Filippov's existence theorem and Pontryagin maximum principle for general Caputo fractional optimal control problems

Maïtine Bergounioux, Loïc Bourdin

FILIPPOV'S EXISTENCE THEOREM AND PONTRYAGIN MAXIMUM PRINCIPLE FOR GENERAL CAPUTO FRACTIONAL OPTIMAL CONTROL PROBLEMS*

MAÏTINE BERGOUNIOUX† AND LOÏC BOURDIN‡

Abstract. In this paper we focus on a general optimal control problem involving a dynamical system described by a nonlinear Caputo fractional differential equation of order $0 < \alpha \leq 1$, associated to a general Bolza cost written as the sum of a standard Mayer cost and a Lagrange cost given by a Riemann-Liouville fractional integral of order $\beta \geq \alpha$. In addition the present work handles general control and mixed initial/final state constraints. Adapting the standard Filippov's approach based on appropriate compactness assumptions and on the convexity of the set of augmented velocities, our first main result ensures the existence of at least one optimal solution. Secondly, the major contribution of this paper is the statement of a Pontryagin maximum principle which provides a first-order necessary optimality condition that can be applied to the fractional framework considered here. In particular, Hamiltonian maximization condition and transversality conditions on the adjoint vector are derived. Our proof is based on the sensitivity analysis of the Caputo fractional state equation with respect to needle-like control perturbations and on Ekeland's variational principle. The paper is concluded with two illustrating examples and with a list of several perspectives for forthcoming works.

Key words. Optimal control; fractional calculus; Riemann-Liouville and Caputo operators; Filippov's existence theorem; Pontryagin maximum principle; needle-like variations; Ekeland's variational principle; adjoint vector; Hamiltonian system; Hamiltonian maximization condition; transversality conditions.

AMS subject classifications. 34K35; 26A33; 34A08; 49J15; 49K15; 49K40; 34H05; 93C15.

1. Introduction.

Optimal control theory is concerned with the analysis of controlled dynamical systems, where one aims at steering such a system from a given configuration to some desired target by minimizing or maximizing some criterion. Most of the literature focuses on dynamical systems driven by ordinary differential equations. In that framework, the Filippov's theorem ensures the existence of at least one optimal trajectory under appropriate compactness and convexity assumptions (see [24], and see, e.g., [15, 39] for recent references and [19, Chapter 9] for some extensions). On the other hand, the Pontryagin Maximum Principle (denoted in short PMP), established at the end of the fifties (see [44], and see [26] for the history of this discovery), is the milestone of the classical optimal control theory. It provides a first-order necessary condition for optimality, by asserting that any optimal trajectory must be the projection of an extremal. The PMP then reduces the search of optimal trajectories to a boundary value problem on extremals. Optimal control theory, and in particular the PMP, have a wide field of applications in various domains. We refer the reader to textbooks such as [1, 15, 17, 18, 28, 32, 37, 48, 50, 51, 53] for theoretical results and/or practical applications, essentially for dynamical systems described by ordinary differential equations.

Before concluding this paragraph, recall that the classical optimal control theory can be seen as a generalization of the historical calculus of variations from the 18th century (see, e.g., [39, Section 3.2]). From this point of view, the PMP corresponds

*Submitted to the editors May 18, 2018.

†Denis Poisson Institute, UMR CNRS 7013, University of Orléans, France (maitine.bergounioux@univ-orleans.fr, maitinebergounioux.net/PagePro/Accueil.html).

‡XLIM Research Institute, UMR CNRS 7252, University of Limoges, France (loic.bourdin@unilim.fr, unilim.fr/pages_perso/loic.bourdin/).

to the extension of the famous Euler-Lagrange equation. Actually, for smooth and unconstrained optimal control problems of Lagrange form, a weak version of the PMP (in which the historical Hamiltonian maximization condition is replaced by a weaker null Hamiltonian gradient condition) can be easily derived from a simple variational approach (see, e.g., [39, Section 3.4]). However, in the general case, obtaining a strong version of the PMP, that moreover handles constraints on the state and/or on the control, requires more sophisticated mathematical tools such as the sensitivity analysis of the state equation with respect to control perturbations (needle-like variations for instance) combined with a Brouwer fixed point argument (see, e.g., [28, 37]) or with Ekeland's variational principle (see, e.g., [22, 38]). Many other variants exist in the literature (based on an implicit function theorem [1], Hahn-Banach separation theorem [15], or Aubin mini-max theorem [53] for example).

1.1. Optimal control theory in a fractional context. The fractional calculus generalizes the classical notions of integral and derivative to any real order. Many famous mathematicians introduced several notions of fractional operators, as Leibniz (1690's), Euler (1730's), Fourier (1820's), Liouville (1830's), Riemann (1840's), Sonin (1860's), Grünwald (1860's), Letnikov (1860's), Caputo (1960's), etc. These a priori different notions are not disconnected. In many cases it can be proved that two different notions actually coincide or are correlated by an explicit formula. The fractional operators are extensively used in many applications. We refer for instance to [29] for a large panorama in physics. The monographs [36, 47] are essential books on fractional calculus and fractional differential equations. In this paper, as commonly in the literature, we only consider the fractional operators of Riemann-Liouville and Caputo types. For the reader who is not familiar with these two notions, we refer to Section 2 for basic recalls and notations.

In [46] Riewe initiates the fractional calculus of variations and derives the first fractional version of the Euler-Lagrange equation, using non-integer order derivatives in order to describe nonconservative systems in mechanics. From there, a large number of publications has been devoted to the minimization of integral functionals involving various fractional operators. Many issues have been addressed and solved, and numerous classical results (such as first- and second-order necessary and/or sufficient optimality conditions, transversality conditions, Noether's theorems, Tonelli's existence theorems, etc.) have been extended to the fractional framework. We refer for example to [2, 5, 6, 8, 11, 20, 40, 42, 56] and references therein.

Compared to the growing literature on fractional calculus of variations, the fractional optimal control theory (where the dynamical system is driven by a fractional differential equation) had at first a slight development at the beginning of the 21th century. We refer the reader to [3, 4, 25, 27, 31] and references therein for some initiating works. These articles constitute a first step in the field and essentially use fractional variational approaches to derive fractional versions of the weak PMP (in which a null Hamiltonian gradient condition is obtained, but no Hamiltonian maximization condition) for smooth and unconstrained fractional optimal control problems. As a second step in the field, we mention the works of Kamocki from 2014. Indeed, a first attempt to establish a strong version of the PMP (with Hamiltonian maximization condition) in the case of a general Riemann-Liouville fractional optimal control problem with a classical Lagrange cost and with control constraints can be found in [34, Theorem 7]. However, several (quite restrictive) hypotheses are assumed, such as the compactness of the control constraint set, the convexity of the set of augmented velocities, the global Lipschitz continuity of the dynamic and some growth conditions on the dy-

namic, the Lagrange cost and its gradient. Moreover, a Riemann-Liouville fractional version of the initial condition is fixed, but no other state constraint can be handled. Hence, many challenging questions remain open in that field. Let us mention that existence results are provided for linear-convex Riemann-Liouville (resp. linear-linear Caputo) fractional optimal control problems associated to a classical Lagrange cost in [33, Theorem 17] (resp. in [35, Theorem 4.2]).

We conclude this section by pointing out that a large number of publications has been dedicated to the numerical study of fractional optimal control problems. We refer for instance to [3, 4, 7, 31, 45].

1.2. Contributions of the present paper. In this paper we work on a fixed real interval $[a, b]$. Our first motivation is to provide a functional framework that is adapted to describe fractional optimal control problems. For example, it is well-known that singularities arise at $t = a$ while using left Riemann-Liouville fractional operators. As a consequence, the class of $\mathcal{C}^1$-functions is too restrictive. In this paper, as in [14], we choose to use the appropriate set $\mathcal{AC}^{\alpha}_{a+}([a, b], \mathbb{R}^n)$ (resp. ${}_{\rm c}\mathcal{AC}^{\alpha}_{a+}([a, b], \mathbb{R}^n)$) of functions possessing Riemann-Liouville (resp. Caputo) fractional derivatives, see Definition 2.5 (resp. Definition 2.9). In order to avoid unboundedness states (which would deprive us of crucial estimations), we choose to work with the well-known Caputo fractional derivative ${}_{\rm c}{\rm D}^{\alpha}_{a+}$ of order $0 < \alpha \leq 1$ for which the corresponding trajectories are continuous. Hence, this paper is dedicated to the study of a general optimal control problem where the dynamical system is driven by the nonlinear Caputo fractional state equation

$$ {}_{\rm c}{\rm D}^{\alpha}_{a+}[x](t) = f(x(t), u(t), t), $$

of order $0 < \alpha \leq 1$.

Our second motivation is to consider a fractional optimal control problem that is sufficiently general in order to handle:

- Mayer and Lagrange costs (classical and fractional). Therefore, we consider the general Bolza cost given by

  $$ \varphi(x(a), x(b)) + {\rm I}^{\beta}_{a+}[F(x, u, \cdot)](b), $$

  where $\beta \geq \alpha$;
- General control constraint:

  $$ u(t) \in {\rm U}, $$

  where U is a nonempty closed set. We refer to Remark 3.16 for a discussion on this closedness assumption.
- General mixed initial/final state constraint:

  $$ g(x(a), x(b)) \in {\rm C}, $$

  where C is a nonempty closed convex set. To the best of our knowledge, no endpoint constraint has never been considered yet in the literature on fractional optimal control theory. Moreover, note that the above consideration of state constraint is very general and allows to encompass a lot of typical situations such as fixed initial and/or final conditions, free initial and/or final conditions, equality and/or inequality constraints, etc. We refer to Remark 3.17 for more details.

Moreover, the regularity assumptions that we require in this paper on the functions f, F, φ and g are reduced as much as possible (as far as we know) to guarantee the applicability of our proofs. In particular, no growth condition and no global Lipschitz continuity are imposed. For the precise definition of the problem investigated in this paper and the corresponding assumptions, we refer the reader to Problem (OCP) in Section 3.1.

The two major contributions of this paper are stated in Section 3:

(i) A Filippov's existence theorem (Theorem 3.1) for Problem (OCP). This result is based on standard compactness and convexity assumptions. Actually, the classical proof is easily extended to the fractional framework considered in this paper;

(ii) A Pontryagin maximum principle (Theorem 3.12) for Problem (OCP). The proof is based on the sensitivity analysis of the Caputo fractional state equation with respect to needle-like control perturbations and on Ekeland's variational principle. In contrast to the proof of Theorem 3.1, the one of Theorem 3.12 requires a lot of technical adjustments from the classical case. Indeed, the nonlocality of the fractional operators induces specific variation vectors (Propositions 3.10 and 3.11). Moreover, the necessary optimality inequalities obtained on the variation vectors require the use of fractional Duhamel formulas derived in [12] to conclude. This (quite long) proof is moved to Appendix A.

Two examples (including endpoint state and control constraints) illustrating the applicability of Theorem 3.12 are provided in Section 4.

Finally, it should be noted that many challenging problems are still open in fractional optimal control theory. We refer to Section 5 for a tentative list of questions to address in future works. In particular, we choose in this paper to deal with a Caputo fractional state equation in order to guarantee the boundedness (even the continuity) of the corresponding trajectories. An important objective would be to adapt the whole framework of the present paper to a Riemann-Liouville fractional state equation.

1.3. Organization of the paper. The paper is organized as follows. Section 2 is devoted to basic recalls and notations from fractional calculus. The main results (Theorems 3.1 and 3.12) are stated in Section 3 with some comments. Two illustrating examples are detailed in Section 4. We conclude in Section 5 with perspectives for future works. The proof of Theorem 3.12 is given in Appendix A. Finally, Appendix B contains two technical results (in particular a new fractional version of the classical Gronwall lemma, see Proposition B.1) that are required in the paper.

2. Basics on fractional calculus. Throughout the paper the abbreviation RL stands for Riemann-Liouville. This section is devoted to basic recalls about RL and Caputo fractional operators. All definitions and results below are standard and are mostly extracted from the monographs [36, 47]. Let $n \in \mathbb{N}^*$ be a positive integer and let $I \subset \mathbb{R}$ be a subinterval of $\mathbb{R}$. In the whole paper we denote by:

- $\mathrm{L}^r(I, \mathbb{R}^n)$ the Lebesgue space of r-integrable functions defined on I with values in $\mathbb{R}^n$, endowed with its usual norm $\|\cdot\|_{\mathrm{L}^r}$, for any $1 \leq r < \infty$;
- $\mathrm{L}^\infty(I, \mathbb{R}^n)$ the Lebesgue space of essentially bounded functions defined on I with values in $\mathbb{R}^n$, endowed with its usual norm $\|\cdot\|_{\mathrm{L}^\infty}$.
- $\mathcal{C}(I, \mathbb{R}^n)$ the space of continuous functions on I with values in $\mathbb{R}^n$, endowed with the uniform norm $\|\cdot\|_{\mathcal{C}}$;
- $\mathcal{AC}(I, \mathbb{R}^n)$ the subspace of $\mathcal{C}(I, \mathbb{R}^n)$ of absolutely continuous functions.

For any $\mathrm{E}(I, \mathbb{R}^n)$ one of the above spaces, we denote by $\mathrm{E}_{\mathrm{loc}}(I, \mathbb{R}^n)$ the space of

functions $x : I \to \mathbb{R}^n$ such that $x \in \mathrm{E}(J, \mathbb{R}^n)$ for every compact subinterval $J \subset I$. In particular, if I is compact, then $\mathrm{E}_{\mathrm{loc}}(I, \mathbb{R}^n) = \mathrm{E}(I, \mathbb{R}^n)$. Finally, for any $a \in I$, we denote by $\mathcal{C}^a(I, \mathbb{R}^n)$ the set of all functions $x \in \mathcal{C}(I, \mathbb{R}^n)$ such that $x(a) = 0_{\mathbb{R}^n}$.

2.1. Left RL and Caputo fractional operators. In this section we fix $a \in \mathbb{R}$ and $I \subset \mathbb{R}$ an interval such that $\{a\} \subsetneq I \subset [a, +\infty)$. Note that I is not necessarily compact. Precisely I can be written either as $I = [a, +\infty)$, or $I = [a, b)$ for some $b > a$, or $I = [a, b]$ for some $b > a$.

2.1.1. Left RL fractional integral. In the sequel Γ denotes the standard Gamma function. We focus in this section on left fractional integrals of RL type.

Definition 2.1 (Left RL fractional integral). *The left Riemann-Liouville fractional integral $\mathrm{I}^{\alpha}_{a+}[x]$ of order $\alpha > 0$ and inferior limit a of a function $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ is defined on I by*

$$\mathrm{I}^{\alpha}_{a+}[x](t) := \frac{1}{\Gamma(\alpha)} \int_a^t (t - \tau)^{\alpha - 1} x(\tau)\, d\tau,$$

provided that the right-hand side term exists. For $\alpha = 0$ we set $\mathrm{I}^0_{a+}[x] := x$.

Proposition 2.2 ([36, Lemma 2.1 p.72]). *If $\alpha \geq 0$ and $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$, then $\mathrm{I}^{\alpha}_{a+}[x] \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$.*

Proposition 2.3 ([36, Lemma 2.3 p.73]). *If $\alpha_1 \geq 0$, $\alpha_2 \geq 0$ and $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$, then the equalities*

$$\mathrm{I}^{\alpha_1}_{a+}\left[\mathrm{I}^{\alpha_2}_{a+}[x]\right] = \mathrm{I}^{\alpha_1+\alpha_2}_{a+}[x] = \mathrm{I}^{\alpha_2+\alpha_1}_{a+}[x] = \mathrm{I}^{\alpha_2}_{a+}\left[\mathrm{I}^{\alpha_1}_{a+}[x]\right]$$

hold true.

Let $\alpha \geq 0$ and $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$. From Proposition 2.2, $\mathrm{I}^{\alpha}_{a+}[x](t)$ exists for almost every $t \in I$. Throughout the paper, if $\mathrm{I}^{\alpha}_{a+}[x]$ is equal almost everywhere on I to a continuous function, then $\mathrm{I}^{\alpha}_{a+}[x]$ is automatically identified to its continuous representative. In that case $\mathrm{I}^{\alpha}_{a+}[x](t)$ is defined for every $t \in I$.

Proposition 2.4 ([47, Theorem 3.6 p.67]). *If $\alpha > 0$ and $x \in \mathrm{L}^{\infty}_{\mathrm{loc}}(I, \mathbb{R}^n)$, then $\mathrm{I}^{\alpha}_{a+}[x] \in \mathcal{C}^a(I, \mathbb{R}^n)$.*

2.1.2. Left RL and Caputo fractional derivatives. This section is devoted to basic recalls on left fractional derivatives of RL and Caputo types.

Definition 2.5 (Left RL fractional derivative). *We say that $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ possesses on I a left RL fractional derivative $\mathrm{D}^{\alpha}_{a+}[x]$ of order $0 \leq \alpha \leq 1$ and inferior limit a if and only if $\mathrm{I}^{1-\alpha}_{a+}[x] \in \mathcal{AC}_{\mathrm{loc}}(I, \mathbb{R}^n)$. In that case $\mathrm{D}^{\alpha}_{a+}[x] \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ is defined on I by*

$$\mathrm{D}^{\alpha}_{a+}[x](t) := \frac{d}{dt}\left[\mathrm{I}^{1-\alpha}_{a+}[x]\right](t).$$

We denote by $\mathcal{AC}^{\alpha}_{a+}(I, \mathbb{R}^n)$ the set of all functions $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ possessing on I a left RL fractional derivative $\mathrm{D}^{\alpha}_{a+}[x]$ of order $0 \leq \alpha \leq 1$ and inferior limit a.

Remark 2.6. If $\alpha = 1$, $\mathcal{AC}^1_{a+}(I, \mathbb{R}^n) = \mathcal{AC}_{\mathrm{loc}}(I, \mathbb{R}^n)$ and $\mathrm{D}^1_{a+}[x] = \dot{x}$ is the usual derivative of x for any $x \in \mathcal{AC}_{\mathrm{loc}}(I, \mathbb{R}^n)$. If $\alpha = 0$, $\mathcal{AC}^0_{a+}(I, \mathbb{R}^n) = \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ and $\mathrm{D}^0_{a+}[x] = x$ for any $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$.

PROPOSITION 2.7. *Let $0 \leq \alpha \leq 1$ and $x \in \mathrm{L}^1_{\mathrm{loc}}(I,\mathbb{R}^n)$. Then $x \in \mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$ if and only if there exist $y \in \mathrm{L}^1_{\mathrm{loc}}(I,\mathbb{R}^n)$ and $x_a \in \mathbb{R}^n$ such that*

$$x(t) = \frac{(t-a)^{\alpha-1}}{\Gamma(\alpha)} x_a + \mathrm{I}^\alpha_{a+}[y](t),$$

for almost every $t \in I$. In that case, it holds that $y = \mathrm{D}^\alpha_{a+}[x]$ and $x_a = \mathrm{I}^{1-\alpha}_{a+}[x](a)$.

Remark 2.8. Let $0 \leq \alpha \leq 1$. In general a function $x \in \mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$ admits a singularity at $t = a$.

DEFINITION 2.9 (Left Caputo fractional derivative). *We say that $x \in \mathcal{C}(I,\mathbb{R}^n)$ possesses on I a left Caputo fractional derivative ${}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x]$ of order $0 \leq \alpha \leq 1$ and inferior limit a if and only if $x - x(a) \in \mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$. In that case ${}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x] \in \mathrm{L}^1_{\mathrm{loc}}(I,\mathbb{R}^n)$ is defined on I by*

$${}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x](t) := \mathrm{D}^\alpha_{a+}[x - x(a)](t).$$

We denote by ${}_{\mathrm{c}}\mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$ the set of all functions $x \in \mathcal{C}(I,\mathbb{R}^n)$ possessing on I a left Caputo fractional derivative ${}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x]$ of order $0 \leq \alpha \leq 1$ and inferior limit a.

Remark 2.10. If $\alpha = 1$, ${}_{\mathrm{c}}\mathcal{AC}^1_{a+}(I,\mathbb{R}^n) = \mathcal{AC}_{\mathrm{loc}}(I,\mathbb{R}^n)$ and ${}_{\mathrm{c}}\mathrm{D}^1_{a+}[x] = \dot{x}$ for any $x \in \mathcal{AC}_{\mathrm{loc}}(I,\mathbb{R}^n)$. If $\alpha = 0$, ${}_{\mathrm{c}}\mathcal{AC}^0_{a+}(I,\mathbb{R}^n) = \mathcal{C}(I,\mathbb{R}^n)$ and ${}_{\mathrm{c}}\mathrm{D}^0_{a+}[x] = x - x(a)$ for any $x \in \mathcal{C}(I,\mathbb{R}^n)$.

Remark 2.11. Let $0 \leq \alpha \leq 1$. Note that ${}_{\mathrm{c}}\mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n) \subset \mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$ and, if $\alpha \neq 1$, it holds that

$$\mathrm{D}^\alpha_{a+}[x](t) = {}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x](t) + \frac{x(a)}{\Gamma(1-\alpha)}(t-a)^{-\alpha},$$

for almost every $t \in I$ and all $x \in {}_{\mathrm{c}}\mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$.

PROPOSITION 2.12. *Let $0 \leq \alpha \leq 1$ and $x \in \mathcal{C}(I,\mathbb{R}^n)$. Then $x \in {}_{\mathrm{c}}\mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$ if and only if there exist $y \in \mathrm{L}^1_{\mathrm{loc}}(I,\mathbb{R}^n)$ and $x_a \in \mathbb{R}^n$ such that*

$$x(t) = x_a + \mathrm{I}^\alpha_{a+}[y](t),$$

for all $t \in I$. In that case, the above relation holds replacing y by ${}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x]$ and x_a by $x(a)$.

