On the strong local optimality for state-constrained control-affine problems

F C Chittaro, L Poggiolini

▶ To cite this version:

F C Chittaro, L Poggiolini. On the strong local optimality for state-constrained control-affine problems. Nonlinear Differential Equations and Applications, 2023, 10.1007/s00030-023-00870-y . hal-04046629

HAL Id: hal-04046629
https://hal.science/hal-04046629

Submitted on 26 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the strong local optimality for state-constrained control-affine problems.

F. C. Chittaro* and L. Poggiolini

Abstract. In this article we establish first and second order sufficient optimality conditions for a class of single-input control-affine problems, in presence of a scalar state constraint. We consider strong-local optimality (that is, the C^0 topology in the state space). The minimum-time and the Mayer problem are addressed. We restrict our analysis to extremals containing a bang arc, a single boundary arc, followed by a finite number of bang arcs.

The sufficient conditions are expressed as a strengthened version of the necessary ones, plus the coerciveness of a suitable finite-dimensional quadratic form. The sufficiency of the given conditions is proven via Hamiltonian methods.

Mathematics Subject Classification (2010). 49K15, 49K30.

Keywords. Optimal control, sufficient optimality conditions, pure state constraints, Pontryagin maximum principle.

1. Introduction

State-constrained optimal control problems are ubiquitous in many applications fields, for instance: optimization of running [19, 24] and driving [11] strategies, aerospace dynamics [8, 12, 28], mechanics [34], model predictive control [14], optimization of medical treatments [13]. For this class of problems, suitable extensions of Pontryagin Maximum Principle provide the main necessary conditions for the optimality of admissible trajectories (for a survey, see for instance [15]). In general, such extensions are quite technical: in particular, the multiplier associated with each inequality state constraint is given by a measure, explicitly appearing in the adjoint equation; among the consequences, we may have discontinuities of the adjoint covector when the trajectory saturates the constraint. However, it is well known that, for autonomous systems and under some conditions on the constraint and the control associated with the candidate trajectory, the multiplier measure is absolutely continuous with respect to the Lebesgue measure; additional conditions may also ensure the continuity of the adjoint covector [8, 15].

Among all control systems, control-affine ones play a major role in applications, being sufficiently general to cover a large amount of practical cases. In this paper, we consider autonomous optimal control problems, defined on a smooth manifold M, in presence of a scalar state constraint (i.e., we assume that some function c must be non-positive along each admissible trajectory) and subject to single-input control-affine dynamics with bounded control.

We deal with the Mayer and the minimum-time problems, with fixed initial point, and constraining the final one to belong to some smooth manifold $\mathcal{N}_f$ of the state space. We assume that we are given a Pontryagin extremal having a special structure; namely, we assume that the extremal trajectory is composed by the concatenation of an internal bang arc (that is, an arc along which the constraint is not saturated, while the control is), a boundary arc (i.e., saturating the constraint), followed by the concatenation of several internal bang arcs. We also assume that, along the boundary arc, the control takes its values in the interior of the set of admissible control values. Our purpose is to provide sufficient conditions assuring that the candidate extremal is a strong local minimizer, according to the following definition.

Definition 1.1 (Strong local optimality). *Let $\widehat{\xi}$ be an admissible trajectory of an optimal control problem. We say that $\widehat{\xi}$ is a strong local minimizer of such problem if there exists a neighbourhood $\mathcal{O}$ of its graph in $\mathbb{R} \times M$ such that $\widehat{\xi}$ is a minimizer among all admissible trajectories whose graphs are contained in $\mathcal{O}$, independently of the values of the associated control function. We say that $\widehat{\xi}$ is a strict strong local minimizer if it is the only minimizing trajectory whose graph is in $\mathcal{O}$.*

We recall that, in literature, other notions of local optimality, involving topologies in the space of *(trajectory,control)*-pairs, are studied; we mention, for instance, the weak local optimality and the so-called Pontryagin optimality (see [6, 7]). Sufficient conditions for Pontryagin optimality in this class of problems have been provided in [6], where concatenations of a finite number of bang, boundary and singular arcs are considered, and in [7]. We also mention [27], where it is shown that strengthened versions of the necessary conditions for optimality guarantee the local embedding of the reference trajectory into a local field of extremals.

