\epsilon(t_f))$, $\widehat{f}(s) := f(t,\overline{X}(s),u(s)) - f(t,\overline{X}(s),\overline{u}(s))$,
$\widehat{l}(s) := l(s,\overline{X}(s),u(s)) - l(t,\overline{X}(s),\overline{u}(s))$.

5.4 Variational Analysis I

By (6) and (8), together with (7) and (9),

$$\begin{aligned} -\sqrt{\epsilon}(|a|+t_f) \leq{}& \frac{1}{J_\epsilon(X_0^\epsilon+\delta a;u^{\epsilon,\delta}(\cdot)) + J_\epsilon(X_0^\epsilon;u^\epsilon(\cdot))} \\ &\times \frac{1}{\delta}\Big(([J(X_0^\epsilon+\delta a;u^{\epsilon,\delta}(\cdot)) - J(\overline{X}_0;\overline{u}(\cdot)) + \epsilon]^+)^2 \\ &- ([J(X_0^\epsilon;u^\epsilon(\cdot)) - J(\overline{X}_0;\overline{u}(\cdot)) + \epsilon]^+)^2 \\ &+ d_F(X_0^\epsilon+\delta a, X^{\epsilon,\delta}(t_f))^2 - d_F(X_0^\epsilon,X^\epsilon(t_f))^2 \\ &+ d_S(\psi(X^{\epsilon,\delta}(\cdot)))^2 - d_S(\psi(X^\epsilon(\cdot)))^2\Big). \end{aligned} \quad (10)$$

Let Z^ϵ and $\widehat{Z}^\epsilon$ be the variational equations related to the pair $(X_0^\epsilon,u^\epsilon(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$ in (8) given by

$$Z^\epsilon(t) = a + \int_{t_0}^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}(\partial_X f^\epsilon(s)Z^\epsilon(s) + \widehat{f}^\epsilon(s))\mathrm{d}s,$$

$$\widehat{Z}^\epsilon(t_f) = \int_{t_0}^{t_f} \frac{(t_f-s)^{\beta-1}}{\Gamma(\beta)}(\partial_X l^\epsilon(s)Z^\epsilon(s) + \widehat{l}^\epsilon(s))\mathrm{d}s + \partial_{X_0} m^\epsilon a + \partial_X m^\epsilon Z^\epsilon(t_f).$$

By continuity of J_ϵ on $\mathbb{R}^n\times\mathcal{U}$, it follows that $\lim_{\delta\downarrow 0} J_\epsilon(X_0^\epsilon + \delta a; u^{\epsilon,\delta}(\cdot)) = J_\epsilon(X_0^\epsilon;u^\epsilon(\cdot))$. Then by the variational analysis of J_ϵ, $\widehat{Z}^\epsilon$ and Z^ϵ, together with Lemmas 3 and 5, as $\delta \downarrow 0$, (10) becomes for any $(a,u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}$,

$$-\sqrt{\epsilon}(|a|+t_f) \leq \lambda^\epsilon \widehat{Z}^\epsilon(t_f) + \langle \xi_1^\epsilon, a\rangle + \langle \xi_2^\epsilon, Z^\epsilon(t_f)\rangle + \langle \mu^\epsilon, \partial_X G(\cdot,X^\epsilon(\cdot))Z^\epsilon(\cdot)\rangle_{C_q^*\times C_q}, \quad (11)$$

where by Lemma 5 and the property of d_F in Section 5.1

$$|\lambda^\epsilon|^2 + |\xi^\epsilon|^2 + \|\mu^\epsilon\|^2_{C([t_0,t_f];\mathbb{R}^q)^*} = 1 \quad (12)$$