Remark 2.13. Let $0 \leq \alpha \leq 1$ and $x(t) = x_a + \mathrm{I}^\alpha_{a+}[y](t)$ for almost every $t \in I$, for some $y \in \mathrm{L}^1_{\mathrm{loc}}(I,\mathbb{R}^n)$ and $x_a \in \mathbb{R}^n$. It might be possible that $x \notin \mathcal{C}(I,\mathbb{R}^n)$ and then Proposition 2.12 cannot be applied. From Proposition 2.4, if $\alpha > 0$ and $y \in \mathrm{L}^\infty_{\mathrm{loc}}(I,\mathbb{R}^n)$, then $x \in \mathcal{C}(I,\mathbb{R}^n)$ with $x(a) = x_a$ and Proposition 2.12 can be applied.

Remark 2.14. Let $0 \leq \alpha \leq 1$ and $x \in \mathcal{C}(I,\mathbb{R}^n)$ such that $x(t) = x_a + \mathrm{I}^\alpha_{a+}[y](t)$ for all $t \in I$, for some $y \in \mathrm{L}^1_{\mathrm{loc}}(I,\mathbb{R}^n)$ and $x_a \in \mathbb{R}^n$. From Proposition 2.12, we know that $x \in {}_{\mathrm{c}}\mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$ and that $x(t) = x(a) + \mathrm{I}^\alpha_{a+}[{}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x]](t)$ for all $t \in I$. However, without any additional assumption, one cannot assert that $y = {}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x]$ and $x_a = x(a)$. From Proposition 2.4, if $\alpha > 0$ and $y \in \mathrm{L}^\infty_{\mathrm{loc}}(I,\mathbb{R}^n)$, then we can conclude that $y = {}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x]$ and $x_a = x(a)$.

From the above definitions and propositions, one can easily recover the following well-known result.

PROPOSITION 2.15 ([36, Theorem 2.1 p.92]). *Let $0 \leq \alpha \leq 1$. The inclusion $\mathcal{AC}_{\mathrm{loc}}(I,\mathbb{R}^n) \subset {}_{\mathrm{c}}\mathcal{AC}^\alpha_{a+}(I,\mathbb{R}^n)$ holds true with ${}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x] = \mathrm{I}^{1-\alpha}_{a+}[\dot{x}]$ for any $x \in \mathcal{AC}_{\mathrm{loc}}(I,\mathbb{R}^n)$.*

2.2. Right RL and Caputo fractional operators. This section is devoted to the right counterparts of the notions recalled in Section 2.1. For this purpose we fix $b \in \mathbb{R}$ and $I \subset \mathbb{R}$ an interval such that $\{b\} \subsetneq I \subset (-\infty, b]$. Precisely, the interval I writes either $I = (-\infty, b]$, or $I = (a, b]$ for some $a < b$, or $I = [a, b]$ for some $a < b$.

DEFINITION 2.16 (Right RL fractional integral). *The right RL fractional integral* $\mathrm{I}^{\alpha}_{b-}[x]$ *of order* $\alpha > 0$ *and superior limit* b *of* $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ *is defined on* I *by*

$$\mathrm{I}^{\alpha}_{b-}[x](t) := \frac{1}{\Gamma(\alpha)} \int_t^b (\tau - t)^{\alpha - 1} x(\tau)\, d\tau,$$

provided that the right-hand side term exists. For $\alpha = 0$ *we define* $\mathrm{I}^0_{b-}[x] := x$.

DEFINITION 2.17 (Right RL fractional derivative). *We say that* $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ *possesses on* I *a right RL fractional derivative* $\mathrm{D}^{\alpha}_{b-}[x]$ *of order* $0 \leq \alpha \leq 1$ *and superior limit* b *if and only if* $\mathrm{I}^{1-\alpha}_{b-}[x] \in \mathcal{AC}_{\mathrm{loc}}(I, \mathbb{R}^n)$. *In that case* $\mathrm{D}^{\alpha}_{b-}[x] \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ *is defined on* I *by*

$$\mathrm{D}^{\alpha}_{b-}[x](t) := -\frac{d}{dt}\Big[\mathrm{I}^{1-\alpha}_{b-}[x]\Big](t).$$

We denote by $\mathcal{AC}^{\alpha}_{b-}(I, \mathbb{R}^n)$ *the set of all functions* $x \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ *possessing on* I *a right RL fractional derivative* $\mathrm{D}^{\alpha}_{b-}[x]$ *of order* $0 \leq \alpha \leq 1$ *and superior limit* b.

DEFINITION 2.18 (Right Caputo fractional derivative). *We say that* $x \in \mathcal{C}(I, \mathbb{R}^n)$ *possesses on* I *a right Caputo fractional derivative* ${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{b-}[x]$ *of order* $0 \leq \alpha \leq 1$ *and superior limit* b *if and only if* $x - x(b) \in \mathcal{AC}^{\alpha}_{b-}(I, \mathbb{R}^n)$. *In that case* ${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{b-}[x] \in \mathrm{L}^1_{\mathrm{loc}}(I, \mathbb{R}^n)$ *is defined on* I *by*

$${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{b-}[x](t) := \mathrm{D}^{\alpha}_{b-}[x - x(b)](t).$$

We denote by ${}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{b-}(I, \mathbb{R}^n)$ *the set of all functions* $x \in \mathcal{C}(I, \mathbb{R}^n)$ *possessing on* I *a right Caputo fractional derivative* ${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{b-}[x]$ *of order* $0 \leq \alpha \leq 1$ *and superior limit* b.

All results recalled in Section 2.1 (for left fractional operators) have each a right counterpart version. We refer the reader to [36, 47] for details.

3. Main results and comments. This section is devoted to the main results (Theorems 3.1 and 3.12) of the present paper.

3.1. Framework, terminology and assumptions. Let $a < b$ be two real numbers. Let m, n, $j \in \mathbb{N}^*$, let $0 < \alpha \leq 1$ and $\beta \geq \alpha$ be fixed. We consider the general Caputo fractional optimal control problem of Bolza form given by

$$\text{(OCP)} \qquad \begin{cases} \text{minimize} & \varphi(x(a), x(b)) + \mathrm{I}^{\beta}_{a+}[F(x, u, \cdot)](b), \\ \text{subject to} & x \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a, b], \mathbb{R}^n),\ u \in \mathrm{L}^{\infty}([a, b], \mathbb{R}^m), \\ & {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}[x](t) = f(x(t), u(t), t), \quad \text{a.e. } t \in [a, b], \\ & g(x(a), x(b)) \in \mathrm{C}, \\ & u(t) \in \mathrm{U}, \quad \text{a.e. } t \in [a, b]. \end{cases}$$

In Problem (OCP), u is the *control* function and x is the *state* function (also called *trajectory*). In this paper we assume that:

- the function $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, that describes the *Mayer cost* $\varphi(x(a), x(b))$, is of class $\mathcal{C}^1$;

- the set $\mathrm{C} \subset \mathbb{R}^j$ is a nonempty closed convex subset of $\mathbb{R}^j$ and the function $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^j$, that describes the *terminal state constraint* $g(x(a), x(b)) \in \mathrm{C}$, is of class $\mathcal{C}^1$;
- the set $\mathrm{U} \subset \mathbb{R}^m$, that describes the *control constraint* $u(t) \in \mathrm{U}$, is a nonempty closed subset of $\mathbb{R}^m$;
- the *dynamic* $f : \mathbb{R}^n \times \mathbb{R}^m \times [a,b] \to \mathbb{R}^n$, that drives the *Caputo fractional state equation* ${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}[x](t) = f(x(t), u(t), t)$, satisfies the following conditions:
  - f is continuous;
  - f is differentiable with respect to its first variable;
  - $\partial_1 f$ is continuous;
  - f is Lipschitz continuous with respect to its first two variables on every compact subset (see Inequality (3.1) for precisions).

  In particular, for every compact subset $\mathrm{K} \subset \mathbb{R}^n \times \mathbb{R}^m \times [a,b]$, there exists a nonnegative constant $L \geq 0$ such that $\|\partial_1 f(x,u,t)\|_{\mathbb{R}^{n\times n}} \leq L$ for all $(x,u,t) \in \mathrm{K}$, and such that

$$\|f(x_2,u_2,t) - f(x_1,u_1,t)\|_{\mathbb{R}^n} \leq L(\|x_2 - x_1\|_{\mathbb{R}^n} + \|u_2 - u_1\|_{\mathbb{R}^m}), \tag{3.1}$$

  for all (x_1,u_1,t), $(x_2,u_2,t) \in \mathrm{K}$.
- the function $F : \mathbb{R}^n \times \mathbb{R}^m \times [a,b] \to \mathbb{R}$, that describes the *fractional Lagrange cost* $\mathrm{I}^{\beta}_{a+}[F(x,u,\cdot)](b)$, is assumed to satisfy the same assumptions than the dynamic f.

Usually the fractional Lagrange cost $\mathrm{I}^{\beta}_{a+}[F(x,u,\cdot)](b)$ is considered with $\beta = \alpha$ or $\beta = 1$ in the literature. Nevertheless, we can always get back to the case where $\beta = \alpha$ by noting that the fractional Lagrange cost can be rewritten as $\mathrm{I}^{\alpha}_{a+}[F^{\beta}(x,u,\cdot)](b)$ where

$$F^{\beta}(x,u,t) := \frac{\Gamma(\alpha)}{\Gamma(\beta)}(b-t)^{\beta-\alpha}F(x,u,t),$$

for all $(x,u,t) \in \mathbb{R}^n \times \mathbb{R}^m \times [a,b]$. Since $\beta \geq \alpha$, note that F^{β} satisfies the same regularity assumptions than F (and thus than f).

3.2. Filippov's existence theorem. We provide hereafter a result stating the existence of at least one optimal solution to Problem (OCP) under some appropriate compactness and convexity assumptions. We follow the standard Filippov's approach (see [15, 19, 24, 39] for example). For this purpose we introduce the usual *set of augmented velocities* defined by

$$(f,F_+)(x,\mathrm{U},t) := \{(f(x,u,t), F(x,u,t)+\gamma) \mid u \in \mathrm{U},\ \gamma \geq 0\} \subset \mathbb{R}^{n+1},$$

for all $(x,t) \in \mathbb{R}^n \times [a,b)$. Moreover let $\mathcal{T} \subset \mathcal{C}$ stand for the set of all trajectories $x \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a,b],\mathbb{R}^n)$ that can be associated to a control $u \in \mathrm{L}^{\infty}([a,b],\mathbb{R}^m)$ such that the couple (x,u) satisfies all the constraints of Problem (OCP). Obviously, if $\mathcal{T}$ is empty, then Problem (OCP) has no solution. Otherwise, the following existence result holds true.

THEOREM 3.1 (Filippov's existence theorem). *If* U *is compact,* $\mathcal{T}$ *is nonempty and bounded in* $\mathcal{C}$*, and if* $(f,F_+)(x,\mathrm{U},t)$ *is convex for all* $(x,t) \in \mathbb{R}^n \times [a,b)$*, then Problem* (OCP) *has at least one optimal solution.*

Proof. First of all, since $(f,F_+)(x,\mathrm{U},t)$ is convex, note that $(f,F^{\beta}_+)(x,\mathrm{U},t)$ is convex for all $(x,t) \in \mathbb{R}^n \times [a,b)$. Since $\mathcal{T}$ is nonempty, let $(x_k)_{k\in\mathbb{N}} \subset \mathcal{T}$ be a minimizing sequence and $(u_k)_{k\in\mathbb{N}} \subset \mathrm{L}^{\infty}([a,b],\mathbb{R}^m)$ be a corresponding control sequence. We

introduce $z_k := \mathrm{I}^{\alpha}_{a+}[F^{\beta}(x_k,u_k,\cdot)] \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a,b],\mathbb{R})$ and $G_k := (f,F^{\beta})(x_k,u_k,\cdot) \in \mathrm{L}^{\infty}([a,b],\mathbb{R}^{n+1})$ for all $k\in\mathbb{N}$. In particular the cost associated to x_k is equal to $\varphi(x_k(a),x_k(b))+z_k(b)$ for all $k\in\mathbb{N}$. We know that $(x_k,z_k)(t) = (x_k(a),0)+\mathrm{I}^{\alpha}_{a+}[G_k](t)$ for all $t\in[a,b]$ and all $k\in\mathbb{N}$. From the hypotheses, the sequence $(G_k)_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^{\infty}([a,b],\mathbb{R}^{n+1})$ and thus, up to a subsequence, weakly* converges to some $G\in\mathrm{L}^{\infty}([a,b],\mathbb{R}^{n+1})$. Similarly, the sequence $(x_k(a))_{k\in\mathbb{N}}$ is bounded in $\mathbb{R}^n$ and thus, up to a subsequence, converges to some $x^*_a\in\mathbb{R}^n$. We deduce that $(x_k,z_k)_{k\in\mathbb{N}}$, up to a subsequence, pointwise converges on $[a,b]$ to (x^*,z^*) where $(x^*,z^*)(t) := (x^*_a,0)+\mathrm{I}^{\alpha}_{a+}[G](t)$ for all $t\in[a,b]$. Note that $(x^*,z^*)\in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a,b],\mathbb{R}^{n+1})$ with ${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}[(x^*,z^*)] = G$ and $(x^*,z^*)(a) = (x^*_a,0)$. From the continuity of φ, the infimum cost of Problem (OCP) is equal to $\varphi(x^*(a),x^*(b))+z^*(b)$. Similarly, from the continuity of g and the closedness of C, we deduce that $g(x^*(a),x^*(b))\in\mathrm{C}$. To conclude the proof, we only need to prove that $G\in\mathcal{W}$ where

$$\mathcal{W} := \{h\in\mathrm{L}^2([a,b],\mathbb{R}^{n+1}) \mid h(\tau)\in(f,F^{\beta}_{+})(x^*(\tau),\mathrm{U},\tau) \text{ for almost every } \tau\in[a,b]\}.$$

Indeed, if $G\in\mathcal{W}$, then there exists $u^*(\tau)\in\mathrm{U}$ and $\gamma^*(\tau)\geq 0$ such that

$$G(\tau) = (f(x^*(\tau),u^*(\tau),\tau), F^{\beta}(x^*(\tau),u^*(\tau),\tau)+\gamma^*(\tau)),$$

for almost every $\tau\in[a,b]$. Moreover, u^* and γ^* can be selected measurable on $[a,b]$ from implicit measurable function theorems (see, e.g., [54, Section 7]). Since U is bounded, we get that $u^*\in\mathrm{L}^{\infty}([a,b],\mathbb{R}^m)$. Thus $x^*\in\mathcal{T}$ associated to the control u^* and the associated cost is equal to

$$\varphi(x^*(a),x^*(b))+\mathrm{I}^{\beta}_{a+}[F(x^*,u^*,\cdot)](b) = \varphi(x^*(a),x^*(b))+\mathrm{I}^{\alpha}_{a+}[F^{\beta}(x^*,u^*,\cdot)](b)$$

$$\leq \varphi(x^*(a),x^*(b))+\mathrm{I}^{\alpha}_{a+}[F^{\beta}(x^*,u^*,\cdot)+\gamma^*](b) = \varphi(x^*(a),x^*(b))+z^*(b),$$

which would prove that the couple (x^*,u^*) is an optimal solution to Problem (OCP). Now, let us prove that $G\in\mathcal{W}$ (in two steps).

First, one can easily deduce from the assumptions that $\mathcal{W}$ is a closed and convex subset of $\mathrm{L}^2([a,b],\mathbb{R}^{n+1})$ with its usual topology, and thus with its weak topology as well. Now let us consider $(H_k)_{k\in\mathbb{N}}\subset\mathcal{W}$ defined by $H_k(\tau) := (f,F^{\beta})(x^*(\tau),u_k(\tau),\tau)$ for almost every $\tau\in[a,b]$ and all $k\in\mathbb{N}$. Since U is compact, the sequence $(H_k)_{k\in\mathbb{N}}$ is bounded in $\mathrm{L}^{\infty}([a,b],\mathbb{R}^{n+1})$ and thus in $\mathrm{L}^2([a,b],\mathbb{R}^{n+1})$. We deduce that, up to a subsequence, $(H_k)_{k\in\mathbb{N}}$ weakly converges in $\mathrm{L}^2([a,b],\mathbb{R}^{n+1})$ to some $H\in\mathcal{W}$.

Next, we prove that $G=H$. From the boundedness of $\mathcal{T}$ and U and from the hypotheses on (f,F^{β}) (Inequality (3.1)), we get that

$$\|G_k(\tau)-H_k(\tau)\|_{\mathbb{R}^{n+1}} \leq L\|x_k(\tau)-x^*(\tau)\|_{\mathbb{R}^n},$$

for almost all $\tau\in[a,b]$ and all $k\in\mathbb{N}$, for some constant $L\geq 0$. Thanks to the pointwise convergence of $(x_k)_{k\in\mathbb{N}}$ to x^*, we deduce that

$$\lim_{k\to\infty}\int_a^b \langle\varphi(\tau),G_k(\tau)-H_k(\tau)\rangle_{\mathbb{R}^{n+1}}\,d\tau = 0$$

for all φ in $\mathrm{L}^2([a,b],\mathbb{R}^{n+1})$. On the other hand, we obtain from the weak star and weak convergences that

$$\lim_{k\to\infty}\int_a^b \langle\varphi(\tau),G_k(\tau)-H_k(\tau)\rangle_{\mathbb{R}^{n+1}}\,d\tau = \int_a^b \langle\varphi(\tau),G(\tau)-H(\tau)\rangle_{\mathbb{R}^{n+1}}\,d\tau$$

for all $\varphi\in\mathrm{L}^2([a,b],\mathbb{R}^{n+1})$. We conclude that $G=H\in\mathcal{W}$. ∎

Remark 3.2. The regularity assumptions on f, F, g and φ introduced in Section 3.1 can be weakened for Theorem 3.1. Indeed only the continuity of f, F, g and φ and Inequality (3.1) (for f and F) are required in the above proof. Similarly, the convexity of C is a useless assumption that can be removed.

Remark 3.3. Theorem 3.1 can be extended to the case where we consider in Problem (OCP) some additional intermediate state constraints (not only at $t = a$ and $t = b$, but also at some times $c_i \in (a, b)$) or even running state constraints (that is, over the whole interval $[a, b]$). Indeed, in the above proof, one can easily see that the pointwise convergence of x_k to x^* allows to preserve such type of state constraints if appropriate continuous and closedness properties are satisfied.

Remark 3.4. This remark is devoted to sufficient conditions ensuring the boundedness of $\mathcal{T}$ in $\mathcal{C}$. If U is compact, if the terminal state constraint $g(x(a), x(b)) \in \mathrm{C}$ allows to bound the initial condition $x(a)$ (if the initial condition is fixed for example) and if f satisfies a global Lipschitz condition of type

$$\|f(x, u, t) - f(0_{\mathbb{R}^n}, u, t)\|_{\mathbb{R}^n} \leq L\|x\|_{\mathbb{R}^n},$$

for all $(x, u, t) \in \mathbb{R}^n \times \mathrm{U} \times [a, b]$, and some $L \geq 0$, then the fractional version of Gronwall lemma given in Proposition B.1 allows to prove that $\mathcal{T}$ is bounded in $\mathcal{C}$.

3.3. Sensitivity analysis of the Caputo fractional state equation. In this section we perform the sensitivity analysis of the Caputo fractional state equation in order to get differentiability results on the trajectory x with respect to perturbations on the control u and on the initial condition x_a. For this purpose, throughout this section, we fix a couple $(u, x_a) \in \mathrm{L}^\infty([a, b], \mathbb{R}^m) \times \mathbb{R}^n$ and we focus on the nonlinear Caputo fractional Cauchy problem (CP) given by

$$\text{(CP)} \qquad \begin{cases} {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}[x](t) = f(x(t), u(t), t), \quad \text{a.e. } t \in [a, b], \\ x(a) = x_a. \end{cases}$$

The results presented in this section are crucial in order to prove our main result (Theorem 3.12) stated in Section 3.4.

3.3.1. Admissible for globality couples. Let us recall some classical definitions of solution to (CP) and some fractional Cauchy-Lipschitz (or Picard-Lindelöf) results. We refer to [12, Section 3.2.2] for details.

Definition 3.5. *A pair (x, I) is said to be a (local) solution to (CP) if*
(i) *$I \subset \mathbb{R}$ is an interval such that $\{a\} \subsetneq I \subset [a, b]$;*
(ii) *$x \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}(I, \mathbb{R}^n)$ and x satisfies*

$$\begin{cases} {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}[x](t) = f(x(t), u(t), t), \quad \text{a.e. } t \in I, \\ x(a) = x_a, \end{cases}$$

or, equivalently, $x \in \mathcal{C}(I, \mathbb{R}^n)$ and x satisfies the integral representation

$$x(t) = x_a + \mathrm{I}^{\alpha}_{a+}[f(x, u, \cdot)](t),$$

for all $t \in I$.

Definition 3.6. *Let (x, I) and (x', I') be two local solutions to (CP). We say that (x', I') is an extension (resp. strict extension) of (x, I) if $I \subset I'$ (resp. $I \subsetneq I'$)*

and $x'(t) = x(t)$ for every $t \in I$. We say that (x, I) is a maximal solution to (CP) if it does not admit any strict extension. Finally, the couple (x, I) is said to be a global solution to (CP) if $I = [a, b]$.

PROPOSITION 3.7 ([12, Theorem 2]). *The Caputo fractional Cauchy problem (CP) admits a unique maximal solution which extends all other local solutions. This unique maximal solution will be denoted by $(x(\cdot, u, x_a), I(u, x_a))$.*

DEFINITION 3.8. *If the maximal solution $(x(\cdot, u, x_a), I(u, x_a))$ to (CP) is global, we say that the couple (u, x_a) is admissible for globality.*

PROPOSITION 3.9 ([12, Theorem 4]). *If the couple (u, x_a) is not admissible for globality, that is, if the maximal solution $(x(\cdot, u, x_a), I(u, x_a))$ of (CP) is not global, then $I(u, x_a)$ is not closed and $x(\cdot, u, x_a)$ is unbounded on $I(u, x_a)$.*

In the sequel we denote by $\mathcal{A}_{\mathcal{G}}$ the set of all couples $(u, x_a) \in \mathrm{L}^\infty([a, b], \mathbb{R}^m) \times \mathbb{R}^n$ that are admissible for globality.