Our approach is based on Hamiltonian methods. These methods, inspired by the technique of fields of extremals in Calculus of Variations, rely on the following steps (see [3, Chapter 17]). Assume that we want to prove strong-local optimality of some admissible trajectory $\widehat{\xi} : [0, T] \to M$; then

i) find a Lagrangian submanifold Λ_ρ of the cotangent bundle, containing the extremal associated with $\widehat{\xi}$ at $t = 0$ and such that, at each time $t \in [0, T]$, its image under the flow generated by the maximized Hamiltonian projects injectively onto a neighborhood of the reference trajectory; in this case, we say then that the maximized Hamiltonian flow is *locally invertible*;

ii) thanks to the local invertibility of this flow, lift all admissible trajectories, with graph belonging to some neighborhood of the graph of $\widehat{\xi}$, to Λ_ρ; we say that we cover a neighborhood of the graph of the reference trajectory with a *field of state extremals* ;

iii) estimate the cost associated with every trajectory, by means of a line integral along the lifts of the state extremals in Λ_ρ.

This scheme applies for quite regular situations, namely: when the maximized Hamiltonian is at least C^2 and its flow is well defined on a neighborhood of the extremal associated with the reference trajectory. In particular, bang-bang extremals or singular extremals (often encountered in Mayer or minimum-time problem, for control-affine systems) do not fit in this frame. Yet, even in regular cases, the scheme does not suggest under which conditions such a manifold Λ_ρ exists.

For bang-bang extremals, some regularity conditions ensure that the maximized Hamiltonian is well defined and piecewise smooth in a tubular neighborhood of the extremal of interest; it is worth noticing that, usually, these regularity conditions are just the strengthened version of some necessary optimality conditions already contained in the PMP. In addition, as shown in [5] and [25] (dealing with bang-bang extremals in the Mayer and minimum-time problem, respectively), the invertibility of the maximized Hamiltonian flow is guaranteed if a suitably defined (finite dimensional) quadratic form (the *second variation*) is coercive.

Singular arcs jeopardize this scheme: in fact, the Hamiltonian vector field associated with the maximized Hamiltonian, in each tubular neighborhood of the singular arc, is well defined only as a multivalued function. As observed for the first time in [30], the problem can be solved by substituting the maximized Hamiltonian with an *overmaximized* one, with some particular features. The "natural" second variation associated with singular arcs is completely degenerate, but it is often possible to extend it (by means of the so-called *Goh transformation*) in order to obtain a non-degenerate functional; the coercivity of such functional can then be used to guarantee the existence of a manifold Λ_ρ with the required properties ([29, 30]).

As already said, here we consider a trajectory composed by the concatenation of internal bang arcs and a boundary arc. Under the assumption that, along the boundary arc, the control is internal, the associated extremal shares some properties with singular arcs, so that we choose to treat boundary arcs as singular ones; this means that we must construct an overmaximized Hamiltonian, and that the natural second variation associated with the problem is infinite-dimensional. Nevertheless, differently from the singular case, we are able to prove our result by studying the coercivity of the restriction of the extended second variation to a finite dimensional space of variations. Also, differently from the case of singular non-constrained extremals, we cannot construct an overmaximised Hamiltonian on a full neighborhood of the boundary arc of the extremal, but only on the lift of admissible arcs, i.e. arcs that satisfy the state constraints.

We provide sufficient condition for strong local optimality of an extremal; these sufficient conditions consist in

- a bunch of regularity conditions on the boundary arc (Assumptions 1-2);
- a set of regularity conditions on the extremal, that boil down to a strengthening of the necessary ones (Assumptions 3-4-5);
- the coercivity of a suitably defined finite-dimensional quadratic form (Assumption 6).

Up to the authors' knowledge, this paper is the first attempt to apply Hamiltonian methods, as first developed in [2], to state-constrained optimal control problems.

In the next section, we will state the problem, introduce Pontryagin Maximum Principle and declare the assumptions; in Section 3, we construct the overmaximized flow and in Section 4 we construct the second variation, both for the minimum-time and the Mayer problem; in Section 5, we prove that the maximized Hamiltonian flow is locally invertible.

In Sections 6 and 7 we prove our main results, the sufficient optimality conditions, for the minimum-time and the Mayer problem, respectively. The results are illustrated by some examples. We conclude the paper with an insight on perspective work.

2. Statement of the problem

Let M be a smooth n-dimensional manifold. We denote with TM and with T^*M the tangent bundle and the cotangent bundle to M, respectively. π denotes the canonical projection of T^*M on M; the elements of T^*M are denoted with ℓ.

In the following, small letters f, g, k denote vector fields on the manifold M, and the corresponding capital letters are used to denote the corresponding Hamiltonian lift, i.e. $F(\ell) = \langle \ell, f(\pi\ell)\rangle$. Given a vector field f on M, the Lie derivative of a smooth function $\varphi\colon M \to \mathbb{R}$ with respect to f, evaluated at the point $q \in M$, is denoted with $L_f\varphi(q) = \langle \mathrm{d}\varphi(q), f(q)\rangle$; analogously, $L^2_f\varphi(q) = L_f\big(L_f\varphi\big)(q)$. The Lie bracket of two vector fields f, g is denoted as commonly with $[f, g]$.