with $\lambda^\epsilon, \xi^\epsilon, \mu^\epsilon$ defined by $\lambda^\epsilon := \frac{[J(X_0^\epsilon;u^\epsilon(\cdot))-J(\overline{X}_0;\overline{u}(\cdot))+\epsilon]^+}{J_\epsilon(X_0^\epsilon;u^\epsilon(\cdot))} \geq 0$, $\xi^\epsilon := \frac{\begin{bmatrix} X_0^\epsilon \\ X^\epsilon(t_f) \end{bmatrix} - P_F(X_0^\epsilon, X^\epsilon(t_f))}{J_\epsilon(X_0^\epsilon;u^\epsilon(\cdot))} \in N_F(P_F(X_0^\epsilon, X^\epsilon(t_f)))$, and $\mu^\epsilon := \frac{d_S(\psi(X^\epsilon(\cdot)))Dd_S(\psi(X^\epsilon(\cdot)))}{J_\epsilon(X_0^\epsilon;u^\epsilon(\cdot))} \in C([t_0,t_f];\mathbb{R}^q)^*$.

5.5 Variational Analysis II

Let Z and $\widehat{Z}$ be the variational equations related to the optimal solution $\overline{u}(\cdot) \in \mathcal{U}$ given by

$$Z(t) = a + \int_{t_0}^{t} \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}(\partial_X \overline{f}(s)Z(s)\mathrm{d}s + \widehat{f}(s))\mathrm{d}s \quad (13)$$

$$\widehat{Z}(t_f) = \int_{t_0}^{t_f} \frac{(t_f-s)^{\beta-1}}{\Gamma(\beta)}(\partial_X \overline{l}(s)Z(s) + \widehat{l}(s))\mathrm{d}s + \partial_{X_0}\overline{m}a + \partial_X \overline{m}Z(t_f). \quad (14)$$

Let $\{\epsilon_k\}$ be a sequence of ϵ such that $\epsilon_k \geq 0$ and $\epsilon_k \downarrow 0$ as $k \to \infty$. By the Banach-Alaoglu theorem (Conway, 2000, page 130), we may extract a subsequence of $\{\epsilon_k\}$, still denoted by $\{\epsilon_k\}$, such that as $k \to \infty$,

$$(\{\lambda^{\epsilon_k}\}, \{\xi^{\epsilon_k}\}, \{\mu^{\epsilon_k}\}) \to (\lambda^0, \xi^0, \mu^0) =: (\lambda, \xi, \mu),$$

where the convergence of $\{\mu^{\epsilon_k}\} \to \mu$ (as $k \to \infty$) is understood in the weak–* sense (Conway, 2000). Hence, we can show that the tuple (λ, ξ, μ) holds

$$\lambda \geq 0,\ \xi \in N_F(P_F(\overline{X}_0, \overline{X}(t_f))),\ \mu \in N_S(\psi(\overline{X}(\cdot))). \quad (15)$$

Therefore, as $k \to \infty$, (11) becomes

$$0 \leq \lambda \widehat{Z}(t_f) + \langle \xi_1, a\rangle + \langle \xi_2, Z(t_f)\rangle + \langle \mu, \partial_X G(\cdot, \overline{X}(\cdot))Z(\cdot)\rangle_{C_q^* \times C_q},\ \forall (a, u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}. \quad (16)$$

5.6 Proof of Theorem 2 (v): Complementary Slackness

Let $\mu = (\mu_1, \ldots, \mu_q) \in C([t_0,t_f];\mathbb{R}^q)^*$. Based on (5),

$$\langle \mu, z - \psi(\overline{X}(\cdot))\rangle_{C_q^* \times C_q} = \sum_{i=1}^{q} \langle \mu_i, z_i - \psi_i(\overline{X}(\cdot))\rangle_{C_1^* \times C_1} \leq 0,\ \forall z \in S. \quad (17)$$