3.3.2. Needle-like perturbation of the control. In this section we assume that $(u, x_a) \in \mathcal{A}_{\mathcal{G}}$ and our objective is to establish the differentiability of the state $x(\cdot, u, x_a)$ with respect to specific perturbations on the control u. For this purpose, we denote in the sequel by $\mathcal{L}[f(x(\cdot, u, x_a), u, \cdot)]$ the set of all Lebesgue points in $[a, b)$ of the essentially bounded function $t \mapsto f(x(t, u, x_a), u(t), t)$. Recall that $\mathcal{L}[f(x(\cdot, u, x_a), u, \cdot)]$ has a full Lebesgue measure equal to $b - a$.

Let $(s, v) \in \mathcal{L}[f(x(\cdot, u, x_a), u, \cdot)] \times \mathbb{R}^m$. For every $\delta \in [0, b - s)$, we introduce the so-called *needle-like perturbation* $u_\delta \in \mathrm{L}^\infty([a, b], \mathbb{R}^m)$ of u associated to (s, v) as

$$u_\delta(\tau) := \begin{cases} v & \text{if} \quad \tau \in [s, s + \delta), \\ u(\tau) & \text{if} \quad \tau \notin [s, s + \delta), \end{cases}$$

for almost every $\tau \in [a, b]$.

PROPOSITION 3.10. *For all $(s, v) \in \mathcal{L}[f(x(\cdot, u, x_a), u, \cdot)] \times \mathbb{R}^m$, there exists $\bar{\delta} > 0$ such that $(u_\delta, x_a) \in \mathcal{A}_{\mathcal{G}}$ for all $0 \leq \delta \leq \bar{\delta}$. Moreover:*

(i) *$x(\cdot, u_\delta, x_a)$ uniformly converges on $[a, b]$ to $x(\cdot, u, x_a)$ when $\delta \to 0$;*

(ii) *the quotient*

$$\frac{x(\cdot, u_\delta, x_a) - x(\cdot, u, x_a)}{\delta},$$

uniformly converges on $[s + \varsigma, b]$, for any $0 < \varsigma \leq b - s$, to $w_{(s,v)}(\cdot, u, x_a)$ when $\delta \to 0$, where $w_{(s,v)}(\cdot, u, x_a)$ is the unique maximal solution, which is moreover global, of the linear left RL fractional Cauchy problem given by

$$\begin{cases} \mathrm{D}^\alpha_{s+}[w](t) = \partial_1 f(x(t, u, x_a), u(t), t) \times w(t), \quad \text{a.e. } t \in [s, b], \\ \mathrm{I}^{1-\alpha}_{s+}[w](s) = f(x(s, u, x_a), v, s) - f(x(s, u, x_a), u(s), s). \end{cases}$$

The function $w_{(s,v)}(\cdot, u, x_a)$ is called the variation vector associated to (u, x_a) and (s, v).

Proof. The technical proof of Proposition 3.10 is detailed in Appendix A.1.2. ∎

The existence, uniqueness and globality of $w_{(s,v)}(\cdot, u, x_a) \in \mathrm{L}^1([s, b], \mathbb{R}^n)$ directly follow from [12, Theorem 1]. Moreover, recall that the product $\rho^\alpha_{s+} w_{(s,v)}(\cdot, u, x_a) \in \mathcal{C}([s, b], \mathbb{R}^n)$ (see [13, Theorem 6]), where $\rho^\alpha_{s+} : [a, b] \to \mathbb{R}$ stands for the *weight function* given by

$$\rho^\alpha_{s+}(t) := \Gamma(\alpha)(t - s)^{1-\alpha},$$

for all $t \in [s,b]$. Finally, the fractional Duhamel formula given in [12, Theorem 5] by

$$w_{(s,v)}(t,u,x_a) = \Phi(t,s) \times (f(x(s,u,x_a),v,s) - f(x(s,u,x_a),u(s),s)), \tag{3.2}$$

holds for all $t \in (s,b]$, where $\Phi(\cdot,\cdot)$ is the left RL fractional state-transition matrix (see [12, Definition 18] for details) associated to the essentially bounded matrix function $t \mapsto \partial_1 f(x(t,u,x_a),u(t),t) \in \mathbb{R}^{n\times n}$.

3.3.3. Perturbation of the initial condition. In this section we assume that $(u,x_a) \in \mathcal{A}_\mathcal{G}$ and our objective is to establish the differentiability of the state $x(\cdot,u,x_a)$ with respect to perturbations of the initial condition x_a.

PROPOSITION 3.11. *For all $y \in \mathbb{R}^n$, there exists $\bar\delta > 0$ such that $(u, x_a+\delta y) \in \mathcal{A}_\mathcal{G}$ for all $0 \le \delta \le \bar\delta$. Moreover:*

(i) *$x(\cdot,u,x_a+\delta y)$ uniformly converges on $[a,b]$ to $x(\cdot,u,x_a)$ when $\delta \to 0$;*

(ii) *the quotient*

$$\frac{x(\cdot,u,x_a+\delta y) - x(\cdot,u,x_a)}{\delta},$$

uniformly converges on $[a,b]$ to $w_y(\cdot,u,x_a)$ when $\delta \to 0$, where $w_y(\cdot,u,x_a)$ is the unique maximal solution, which is moreover global, of the linear left Caputo fractional Cauchy problem given by

$$\begin{cases} {}_{\rm c}\mathrm{D}^\alpha_{a+}[w](t) = \partial_1 f(x(t,u,x_a),u(t),t) \times w(t), \quad \text{a.e. } t \in [a,b], \\ w(a) = y. \end{cases}$$

The function $w_y(\cdot,u,x_a)$ is called the variation vector associated to (u,x_a) and y.

Proof. The technical proof of Proposition 3.11 is detailed in Appendix A.1.3. ∎

The existence, uniqueness and globality of $w_y(\cdot,u,x_a) \in \mathcal{C}([a,b],\mathbb{R}^n)$ directly follow from [12, Theorem 3]. Considering the function $\xi := w_y(\cdot,u,x_a) - y$, one can easily see that ξ is the unique maximal solution, which is moreover global, of the (nonhomogeneous) linear left RL fractional Cauchy problem given by

$$\begin{cases} \mathrm{D}^\alpha_{a+}[\xi](t) = \partial_1 f(x(t,u,x_a),u(t),t) \times (\xi(t)+y), \quad \text{a.e. } t \in [a,b], \\ \mathrm{I}^{1-\alpha}_{a+}[\xi](a) = 0_{\mathbb{R}^n}. \end{cases}$$

The existence, uniqueness and globality of $\xi \in \mathrm{L}^1([a,b],\mathbb{R}^n)$ directly follow from [12, Theorem 1]. Recall that the product $\rho^\alpha_{a+}\xi \in \mathcal{C}^a([a,b],\mathbb{R}^n)$ (see [13, Theorem 6]) and note that the fractional Duhamel formula given in [12, Theorem 5] applied to ξ allows to get that

$$w_y(t,u,x_a) = y + \int_a^t \Phi(t,\tau) \times \partial_1 f(x(\tau,u,x_a),u(\tau),\tau)\, d\tau \times y, \tag{3.3}$$

for every $t \in [a,b]$, where the notations ρ^α_{a+} and $\Phi(\cdot,\cdot)$ are introduced in Section 3.3.2.

3.4. Pontryagin Maximum Principle (PMP). Let us recall that the *normal cone* to C at a point $x \in \mathrm{C}$ is defined by

$$\mathcal{N}_\mathrm{C}[x] := \{z \in \mathbb{R}^j \mid \forall x' \in \mathrm{C},\ \langle z, x'-x\rangle_{\mathbb{R}^j} \le 0\}.$$

It is a nonempty closed convex cone containing $0_{\mathbb{R}^j}$. Note that several basics on convexity notions (such as *distance function*, *projection*, etc.) are recalled in Appendix A.2.2. Also recall that $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^j$ is said to be *submersive* at a point

$(x_a, x_b) \in \mathbb{R}^n \times \mathbb{R}^n$ if its differential at this point is surjective. We may now formulate the main result of the present paper.

THEOREM 3.12 (PMP). *If $(x^*, u^*) \in {}_c\mathcal{AC}^\alpha_{a+}([a,b],\mathbb{R}^n) \times \mathrm{L}^\infty([a,b],\mathbb{R}^m)$ is an optimal solution to Problem (OCP), then there exists a nontrivial couple (p, p^0), where $p \in \mathcal{AC}^\alpha_{b-}([a,b],\mathbb{R}^n)$ (called adjoint vector) and $p^0 \leq 0$, such that the following conditions hold:*

(i) *Fractional Hamiltonian system (or extremal equations):*

$$
\begin{aligned}
{}_c\mathrm{D}^\alpha_{a+}[x^*](t) &= \partial_3 H(x^*(t), u^*(t), p(t), p^0, t),\\
\mathrm{D}^\alpha_{b-}[p](t) &= \partial_1 H(x^*(t), u^*(t), p(t), p^0, t),
\end{aligned}
$$

for almost every $t \in [a,b]$, where the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times [a,b) \to \mathbb{R}$ associated to Problem (OCP) is defined by

$$
H(x,u,p,p^0,t) := \langle p, f(x,u,t)\rangle_{\mathbb{R}^n} + \frac{p^0}{\Gamma(\beta)}(b-t)^{\beta-1}F(x,u,t),
$$

for all $(x,u,p,p^0,t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times [a,b)$;

(ii) *Hamiltonian maximization condition:*

$$
u^*(t) \in \operatorname*{arg\,max}_{v \in \mathrm{U}} H(x^*(t), v, p(t), p^0, t),
$$

for almost every $t \in [a,b]$;

(iii) *Transversality conditions on the adjoint vector: if the function g is submersive at $(x^*(a), x^*(b))$, then the nontrivial couple (p, p^0) can be selected to satisfy*

$$
\begin{aligned}
\mathrm{I}^{1-\alpha}_{b-}[p](a) &= -p^0\partial_1\varphi(x^*(a),x^*(b)) - \partial_1 g(x^*(a),x^*(b))^\top \times \Psi,\\
\mathrm{I}^{1-\alpha}_{b-}[p](b) &= p^0\partial_2\varphi(x^*(a),x^*(b)) + \partial_2 g(x^*(a),x^*(b))^\top \times \Psi,
\end{aligned}
$$

where $-\Psi \in \mathcal{N}_{\mathrm{C}}[g(x^*(a),x^*(b))]$.

The technical (and quite long) proof of Theorem 3.12 (in its general form) is detailed in Appendix A.2. From a standard change of variable, we reduce Problem (OCP) to the case where there is no Lagrange cost (that is, with $F = 0$). Then, the proof is based on the sensitivity analysis performed in Section 3.3 and, in order to take into account the terminal state constraint $g(x(a), x(b)) \in \mathrm{C}$, from the application of Ekeland's variational principle (recalled in Proposition A.7) on a penalized functional. We feel that it is of interest to provide here a simple proof of Theorem 3.12 in two particular and simpler cases.

Proof of Theorem 3.12 in two particular and simpler cases. In this proof we will assume that there is no Lagrange cost (that is, $F = 0$) in Problem (OCP), that g is the identity function (in particular $j = 2n$ and g is submersive at any point) and $\mathrm{C} = \mathrm{C}_1 \times \mathbb{R}^n$ where:

(i) either $\mathrm{C}_1 = \{x_a\}$ for some $x_a \in \mathbb{R}^n$ fixed (corresponding to the case where the initial condition is fixed and the final condition is free in Problem (OCP));
(ii) either $\mathrm{C}_1 = \mathbb{R}^n$ (which corresponds to the case where the initial and final conditions are let free in Problem (OCP)).

Let $(x^*,u^*) \in {}_c\mathcal{AC}^\alpha_{a+}([a,b],\mathbb{R}^n) \times \mathrm{L}^\infty([a,b],\mathbb{R}^m)$ be an optimal solution to Problem (OCP). With the notations introduced in Section 3.3, note that $(u^*,x^*(a)) \in \mathcal{A}_\mathcal{G}$ and $x^* = x(\cdot,u^*,x^*(a))$. Let $(s,v) \in \mathcal{L}(f(x^*,u^*,\cdot)) \times \mathrm{U}$. From optimality of (x^*,u^*) and using Proposition 3.10, we know that

$$\frac{\varphi(x(a,u_\delta,x^*(a)),x(b,u_\delta,x^*(a))) - \varphi(x^*(a),x^*(b))}{\delta} \geq 0,$$

for all $\delta > 0$ sufficiently small, where u_δ stands for the needle-like perturbation of u^* associated to (s,v). From Proposition 3.10 and letting $\delta \to 0$, we get that

$$\big\langle -\partial_2\varphi(x^*(a),x^*(b)), w_{(s,v)}(b,u^*,x^*(a))\big\rangle_{\mathbb{R}^n} \leq 0.$$

From the Duhamel formula (3.2), recall that

$$w_{(s,v)}(b,u^*,x^*(a)) = \Phi(b,s) \times (f(x^*(s),v,s) - f(x^*(s),u^*(s),s)).$$

Hence, defining

$$p(s) := \Phi(b,s)^\top \times (-\partial_2\varphi(x^*(a),x^*(b))),$$

for almost all $s \in [a,b)$, we easily deduce from the above inequality the Hamiltonian maximization condition of Theorem 3.12. From the duality theorem given in [12, Theorem 7], we deduce that $p \in \mathcal{AC}^\alpha_{b-}([a,b],\mathbb{R}^n)$ is the unique maximal solution, which is moreover global, of the linear right RL fractional Cauchy problem given by

$$\begin{cases} \mathrm{D}^\alpha_{b-}[p](t) = \partial_1 f(x^*(t),u^*(t),t)^\top \times p(t), \quad \text{a.e. } t \in [a,b], \\ \mathrm{I}^{1-\alpha}_{b-}[p](b) = -\partial_2\varphi(x^*(a),x^*(b)). \end{cases}$$

Recall that the existence, uniqueness and globality of $p \in \mathrm{L}^1([a,b],\mathbb{R}^n)$ directly follow from the right counterpart of [12, Theorem 1]. Recall that the product $\rho^\alpha_{b-}p \in \mathcal{C}([a,b],\mathbb{R}^n)$ from the right counterpart of [13, Theorem 6], where $\rho^\alpha_{b-} : [a,b] \to \mathbb{R}$ stands for the weight function given by $\rho^\alpha_{b-}(t) := \Gamma(\alpha)(b-t)^{1-\alpha}$ for all $t \in [a,b]$. Then, from the above Cauchy problem, one can easily see that the fractional Hamiltonian system in Theorem 3.12 is satisfied. Then, let us define $p^0 := -1$ (in particular, the nontriviality of (p,p^0) is guaranteed). If $\mathrm{C}_1 = \{x_a\}$ for some $x_a \in \mathbb{R}^n$ fixed, the transversality conditions in Theorem 3.12 are both satisfied by considering $\Psi = (\Psi_1,\Psi_2)$ with $\Psi_1 := \partial_1\varphi(x^*(a),x^*(b)) - \mathrm{I}^{1-\alpha}_{b-}[p](a)$ and $\Psi_2 := 0_{\mathbb{R}^n}$. Indeed the normal cone to the entire space $\mathbb{R}^n$ (at any point) is reduced to the singleton $\{0_{\mathbb{R}^n}\}$ and the normal cone to the singleton $\{x_a\}$ (at x_a) is the entire space $\mathbb{R}^n$. So the proof is complete in the case where $\mathrm{C}_1 = \{x_a\}$ for some $x_a \in \mathbb{R}^n$ fixed. Now, let us assume that $\mathrm{C}_1 = \mathbb{R}^n$ and let $y \in \mathbb{R}^n$. From the optimality of (x^*,u^*) and using Proposition 3.11, we know that

$$\frac{\varphi(x(a,u^*,x^*(a)+\delta y),x(b,u^*,x^*(a)+\delta y)) - \varphi(x^*(a),x^*(b))}{\delta} \geq 0,$$

for all $\delta > 0$ sufficiently small. From Proposition 3.11 and letting $\delta \to 0$, we get that

$$\langle -\partial_1\varphi(x^*(a),x^*(b)),y\rangle_{\mathbb{R}^n} + \langle -\partial_2\varphi(x^*(a),x^*(b)), w_y(b,u^*,x^*(a))\rangle_{\mathbb{R}^n} \leq 0.$$

From the Duhamel formula (3.3), we know that

$$w_y(b,u^*,x^*(a)) = y + \int_a^b \Phi(b,\tau) \times \partial_1 f(x^*(\tau),u^*(\tau),\tau)\, d\tau \times y.$$

Thus we can rewrite the above inequality as

$$\left\langle -\partial_1\varphi(x^*(a),x^*(b)) - \partial_2\varphi(x^*(a),x^*(b)) + \int_a^b \mathrm{D}^{\alpha}_{b-}[p](\tau)\, d\tau, y \right\rangle_{\mathbb{R}^n} \leq 0.$$

From the equalities $\mathrm{D}^{\alpha}_{b-}[p] = -\frac{d}{dt}[\mathrm{I}^{1-\alpha}_{b-}[p]]$ and $\mathrm{I}^{1-\alpha}_{b-}[p](b) = -\partial_2\varphi(x^*(a),x^*(b))$, we deduce that

$$\left\langle -\partial_1\varphi(x^*(a),x^*(b)) + \mathrm{I}^{1-\alpha}_{b-}[p](a), y \right\rangle_{\mathbb{R}^n} \leq 0.$$

Since the above inequality is true for all $y \in \mathbb{R}^n$, it is clear that the conditon $\mathrm{I}^{1-\alpha}_{b-}[p](a) = \partial_1\varphi(x^*(a),x^*(b))$ follows. Hence, the transversality conditions in Theorem 3.12 are both satisfied by considering $\Psi = (\Psi_1, \Psi_2)$ with $\Psi_1 := 0_{\mathbb{R}^n}$ and $\Psi_2 := 0_{\mathbb{R}^n}$. This completes the proof in the case where $\mathrm{C}_1 = \mathbb{R}^n$. ∎

We end this section with some comments.

Remark 3.13. The nontrivial couple (p, p^0) in Theorem 3.12, which is a Lagrange multiplier, is defined up to a positive multiplicative scalar. Defining as usual an *extremal* as a quadruple (x, u, p, p^0) solution to the extremal equations, an extremal is said to be *normal* whenever $p^0 \neq 0$ and *abnormal* whenever $p^0 = 0$. In the normal case $p^0 \neq 0$, it is usual to normalize the Lagrange multiplier so that $p^0 = -1$.

Remark 3.14. Theorem 3.12 encompasses the historical Pontryagin maximum principle by considering the case where $\alpha = \beta = 1$.

Remark 3.15. In Problem (OCP), the state fractional equation ${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}[x](t) = f(x(t),u(t),t)$ involves the left Caputo fractional operator ${}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}$, while the adjoint equation $\mathrm{D}^{\alpha}_{b-}[p](t) = \partial_1 H(x^*(t),u^*(t),p(t),p^0,t)$ in Theorem 3.12 depends on the right RL fractional derivative D^{α}_{b-}. Accordingly, Problem (OCP) depends on the terminal conditions $x(a)$ and $x(b)$, while the transversality conditions on the adjoint vector in Theorem 3.12 involve $\mathrm{I}^{1-\alpha}_{b-}[p](a)$ and $\mathrm{I}^{1-\alpha}_{b-}[p](b)$. Finally, note that the adjoint vector $p \in \mathcal{AC}^{\alpha}_{b-}([a,b],\mathbb{R}^n)$ may admit a singularity at $t = b$ (see the right counterpart of Remark 2.8 and see Section 4 for two examples).

Remark 3.16. Our strategy in order to prove Theorem 3.12 in its general form is based on Ekeland's variational principle (Proposition A.7) and thus requires the closedness of U in order to define the corresponding penalized functional on a complete metric space (see details in Appendix A.2). In the two particular and simpler cases considered above, Ekeland's variational principle is not required and the closedness of U is a useless assumption that can be removed.

Remark 3.17. Let us describe some typical situations of terminal state constraint $g(x(a),x(b)) \in \mathrm{C}$ in Problem (OCP), and the corresponding transversality conditions in Theorem 3.12:

- If the terminal points are fixed in Problem (OCP), one may consider g as the identity function and $\mathrm{C} = \{x_a\} \times \{x_b\}$ where x_a, $x_b \in \mathbb{R}^n$ are the fixed terminal points. In that case, the transversality conditions in Theorem 3.12 do not provide any additional information.
- If the initial point is fixed and if the final point is free in Problem (OCP), one may consider g as the identity function and $\mathrm{C} = \{x_a\} \times \mathbb{R}^n$ where $x_a \in \mathbb{R}^n$ is the fixed initial point. In that case, the nontriviality of the couple (p, p^0) and the transversality conditions in Theorem 3.12 imply that $p^0 \neq 0$ (which we normalize to $p^0 = -1$, see Remark 3.13) and $\mathrm{I}^{1-\alpha}_{b-}[p](b) = -\partial_2\varphi(x^*(a),x^*(b))$.

- If the initial point is fixed and the final point is subject to inequality constraints $G_i(x(b)) \geq 0$ for $i = 1, \ldots, k$, one may consider $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n+k}$, $g(x_a, x_b) = (x_a, G(x_b))$ where $G = (G_1, \ldots, G_k) : \mathbb{R}^n \to \mathbb{R}^k$ and $\mathrm{C} = \{x_a\} \times (\mathbb{R}_+)^k$. If G is of class $\mathcal{C}^1$ and is submersive at any point $x_b \in G^{-1}((\mathbb{R}_+)^k)$, then the transversality conditions in Theorem 3.12 can be written as

$$\mathrm{I}_{b-}^{1-\alpha}[p](b) = p^0 \partial_2 \varphi(x^*(a), x^*(b)) + \sum_{i=1}^{k} \lambda_i \nabla G_i(x^*(b)),$$

  for some $\lambda_i \geq 0$, $i = 1, \ldots, k$.
- If there is no Mayer cost (that is, $\varphi = 0$) and the periodic condition $x(a) = x(b)$ is considered in Problem (OCP), one may consider $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $g(x_a, x_b) = x_b - x_a$ and $\mathrm{C} = \{0_{\mathbb{R}^n}\}$. In that case, the transversality conditions in Theorem 3.12 yield that $\mathrm{I}_{b-}^{1-\alpha}[p](a) = \mathrm{I}_{b-}^{1-\alpha}[p](b)$.