The symbol ς denotes the Poincaré-Cartan invariant on T^*M, defined as $\varsigma_\ell = \ell \circ \pi_* \ \forall \ell \in T^*M$. The symbol $\boldsymbol{\sigma}_\ell = \mathrm{d}\varsigma_\ell$ denotes the canonical symplectic form on T^*M. With each Hamiltonian function F we associate the Hamiltonian vector field $\vec{F}$ on T^*M defined by

$$\langle \mathrm{d}F(\ell)\,,\,\cdot\rangle = \boldsymbol{\sigma}_\ell(\cdot, \vec{F}(\ell)).$$

Let $x_0 \in M$ be a fixed point and let $\mathcal{N}_f \subset M$ be a smooth submanifold of positive codimension, with $x_0 \notin \mathcal{N}_f$. We consider the following state-constrained control system

$$\begin{aligned}
&\dot{\xi}(t) = f_0(\xi(t)) + u(t) f_1(\xi(t)) \quad \text{a.e. } t \in [0,T], && \text{(2.1a)}\\
&\xi(0) = x_0, \quad \xi(T) \in \mathcal{N}_f, && \text{(2.1b)}\\
&c(\xi(t)) \leq 0 \qquad \forall t \in [0,T], && \text{(2.1c)}\\
&|u(t)| \leq 1 \qquad \text{a.e. } t \in [0,T], && \text{(2.1d)}
\end{aligned}$$

where the function $c\colon M \to \mathbb{R}$ defining the state constraint is assumed to be smooth on a neighborhood of its zero-level set; f_0, f_1 are smooth vector fields on M. We denote the boundary of the admissible region as

$$\mathcal{C} = \{x \in M\colon c(x) = 0\}.$$

We associate with the problem (2.1) a cost J to be minimized, that can be either in Mayer form or the minimum time to reach $\mathcal{N}_f$. More precisely, we deal either with the problem

$$\text{minimize} \quad \psi(\xi(T)), \quad T > 0 \text{ fixed}, \tag{\textbf{M}}$$

where $\psi\colon M \to \mathbb{R}$ is a smooth function, or with the problem

$$\text{minimize} \quad T, \quad T > 0 \text{ free}. \tag{\textbf{T}}$$

If we consider problem (2.1)-(**M**), we assume we are given a candidate trajectory-control pair $(\widehat{\xi}, \widehat{u})$ where the reference trajectory $\widehat{\xi}$ is a concatenation of an *internal* arc, a *boundary* arc and another internal arc, i.e. there exist $0 < \widehat{t}_1 < \widehat{t}_2 < T$ such that

$$\begin{aligned}
c(\widehat{\xi}(t)) < 0 \quad \forall t \in [0, \widehat{t}_1) \cup (\widehat{t}_2, T],\\
c(\widehat{\xi}(t)) = 0 \quad \forall t \in [\widehat{t}_1, \widehat{t}_2].
\end{aligned} \tag{2.2}$$

In particular, we say that the constraint c is *saturated* for $t \in [t_1, t_2]$. The time $\widehat{t}_1$ is called an *entry time*, while the time $\widehat{t}_2$ is called an *exit time*; taken together, they are called *junction times.* The intervals $[0, \widehat{t}_1)$ and $(\widehat{t}_2, T]$ are called *internal intervals* and the interval $[\widehat{t}_1, \widehat{t}_2]$ is called a *boundary interval.*

We assume that the first internal arc is bang, while the second internal arc is given by the concatenation of several bang arcs; more precisely, we assume that there exist $\widehat{t}_i$, $i = 3, \ldots, N$ such that $\widehat{t}_2 < \widehat{t}_3 < \ldots < \widehat{t}_N < \widehat{t}_{N+1} = T$ and such that the control $\widehat{u}$ has the following structure:

$$\widehat{u}(t) = \begin{cases} u_1 & t \in [0, \widehat{t}_1),\\ \widehat{u}_s(t) & t \in [\widehat{t}_1, \widehat{t}_2],\\ u_i & t \in (t_{i-1}, t_i), \ i = 3, \ldots, N+1, \end{cases} \tag{2.3}$$

where $|u_i| = 1$ and $u_i u_{i+1} = -1$. We also introduce the following notations

$$\widehat{x}_i = \widehat{\xi}(\widehat{t}_i), \quad i = 1, \ldots, N, \qquad \widehat{x}_f = \widehat{\xi}(T).$$

In particular, $\widehat{x}_1$ and $\widehat{x}_2$ are called entry and exit points, respectively, and the points $\widehat{x}_i$, $i \geq 3$ are called switching points.