By choosing z appropriately, we can show that (17) is equivalent to $\langle \mu_i, \psi_i(\overline{X}(\cdot))\rangle_{C_1^* \times C_1} = 0$ for $i = 1, \ldots, q$. Then by the Riesz representation theorem (see (Conway, 2000, page 75 and page 382)), there is a unique $\theta(\cdot) = (\theta_1(\cdot), \ldots, \theta_q(\cdot)) \in \mathrm{NBV}([t_0,t_f];\mathbb{R}^q)$ with $\theta_i(\cdot) \in \mathrm{NBV}([t_0,t_f];\mathbb{R})$, i.e., θ_i, $i = 1, \ldots, m$, are normalized functions of bounded variation on $[t_0, t_f]$, such that every θ_i is finite, nonnegative, and monotonically nondecreasing on $[t_0,t_f]$ with $\theta_i(0) = 0$. Moreover, the Riesz representation theorem implies that for $i = 1, \ldots, q$, $\langle \mu_i, \psi_i(\overline{X}(\cdot))\rangle_{C_1^* \times C_1} = \int_{t_0}^{t_f} \overline{G}_i(s)\mathrm{d}\theta_i(s) = 0$. This shows the complementary slackness condition in Theorem 2.

5.7 Proof of Theorem 2 (i) and (ii): Nontriviality and Nonnegativity Conditions

Recall (17). By the Riesz representation theorem (see (Conway, 2000, page 75 and page 382)) and the fact that θ_i, $i = 1, \ldots, q$, is finite, nonnegative, and monotonically nondecreasing on $[t_0,t_f]$ with $\theta_i(0) = 0$ (see Section 5.6), it follows that $\|\mu_i\|_{C([t_0,t_f];\mathbb{R})^*} = \|\theta_i(\cdot)\|_{\mathrm{NBV}} = \theta_i(t_f) \geq 0$ for $i = 1, \ldots, q$. In addition, as θ_i is monotonically nondecreasing, we have $\mathrm{d}\theta_i(s) \geq 0$ for $s \in [t_0, t_f]$.

By (15) and the fact that $(\overline{X}_0, \overline{X}(t_f)) \in F$ implies $P_F(\overline{X}_0, \overline{X}(t_f)) = (\overline{X}_0, \overline{X}(t_f))$ (see Section 5.1), we have $\xi = \begin{bmatrix} \xi_1^\top & \xi_2^\top \end{bmatrix}^\top \in N_F(\overline{X}_0, \overline{X}(t_f))$.

In addition, from the fact that $S = C([t_0,t_f];\mathbb{R}^q_-)$ has an nonempty interior, there are $z' \in S$ and $\sigma > 0$ such that $z' + \sigma z \in S$ for all $z \in \overline{B}_{(C([t_0,t_f];\mathbb{R}^q), \|\cdot\|_{C([t_0,t_f];\mathbb{R}^q)})}(0,1)$ (the closure of the unit ball in $C([t_0,t_f];\mathbb{R}^q)$). Then using (15) and (12), we can show that $\sigma\|\mu\|_{C([t_0,t_f];\mathbb{R}^q)^*} = \sigma\sqrt{1 - |\lambda|^2 - |\xi|^2} \leq \langle \mu, \psi(\overline{X}(\cdot)) - z'\rangle_{C_q^* \times C_q}$, $z' \in S$. Now, note that $\sigma > 0$. When $\mu = 0 \in C([t_0,t_f];\mathbb{R}^q)^*$ and $\xi = 0$, we have $\lambda = 1$. When $\lambda = 0$ and $\mu = 0 \in C([t_0,t_f];\mathbb{R}^q)^*$, we have $|\xi| = 1$. When $\lambda = 0$ and $\xi = 0$, we have $\mu \neq 0 \in C([t_0,t_f];\mathbb{R}^q)^*$. This implies $(\lambda, \xi, \theta) \neq 0$, i.e., they cannot vanish simultaneously. This shows the nontriviality and nonnegativity conditions in Theorem 2.