We point out that, in all examples above, the function g is indeed a submersion.

Remark 3.18. Let us assume that the Hamiltonian H considered in Theorem 3.12 is differentiable with respect to its second variable (for example, if f and F are). In that case, and if U is convex, the Hamiltonian maximization condition in Theorem 3.12 implies the (weaker) nonnegative Hamiltonian gradient condition given by

$$\langle \partial_2 H(x^*(t), u^*(t), p(t), p^0, t), v - u^*(t) \rangle_{\mathbb{R}^m} \leq 0,$$

for all $v \in \mathrm{U}$ and for almost every $t \in [a, b]$. Similarly, if $\mathrm{U} = \mathbb{R}^m$ (that is, no control constraint in Problem (OCP)), then the Hamiltonian maximization condition in Theorem 3.12 implies the (weaker) null Hamiltonian gradient condition given by

$$\partial_2 H(x^*(t), u^*(t), p(t), p^0, t) = 0_{\mathbb{R}^m},$$

for almost every $t \in [a, b]$.

Remark 3.19. Note that an extension of Theorem 3.12 for a parameterized version of Problem (OCP) (that is, depending on a vectorial parameter z to be optimized) can be easily derived by adding the fractional state equation ${}_{\mathrm{c}}\mathrm{D}_{a+}^{\alpha}[z](t) = 0$.

Remark 3.20. If $\beta = 1$ in Problem (OCP), then we recover in Theorem 3.12 the very standard Hamiltonian given by $H(x, u, p, p^0, t) := \langle p, f(x, u, t) \rangle_{\mathbb{R}^n} + p^0 F(x, u, t)$. On the other hand, if $\beta \neq 1$ in Problem (OCP), then the Hamiltonian is not standard since it is given by $H(x, u, p, p^0, t) := \langle p, f(x, u, t) \rangle_{\mathbb{R}^n} + \frac{p^0}{\Gamma(\beta)}(b - t)^{\beta - 1} F(x, u, t)$. This phenomenon is due to the nonlocality of the fractional operator I_{a+}^{β} but it is natural since the fractional Lagrange cost can be rewritten as

$$\mathrm{I}_{a+}^{\beta}[F(x, u, \cdot)](b) = \mathrm{I}_{a+}^{1}\left[\frac{(b - \cdot)^{\beta - 1}}{\Gamma(\beta)} F(x, u, \cdot)\right](b).$$

In particular, if $\beta \neq 1$, the Hamiltonian considered in Theorem 3.12 may be not autonomous, even if f and F are.

4. Two examples. This section is devoted to the application of Theorem 3.12 to solve simple examples. We focus on the fractional versions ($0 < \alpha \leq 1$ and $\beta \geq \alpha$) of two basic problems from the classical theory.

4.1. A fractional linear-quadratic problem. In this section we consider the fractional linear-quadratic problem given by

$$\left\{\begin{array}{ll} \text{minimize} & \dfrac{1}{2}\mathrm{I}^{\beta}_{0+}[u^2](T), \\ \text{subject to} & x=(x_1,x_2)\in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{0+}([0,T],\mathbb{R}^2),\ u\in \mathrm{L}^{\infty}([0,T],\mathbb{R}), \\ & {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{0+}[x_1](t)=x_2(t), \quad \text{a.e. } t\in[0,T], \\ & {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{0+}[x_2](t)=u(t), \quad \text{a.e. } t\in[0,T], \\ & x(0)=(a,b), \\ & x(T)=(0,0), \end{array}\right. \tag{4.1}$$

where $T>0$ and $(a,b)\neq(0,0)$. Problem (4.1) corresponds to a fractional version of the classical *parking problem* (or *double integrator problem*). Let us assume that Problem (4.1) admits an optimal solution denoted by (x^*,u^*). Then, there exists a nontrivial couple $(p,p^0)\in \mathcal{AC}^{\alpha}_{T-}([0,T],\mathbb{R}^2)\times\mathbb{R}$ such that all necessary conditions provided in Theorem 3.12 are satisfied, where the Hamiltonian is given by

$$H(x_1,x_2,u,p_1,p_2,p^0,t)=p_1x_2+p_2u+\frac{p^0}{2}\frac{(T-t)^{\beta-1}}{\Gamma(\beta)}u^2.$$

The fractional Hamiltonian system gives $\mathrm{D}^{\alpha}_{T-}[p_1]=0$ and $\mathrm{D}^{\alpha}_{T-}[p_2]=p_1$ leading to

$$p_1(t)=\frac{(T-t)^{\alpha-1}}{\Gamma(\alpha)}c_2 \quad\text{and}\quad p_2(t)=\frac{(T-t)^{\alpha-1}}{\Gamma(\alpha)}c_1+\frac{(T-t)^{2\alpha-1}}{\Gamma(2\alpha)}c_2,$$

for every $t\in[0,T)$, for some constants c_1, $c_2\in\mathbb{R}$. Since there is no control constraint in Problem (OCP) and from Remark 3.18, the null Hamiltonian gradient condition gives

$$p_2(t)=-p^0\frac{(T-t)^{\beta-1}}{\Gamma(\beta)}u^*(t),$$

for almost all $t\in[0,T]$. From the nontriviality of the couple (p,p^0), one can easily see that $p^0\neq 0$ and normalize the Lagrange multiplier so that $p^0=-1$ (Remark 3.13). Thus

$$u^*(t)=\Gamma(\beta)(T-t)^{1-\beta}p_2(t)=c_1\xi_1(t)+c_2\xi_2(t),$$

where $\xi_k(t):=\frac{\Gamma(\beta)}{\Gamma(k\alpha)}(T-t)^{k\alpha-\beta}$ for all $t\in[0,T]$ and $k=1,2$. Our objective is now to determine the real constants c_1, c_2. From the fractional state equation and the initial condition $x^*(0)=(a,b)$, we get $x_2^*=b+\mathrm{I}^{\alpha}_{0+}[u^*]$ and $x_1^*=a+b\mathrm{I}^{\alpha}_{0+}[1]+\mathrm{I}^{2\alpha}_{0+}[u^*]$. From the final condition $x^*(T)=(0,0)$, we obtain that the couple (c_1,c_2) is solution to the linear system given by

$$\begin{pmatrix} \mathrm{I}^{2\alpha}_{0+}[\xi_1](T) & \mathrm{I}^{2\alpha}_{0+}[\xi_2](T) \\ \mathrm{I}^{\alpha}_{0+}[\xi_1](T) & \mathrm{I}^{\alpha}_{0+}[\xi_2](T) \end{pmatrix}\times\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}=\begin{pmatrix} -a-b\mathrm{I}^{\alpha}_{0+}[1](T) \\ -b \end{pmatrix}.$$

4.2. A fractional Zermelo problem. In this section we consider the fractional version of the classical *Zermelo problem* given by

$$
\left\{
\begin{array}{ll}
\text{minimize} & -x_1(T), \\
\\
\text{subject to} & x = (x_1, x_2) \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{0+}([0,T],\mathbb{R}^2),\ u \in \mathrm{L}^{\infty}([0,T],\mathbb{R}), \\
& {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{0+}[x_1](t) = x_2(t) + \cos(u(t)), \quad \text{a.e. } t \in [0,T], \\
& {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{0+}[x_2](t) = \sin(u(t)), \quad \text{a.e. } t \in [0,T], \\
& x(0) = (0,0), \\
& u(t) \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}], \quad \text{a.e. } t \in [0,T],
\end{array}
\right. \tag{4.2}
$$

where $T > 0$. Let us assume that Problem (4.2) admits an optimal solution denoted by (x^*, u^*). Then, there exists a nontrivial couple $(p, p^0) \in \mathcal{AC}^{\alpha}_{T-}([0,T],\mathbb{R}^2) \times \mathbb{R}$ such that all necessary conditions provided in Theorem 3.12 are satisfied, where the Hamiltonian is given by

$$H(x_1, x_2, u, p_1, p_2, t) = p_1(x_2 + \cos(u)) + p_2 \sin(u).$$

The fractional Hamiltonian system gives $\mathrm{D}^{\alpha}_{T-}[p_1] = 0$ and $\mathrm{D}^{\alpha}_{T-}[p_2] = p_1$ leading to

$$p_1(t) = \frac{(T-t)^{\alpha-1}}{\Gamma(\alpha)} c_2 \quad \text{and} \quad p_2(t) = \frac{(T-t)^{\alpha-1}}{\Gamma(\alpha)} c_1 + \frac{(T-t)^{2\alpha-1}}{\Gamma(2\alpha)} c_2,$$

for every $t \in [0,T)$, where $c_1 = \mathrm{I}^{1-\alpha}_{T-}[p_2](T) = 0$ and $c_2 = \mathrm{I}^{1-\alpha}_{T-}[p_1](T) = -p^0$ from the transversality conditions. From the nontriviality of the couple (p, p^0), one can easily see that $p^0 \neq 0$ and normalize the Lagrange multiplier so that $p^0 = -1$ (Remark 3.13). The Hamiltonian maximization condition gives

$$u^*(t) \in \underset{v \in [-\frac{\pi}{2}, \frac{\pi}{2}]}{\arg\max} \left\langle \begin{pmatrix} \cos(v) \\ \sin(v) \end{pmatrix}, \begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix} \right\rangle_{\mathbb{R}^2},$$

for almost all $t \in [0,T]$. With the classical Cauchy-Schwarz inequality we get that

$$\begin{pmatrix} \cos(u^*(t)) \\ \sin(u^*(t)) \end{pmatrix} = \frac{1}{\|(p_1(t), p_2(t))\|_{\mathbb{R}^2}} \begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix},$$

leading to $\tan(u^*(t)) = \frac{p_2(t)}{p_1(t)}$ which gives

$$u^*(t) = \arctan\left(\frac{p_2(t)}{p_1(t)}\right) = \arctan\left(\frac{\Gamma(\alpha)}{\Gamma(2\alpha)}(T-t)^{\alpha}\right),$$

for almost all $t \in [0,T]$.

5. Conclusion. As a conclusion we present perspectives and forthcoming works which may follow the present paper.

5.1. Works in progress. We first consider the framework of Theorem 3.12 in the classical case $\alpha = \beta = 1$ and we recall the standard notion of *maximized Hamiltonian* $\mathcal{H} : [a,b] \to \mathbb{R}$ given by

$$\mathcal{H}(t) := H(x^*(t), u^*(t), p(t), p^0, t),$$

for almost every $t \in [a,b]$. If H is differentiable with respect to t with $\partial_t H$ continuous (for example if f and F are), it is well-known that $\mathcal{H}$ is equal almost everywhere on $[a,b]$ to an absolutely continuous function (denoted similarly) which satisfies

$$\dot{\mathcal{H}}(t) = \partial_t H(x^*(t), u^*(t), p(t), p^0, t),$$

for almost every $t \in [a,b]$. This property is known as the *Hamiltonian (absolute) continuity* and we roughly say that the *total derivative of the Hamiltonian is equal to its partial derivative*. In particular, if the problem is moreover autonomous, then $\mathcal{H}$ is constant. We refer for instance to [23, Theorem 2.6.3 p.73] for details in the classical theory. This property constitutes an additional necessary optimality condition and it is particularly interesting in order to deal with optimal control problems with free final time (which encompass minimal time problems for example). Indeed, it is well-known that a change of time variable allows to convert a free final time problem into an autonomous fixed final time problem. Then, from the constancy of the corresponding maximized Hamiltonian, combined with a parameterized version of the classical PMP, the classical transversality condition on the optimal free final time is derived. We refer for instance to [30, Chapter 14] for details in the classical theory.

To the best of our knowledge, no Hamiltonian continuity in the fractional case $0 < \alpha \leq 1$ and $\beta \geq \alpha$ has been announced, proved or even refuted in the literature. The priority for the authors of the present paper is to deal with this issue. Preliminary results have been obtained and a complete study will be published in the near future. Our final objective towards this project, with the help of a parameterized version of Theorem 3.12 (Remark 3.19), is to establish a version of the Pontryagin maximum principle that handles Problem (OCP) with a free final time. Let us point out that some earlier works like [9, 41, 45, 52] already deal with fractional optimal control problems with free final time.

We conclude by mentioning that, similarly to the classical case, the Pontryagin maximum principle stated in Theorem 3.12 only allows to solve explicitly a few number of basic Caputo fractional optimal control problems (see Section 4 for two simple examples). Nevertheless, like in the classical theory, Theorem 3.12 clearly induces a numerical way in order to solve them by adapting to the fractional framework the well-known *shooting methods* (which are indirect numerical methods reducing classical optimal control problems into two-point boundary value problems that can be solved by Newton's methods for example). We refer for instance to [10, Section 3.3] for details on shooting methods in the classical case. In order to extend this method to the fractional framework considered in this paper, one may consider Grünwald-Letnikov discretizations of the RL and Caputo fractional derivatives (see, e.g., [43, p.43 and p.200] or [49] for details). This challenge will also be the subject of a forthcoming paper. We refer to [3, 4, 7, 31, 45] for previous numerical studies on fractional optimal control problems.

5.2. Other possible extensions. In Theorem 3.1 we provide an existence result for Problem (OCP) based on the classical Filippov's approach (see, e.g., [15, 24, 39]). Extensions and other approaches are well-known in the classical theory (see, e.g., [19, Chapter 9]). Fractional versions of these results can be considered as perspectives. Let us mention here the works of Kamocki in [33, Theorem 17] and [35, Theorem 4.2] that provide existence results for some linear RL and Caputo fractional optimal control problems.

Concerning the Pontryagin maximum principle, many extensions of Theorem 3.12 are possible. We may for instance:

- consider $\alpha > 1$ and/or $\beta < \alpha$;
- consider fractional multi-order $\alpha = (\alpha_i)_{i=1,\dots,n}$ as in [12];
- rule out the closedness assumption on U (with a different approach than Ekeland's variational principle, see Remark 3.16);
- introduce time dependence $\mathrm{U} = \mathrm{U}(t)$ on the control constraint set;
- consider intermediate state constraints (that is, not only at $t = a$ or $t = b$, but also at some times $c_i \in (a, b)$) or running state constraints (over the whole interval $[a, b]$);
- etc.

We conclude this section by pointing out that we deal in this paper with a general optimal control problem involving a Caputo fractional state equation. In our opinion, a relevant extension to Theorem 3.12 would be to consider a RL fractional state equation

$$\mathrm{D}^{\alpha}_{a+}[x](t) = f(x(t), u(t), t),$$

driven by the RL fractional operator D^{α}_{a+}. A lot of difficulties are expected in that framework since the corresponding trajectories $x \in \mathcal{AC}^{\alpha}_{a+}([a, b], \mathbb{R}^n)$ are not bounded since singularities occur at $t = a$. As a consequence, most of estimates used in this paper cannot be extended. In that framework, one should consider a Mayer cost and a state constraint that involve the RL initial and final conditions $\mathrm{I}^{1-\alpha}_{a+}[x](a)$ and $\mathrm{I}^{1-\alpha}_{a+}[x](b)$.

Appendix A. Proofs of Section 3. This appendix is devoted to the detailed proofs of Section 3 (precisely Propositions 3.10 and 3.11 and Theorem 3.12). For this purpose we have to fix for the whole section some $1 \leq r_\alpha < \infty$ such that $r_\alpha = 1$ if $\alpha = 1$, and $r_\alpha > \frac{1}{\alpha}$ if $\alpha \in (0, 1)$ and to introduce the notations

$$M^1_\alpha := \begin{cases} 1 & \text{if} \quad \alpha = 1, \\ \dfrac{1}{\Gamma(\alpha)} \left(\dfrac{(b-a)^{r'_\alpha(\alpha-1)+1}}{r'_\alpha(\alpha-1)+1} \right)^{1/r'_\alpha} & \text{if} \quad \alpha \in (0,1), \end{cases}$$

and

$$M^2_\alpha := \begin{cases} 1 & \text{if} \quad \alpha = 1, \\ (b-a)^{\alpha-(1/r_\alpha)} \left(\dfrac{\Gamma(r'_\alpha(\alpha-1)+1)^2}{\Gamma(2r'_\alpha(\alpha-1)+2)} \right)^{1/r'_\alpha} & \text{if} \quad \alpha \in (0,1), \end{cases}$$

where $r'_\alpha := \frac{r_\alpha}{r_\alpha - 1}$ denotes the Lebesgue conjugate of r_α (satisfying $\frac{1}{r_\alpha} + \frac{1}{r'_\alpha} = 1$).

A.1. Proofs of Section 3.3. This section is devoted to the proofs of Propositions 3.10 and 3.11 detailed respectively in Appendices A.1.2 and A.1.3. The notations introduced in the next preliminary Appendix A.1.1 will be required.

A.1.1. Preliminaries on stability and continuity results. We first prove that the set $\mathcal{A}_\mathcal{G}$ is open (Proposition A.1) and establish the continuity of the state $x(\cdot, u, x_a)$ with respect to the couple $(u, x_a) \in \mathcal{A}_\mathcal{G}$ (Proposition A.2). For this purpose, we fix $(u, x_a) \in \mathcal{A}_\mathcal{G}$ in the whole section and for every $R \geq \|u\|_{\mathrm{L}^\infty}$, we introduce the set

$$\mathrm{K}_R := \{(y, v, t) \in \mathbb{R}^n \times \mathbb{R}^m \times [a, b] \mid \|y - x(t, u, x_a)\|_{\mathbb{R}^n} \leq 1 \text{ and } \|v\|_{\mathbb{R}^m} \leq R\}.$$

Note that any convex combination of two elements (y_1, v_1, t), $(y_2, v_2, t) \in \mathrm{K}_R$ belongs to K_R. Moreover, from continuity of $x(\cdot, u, x_a)$ on $[a, b]$, K_R is a compact subset of $\mathbb{R}^n \times \mathbb{R}^m \times [a, b]$. From the assumptions on the dynamic f (Section 3.1), there exists a nonnegative constant $L_R \geq 0$ such that

$$\|\partial_1 f(y, v, t)\|_{\mathbb{R}^{n\times n}} \leq L_R, \tag{A.1}$$

for all $(y, v, t) \in \mathrm{K}_R$, and

$$\|f(y_2, v_2, t) - f(y_1, v_1, t)\|_{\mathbb{R}^n} \leq L_R(\|y_2 - y_1\|_{\mathbb{R}^n} + \|v_2 - v_1\|_{\mathbb{R}^m}), \tag{A.2}$$

for all (y_1, v_1, t), $(y_2, v_2, t) \in \mathrm{K}_R$.

Proposition A.1. *For every $R \geq \|u\|_{\mathrm{L}^\infty}$, there exists $\eta_R > 0$ such that the neighborhood of (u, x_a) given by*

$$\mathrm{N}_R := \Big[\overline{\mathrm{B}}_{\mathrm{L}^1}(u, \eta_R) \cap \overline{\mathrm{B}}_{\mathrm{L}^\infty}(0, R)\Big] \times \overline{\mathrm{B}}_{\mathbb{R}^n}(x_a, \eta_R),$$

is contained in $\mathcal{A}_\mathcal{G}$. Moreover, it holds that $(x(\tau, u', x'_a), u'(\tau), \tau) \in \mathrm{K}_R$ for almost every $\tau \in [a, b]$ and all $(u', x'_a) \in \mathrm{N}_R$.

Proof. Let $R \geq \|u\|_{\mathrm{L}^\infty}$ and let $0 < \eta_R < 1$ be such that

$$\Big(\eta_R + M^1_\alpha L_R (2R)^{1/r'_\alpha} \eta_R^{1/r_\alpha}\Big) E_{\alpha,1}(L_R(b-a)^\alpha) < 1,$$

where r_α and M^1_α are defined at the beginning of Appendix A.1, and $E_{\alpha,1}$ denotes the classical Mittag-Leffler function (see Appendix B for details). Let $(u', x'_a) \in \mathrm{N}_R$. Our aim is to prove that $(u', x'_a) \in \mathcal{A}_\mathcal{G}$, that is, $b \in I(u', x'_a)$. By contradiction, let us assume that the set

$$A := \{t \in I(u', x'_a) \mid \|x(t, u', x'_a) - x(t, u, x_a)\|_{\mathbb{R}^n} > 1\},$$

is not empty and let $t_0 := \inf A$. It holds that $\|x(t_0, u', x'_a) - x(t_0, u, x_a)\|_{\mathbb{R}^n} \geq 1$ by continuity. Since $\|x(a, u', x'_a) - x(a, u, x_a)\|_{\mathbb{R}^n} = \|x'_a - x_a\|_{\mathbb{R}^n} \leq \eta_R < 1$, we get that $t_0 > a$. Moreover, one has $\|x(t, u', x'_a) - x(t, u, x_a)\|_{\mathbb{R}^n} \leq 1$ for every $t \in [a, t_0]$. Therefore $(x(\tau, u', x'_a), u'(\tau), \tau)$ and $(x(\tau, u, x_a), u(\tau), \tau)$ belong to K_R for almost every $\tau \in [a, t_0]$. From integral representations we obtain

$$x(t, u', x'_a) - x(t, u, x_a) = x'_a - x_a + \mathrm{I}^\alpha_{a+}[f(x(\cdot, u', x'_a), u', \cdot) - f(x(\cdot, u, x_a), u, \cdot)](t),$$

for every $t \in [a, t_0]$. It follows from Inequality (A.2) that

$$\begin{aligned}\|x(t, u', x'_a) - x(t, u, x_a)\|_{\mathbb{R}^n} \leq & \|x'_a - x_a\|_{\mathbb{R}^n} \\ & + \frac{L_R}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \|u'(\tau) - u(\tau)\|_{\mathbb{R}^m}\, d\tau \\ & + \frac{L_R}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \|x(\tau, u', x'_a) - x(\tau, u, x_a)\|_{\mathbb{R}^n}\, d\tau,\end{aligned}$$

for every $t \in [a, t_0]$. Then, from the classical Hölder inequality applied to the second right-hand side term and since $(u', x'_a) \in \mathrm{N}_R$, we get that

$$\begin{aligned}\|x(t, u', x'_a) - x(t, u, x_a)\|_{\mathbb{R}^n} \leq & \eta_R + M^1_\alpha L_R (2R)^{1/r'_\alpha} \eta_R^{1/r_\alpha} \\ & + L_R \mathrm{I}^\alpha_{a+}\Big[\|x(\cdot, u', x'_a) - x(\cdot, u, x_a)\|_{\mathbb{R}^n}\Big](t),\end{aligned}$$

for every $t \in [a, t_0]$. From the fractional version of the Gronwall lemma given in Proposition B.1, we deduce that

$$\|x(t,u',x'_a) - x(t,u,x_a)\|_{\mathbb{R}^n} \leq \Big(\eta_R + M^1_\alpha L_R (2R)^{1/r'_\alpha} \eta_R^{1/r_\alpha}\Big) E_{\alpha,1}(L_R(b-a)^\alpha) < 1,$$

for every $t \in [a, t_0]$, which raises a contradiction at $t = t_0$. Therefore A is empty. We conclude that $x(\cdot, u', x'_a)$ is bounded on $I(u', x'_a)$, then $b \in I(u', x'_a)$ from Proposition 3.9. Moreover, since A is empty, we also conclude that

$$\|x(t,u',x'_a) - x(t,u,x_a)\|_{\mathbb{R}^n} \leq 1$$

for every $t \in [a, b]$, and thus $(x(\tau, u', x'_a), u'(\tau), \tau) \in \mathrm{K}_R$ for almost every $\tau \in [a, b]$. ∎

For every $R \geq \|u\|_{\mathrm{L}^\infty}$, we endow the neighborhood N_R of (u, x_a) with the basic distance

$$\mathrm{d}_{\mathrm{L}^1\times\mathbb{R}^n}\Big((u',x'_a),(u'',x''_a)\Big) := \|u'' - u'\|_{\mathrm{L}^1} + \|x''_a - x'_a\|_{\mathbb{R}^n}, \tag{A.3}$$

for all (u', x'_a), $(u'', x''_a) \in \mathrm{N}_R$. We conclude this section with the following continuity result.

Proposition A.2. *For every $R \geq \|u\|_{\mathrm{L}^\infty}$, there exists $C_R \geq 0$ such that*

$$\|x(\cdot,u'',x''_a) - x(\cdot,u',x'_a)\|_{\mathcal{C}} \leq C_R(\|u'' - u'\|^{1/r_\alpha}_{\mathrm{L}^1} + \|x''_a - x'_a\|_{\mathbb{R}^n}),$$

for all (u', x'_a), $(u'', x''_a) \in \mathrm{N}_R$. In particular, the mapping

$$(u', x'_a) \in \mathrm{N}_R \subset \mathcal{A}_\mathcal{G} \longmapsto x(\cdot, u', x'_a) \in \mathcal{C}([a,b], \mathbb{R}^n),$$

is continuous.

Proof. Let (u', x'_a), $(u'', x''_a) \in \mathrm{N}_R$. We know that $(x(\tau, u'', x''_a), u''(\tau), \tau)$ and $(x(\tau, u', x'_a), u'(\tau), \tau)$ are elements of K_R for almost every $\tau \in [a, b]$ (Proposition A.1). Following the same arguments as in the proof of Proposition A.1, it follows that

$$\begin{aligned}\|x(t,u'',x''_a) - x(t,u',x'_a)\|_{\mathbb{R}^n} \\ \leq \Big(\|x''_a - x'_a\|_{\mathbb{R}^n} + M^1_\alpha L_R(2R)^{1/r'_\alpha}\|u'' - u'\|^{1/r_\alpha}_{\mathrm{L}^1}\Big) E_{\alpha,1}(L_R(b-a)^\alpha),\end{aligned}$$

for every $t \in [a, b]$. Setting $C_R := (1 + M^1_\alpha L_R(2R)^{1/r'_\alpha}) E_{\alpha,1}(L_R(b-a)^\alpha)$ concludes the proof. ∎

A.1.2. Proof of Proposition 3.10. Let $(s, v) \in \mathcal{L}[f(x(\cdot, u, x_a), u, \cdot)] \times \mathbb{R}^m$ and let $R = \max(\|u\|_{\mathrm{L}^\infty}, \|v\|_{\mathbb{R}^m})$. Since $\|v\|_{\mathbb{R}^m} \leq R$ and since $\|u_\delta - u\|_{\mathrm{L}^1} \leq 2R\delta$ for all $\delta \in [0, b - s)$, we deduce from Proposition A.1 that there exists $\bar\delta > 0$ such that $(u_\delta, x_a) \in \mathrm{N}_R \subset \mathcal{A}_\mathcal{G}$ for all $0 \leq \delta \leq \bar\delta$. Then the first item of Proposition 3.10 easily follows from Proposition A.2. Before proving the second item of Proposition 3.10, we need the following Lemmas A.3, A.4 and A.5. For the ease of notations, in this section, we denote by $x_\delta := x(\cdot, u_\delta, x_a)$ for all $0 \leq \delta \leq \bar\delta$, by $x := x(\cdot, u, x_a)$ and by $w := w_{(s,v)}(\cdot, u, x_a)$.

Lemma A.3. *There exists $\mu^1_R \geq 0$ such that*

$$\|x_\delta - x\|_{\mathcal{C}} \leq \mu^1_R \delta^\alpha,$$

for all $0 \leq \delta \leq \bar\delta$.

Proof. First, note that $\|x_\delta(t) - x(t)\|_{\mathbb{R}^n} = 0$ for all $t \in [a, s]$ and all $0 \leq \delta \leq \bar{\delta}$. Then, from integral representations, one has

$$x_\delta(t) - x(t) = \mathrm{I}^\alpha_{s+}[f(x_\delta, u_\delta, \cdot) - f(x, u, \cdot)](t),$$

for all $t \in [s, b]$ and all $0 \leq \delta \leq \bar{\delta}$. Since $(x_\delta(\tau), u_\delta(\tau), \tau)$ and $(x(\tau), u(\tau), \tau)$ belong to K_R for almost every $\tau \in [s, b]$ and all $0 \leq \delta \leq \bar{\delta}$ (since $(u_\delta, x_a) \in \mathrm{N}_R$ and from Proposition A.1), we get from Inequality (A.2) that

$$\begin{aligned}\|x_\delta(t) - x(t)\|_{\mathbb{R}^n} \leq & \frac{L_R}{\Gamma(\alpha)} \int_s^t (t-\tau)^{\alpha-1} \|u_\delta(\tau) - u(\tau)\|_{\mathbb{R}^m} \, d\tau \\ & + \frac{L_R}{\Gamma(\alpha)} \int_s^t (t-\tau)^{\alpha-1} \|x_\delta(\tau) - x(\tau)\|_{\mathbb{R}^n} \, d\tau,\end{aligned}$$

for all $t \in [s, b]$ and all $0 \leq \delta \leq \bar{\delta}$. Moreover

$$\frac{L_R}{\Gamma(\alpha)} \int_s^t (t-\tau)^{\alpha-1} \|u_\delta(\tau) - u(\tau)\|_{\mathbb{R}^m} \, d\tau \leq \frac{2RL_R}{\Gamma(\alpha+1)} \delta^\alpha,$$

and thus,

$$\|x_\delta(t) - x(t)\|_{\mathbb{R}^n} \leq \frac{2RL_R}{\Gamma(\alpha+1)} \delta^\alpha + L_R \mathrm{I}^\alpha_{s+}[\|x_\delta - x\|_{\mathbb{R}^n}](t),$$

for all $t \in [s, b]$ and all $0 \leq \delta \leq \bar{\delta}$. Using the fractional version of the Gronwall lemma given in Proposition B.1 gives

$$\|x_\delta(t) - x(t)\|_{\mathbb{R}^n} \leq \frac{2RL_R}{\Gamma(\alpha+1)} \delta^\alpha E_{\alpha,1}(L_R(b-a)^\alpha),$$

for every $t \in [s, b]$ and all $0 \leq \delta \leq \bar{\delta}$. Setting $\mu^1_R := \frac{2RL_R}{\Gamma(\alpha+1)} E_{\alpha,1}(L_R(b-a)^\alpha)$ ends the proof. ∎

Lemma A.4. *There exists $\mu^2_R \geq 0$ such that*

$$\|x_\delta(t) - x(t)\|_{\mathbb{R}^n} \leq \mu^2_R \delta \Big(t - (s+\delta)\Big)^{\alpha-1},$$

for all $t \in (s+\delta, b]$ and all $0 \leq \delta \leq \bar{\delta}$.

Proof. Similarly to the proof of Lemma A.3, we get that

$$\begin{aligned}\|x_\delta(t) - x(t)\|_{\mathbb{R}^n} \leq & \frac{L_R}{\Gamma(\alpha)} \int_s^{s+\delta} (t-\tau)^{\alpha-1} \|v - u(\tau)\|_{\mathbb{R}^m} \, d\tau \\ & + \frac{L_R}{\Gamma(\alpha)} \int_s^{s+\delta} (t-\tau)^{\alpha-1} \|x_\delta(\tau) - x(\tau)\|_{\mathbb{R}^n} \, d\tau \\ & + \frac{L_R}{\Gamma(\alpha)} \int_{s+\delta}^t (t-\tau)^{\alpha-1} \|x_\delta(\tau) - x(\tau)\|_{\mathbb{R}^n} \, d\tau,\end{aligned}$$

for all $t \in (s+\delta, b]$ and all $0 \leq \delta \leq \bar{\delta}$. Using Lemma A.3 and the inequality $\int_s^{s+\delta} (t-\tau)^{\alpha-1} d\tau \leq \delta(t-(s+\delta))^{\alpha-1}$, we obtain that

$$\begin{aligned}\|x_\delta(t) - x(t)\|_{\mathbb{R}^n} \leq & \frac{L_R}{\Gamma(\alpha)} (2R + \mu^1_R \delta^\alpha) \delta \Big(t - (s+\delta)\Big)^{\alpha-1} \\ & + \frac{L_R}{\Gamma(\alpha)} \int_{s+\delta}^t (t-\tau)^{\alpha-1} \|x_\delta(\tau) - x(\tau)\|_{\mathbb{R}^n} d\tau,\end{aligned}$$

and thus,

$$\|x_\delta(t)-x(t)\|_{\mathbb{R}^n} \le \frac{L_R}{\Gamma(\alpha)}(2R+\mu_R^1\bar\delta^\alpha)\delta\Big(t-(s+\delta)\Big)^{\alpha-1} + L_R \mathrm{I}^\alpha_{(s+\delta)+}[\|x_\delta - x\|_{\mathbb{R}^n}](t),$$

for all $t \in (s+\delta, b]$ and all $0 \le \delta \le \bar\delta$. Once again, the fractional version of the Gronwall lemma given in Proposition B.1 concludes the proof by setting $\mu_R^2 := L_R(2R+\mu_R^1\bar\delta^\alpha)E_{\alpha,\alpha}(L_R(b-a)^\alpha)$. ∎

Lemma A.5. *Let us define*

$$\zeta_\delta(\tau) := \int_0^1 \Big\|\partial_1 f\Big(x(\tau)+\theta(x_\delta(\tau)-x(\tau)),u(\tau),\tau\Big) - \partial_1 f(x(\tau),u(\tau),\tau)\Big\|_{\mathbb{R}^{n\times n}}\, d\theta,$$

for almost every $\tau \in [a,b]$ and every $0 \le \delta \le \bar\delta$. Then $\|\zeta_\delta\|_{\mathrm{L}^{r_\alpha}}$ tends to zero when $\delta \to 0$.

Proof. Since $(u_\delta, x_a) \in \mathrm{N}_R$ and $\partial_1 f$ is bounded by $L_R \ge 0$ on K_R (Inequality (A.1)), it holds from Proposition A.1 that $0 \le \zeta_\delta(\tau) \le 2L_R$ for almost every $\tau \in [a,b]$, and thus $\zeta_\delta \in \mathrm{L}^\infty([a,b],\mathbb{R})$ for all $0 \le \delta \le \bar\delta$. From Lemma A.3, recall that $\|x_\delta - x\|_{\mathcal{C}} \le \mu_R^1\delta^\alpha$ for all $0 \le \delta \le \bar\delta$. Thus Lemma A.5 easily follows from Lebesgue's dominated convergence theorem. ∎

Now we can prove the second item of Proposition 3.10. Let us define $z_\delta(t) := \frac{x_\delta(t)-x(t)}{\delta} - w(t)$ for all $t \in (s+\delta, b]$ and all $0 < \delta \le \bar\delta$. Let $0 < \varsigma \le b-s$. Our aim is to prove that z_δ uniformly converges to zero over $[s+\varsigma, b]$ as $\delta \to 0$. The integral representations give

$$z_\delta(t) = \frac{1}{\Gamma(\alpha)}\int_s^{s+\delta}(t-\tau)^{\alpha-1}\frac{f(x_\delta(\tau),v,\tau)-f(x(\tau),v,\tau)}{\delta}\,d\tau \tag{A.4}$$

$$+\frac{1}{\Gamma(\alpha)}\int_s^{s+\delta}(t-\tau)^{\alpha-1}\frac{f(x(\tau),v,\tau)-f(x(\tau),u(\tau),\tau)}{\delta}\,d\tau \tag{A.5}$$

$$-\frac{1}{\Gamma(\alpha)}(t-s)^{\alpha-1}\Big(f(x(s),v,s)-f(x(s),u(s),s)\Big) \tag{A.6}$$

$$+\frac{1}{\Gamma(\alpha)}\int_{s+\delta}^{t}(t-\tau)^{\alpha-1}\left(\frac{f(x_\delta(\tau),u(\tau),\tau)-f(x(\tau),u(\tau),\tau)}{\delta} - \partial_1 f(x(\tau),u(\tau),\tau)\times\frac{x_\delta(\tau)-x(\tau)}{\delta}\right)d\tau \tag{A.7}$$

$$-\frac{1}{\Gamma(\alpha)}\int_s^{s+\delta}(t-\tau)^{\alpha-1}\partial_1 f(x(\tau),u(\tau),\tau)\times w(\tau)\,d\tau \tag{A.8}$$

$$+\frac{1}{\Gamma(\alpha)}\int_{s+\delta}^{t}(t-\tau)^{\alpha-1}\partial_1 f(x(\tau),u(\tau),\tau)\times z_\delta(\tau)\,d\tau, \tag{A.9}$$

for every $t \in (s+\delta, b]$ and all $0 < \delta \le \bar\delta$. Let us deal with the six above right-hand side terms:

- From Proposition A.1, Inequality (A.2) and Lemma A.3, the norm of the term (A.4) is bounded by

$$\frac{L_R}{\Gamma(\alpha)}\int_s^{s+\delta}(t-\tau)^{\alpha-1}\frac{\|x_\delta(\tau)-x(\tau)\|_{\mathbb{R}^n}}{\delta}\,d\tau \le \frac{L_R\mu_R^1}{\Gamma(\alpha)}\delta^\alpha\Big(t-(s+\delta)\Big)^{\alpha-1},$$

for every $t \in (s+\delta, b]$ and all $0 < \delta \leq \bar{\delta}$.

- The sum of (A.5) and (A.6) terms is equal to $\frac{1}{\delta}\int_s^{s+\delta}(t-\tau)^{\alpha-1}\chi(\tau)d\tau - (t-s)^{\alpha-1}\chi(s)$ where $\chi(\tau) := \frac{1}{\Gamma(\alpha)}(f(x(\tau),v,\tau) - f(x(\tau),u(\tau),\tau))$ for almost every $\tau \in [a,b]$. Note that $s \in \mathcal{L}[\chi]$ (Lebesgue point) and from Inequality (A.2), it holds that $\|\chi\|_{\mathrm{L}^\infty} \leq \frac{2RL_R}{\Gamma(\alpha)}$. Hence the norm of the sum of (A.5) and (A.6) terms is bounded by

$$\begin{aligned}
&\left\|\frac{1}{\delta}\int_s^{s+\delta}\Big((t-\tau)^{\alpha-1} - (t-s)^{\alpha-1}\Big)\chi(\tau)\,d\tau\right\|_{\mathbb{R}^n} \\
&\quad + \left\|\frac{(t-s)^{\alpha-1}}{\delta}\int_s^{s+\delta}\chi(\tau)-\chi(s)\,d\tau\right\|_{\mathbb{R}^n} \\
&\quad \leq \frac{2RL_R}{\Gamma(\alpha)}\left(\frac{1}{\delta}\int_s^{s+\delta}(t-\tau)^{\alpha-1}\,d\tau - (t-s)^{\alpha-1}\right) \\
&\quad + \Big(t-(s+\delta)\Big)^{\alpha-1}\left\|\frac{1}{\delta}\int_s^{s+\delta}\chi(\tau)\,d\tau - \chi(s)\right\|_{\mathbb{R}^n},
\end{aligned}$$

for every $t \in (s+\delta, b]$ and all $0 < \delta \leq \bar{\delta}$. Considering some $0 < \alpha' < \alpha \leq 1$, one obtains

$$\begin{aligned}
&\left|\frac{1}{\delta}\int_s^{s+\delta}(t-\tau)^{\alpha-1}\,d\tau - (t-s)^{\alpha-1}\right| \\
&= \left|\frac{(t-s)^\alpha - (t-s-\delta)^\alpha}{\delta\alpha} - (t-s)^{\alpha-1}\right| \\
&= \delta^{\alpha-1}\left|\frac{\left(\frac{t-s}{\delta}\right)^\alpha - \left(\frac{t-s}{\delta}-1\right)^\alpha}{\alpha} - \left(\frac{t-s}{\delta}\right)^{\alpha-1}\right| \\
&= \delta^{\alpha-1}\left(\frac{t-s}{\delta}-1\right)^{\alpha'-1}\sigma_{\alpha,\alpha'}\left(\frac{t-s}{\delta}\right) \\
&= \Big(t-(s+\delta)\Big)^{\alpha'-1}\delta^{\alpha-\alpha'}\sigma_{\alpha,\alpha'}\left(\frac{t-s}{\delta}\right),
\end{aligned}$$

for every $t \in (s+\delta, b]$ and all $0 < \delta \leq \bar{\delta}$, where the bounded real function $\sigma_{\alpha,\alpha'}$ is defined in Lemma B.2. We conclude that the norm of the sum of (A.5) and (A.6) terms is bounded by

$$\begin{aligned}
&\frac{2RL_R}{\Gamma(\alpha)}\|\sigma_{\alpha,\alpha'}\|_{\mathcal{C}}\delta^{\alpha-\alpha'}\Big(t-(s+\delta)\Big)^{\alpha'-1} \\
&\qquad + \left\|\frac{1}{\delta}\int_s^{s+\delta}\chi(\tau)\,d\tau - \chi(s)\right\|_{\mathbb{R}^n}\Big(t-(s+\delta)\Big)^{\alpha-1},
\end{aligned}$$

for every $t \in (s+\delta, b]$ and all $0 < \delta \leq \bar{\delta}$.

- Using the classical Taylor expansion formula with integral rest, Lemmas A.4 and A.5 and the classical Hölder inequality, the norm of the term (A.7) is

bounded by

$$\frac{\mu_R^2}{\Gamma(\alpha)}\|\zeta_\delta\|_{\mathrm{L}^{r_\alpha}}\left(\int_{s+\delta}^{t}(t-\tau)^{r'_\alpha(\alpha-1)}\Big(\tau-(s+\delta)\Big)^{r'_\alpha(\alpha-1)}d\tau\right)^{1/r'_\alpha}$$
$$\leq \frac{\mu_R^2\|\zeta_\delta\|_{\mathrm{L}^{r_\alpha}}M_\alpha^2}{\Gamma(\alpha)}\Big(t-(s+\delta)\Big)^{\alpha-1},$$

for every $t\in(s+\delta,b]$ and all $0<\delta\leq\bar\delta$.

- Since the product $\rho^\alpha_{s+}w\in\mathcal{C}([s,b],\mathbb{R}^n)$ (see [13, Theorem 6]), the norm of the term (A.8) is bounded by

$$\frac{L_R\|\rho^\alpha_{s+}w\|_{\mathcal{C}}}{\Gamma(\alpha)^2}\int_s^{s+\delta}(t-\tau)^{\alpha-1}(\tau-s)^{\alpha-1}\,d\tau\leq\frac{L_R\|\rho^\alpha_{s+}w\|_{\mathcal{C}}}{\Gamma(\alpha)\Gamma(\alpha+1)}\delta^\alpha\Big(t-(s+\delta)\Big)^{\alpha-1},$$

for every $t\in(s+\delta,b]$ and all $0<\delta\leq\bar\delta$.

Finally we have proved that

$$\|z_\delta(t)\|_{\mathbb{R}^n}\leq\frac{\Theta^1_\delta}{\Gamma(\alpha)}\Big(t-(s+\delta)\Big)^{\alpha-1}+\frac{\Theta^2_\delta}{\Gamma(\alpha')}\Big(t-(s+\delta)\Big)^{\alpha'-1}+L_R\mathrm{I}^\alpha_{(s+\delta)+}[\|z_\delta\|_{\mathbb{R}^n}](t),$$

for every $t\in(s+\delta,b]$ and all $0<\delta\leq\bar\delta$, where

$$\Theta^1_\delta:=L_R\delta^\alpha\left(\mu^1_R+\frac{\|\rho^\alpha_{s+}w\|_{\mathcal{C}}}{\Gamma(\alpha+1)}\right)+\Gamma(\alpha)\left\|\frac{1}{\delta}\int_s^{s+\delta}\chi(\tau)\,d\tau-\chi(s)\right\|_{\mathbb{R}^n}+\mu_R^2\|\zeta_\delta\|_{\mathrm{L}^{r_\alpha}}M_\alpha^2,$$

$$\Theta^2_\delta:=2RL_R\frac{\Gamma(\alpha')}{\Gamma(\alpha)}\|\sigma_{\alpha,\alpha'}\|_{\mathcal{C}}\delta^{\alpha-\alpha'}.$$

Since $s\in\mathcal{L}[\chi]$ (Lebesgue point), $\alpha-\alpha'>0$ and $\|\zeta_\delta\|_{\mathrm{L}^{r_\alpha}}$ tends to zero (Lemma A.5), we know that $\Theta^1_\delta\to0$ and $\Theta^2_\delta\to0$ when $\delta\to0$. Finally, from the fractional version of the Gronwall lemma given in Proposition B.1, we get that

$$\|z_\delta(t)\|_{\mathbb{R}^n}\leq\Theta^1_\delta\Big(t-(s+\delta)\Big)^{\alpha-1}E_{\alpha,\alpha}\Big(L_R(b-a)^\alpha\Big)$$
$$+\Theta^2_\delta\Big(t-(s+\delta)\Big)^{\alpha'-1}E_{\alpha,\alpha'}\Big(L_R(b-a)^\alpha\Big),$$

for every $t\in(s+\delta,b]$ and all $0<\delta\leq\bar\delta$. Since $t-(s+\delta)\geq\varsigma-\delta$ for all $t\in[s+\varsigma,b]$, the second item of Proposition 3.10 is achieved by letting $\delta\to0$.