5.8 Proof of Theorem 2 (iii): Adjoint Equation and Duality Analysis

We can show that the unique solution of the adjoint equation in (3) can be written as

$$p(t) = \Pi(t_f, t)^\top(\xi_2 + \lambda\partial_X \overline{m}^\top) + \lambda \int_t^{t_f} \Pi(\tau, t)^\top \frac{(t_f - \tau)^{\beta-1}}{\Gamma(\beta)} \partial_X \overline{l}(\tau)^\top \mathrm{d}\tau + \int_t^{t_f} \Pi(\tau,t)^\top \sum_{i=1}^{q} \partial_X \overline{G}_i(\tau)^\top \Theta_i(\tau)\mathrm{d}\tau, \quad (18)$$

where Π is the RL state-transition matrix associated with $\partial_X \overline{f}$. Moreover, Z in (13) can be written as

$$Z(t) = a + \int_{t_0}^{t} \Pi(t,s)(\partial_X \overline{f}(s)a + \widehat{f}(s))\mathrm{d}s. \quad (19)$$

Using (14), (18) and (19) with Fubini's formula, (16) becomes

$$0 \leq \langle \xi_1 + \lambda\partial_{X_0}\overline{m}^\top, a\rangle + \langle \xi_2 + \lambda\partial_X \overline{m}^\top, a\rangle + \int_{t_0}^{t_f} \sum_{i=1}^{q} \partial_X \overline{G}_i(s)\mathrm{d}\theta_i(s)a + \int_{t_0}^{t_f} \langle p(s), \partial_X \overline{f}(s)a\rangle\mathrm{d}s + \lambda \int_{t_0}^{t_f} \frac{(t_f-s)^{\beta-1}}{\Gamma(\beta)} \partial_X \overline{l}(s)\mathrm{d}sa + \int_{t_0}^{t_f} \langle p(s), \widehat{f}(s)\rangle\mathrm{d}s + \lambda \int_{t_0}^{t_f} \frac{(t_f-s)^{\beta-1}}{\Gamma(\beta)} \widehat{l}(s)\mathrm{d}s. \quad (20)$$

5.9 Proof of Theorem 2 (iv): Transversality Condition

When $u = \overline{u}$, by the adjoint equation in (3), (20) becomes

$$0 \leq \langle \xi_1 + \lambda\partial_{X_0}\overline{m}^\top, a\rangle + \langle \xi_2 + \lambda\partial_X \overline{m}^\top, a\rangle + \langle \int_{t_0}^{t_f} \mathrm{D}^\alpha_{t_f-}[p](s)\mathrm{d}s, a\rangle. \quad (21)$$

Recall from (18) that $\mathrm{I}^{1-\alpha}_{t_f-}[p](t_f) = \xi_2 + \lambda\partial_X \overline{m}^\top$. Then (21) becomes $0 \leq \langle \xi_1 + \lambda\partial_{X_0}\overline{m}^\top + \mathrm{I}^{1-\alpha}_{t_f-}[p](t_0), a\rangle$. As this holds for $a, -a \in \mathbb{R}^n$, we must have $\mathrm{I}^{1-\alpha}_{t_f-}[p](t_0) = -(\xi_1 + \lambda\partial_{X_0}\overline{m}^\top)$. This shows the transversality condition.

5.10 Proof of Theorem 2 (vi): Hamiltonian Minimization Condition

When $a = 0$, (20) becomes $0 \leq \int_{t_0}^{t_f} \langle p(s), \widehat{f}(s) \rangle \mathrm{d}s + \lambda \int_{t_0}^{t_f} \frac{(t_f - s)^{\beta-1}}{\Gamma(\beta)} \widehat{l}(s) \mathrm{d}s$. By the definition of $\mathcal{H}$, we have $\int_{t_0}^{t_f} \mathcal{H}(\overline{X}(s), p(s); u(s)) \mathrm{d}s \leq \int_{t_0}^{t_f} \mathcal{H}(\overline{X}(s), p(s); \overline{u}(s)) \mathrm{d}s$. As $\mathcal{H}$ is continuous in $u \in U$, and U is separable, for any $u \in U$, we have $\mathcal{H}(\overline{X}(t), p(t); u) \leq \mathcal{H}(\overline{X}(t), p(t); \overline{u}(t))$, a.e. $t \in [t_0, t_f]$, which proves the Hamiltonian minimization condition in Theorem 2.