A.1.3. Proof of Proposition 3.11. Let $y\in\mathbb{R}^n$ and let $R=\|u\|_{\mathrm{L}^\infty}$. Since $\|x_a+\delta y-x_a\|_{\mathbb{R}^n}\leq\delta\|y\|_{\mathbb{R}^n}$ for all $\delta\geq0$, we deduce from Proposition A.1 that there exists $\bar\delta>0$ such that $(u,x_a+\delta y)\in\mathrm{N}_R\subset\mathcal{A}_\mathcal{G}$ for all $0\leq\delta\leq\bar\delta$. Then the first item of Proposition 3.11 easily follows from Proposition A.2. Before proving the second item of Proposition 3.11, we first state the following Lemma A.6. For the ease of notations, in this section, we denote by $x_\delta:=x(\cdot,u,x_a+\delta y)$ for all $0\leq\delta\leq\bar\delta$, by $x:=x(\cdot,u,x_a)$ and by $w:=w_y(\cdot,u,x_a)$.

Lemma A.6. *We define*

$$\zeta_\delta(\tau):=\int_0^1\Big\|\partial_1 f\Big(x(\tau)+\theta(x_\delta(\tau)-x(\tau)),u(\tau),\tau\Big)-\partial_1 f(x(\tau),u(\tau),\tau)\Big\|_{\mathbb{R}^{n\times n}}\,d\theta,$$

for almost every $\tau\in[a,b]$ and all $0\leq\delta\leq\bar\delta$. Then $\|\zeta_\delta\|_{\mathrm{L}^{r_\alpha}}$ tends to zero when $\delta\to0$.

Proof. Since $(u, x_a + \delta y) \in \mathrm{N}_R$ and $\partial_1 f$ is bounded by $L_R \geq 0$ on K_R (Inequality (A.1)), it holds from Proposition A.1 that $0 \leq \zeta_\delta(\tau) \leq 2L_R$ for almost every $\tau \in [a,b]$, and thus $\zeta_\delta \in \mathrm{L}^\infty([a,b],\mathbb{R})$ for all $0 \leq \delta \leq \bar{\delta}$. From Proposition A.2, recall that $\|x_\delta - x\|_{\mathcal{C}} \leq \delta C_R \|y\|_{\mathbb{R}^n}$ for all $0 \leq \delta \leq \bar{\delta}$. Thus Lemma A.6 easily follows from Lebesgue's dominated convergence theorem. ∎

We may now prove the second item of Proposition 3.11. Let us we define $z_\delta(t) := \frac{x_\delta(t) - x(t)}{\delta} - w(t)$ for all $t \in [a,b]$ and all $0 < \delta \leq \bar{\delta}$. Our aim is to show that z_δ uniformly converges to zero over $[a,b]$ as $\delta \to 0$. The integral representations give

$$z_\delta(t) = \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \Big(\frac{f(x_\delta(\tau), u(\tau), \tau) - f(x(\tau), u(\tau), \tau)}{\delta} \\ - \partial_1 f(x(\tau), u(\tau), \tau) \times w(\tau) \Big) \, d\tau,$$

for every $t \in [a,b]$ and all $0 < \delta \leq \bar{\delta}$. Recall that $\|x_\delta - x\|_{\mathcal{C}} \leq \delta C_R \|y\|_{\mathbb{R}^n}$ for all $0 \leq \delta \leq \bar{\delta}$ thanks to Proposition A.2. With the classical Taylor expansion with integral rest, we obtain that

$$\|z_\delta(t)\|_{\mathbb{R}^n} \leq \frac{C_R \|y\|_{\mathbb{R}^n}}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \zeta_\delta(\tau) \, d\tau \\ + \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \|\partial_1 f(x(\tau), u(\tau), \tau) \times z_\delta(\tau)\|_{\mathbb{R}^n} \, d\tau,$$

for every $t \in [a,b]$ and all $0 < \delta \leq \bar{\delta}$. Using Hölder inequality and the fact that $\partial_1 f$ is bounded by $L_R \geq 0$ on K_R, we get from the fractional version of the Gronwall lemma (Proposition B.1) that

$$\|z_\delta(t)\|_{\mathbb{R}^n} \leq M_\alpha^1 C_R \|y\|_{\mathbb{R}^n} \|\zeta_\delta\|_{\mathrm{L}^{r_\alpha}} E_{\alpha,1}(L_R (b-a)^\alpha),$$

for every $t \in [a,b]$ and all $0 < \delta \leq \bar{\delta}$. The second item of Proposition 3.11 is achieved from Lemma A.6.

A.2. Proof of Theorem 3.12. Our strategy in order to prove Theorem 3.12 follows the five following steps:

(i) We first reduce Problem (OCP) to a Mayer problem, that is, with no Lagrange cost ($F = 0$). For this purpose we use a standard change of variable. We refer to Appendix A.2.1 for details. As a consequence, from this first step until the end of this appendix, we assume that there is no Lagrange cost ($F = 0$) in Problem (OCP);

(ii) We give some preliminaries and notations in Appendix A.2.2;

(iii) Using the sensitivity analysis of the fractional state equation performed in Section 3.3, we compute derivatives of the trajectory x with respect to perturbations of the control u and of the initial condition x_a;

(iv) We apply Ekeland's variational principle (Proposition A.7) to take into account the terminal state constraint $g(x(a), x(b)) \in \mathrm{C}$. We derive crucial inequalities on variation vectors. We refer to Appendices A.2.3, A.2.4 and A.2.5 for details;

(v) We conclude the proof by introducing the adjoint vector p in Appendix A.2.6. This adjoint vector allows to derive the transversality conditions and the Hamiltonian maximization condition of Theorem 3.12 from the inequalities obtained in the previous step.

For the reader's convenience, we recall hereafter a simplified version (but sufficient for our purposes) of Ekeland's variational principle.

PROPOSITION A.7 (Ekeland's variational principle [22]). *Let $(\mathcal{M}, \mathrm{d}_{\mathcal{M}})$ be a complete metric space and let $\mathcal{J} : \mathcal{M} \to \mathbb{R}_+$ be a nonnegative continuous mapping. Let $\varepsilon > 0$ and $x^* \in \mathcal{M}$ such that $J(x^*) \leq \varepsilon$. Then, there exists $x_\varepsilon \in \mathcal{M}$ such that $\mathrm{d}_{\mathcal{M}}(x_\varepsilon, x^*) \leq \sqrt{\varepsilon}$ and $-\sqrt{\varepsilon}\,\mathrm{d}_{\mathcal{M}}(x, x_\varepsilon) \leq \mathcal{J}(x) - \mathcal{J}(x_\varepsilon)$ for all $x \in \mathcal{M}$.*

A.2.1. Reduction to Mayer problem from a change of variable. Let us assume that Theorem 3.12 is already proved in the case of a Mayer problem, that is, with no Lagrange cost ($F = 0$). Our aim in this section is to derive Theorem 3.12 in the general Bolza case.

Let $(x^*, u^*) \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a,b], \mathbb{R}^n) \times \mathrm{L}^\infty([a,b], \mathbb{R}^m)$ be an optimal solution to Problem (OCP) and let us introduce $z^* := \mathrm{I}^{\alpha}_{a+}[F^\beta(x^*, u^*, \cdot)] \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a,b], \mathbb{R})$. One can easily see that the couple $((x^*, z^*), u^*)$ is an optimal solution to the augmented Caputo fractional optimal control problem of Mayer form given by

$$
\text{(OCPAM)} \quad \left\{ \begin{array}{ll} \text{minimize} & \tilde{\varphi}((x,z)(a),(x,z)(b)), \\ \\ \text{subject to} & (x,z) \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a,b], \mathbb{R}^{n+1}),\ u \in \mathrm{L}^\infty([a,b], \mathbb{R}^m), \\ & {}_{\mathrm{c}}\mathrm{D}^{\alpha}_{a+}[(x,z)](t) = (f, F^\beta)(x(t), u(t), t), \quad \text{a.e. } t \in [a,b], \\ & \tilde{g}((x,z)(a),(x,z)(b)) \in \tilde{\mathrm{C}}, \\ & u(t) \in \mathrm{U}, \quad \text{a.e. } t \in [a,b], \end{array} \right.
$$

where $\tilde{\varphi} : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$ is defined by $\tilde{\varphi}((x_a, z_a), (x_b, z_b)) := \varphi(x_a, x_b) + z_b$ for all $((x_a, z_a), (x_b, z_b)) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$. Moreover:

(i) if g is submersive at $(x^*(a), x^*(b))$, then we choose $\tilde{\mathrm{C}} := \mathrm{C} \times \{0\} \times \mathbb{R}$ and $\tilde{g} : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}^{j+2}$ defined by $\tilde{g}((x_a, z_a), (x_b, z_b)) := (g(x_a, x_b), z_a, z_b)$ for all $((x_a, z_a), (x_b, z_b)) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$;

(ii) if g is not submersive at $(x^*(a), x^*(b))$, then we choose $\tilde{\mathrm{C}} := \{(x^*(a), x^*(b))\} \times \{0\} \times \mathbb{R}$ and $\tilde{g} : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}^{2n+2}$ defined by $\tilde{g}((x_a, z_a), (x_b, z_b)) := (x_a, x_b, z_a, z_b)$ for all $((x_a, z_a), (x_b, z_b)) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$.

In both cases, note that $\tilde{g}$ is submersive at $((x^*, z^*)(a), (x^*, z^*)(b))$. We deduce the existence of a nontrivial augmented couple $((p, q), p^0) \in \mathcal{AC}^{\alpha}_{b-}([a,b], \mathbb{R}^{n+1}) \times \mathbb{R}$ satisfying all necessary conditions listed in Theorem 3.12 adapted to the augmented Problem (OCPAM) of Mayer form. In particular, we get that $\mathrm{D}^{\alpha}_{b-}[q] = 0$ with $\mathrm{I}^{1-\alpha}_{b-}[q](b) = p^0$ and thus $q(t) = \frac{p^0}{\Gamma(\alpha)}(b-t)^{\alpha-1}$ for all $t \in [a,b)$. The rest of the proof is staightforward from all necessary conditions provided in Theorem 3.12 (in the case of a Mayer problem).

Thanks to this section, from now and until the end of the proof of Theorem 3.12 in its general form, we may assume that there is no Lagrange cost ($F = 0$) in Problem (OCP).

A.2.2. Some preliminaries and notations. We denote by d_{C} the *distance function* to the nonempty closed convex subset $\mathrm{C} \subset \mathbb{R}^j$ defined as usual by $\mathrm{d}_{\mathrm{C}}(x) := \inf_{x' \in \mathrm{C}} \|x - x'\|_{\mathbb{R}^j}$ for all $x \in \mathbb{R}^j$. Recall that, for every $x \in \mathbb{R}^j$, there exists a unique element $\mathrm{P}_{\mathrm{C}}(x) \in \mathrm{C}$ (*projection* of x on C) such that $\mathrm{d}_{\mathrm{C}}(x) = \|x - \mathrm{P}_{\mathrm{C}}(x)\|_{\mathbb{R}^j}$. It is characterized by the property $\langle x - \mathrm{P}_{\mathrm{C}}(x), x' - \mathrm{P}_{\mathrm{C}}(x)\rangle_{\mathbb{R}^j} \leq 0$ for every $x' \in \mathrm{C}$. In particular, $x - \mathrm{P}_{\mathrm{C}}(x) \in \mathcal{N}_{\mathrm{C}}[\mathrm{P}_{\mathrm{C}}(x)]$. The function $\mathrm{P}_{\mathrm{C}} : \mathbb{R}^j \to \mathrm{C}$ is 1-Lipschitz continuous. We refer to [16, p.131] for details. The proof of Theorem 3.12 requires the two following lemmas.

LEMMA A.8. *The function* $\mathrm{d}_{\mathrm{C}}^2 : x \mapsto \mathrm{d}_{\mathrm{C}}(x)^2$ *is differentiable on* $\mathbb{R}^j$*, and it holds that* $\mathrm{D}\mathrm{d}_{\mathrm{C}}^2(x)(x') = 2\langle x - \mathrm{P}_{\mathrm{C}}(x), x'\rangle_{\mathbb{R}^j}$ *for all* x, $x' \in \mathbb{R}^j$.

LEMMA A.9. *Let* $(x_k)_{k\in\mathbb{N}}$ *be a sequence of points in* $\mathbb{R}^j$ *and let* $(\varsigma_k)_{k\in\mathbb{N}}$ *be a sequence of positive real numbers such that* $x_k \to x \in \mathrm{C}$ *and* $\varsigma_k(x_k - \mathrm{P}_{\mathrm{C}}(x_k)) \to z \in \mathbb{R}^j$ *as* $k \to +\infty$*. Then* $z \in \mathcal{N}_{\mathrm{C}}[x]$.

The *positive part function* is defined on $\mathbb{R}$ by $\mathrm{Pos} : x \mapsto x^+$ where $x^+ = 0$ if $x \leq 0$ and $x^+ = x$ if $x \geq 0$. The proof of our main result also needs the following classical result.

LEMMA A.10. *The function* $\mathrm{Pos}^2 : x \mapsto x^{+2}$ *is differentiable on* $\mathbb{R}$*, and it holds that* $\mathrm{D}\mathrm{Pos}^2(x)(x') = 2x^+x'$ *for all* x, $x' \in \mathbb{R}$.

Let $(x^*, u^*) \in {}_{\mathrm{c}}\mathcal{AC}^{\alpha}_{a+}([a,b],\mathbb{R}^n) \times \mathrm{L}^\infty([a,b],\mathbb{R}^m)$ be an optimal solution to Problem (OCP). With the notations introduced in Section 3.3, $(u^*, x^*(a)) \in \mathcal{A}_{\mathcal{G}}$ and $x^* = x(\cdot, u^*, x^*(a))$. In the sequel, for all $R \in \mathbb{N}$ such that $R \geq \|u^*\|_{\mathrm{L}^\infty}$, we denote by K^*_R the corresponding compact subset of $\mathbb{R}^n \times \mathbb{R}^m \times [a,b]$ defined as in Appendix 3.3.1 and by L^*_R the corresponding nonnegative constant (Inequality (A.2)). Similarly, we denote by $\mathrm{N}^*_R \subset \mathcal{A}_{\mathcal{G}}$ the neighborhood of $(u^*, x^*(a))$ defined as in Proposition A.1 and by $\eta^*_R > 0$ the corresponding positive radius. Finally we endow N^*_R with the same basic distance $\mathrm{d}_{\mathrm{L}^1\times\mathbb{R}^n}$ defined as in Equality (A.3).

In order to take into account the control constraint set U in Problem (OCP), we set

$$\mathrm{N}^{*,\mathrm{U}}_R := \{(u, x_a) \in \mathrm{N}^*_R \mid u(t) \in \mathrm{U} \text{ a.e. } t \in [a,b]\}.$$

For example, we know that $(u^*, x^*(a)) \in \mathrm{N}^{*,\mathrm{U}}_R$. Since $\mathrm{U} \subset \mathbb{R}^m$ is a nonempty closed subset of $\mathbb{R}^m$, one can easily get from the (partial) converse of Lebesgue's dominated convergence theorem that $(\mathrm{N}^{*,\mathrm{U}}_R, \mathrm{d}_{\mathrm{L}^1\times\mathbb{R}^n})$ is a complete metric space. Note that the closedness assumption on the control constraint set U is crucial here in order to apply Ekeland's variational principle on a complete metric space.

A.2.3. A penalized functional with $R \in \mathbb{N}$ fixed. In the whole section we fix some $R \in \mathbb{N}$ such that $R \geq \|u^*\|_{\mathrm{L}^\infty}$. Let us consider a positive sequence $(\varepsilon^R_k)_{k\in\mathbb{N}}$ converging to zero as $k \to \infty$ such that

$$0 < \sqrt{\varepsilon^R_k} < \eta^*_R, \tag{A.10}$$

for all $k \in \mathbb{N}$. Then, we define the *penalized functional* $\mathcal{J}^R_k : \mathrm{N}^{*,\mathrm{U}}_R \to \mathbb{R}_+$ by

$$\mathcal{J}^R_k(u, x_a) := \sqrt{\Big(\varphi(x_a, x(b,u,x_a)) - \varphi(x^*(a), x^*(b)) + \varepsilon^R_k\Big)^{+2} + \mathrm{d}^2_{\mathrm{C}}\Big(g(x_a, x(b,u,x_a))\Big)},$$

for all $(u, x_a) \in \mathrm{N}^{*,\mathrm{U}}_R$ and all $k \in \mathbb{N}$.

• *Fix* $k \in \mathbb{N}$. From Proposition A.1, we know that $\mathrm{N}^{*,\mathrm{U}}_R \subset \mathrm{N}^*_R \subset \mathcal{A}_{\mathcal{G}}$ and thus $\mathcal{J}^R_k(u, x_a)$ is well-defined for all $(u, x_a) \in \mathrm{N}^{*,\mathrm{U}}_R$. Thus, from the optimality of the couple (x^*, u^*), one gets by contradiction that $\mathcal{J}^R_k(u, x_a) > 0$ for all $(u, x_a) \in \mathrm{N}^{*,\mathrm{U}}_R$. From the continuity result of Proposition A.2 and since φ, g and $\mathrm{d}^2_{\mathrm{C}}$ are continuous, it follows that $\mathcal{J}^R_k$ is a positive continuous map. Moreover one has $\mathcal{J}^R_k(u^*, x^*(a)) = \varepsilon^R_k$. Then Ekeland's variational principle (Proposition A.7) yields that there exists $(u^R_k, x^R_{a,k}) \in \mathrm{N}^{*,\mathrm{U}}_R$ such that

$$\mathrm{d}_{\mathrm{L}^1\times\mathbb{R}^n}\Big((u^R_k, x^R_{a,k}), (u^*, x^*(a))\Big) \leq \sqrt{\varepsilon^R_k}, \tag{A.11}$$

and

$$-\sqrt{\varepsilon_k^R}\,\mathrm{d}_{\mathrm{L}^1\times\mathbb{R}^n}\Big((u,x_a),(u_k^R,x_{a,k}^R)\Big)\leq \mathcal{J}_k^R(u,x_a)-\mathcal{J}_k^R(u_k^R,x_{a,k}^R), \tag{A.12}$$

for all $(u,x_a)\in \mathrm{N}_R^{*,\mathrm{U}}$. Finally we introduce the terms

$$\Psi_k^{0R}:=\frac{-1}{\mathcal{J}_k^R(u_k^R,x_{a,k}^R)}\Big(\varphi(x_{a,k}^R,x(b,u_k^R,x_{a,k}^R))-\varphi(x^*(a),x^*(b))+\varepsilon_k^R\Big)^+\in\mathbb{R},$$

and

$$\Psi_k^{R}:=\frac{-1}{\mathcal{J}_k^R(u_k^R,x_{a,k}^R)}\Big(g(x_{a,k}^R,x(b,u_k^R,x_{a,k}^R))-\mathrm{P}_{\mathrm{C}}\Big(g(x_{a,k}^R,x(b,u_k^R,x_{a,k}^R))\Big)\Big)\in\mathbb{R}^j.$$

Note that $\Psi_k^{0R}\leq 0$ and that $|\Psi_k^{0R}|^2+\|\Psi_k^R\|_{\mathbb{R}^j}^2=1$ from the definition of $\mathcal{J}_k^R(u_k^R,x_{a,k}^R)$ and that $-\Psi_k^R\in\mathcal{N}_{\mathrm{C}}[\mathrm{P}_{\mathrm{C}}(g(x_{a,k}^R,x(b,u_k^R,x_{a,k}^R)))]$.

• *Compactness arguments.* Recall that $(\varepsilon_k^R)_{k\in\mathbb{N}}$ converges to zero as $k\to\infty$. Using Inequality (A.11), some compactness arguments from the equality $|\Psi_k^{0R}|^2+\|\Psi_k^R\|_{\mathbb{R}^j}^2=1$, the (partial) converse of Lebesgue's dominated convergence theorem and extracting some subsequences (denoted similarly), we get that:

- $(x_{a,k}^R)_{k\in\mathbb{N}}$ converges to $x^*(a)$ in $\mathbb{R}^n$;
- $(u_k^R)_{k\in\mathbb{N}}$ converges to u^* in $\mathrm{L}^1([a,b],\mathbb{R}^m)$;
- $(u_k^R(t))_{k\in\mathbb{N}}$ converges to $u^*(t)$ in $\mathbb{R}^m$ for almost every $t\in[a,b]$;
- $(\Psi_k^{0R})_{k\in\mathbb{N}}$ converges to some Ψ^{0R} in $\mathbb{R}$;
- $(\Psi_k^{R})_{k\in\mathbb{N}}$ converges to some Ψ^{R} in $\mathbb{R}^j$.

Moreover it holds that $\Psi^{0R}\leq 0$, $|\Psi^{0R}|^2+\|\Psi^R\|_{\mathbb{R}^j}^2=1$ and $-\Psi^R\in\mathcal{N}_{\mathrm{C}}[g(x^*(a),x^*(b))]$ (from the continuity result claimed in Proposition A.2, the continuity of g and from Lemma A.9).