This is the end of the short proof for Theorem 2.

6. CONCLUSIONS

In this paper, we have obtained the maximum principle for the fractional optimal control problem with terminal and running state constraints. The new proof has to be developed due to the inherent complex nature of the fractional control problem, the presence of the terminal and running state constraints, and the generalized standing assumptions. Some potential future research problems are (i) the generalization to other fractional equations such as the Caputo–Katugampola differential equation and (ii) studying the fractional dynamic programming principle.

REFERENCES

Ali, H.M., Pereira, F.L., and Gamma, S.M.A. (2016). A new approach to the Pontryagin maximum principle for nonlinear fractional optimal control problems. *Mathematical Methods in the Applied Sciences*, 39(13), 3640–3649.

Almeida, R., Kamocki, R., Malinowska, A.B., and Odzijewicz, T. (2021). On the necessary optimality conditions for the fractional Cucker–Smale optimal control problem. *Communications in Nonlinear Science and Numerical Simulation*, 96, 105678.

Bergounioux, M. and Bourdin, L. (2020). Pontryagin maximum principle for general Caputo fractional optimal control problems with Bolza cost and terminal constraints. *ESAIM Control Optimisation and Calculus of Variations*, 26(35), 1–38.

Clarke, F.H. (1990). *Optimization and Nonsmooth Analysis.* SIAM.

Conway, J.B. (2000). *A Course in Functional Analysis.* Springer.

Diethelm, K. (2010). *The Analysis of Fractional Differential Equations.* Springer.

Ekeland, I. (1974). On the variational principle. *Journal of Mathematical Analysis and Applications*, 47, 324–353.

Gomoyunov, M.I. (2020). Optimal control problems with a fixed terminal time in linear fractional-order systems. *Archives of Control Sciences*, 4, 721–744.

Gomoyunov, M.I. (2022). Minimax solutions of Hamilton–Jacobi equations with fractional coinvariant derivatives. *ESAIM: Control, Optimisation and Calculus of Variations*, 28(23), 1–36.

Gong, Z., Liu, C., Teo, K.L., Wang, S., and Wu, Y. (2021). Numerical solution of free final time fractional optimal control problems. *Applied Mathematics and Computation*, 405, 126270.

Jelicic, Z.D. and Petrovacki, N. (2008). Optimality conditions and a solution scheme for fractional optimal control problems. *Structural and Multidisciplinary Optimization*, 38, 571–581.

Kamocki, R. (2014). Pontryagin maximum principle for fractional ordinary optimal control problems. *Mathematical Methods in the Applied Sciences*, 37(11), 1668–1686.

Kilbas, A.A., Srivastava, H.M., and Trujillo, J.J. (2006). *Theory and Applications of Fractional Differential Equations.* Elsevier.

Liu, C., Sun, W., and Yi, X. (2022). Optimal control of a fractional smoking systems. *Journal of Industrial and Management Optimization*, 1–19. doi:10.3934/jimo.2022071.

Malinowska, A.B., Odzijewicz, T., and Torres, D.F. (2015). *Advanced Methods in the Fractional Calculus of Variations.* Springer.

Mordukhovich, B.S. (2006). *Variational Analysis and Generalized Differentiation I.* Springer.

Pooseh, S., Almeida, R., and Torres, D.F. (2014). Fractional order optimal control problems with free terminal time. *Journal of Industrial and Management Optimization*, 10(2), 363–381.

Rahimkhani, P. and Ordokhani, Y. (2021). Numerical investigation of distributed-order fractional optimal control problems via Bernstein wavelets. *Optimal Control Applications and Methods*, 42(1), 355–373.

Ruszczynski, A. (2006). *Nonlinear Optimization.* Princeton University Press.

Yusubov, S.S. and Mahmudov, E.N. (2021). Optimality conditions of singular controls for systems with Caputo fractional derivatives. *Journal of Industrial and Management Optimization*, 1–19. doi:10.3934/jimo.2021182.