A.2.4. Two crucial inequalities depending on $R\in\mathbb{N}$ fixed. In the whole section we fix some $R\in\mathbb{N}$ such that $R\geq\|u^*\|_{\mathrm{L}^\infty}$. We introduce

$$\mathcal{L}^R:=\Big\{s\in[a,b)\mid u_k^R(s)\to u^*(s)\text{ as }k\to\infty\Big\}\cap\bigcap_{k\in\mathbb{N}}\mathcal{L}\Big[f\Big(x(\cdot,u_k^R,x_{a,k}^R),u_k^R,\cdot\Big)\Big].$$

Note that $\mathcal{L}^R$ has a full Lebesgue measure equal to $b-a$. Let $(s,v)\in\mathcal{L}^R\times(\mathrm{U}\cap\overline{\mathrm{B}}_{\mathbb{R}^m}(0,R))$ and $y\in\mathbb{R}^n$.

• *Perturbation of the control u_k^R with $k\in\mathbb{N}$ fixed.* For all $\delta\in[0,b-s)$, let us consider $u_{k,\delta}^R$ the needle-like perturbation of u_k^R associated to (s,v) defined as in Section 3.3.2. From Inequality (A.11), since $v\in\overline{\mathrm{B}}_{\mathbb{R}^m}(0,R)$, we get that

$$\|u_{k,\delta}^R-u^*\|_{\mathrm{L}^1}\leq\|u_{k,\delta}^R-u_k^R\|_{\mathrm{L}^1}+\|u_k^R-u^*\|_{\mathrm{L}^1}\leq 2R\delta+\sqrt{\varepsilon_k^R},$$

for all $\delta\in[0,b-s)$. With Inequality (A.10) and since $v\in\mathrm{U}\cap\overline{\mathrm{B}}_{\mathbb{R}^m}(0,R)$, we deduce that Inequality (A.12) can be applied to $(u,x_a)=(u_{k,\delta}^R,x_{a,k}^R)\in\mathrm{N}_R^{*,\mathrm{U}}$ for all $\delta\in[0,b-s)$ sufficiently small. We obtain that

$$-2R\sqrt{\varepsilon_k^R}\leq\frac{1}{\mathcal{J}_k^R(u_{k,\delta}^R,x_{a,k}^R)+\mathcal{J}_k^R(u_k^R,x_{a,k}^R)}\frac{\mathcal{J}_k^R(u_{k,\delta}^R,x_{a,k}^R)^2-\mathcal{J}_k^R(u_k^R,x_{a,k}^R)^2}{\delta},$$

for all $\delta \in (0, b-s)$ small enough. Passing to the limit as $\delta \to 0$, using the differentiability results of Proposition 3.10 and Lemmas A.8, A.10, the differentiabilities of φ and g, we get that

$$\Big\langle \Psi_k^{0R} \partial_2 \varphi\Big(x_{a,k}^R, x(b, u_k^R, x_{a,k}^R)\Big) + \partial_2 g\Big(x_{a,k}^R, x(b, u_k^R, x_{a,k}^R)\Big)^\top \times \Psi_k^R,\, w_{(s,v)}(b, u_k^R, x_{a,k}^R)\Big\rangle_{\mathbb{R}^n} \leq 2R\sqrt{\varepsilon_k^R}, \tag{A.13}$$

where $w_{(s,v)}(\cdot, u_k^R, x_{a,k}^R)$ is the variation vector associated to $(u_k^R, x_{a,k}^R)$ and (s,v) defined as in Section 3.3.2. Note that we used the fact that $s \in \mathcal{L}[f(x(\cdot, u_k^R, x_{a,k}^R), u_k^R, \cdot)]$ in order to apply Proposition 3.10.

• *Passing to the limit $k \to \infty$.* We first state the following lemma requiring in particular that $u_k^R(s) \to u^*(s)$ as $k \to \infty$.

LEMMA A.11. *It holds that the sequence $w_{(s,v)}(\cdot, u_k^R, x_{a,k}^R)$ uniformly converges on $[s+\varsigma, b]$, for any $0 < \varsigma \leq b-s$, to $w_{(s,v)}(\cdot, u^*, x^*(a))$ when $k \to \infty$.*

Proof. Let $0 < \varsigma \leq b-s$. For the ease of notations, we denote by $x_k^R := x(\cdot, u_k^R, x_{a,k}^R)$, $w_k := w_{(s,v)}(\cdot, u_k^R, x_{a,k}^R)$, $A_k := \partial_1 f(x_k^R, u_k^R, \cdot)$ for all $k \in \mathbb{N}$, and $w := w_{(s,v)}(\cdot, u^*, x^*(a))$, $A := \partial_1 f(x^*, u^*, \cdot)$. Since $(x_k^R)_{k\in\mathbb{N}}$ and $(u_k^R)_{k\in\mathbb{N}}$ pointwise converge on $[a,b]$ to x^* and u^* (Proposition A.2) when $k \to \infty$ and $\partial_1 f$ is continuous, bounded by L_R^* on K_R^*, one can deduce from Lebesgue's dominated convergence theorem that $(A_k)_{k\in\mathbb{N}}$ converges in $\mathrm{L}^{r_\alpha}([a,b], \mathbb{R}^{n\times n})$ to A. We also denote by

$$\nu_k := f(x_k^R(s), v, s) - f(x_k^R(s), u_k^R(s), s) - \Big(f(x^*(s), v, s) - f(x^*(s), u^*(s), s)\Big) \in \mathbb{R}^n,$$

for all $k \in \mathbb{N}$. Since $s \in \mathcal{L}^R$, we know that $u_k^R(s) \to u^*(s)$ when $k \to \infty$ and thus we deduce that $\|\nu_k\|_{\mathbb{R}^n} \to 0$ when $k \to \infty$. Integral representations give

$$\begin{aligned} w_k(t) - w(t) = \frac{1}{\Gamma(\alpha)}(t-s)^{\alpha-1}\nu_k + \frac{1}{\Gamma(\alpha)}\int_s^t (t-\tau)^{\alpha-1} A_k(\tau) \times (w_k(\tau) - w(\tau))\, d\tau \\ + \frac{1}{\Gamma(\alpha)}\int_s^t (t-\tau)^{\alpha-1}(A_k(\tau) - A(\tau)) \times w(\tau)\, d\tau, \end{aligned}$$

for all $t \in (s,b]$ and all $k \in \mathbb{N}$. Since the product $\rho_{s+}^\alpha w \in \mathcal{C}([s,b], \mathbb{R}^n)$ (see [13, Theorem 6]), we get from Hölder inequality that

$$\begin{aligned} \|w_k(t) - w(t)\|_{\mathbb{R}^n} \leq \frac{1}{\Gamma(\alpha)}\left(\|\nu_k\|_{\mathbb{R}^n} + \frac{\|\rho_{s+}^\alpha w\|_{\mathcal{C}}}{\Gamma(\alpha)}\|A_k - A\|_{\mathrm{L}^{r_\alpha}} M_\alpha^2\right)(t-s)^{\alpha-1} \\ + \frac{L_R^*}{\Gamma(\alpha)}\int_s^t (t-\tau)^{\alpha-1}\|w_k(\tau) - w(\tau)\|_{\mathbb{R}^n}\, d\tau, \end{aligned}$$

for all $t \in (s,b]$ and all $k \in \mathbb{N}$. The fractional version of the Gronwall lemma given in Proposition B.1 yields to

$$\|w_k(t) - w(t)\|_{\mathbb{R}^n} \leq \left(\|\nu_k\|_{\mathbb{R}^n} + \frac{\|\rho_{s+}^\alpha w\|_{\mathcal{C}}}{\Gamma(\alpha)}\|A_k - A\|_{\mathrm{L}^{r_\alpha}} M_\alpha^2\right)(t-s)^{\alpha-1} E_{\alpha,\alpha}(L_R^*(b-a)^\alpha),$$

for every $t \in (s,b]$ and all $k \in \mathbb{N}$. Since $t - s \geq \varsigma$ for all $t \in [s+\varsigma, b]$, the proof is achieved by letting $k \to \infty$. ∎

Thanks to the continuity result of Proposition A.2, the fact that φ and g are $\mathcal{C}^1$ and $w_{(s,v)}(b, u_k^R, x_{a,k}^R)$ tends to $w_{(s,v)}(b, u^*, x^*(a))$ when $k \to \infty$ (Lemma A.11), we get from Inequality (A.13) that

(A.14)
$$\left\langle \Psi^{0R} \partial_2 \varphi(x^*(a), x^*(b)) + \partial_2 g(x^*(a), x^*(b))^\top \times \Psi^R, w_{(s,v)}(b, u^*, x^*(a)) \right\rangle_{\mathbb{R}^n} \leq 0.$$

The above inequality constitutes the first crucial inequality depending on $R \in \mathbb{N}$ fixed.

• *Perturbation of the initial condition $x_{a,k}^R$ with $k \in \mathbb{N}$ fixed.* From Inequality (A.11), we have

$$\|x_{a,k}^R + \delta y - x^*(a)\|_{\mathbb{R}^n} \leq \delta \|y\|_{\mathbb{R}^n} + \|x_{a,k}^R - x^*(a)\|_{\mathbb{R}^n} \leq \delta \|y\|_{\mathbb{R}^n} + \sqrt{\varepsilon_k^R},$$

for all $\delta \geq 0$. From Inequality (A.10), we deduce that Inequality (A.12) can be applied to $(u, x_a) = (u_k^R, x_{a,k}^R + \delta y) \in \mathrm{N}_R^{*,\mathrm{U}}$ for every $\delta \geq 0$ sufficiently small. We obtain

$$-\|y\|_{\mathbb{R}^n} \sqrt{\varepsilon_k^R} \leq \frac{1}{\mathcal{J}_k^R(u_k^R, x_{a,k}^R + \delta y) + \mathcal{J}_k^R(u_k^R, x_{a,k}^R)} \frac{\mathcal{J}_k^R(u_k^R, x_{a,k}^R + \delta y)^2 - \mathcal{J}_k^R(u_k^R, x_{a,k}^R)^2}{\delta},$$

for all $\delta > 0$ sufficiently small. Passing to the limit $\delta \to 0$, using the differentiability results of Proposition 3.11 and Lemmas A.8, A.10, the differentiabilities of φ and g, we get that

$$\begin{aligned} (\mathrm{A.15}) \qquad & \left\langle \Psi_k^{0R} \partial_1 \varphi\Big(x_{a,k}^R, x(b, u_k^R, x_{a,k}^R)\Big) + \partial_1 g\Big(x_{a,k}^R, x(b, u_k^R, x_{a,k}^R)\Big)^\top \times \Psi_k^R, y \right\rangle_{\mathbb{R}^n} + \\ & \left\langle \Psi_k^{0R} \partial_2 \varphi\Big(x_{a,k}^R, x(b, u_k^R, x_{a,k}^R)\Big) + \partial_2 g\Big(x_{a,k}^R, x(b, u_k^R, x_{a,k}^R)\Big)^\top \times \Psi_k^R, w_y(b, u_k^R, x_{a,k}^R) \right\rangle_{\mathbb{R}^n} \\ & \leq \|y\|_{\mathbb{R}^n} \sqrt{\varepsilon_k^R}, \end{aligned}$$

where $w_y(\cdot, u_k^R, x_{a,k}^R)$ is the variation vector associated to $(u_k^R, x_{a,k}^R)$ and y defined as in Section 3.3.3.

• *Passing to the limit $k \to \infty$.* We first state the following lemma.

Lemma A.12. *It holds that the sequence $w_y(\cdot, u_k^R, x_{a,k}^R)$ uniformly converges on $[a, b]$ to $w_y(\cdot, u^*, x^*(a))$ when $k \to \infty$.*

Proof. For the ease of notations, we denote by $x_k^R := x(\cdot, u_k^R, x_{a,k}^R)$, by $w_k := w_y(\cdot, u_k^R, x_{a,k}^R)$, $A_k := \partial_1 f(x_k^R, u_k^R, \cdot)$ for all $k \in \mathbb{N}$, and $w := w_y(\cdot, u^*, x^*(a))$, $A := \partial_1 f(x^*, u^*, \cdot)$. Since $(x_k^R)_{k \in \mathbb{N}}$ and $(u_k^R)_{k \in \mathbb{N}}$ pointwise converge on $[a, b]$ to x^* and u^* (Proposition A.2) when $k \to \infty$ and $\partial_1 f$ is continuous, bounded by L_R^* on K_R^*, then we get from Lebesgue's dominated convergence theorem that $(A_k)_{k \in \mathbb{N}}$ converges to A in $\mathrm{L}^{r_\alpha}([a, b], \mathbb{R}^{n \times n})$. Integral representations yield

$$\begin{aligned} w_k(t) - w(t) = & \frac{1}{\Gamma(\alpha)} \int_a^t (t - \tau)^{\alpha - 1} A_k(\tau) \times (w_k(\tau) - w(\tau)) \, d\tau \\ & + \frac{1}{\Gamma(\alpha)} \int_a^t (t - \tau)^{\alpha - 1} (A_k(\tau) - A(\tau)) \times w(\tau) \, d\tau, \end{aligned}$$

for all $t \in [a,b]$ and all $k \in \mathbb{N}$. With Hölder inequality, this gives

$$\|w_k(t) - w(t)\|_{\mathbb{R}^n} \leq M_\alpha^1 \|w\|_{\mathcal{C}} \|A_k - A\|_{\mathrm{L}^{r_\alpha}} + \frac{L_R^*}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \|w_k(\tau) - w(\tau)\|_{\mathbb{R}^n} \, d\tau,$$

for all $t \in [a,b]$ and all $k \in \mathbb{N}$. Using the fractional version of the Gronwall lemma given in Proposition B.1 implies that

$$\|w_k(t) - w(t)\|_{\mathbb{R}^n} \leq M_\alpha^1 \|w\|_{\mathcal{C}} \|A_k - A\|_{\mathrm{L}^{r_\alpha}} E_{\alpha,1}(L_R^*(b-a)^\alpha),$$

for every $t \in [a,b]$ and all $k \in \mathbb{N}$. The proof is achieved by letting $k \to \infty$. ∎

Now, using the continuity result of Proposition A.2, the fact that φ and g are of class $\mathcal{C}^1$ and that $w_y(b, u_k^R, x_{a,k}^R)$ tends to $w_y(b, u^*, x^*(a))$ when $k \to \infty$ (Lemma A.12), we get from Inequality (A.15) that

$$\begin{aligned} &\big\langle \Psi^{0R} \partial_1 \varphi(x^*(a), x^*(b)) + \partial_1 g(x^*(a), x^*(b))^\top \times \Psi^R, y \big\rangle_{\mathbb{R}^n} \\ &+ \big\langle \Psi^{0R} \partial_2 \varphi(x^*(a), x^*(b)) + \partial_2 g(x^*(a), x^*(b))^\top \times \Psi^R, w_y(b, u^*, x^*(a)) \big\rangle_{\mathbb{R}^n} \leq 0. \end{aligned} \tag{A.16}$$

The above inequality constitutes the second crucial inequality depending on $R \in \mathbb{N}$ fixed.

A.2.5. Two crucial inequalities independent of $R \in \mathbb{N}$. Inequalities (A.14) and (A.16) depend both on $R \in \mathbb{N}$ such that $R \geq \|u^*\|_{\mathrm{L}^\infty}$. In particular, Inequality (A.14) is satisfied only for all $(s,v) \in \mathcal{L}^R \times (\mathrm{U} \cap \overline{\mathrm{B}}_{\mathbb{R}^m}(0,R))$. Our goal in this section is to get rid of the dependence in R. Let us define

$$\mathcal{L} := \bigcap_{\substack{R \in \mathbb{N} \\ R \geq \|u^*\|_{\mathrm{L}^\infty}}} \mathcal{L}^R.$$

Note that $\mathcal{L}$ has a full Lebesgue measure equal to $b-a$. From the equality $|\Psi^{0R}|^2 + \|\Psi^R\|_{\mathbb{R}^j}^2 = 1$ and extracting some subsequences (denoted similarly), $(\Psi^{0R})_{R \in \mathbb{N}}$ converges to some Ψ^0 in $\mathbb{R}$ and $(\Psi^R)_{R \in \mathbb{N}}$ converges to some Ψ in $\mathbb{R}^j$ when $R \to \infty$. Moreover, it holds that $\Psi^0 \leq 0$, $|\Psi^0|^2 + \|\Psi\|_{\mathbb{R}^j}^2 = 1$ and $-\Psi \in \mathcal{N}_{\mathrm{C}}[g(x^*(a), x^*(b))]$ (since $\mathcal{N}_{\mathrm{C}}[g(x^*(a), x^*(b))]$ is closed).

Let $y \in \mathbb{R}^n$ and $(s,v) \in \mathcal{L} \times \mathrm{U}$. For all $R \in \mathbb{N}$ such that $R \geq \max(\|u^*\|_{\mathrm{L}^\infty}, \|v\|_{\mathbb{R}^m})$, we have $s \in \mathcal{L}^R$ and $v \in \mathrm{U} \cap \overline{\mathrm{B}}_{\mathbb{R}^m}(0,R)$ and thus Inequalities (A.14) and (A.16) are satisfied. Letting $R \to \infty$ gives

$$\big\langle \Psi^0 \partial_2 \varphi(x^*(a), x^*(b)) + \partial_2 g(x^*(a), x^*(b))^\top \times \Psi, w_{(s,v)}(b, u^*, x^*(a)) \big\rangle_{\mathbb{R}^n} \leq 0, \tag{A.17}$$

and

$$\begin{aligned} &\big\langle \Psi^0 \partial_1 \varphi(x^*(a), x^*(b)) + \partial_1 g(x^*(a), x^*(b))^\top \times \Psi, y \big\rangle_{\mathbb{R}^n} \\ &+ \big\langle \Psi^0 \partial_2 \varphi(x^*(a), x^*(b)) + \partial_2 g(x^*(a), x^*(b))^\top \times \Psi, w_y(b, u^*, x^*(a)) \big\rangle_{\mathbb{R}^n} \leq 0. \end{aligned} \tag{A.18}$$

Hence, the two crucial Inequalities (A.17) and (A.18) are satisfied for all $y \in \mathbb{R}^n$, almost every $s \in [a,b]$ and all $v \in \mathrm{U}$.

A.2.6. End of the proof. Now we can end the proof of Theorem 3.12 with the introduction of the adjoint vector p.

• We first consider the case where g is submersive at $(x^*(a), x^*(b))$. The other case will be discussed next. Let us define $p^0 := \Psi^0 \in \mathbb{R}$ and

$$p(t) := \Phi(b,t)^\top \times \Big[p^0 \partial_2 \varphi(x^*(a), x^*(b)) + \partial_2 g(x^*(a), x^*(b))^\top \times \Psi\Big] \in \mathbb{R}^n,$$

for every $t \in [a,b)$, where $\Phi(\cdot,\cdot)$ denotes the left RL state-transition matrix (see [12, Definition 18] for details) associated to the essentially bounded matrix function $t \mapsto \partial_1 f(x^*(t), u^*(t), t) \in \mathbb{R}^{n\times n}$. Note that $p^0 \leq 0$ and, from the duality theorem in [12, Theorem 7], one can easily get that $p \in \mathcal{AC}^\alpha_{b-}([a,b],\mathbb{R}^n)$ is the unique maximal solution, which is moreover global, of the linear right RL fractional Cauchy problem given by

$$\text{(A.19)} \qquad \begin{cases} \mathrm{D}^\alpha_{b-}[p](t) = \partial_1 f(x^*(t), u^*(t), t)^\top \times p(t), \quad \text{a.e. } t \in [a,b], \\ \mathrm{I}^{1-\alpha}_{b-}[p](b) = p^0 \partial_2 \varphi(x^*(a), x^*(b)) + \partial_2 g(x^*(a), x^*(b))^\top \times \Psi. \end{cases}$$

The existence, uniqueness and globality of $p \in \mathrm{L}^1([a,b],\mathbb{R}^n)$ directly follow from the right counterpart of [12, Theorem 1]. Moreover, from the right version of [13, Theorem 6], recall that the product $\rho^\alpha_{b-} p \in \mathcal{C}([a,b],\mathbb{R}^n)$.

Fractional Hamiltonian system. From the definition of the Hamiltonian H and since the optimal pair (x^*, u^*) satisfies the state equation of Problem (OCP), it holds that

$${}_{\mathrm{c}}\mathrm{D}^\alpha_{a+}[x^*](t) = f(x^*(t), u^*(t), t) = \partial_3 H(x^*(t), u^*(t), p(t), p^0, t),$$

for almost every $t \in [a,b]$. Moreover, from (A.19), we have

$$\mathrm{D}^\alpha_{b-}[p](t) = \partial_1 f(x^*(t), u^*(t), t)^\top \times p(t) = \partial_1 H(x^*(t), u^*(t), p(t), p^0, t),$$

for almost every $t \in [a,b]$.

Hamiltonian maximization condition. We know that Inequality (A.17) is satisfied for almost every $s \in [a,b]$ and all $v \in \mathrm{U}$. From the fractional Duhamel formula (3.2), we can use the equality

$$w_{(s,v)}(b, u^*, x^*(a)) = \Phi(b,s) \times (f(x^*(s), v, s) - f(x^*(s), u^*(s), s)),$$

in Inequality (A.17) to infer that

$$\langle p(s), f(x^*(s), v, s) - f(x^*(s), u^*(s), s)\rangle_{\mathbb{R}^n} \leq 0,$$

for almost every $s \in [a,b]$ and all $v \in \mathrm{U}$. We conclude that

$$u^*(s) \in \operatorname*{arg\,max}_{v\in \mathrm{U}} H(x^*(s), v, p(s), p^0, s),$$

for almost every $s \in [a,b]$.

Transversality conditions on the adjoint vector. For the ease of notations, we introduce

$$p_a := p^0 \partial_1 \varphi(x^*(a), x^*(b)) + \partial_1 g(x^*(a), x^*(b))^\top \times \Psi,$$

$$p_b := p^0 \partial_2 \varphi(x^*(a), x^*(b)) + \partial_2 g(x^*(a), x^*(b))^\top \times \Psi,$$

and we recall that $-\Psi \in \mathcal{N}_{\mathrm{C}}[g(x^*(a), x^*(b))]$. From (A.19), we get that $\mathrm{I}_{b-}^{1-\alpha}[p](b) = p_b$. In order to derive the second transversality condition, we use Inequality (A.18) that is satisfied for all $y \in \mathbb{R}^n$. From the fractional Duhamel formula (3.3), we can use the equality

$$w_y(b, u^*, x^*(a)) = y + \int_a^b \Phi(b, \tau) \times \partial_1 f(x^*(\tau), u^*(\tau), \tau)\, d\tau \times y,$$

in Inequality (A.18) to obtain that

$$\left\langle p_a + p_b + \int_a^b \partial_1 f(x^*(\tau), u^*(\tau), \tau)^\top \times \Phi(b, \tau)^\top \times p_b\, d\tau, y \right\rangle_{\mathbb{R}^n} \leq 0,$$

that is,

$$\left\langle p_a + p_b + \int_a^b \mathrm{D}_{b-}^{\alpha}[p](\tau)\, d\tau, y \right\rangle_{\mathbb{R}^n} \leq 0,$$

for all $y \in \mathbb{R}^n$. From the equalities $\mathrm{D}_{b-}^{\alpha}[p] = -\frac{d}{dt}[\mathrm{I}_{b-}^{1-\alpha}[p]]$ and $\mathrm{I}_{b-}^{1-\alpha}[p](b) = p_b$, we deduce that

$$\left\langle p_a + \mathrm{I}_{b-}^{1-\alpha}[p](a), y \right\rangle_{\mathbb{R}^n} \leq 0,$$

for all $y \in \mathbb{R}^n$. We conclude that $\mathrm{I}_{b-}^{1-\alpha}[p](a) = -p_a$.

Nontriviality of the couple (p, p^0). By contradiction, let us assume that the pair (p, p^0) is zero. In that case, we have $p^0 = \Psi^0 = 0$ and $\mathrm{I}_{b-}^{1-\alpha}[p](a) = \mathrm{I}_{b-}^{1-\alpha}[p](b) = 0_{\mathbb{R}^n}$. The transversality conditions give $\partial_1 g(x^*(a), x^*(b))^\top \times \Psi = \partial_2 g(x^*(a), x^*(b))^\top \times \Psi = 0_{\mathbb{R}^n}$. Since g is submersive at $(x^*(a), x^*(b))$, we deduce that $\Psi = 0_{\mathbb{R}^j}$ which raises a contradiction with the equality $|\Psi^0|^2 + \|\Psi\|_{\mathbb{R}^j}^2 = 1$.

• If g is not submersive at $(x^*(a), x^*(b))$, we have to see that if a couple (x^*, u^*) is an optimal solution to Problem (OCP), then (x^*, u^*) is also an optimal solution to a similar problem where the function g is replaced by the identity function $\tilde{g} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \to \mathbb{R}^n$, $\tilde{g}(x_a, x_b) := (x_a, x_b)$, which is submersive at any point, and where the closed convex set C is replaced by the singleton $\tilde{\mathrm{C}} := \{(x^*(a), x^*(b))\}$. So we get back to the submersive case, but with a different function g and a different closed convex set C that would impact only the transversality conditions in Theorem 3.12. With the use of $\tilde{g}$ and $\tilde{\mathrm{C}}$, note that the transversality conditions do not provide any information.

Appendix B. Two technical results. This appendix is devoted to the statements and proofs of two technical results that are required in this paper. First we provide a new fractional version of the Gronwall lemma (Proposition B.1). Let us recall that a lot of fractional versions of this classical result (which allows to derive estimations on trajectories of differential equations) have already been established in the literature (see, e.g., [21, 55] and references therein). For our purposes, we recall that the classical Mittag-Leffler function $E_{\alpha,\alpha'}$ associated to $\alpha > 0$ and $\alpha' > 0$ is usually defined by

$$E_{\alpha,\alpha'}(t) := \sum_{k \in \mathbb{N}} \frac{t^k}{\Gamma(\alpha k + \alpha')},$$

for all $t \geq 0$. Note that $E_{\alpha,\alpha'}$ is continuous and monotonically increasing. Let $a < b$ be two real numbers.

PROPOSITION B.1 (Fractional Gronwall lemma). *Let $\alpha > 0$, $\alpha_1' > 0$, $\alpha_2' > 0$, $\kappa \in (0,1]$ and let $x \in \mathrm{L}^1([a,b],\mathbb{R})$ such that the product $\rho_{a+}^{\kappa} x \in \mathcal{C}([a,b],\mathbb{R})$. If there exist some nonnegative constants $c_1 \geq 0$, $c_2 \geq 0$ and $c_3 \geq 0$ such that*

$$0 \leq x(t) \leq \frac{c_1}{\Gamma(\alpha_1')}(t-a)^{\alpha_1'-1} + \frac{c_2}{\Gamma(\alpha_2')}(t-a)^{\alpha_2'-1} + c_3 \mathrm{I}_{a+}^{\alpha}[x](t),$$

for every $t \in (a,b]$, then it holds that

$$0 \leq x(t) \leq c_1(t-a)^{\alpha_1'-1} E_{\alpha,\alpha_1'}\Big(c_3(t-a)^{\alpha}\Big) + c_2(t-a)^{\alpha_2'-1} E_{\alpha,\alpha_2'}\Big(c_3(t-a)^{\alpha}\Big),$$

for every $t \in (a,b]$.

Proof. One can easily prove by induction that

$$0 \leq x(t) \leq c_1(t-a)^{\alpha_1'-1}\left(\sum_{k=0}^{N} \frac{\Big(c_3(t-a)^{\alpha}\Big)^k}{\Gamma(\alpha k + \alpha_1')}\right) + c_2(t-a)^{\alpha_2'-1}\left(\sum_{k=0}^{N} \frac{\Big(c_3(t-a)^{\alpha}\Big)^k}{\Gamma(\alpha k + \alpha_2')}\right) + c_3^{N+1}\mathrm{I}_{a+}^{(N+1)\alpha}[x](t),$$

for every $t \in (a,b]$ and for all $N \in \mathbb{N}$. The proof is concluded by noting that, since the product $\rho_{a+}^{\kappa} x \in \mathcal{C}([a,b],\mathbb{R})$, the last term tends to zero when $N \to \infty$. ∎

We conclude this appendix with the following technical lemma required in Appendix A.1.2 for the proof of Proposition 3.10.

LEMMA B.2. *Let $0 < \alpha' < \alpha \leq 1$ and let*

$$\begin{array}{rccl} \sigma_{\alpha,\alpha'} : & [1,+\infty) & \longrightarrow & \mathbb{R}_+ \\ & t & \longmapsto & \sigma_{\alpha,\alpha'}(t) := \left|(t-1)^{1-\alpha'}\left(\dfrac{t^{\alpha}-(t-1)^{\alpha}}{\alpha} - t^{\alpha-1}\right)\right|. \end{array}$$

The real function $\sigma_{\alpha,\alpha'}$ is bounded.

Proof. From the classical Taylor expansion

$$(1-X)^{\alpha} = 1 - \alpha X + \frac{\alpha(\alpha-1)}{2}X^2 + X^2\varepsilon(X),$$

where $\varepsilon(X) \to 0$ when $X \to 0$, it holds that

$$\begin{aligned} \sigma_{\alpha,\alpha'}(t) &= \left|\left(1-\frac{1}{t}\right)^{1-\alpha'} t^{1+\alpha-\alpha'}\left(\frac{1-\left(1-\frac{1}{t}\right)^{\alpha}}{\alpha} - \frac{1}{t}\right)\right| \\ &= \left|\left(1-\frac{1}{t}\right)^{1-\alpha'} t^{\alpha-\alpha'-1}\left(-\frac{\alpha-1}{2} - \frac{1}{\alpha}\varepsilon\left(\frac{1}{t}\right)\right)\right|, \end{aligned}$$

for all $t \geq 1$. Since $\alpha - \alpha' - 1 < 0$, it holds that $\sigma_{\alpha,\alpha'}(t) \to 0$ when $t \to +\infty$. The proof is complete from the continuity of $\sigma_{\alpha,\alpha'}$. ∎

REFERENCES

[1] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*, vol. 87 of Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 2004, https://doi.org/10.1007/978-3-662-06404-7, https://doi-org.ezproxy.unilim.fr/10.1007/978-3-662-06404-7. Control Theory and Optimization, II.

[2] O. P. Agrawal, *Formulation of Euler-Lagrange equations for fractional variational problems*, J. Math. Anal. Appl., 272 (2002), pp. 368–379, https://doi.org/10.1016/S0022-247X(02)00180-4, https://doi-org.ezproxy.unilim.fr/10.1016/S0022-247X(02)00180-4.

[3] O. P. Agrawal, *A general formulation and solution scheme for fractional optimal control problems*, Nonlinear Dynam., 38 (2004), pp. 323–337, https://doi.org/10.1023/B:NODY.0000045544.96418.bf, https://doi-org.ezproxy.unilim.fr/10.1023/B:NODY.0000045544.96418.bf.

[4] O. P. Agrawal, O. Defterli, and D. Baleanu, *Fractional optimal control problems with several state and control variables*, J. Vib. Control, 16 (2010), pp. 1967–1976, https://doi.org/10.1177/1077546309353361, https://doi-org.ezproxy.unilim.fr/10.1177/1077546309353361.

[5] R. Almeida and D. F. M. Torres, *Calculus of variations with fractional derivatives and fractional integrals*, Appl. Math. Lett., 22 (2009), pp. 1816–1820, https://doi.org/10.1016/j.aml.2009.07.002, https://doi-org.ezproxy.unilim.fr/10.1016/j.aml.2009.07.002.

[6] R. Almeida and D. F. M. Torres, *Necessary and sufficient conditions for the fractional calculus of variations with Caputo derivatives*, Commun. Nonlinear Sci. Numer. Simul., 16 (2011), pp. 1490–1500, https://doi.org/10.1016/j.cnsns.2010.07.016, https://doi-org.ezproxy.unilim.fr/10.1016/j.cnsns.2010.07.016.

[7] R. Almeida and D. F. M. Torres, *A discrete method to solve fractional optimal control problems*, Nonlinear Dynam., 80 (2015), pp. 1811–1816, https://doi.org/10.1007/s11071-014-1378-1, https://doi-org.ezproxy.unilim.fr/10.1007/s11071-014-1378-1.

[8] D. Baleanu and S. I. Muslih, *Lagrangian formulation of classical fields within Riemann-Liouville fractional derivatives*, Phys. Scripta, 72 (2005), pp. 119–121, https://doi.org/10.1238/Physica.Regular.072a00119, https://doi-org.ezproxy.unilim.fr/10.1238/Physica.Regular.072a00119.

[9] R. K. Biswas and S. Sen, *Free final time fractional optimal control problems*, J. Franklin Inst., 351 (2014), pp. 941–951, https://doi.org/10.1016/j.jfranklin.2013.09.024, https://doi-org.ezproxy.unilim.fr/10.1016/j.jfranklin.2013.09.024.

[10] J. Bonnans, *The shooting approach to optimal control problems*, IFAC Proceedings Volumes, 46 (2013), pp. 281–292.

[11] L. Bourdin, *Existence of a weak solution for fractional Euler-Lagrange equations*, J. Math. Anal. Appl., 399 (2013), pp. 239–251, https://doi.org/10.1016/j.jmaa.2012.10.008, https://doi-org.ezproxy.unilim.fr/10.1016/j.jmaa.2012.10.008.

[12] L. Bourdin, *Cauchy-Lipschitz theory for fractional multi-order dynamics - State-transition matrices, Duhamel formulas and duality theorems*, Differential Integral Equations, 31 (2018), pp. 559–594.

[13] L. Bourdin, *Weighted Hölder continuity of Riemann-Liouville fractional integrals - Application to regularity of solutions to fractional Cauchy problems with Carathéodory dynamics*, submitted, hal-01726069 (2018).

[14] L. Bourdin and D. Idczak, *A fractional fundamental lemma and a fractional integration by parts formula—Applications to critical points of Bolza functionals and to linear boundary value problems*, Adv. Differential Equations, 20 (2015), pp. 213–232, http://projecteuclid.org.ezproxy.unilim.fr/euclid.ade/1423055200.

[15] A. Bressan and B. Piccoli, *Introduction to the mathematical theory of control*, vol. 2 of AIMS Series on Applied Mathematics, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

[16] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.

[17] A. E. Bryson, Jr. and Y. C. Ho, *Applied optimal control*, Hemisphere Publishing Corp. Washington, D. C.; distributed by Halsted Press [John Wiley & Sons], New York-London-Sydney, 1975. Optimization, estimation, and control, Revised printing.

[18] F. Bullo and A. D. Lewis, *Geometric control of mechanical systems*, vol. 49 of Texts in Applied Mathematics, Springer-Verlag, New York, 2005, https://doi.org/10.1007/978-1-4899-7276-7, https://doi-org.ezproxy.unilim.fr/10.1007/978-1-4899-7276-7. Modeling, analysis, and design for simple mechanical control systems.

[19] L. Cesari, *Optimization—theory and applications*, vol. 17 of Applications of Mathematics (New York), Springer-Verlag, New York, 1983, https://doi.org/10.1007/978-1-4613-8165-5,

https://doi-org.ezproxy.unilim.fr/10.1007/978-1-4613-8165-5. Problems with ordinary differential equations.

[20] J. Cresson, *Fractional embedding of differential operators and Lagrangian systems*, J. Math. Phys., 48 (2007), pp. 033504, 34, https://doi.org/10.1063/1.2483292, https://doi-org.ezproxy.unilim.fr/10.1063/1.2483292.

[21] K. Diethelm, *The analysis of fractional differential equations*, vol. 2004 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2010, https://doi.org/10.1007/978-3-642-14574-2, https://doi-org.ezproxy.unilim.fr/10.1007/978-3-642-14574-2. An application-oriented exposition using differential operators of Caputo type.

[22] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl., 47 (1974), pp. 324–353.

[23] H. O. Fattorini, *Infinite-dimensional optimization and control theory*, vol. 62 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1999, https://doi.org/10.1017/CBO9780511574795, https://doi-org.ezproxy.unilim.fr/10.1017/CBO9780511574795.

[24] A. F. Filippov, *On some questions in the theory of optimal regulation: existence of a solution of the problem of optimal regulation in the class of bounded measurable functions*, Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him., 1959 (1959), pp. 25–32.

[25] G. a. S. F. Frederico and D. F. M. Torres, *Fractional optimal control in the sense of Caputo and the fractional Noether's theorem*, Int. Math. Forum, 3 (2008), pp. 479–493.

[26] R. V. Gamkrelidze, *Discovery of the maximum principle*, in Mathematical events of the twentieth century, Springer, Berlin, 2006, pp. 85–99, https://doi.org/10.1007/3-540-29462-7_5, https://doi-org.ezproxy.unilim.fr/10.1007/3-540-29462-7_5.

[27] T. L. Guo, *The necessary conditions of fractional optimal control in the sense of Caputo*, J. Optim. Theory Appl., 156 (2013), pp. 115–126, https://doi.org/10.1007/s10957-012-0233-0, https://doi-org.ezproxy.unilim.fr/10.1007/s10957-012-0233-0.

[28] M. R. Hestenes, *Calculus of variations and optimal control theory*, Robert E. Krieger Publishing Co., Inc., Huntington, N.Y., 1980. Corrected reprint of the 1966 original.

[29] R. Hilfer, *Applications of Fractional Calculus in Physics*, Applications of Fractional Calculus in Physics, World Scientific, 2000, https://books.google.fr/books?id=lSCIQgAACAAJ.

[30] D. G. Hull, *Optimal control theory for applications*, Mechanical Engineering Series, Springer-Verlag, New York, 2003, https://doi.org/10.1007/978-1-4757-4180-3, https://doi-org.ezproxy.unilim.fr/10.1007/978-1-4757-4180-3.

[31] Z. D. Jelicic and N. Petrovacki, *Optimality conditions and a solution scheme for fractional optimal control problems*, Struct. Multidiscip. Optim., 38 (2009), pp. 571–581, https://doi.org/10.1007/s00158-008-0307-7, https://doi-org.ezproxy.unilim.fr/10.1007/s00158-008-0307-7.

[32] V. Jurdjevic, *Geometric control theory*, vol. 52 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1997.

[33] R. Kamocki, *On the existence of optimal solutions to fractional optimal control problems*, Appl. Math. Comput., 235 (2014), pp. 94–104, https://doi.org/10.1016/j.amc.2014.02.086, https://doi-org.ezproxy.unilim.fr/10.1016/j.amc.2014.02.086.

[34] R. Kamocki, *Pontryagin maximum principle for fractional ordinary optimal control problems*, Math. Methods Appl. Sci., 37 (2014), pp. 1668–1686, https://doi.org/10.1002/mma.2928, https://doi-org.ezproxy.unilim.fr/10.1002/mma.2928.

[35] R. Kamocki and M. Majewski, *Fractional linear control systems with Caputo derivative and their optimization*, Optimal Control Appl. Methods, 36 (2015), pp. 953–967, https://doi.org/10.1002/oca.2150, https://doi-org.ezproxy.unilim.fr/10.1002/oca.2150.

[36] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and applications of fractional differential equations*, vol. 204 of North-Holland Mathematics Studies, Elsevier Science B.V., Amsterdam, 2006.

[37] E. B. Lee and L. Markus, *Foundations of optimal control theory*, Robert E. Krieger Publishing Co., Inc., Melbourne, FL, second ed., 1986.

[38] X. J. Li and J. M. Yong, *Optimal control theory for infinite-dimensional systems*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 1995, https://doi.org/10.1007/978-1-4612-4260-4, https://doi-org.ezproxy.unilim.fr/10.1007/978-1-4612-4260-4.

[39] D. Liberzon, *Calculus of variations and optimal control theory*, Princeton University Press, Princeton, NJ, 2012. A concise introduction.

[40] A. B. Malinowska and D. F. M. Torres, *Introduction to the fractional calculus of variations*, Imperial College Press, London, 2012, https://doi.org/10.1142/p871, https://doi-org.ezproxy.unilim.fr/10.1142/p871.

[41] I. Matychyn and V. Onyshchenko, *Time-optimal control of fractional-order linear systems*,

Fract. Calc. Appl. Anal., 18 (2015), pp. 687–696, https://doi.org/10.1515/fca-2015-0042, https://doi-org.ezproxy.unilim.fr/10.1515/fca-2015-0042.

[42] T. Odzijewicz, A. B. Malinowska, and D. F. M. Torres, *Fractional variational calculus with classical and combined Caputo derivatives*, Nonlinear Anal., 75 (2012), pp. 1507–1515, https://doi.org/10.1016/j.na.2011.01.010, https://doi-org.ezproxy.unilim.fr/10.1016/j.na.2011.01.010.

[43] I. Podlubny, *Fractional differential equations*, vol. 198 of Mathematics in Science and Engineering, Academic Press, Inc., San Diego, CA, 1999. An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications.

[44] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt, Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962.

[45] S. Pooseh, R. Almeida, and D. F. M. Torres, *Fractional order optimal control problems with free terminal time*, J. Ind. Manag. Optim., 10 (2014), pp. 363–381, https://doi.org/10.3934/jimo.2014.10.363, https://doi-org.ezproxy.unilim.fr/10.3934/jimo.2014.10.363.

[46] F. Riewe, *Mechanics with fractional derivatives*, Phys. Rev. E (3), 55 (1997), pp. 3581–3592, https://doi.org/10.1103/PhysRevE.55.3581, https://doi-org.ezproxy.unilim.fr/10.1103/PhysRevE.55.3581.

[47] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives*, Gordon and Breach Science Publishers, Yverdon, 1993. Theory and applications, Edited and with a foreword by S. M. Nikolskiui, Translated from the 1987 Russian original, Revised by the authors.

[48] H. Schättler and U. Ledzewicz, *Geometric optimal control*, vol. 38 of Interdisciplinary Applied Mathematics, Springer, New York, 2012, https://doi.org/10.1007/978-1-4614-3834-2, https://doi-org.ezproxy.unilim.fr/10.1007/978-1-4614-3834-2. Theory, methods and examples.

[49] R. Scherer, S. L. Kalla, Y. Tang, and J. Huang, *The Grünwald-Letnikov method for fractional differential equations*, Comput. Math. Appl., 62 (2011), pp. 902–917, https://doi.org/10.1016/j.camwa.2011.03.054, https://doi-org.ezproxy.unilim.fr/10.1016/j.camwa.2011.03.054.

[50] S. P. Sethi and G. L. Thompson, *Optimal control theory*, Kluwer Academic Publishers, Boston, MA, second ed., 2000. Applications to management science and economics.

[51] E. Trélat, *Contrôle optimal*, Mathématiques Concrètes. [Concrete Mathematics], Vuibert, Paris, 2005. Théorie & applications. [Theory and applications].

[52] C. Tricaud and Y. Chen, *Time-optimal control of systems with fractional dynamics*, Int. J. Differ. Equ., (2010), pp. Art. ID 461048, 16.

[53] R. Vinter, *Optimal control*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 2000.

[54] D. H. Wagner, *Survey of measurable selection theorems*, SIAM J. Control Optimization, 15 (1977), pp. 859–903, https://doi.org/10.1137/0315056, https://doi-org.ezproxy.unilim.fr/10.1137/0315056.

[55] H. Ye, J. Gao, and Y. Ding, *A generalized Gronwall inequality and its application to a fractional differential equation*, J. Math. Anal. Appl., 328 (2007), pp. 1075–1081, https://doi.org/10.1016/j.jmaa.2006.05.061, https://doi-org.ezproxy.unilim.fr/10.1016/j.jmaa.2006.05.061.

[56] S. A. Yousefi, M. Dehghan, and A. Lotfi, *Generalized Euler-Lagrange equations for fractional variational problems with free boundary conditions*, Comput. Math. Appl., 62 (2011), pp. 987–995, https://doi.org/10.1016/j.camwa.2011.03.064, https://doi-org.ezproxy.unilim.fr/10.1016/j.camwa.2011.03.064.