Analogously, when considering the problem (2.1)-(**T**), we assume we are given a candidate triple $(\widehat{T}, \widehat{\xi}, \widehat{u})$, where $\widehat{\xi}$ and $\widehat{u}$ are defined on $[0, \widehat{T}]$, and satisfy (2.2)-(2.3), with $\widehat{t}_{N+1} = \widehat{T}$.

We start by making some assumptions on the reference pair $(\widehat{\xi}, \widehat{u})$ along the boundary arc. As anticipated, these assumptions allows us to simplify the statement of the Pontryagin Maximum Principle.

Assumption 1 (internal control). *We assume that* $\operatorname*{ess\,sup}_{t\in[\widehat{t}_1,\widehat{t}_2]}|\widehat{u}_s(t)| < 1.$

Assumption 1 says that $\widehat{u}$ is discontinuous at the junction points.

Assumption 2 (Order of the constraint). *The boundary constraint is of first order, that is*

$$L_{f_1}c(\widehat{\xi}(t)) \neq 0 \quad \forall t \in [\widehat{t}_1, \widehat{t}_2].$$

We notice that Assumptions 1-2 imply that the entry and exit points are regular, that is

$$\lim_{t\to\widehat{t}_1^-} L_{f_0+\widehat{u}(t)f_1}c(\widehat{\xi}(t)) \neq 0, \qquad \lim_{t\to\widehat{t}_2^+} L_{f_0+\widehat{u}(t)f_1}c(\widehat{\xi}(t)) \neq 0.$$

Equivalently

$$L_{f_0+u_1f_1}c(\widehat{x}_1) > 0, \qquad L_{f_0+u_3f_1}c(\widehat{x}_2) < 0,$$

which imply that

$$u_1L_{f_1}c(\widehat{\xi}(t)) > 0, \qquad u_3L_{f_1}c(\widehat{\xi}(t)) < 0 \qquad \forall t \in [\widehat{t}_1, \widehat{t}_2],$$

so that $u_1u_3 = -1$.

Remark 2.1. *By continuity, if Assumption 2 holds true, then there exists a neighborhood* $U_{\mathcal{C}}$ *in* M *of* $\widehat{\xi}([\widehat{t}_1, \widehat{t}_2])$ *such that* $f_1(x) \neq 0$ *and* $L_{f_1}c(x) \neq 0$ *for any* $x \in U_{\mathcal{C}}$*. Moreover, the level set* $\mathcal{C} \cap U_{\mathcal{C}}$ *is a codimension one smooth submanifold of* M*, and, for every* $x \in \mathcal{C} \cap U_{\mathcal{C}}$*, the tangent space* T_xM *splits as the direct sum* $T_x\mathcal{C} \oplus \mathbb{R}f_1(x)$*.*

Remark 2.2. *Thanks to Assumption 2, we can recover the value of* $\widehat{u}_s$ *as a feedback control. Indeed, from* $c(\widehat{\xi}(t)) = 0\ \forall t \in [\widehat{t}_1, \widehat{t}_2]$*, we get that*

$$\widehat{u}_s(t) = -\frac{L_{f_0}c(\widehat{\xi}(t))}{L_{f_1}c(\widehat{\xi}(t))} \qquad \forall t \in [\widehat{t}_1, \widehat{t}_2]. \tag{2.4}$$

In particular, we obtain that $\widehat{u}_s(\cdot)$ *is continuous and, by induction, that both* $\widehat{\xi}$ *and* $\widehat{u}_s$ *are* C^∞ *on* $[\widehat{t}_1, \widehat{t}_2]$*.*

In view of the preceding remark, we define the following feedback control

$$v_s(x) = -\frac{L_{f_0}c(x)}{L_{f_1}c(x)}, \qquad x \in U_{\mathcal{C}},$$

and the piecewise time-dependent vector field:

$$\widehat{f}_t = \begin{cases} h_1 = f_0 + u_1 f_1 & t \in [0, \widehat{t}_1), \\ h_2 = f_0 + v_s f_1 & t \in (\widehat{t}_1, \widehat{t}_2), \\ h_i = f_0 + u_i f_1 & t \in (\widehat{t}_{i-1}, t_i),\ i = 3, \dots, N, \\ h_{N+1} = f_0 + u_{N+1} f_1 & t \in (\widehat{t}_N, T]. \end{cases}$$

Notice that c is constant along any integral line of h_2.

In this paper, a special role is played by the exit time $\widehat{t}_2$ (see Remark 5.4); in particular, we denote with $\widehat{S}_t$ the flow of the vector field $\widehat{f}_t$ starting from time $\widehat{t}_2$, evolving backwards in time up to time $t = 0$, or forward in time up to time $t = T$; equivalently, $\widehat{S}_t(q)$ denotes the solution at time t of the Cauchy problem

$$\begin{cases} \dot{\xi}(t) = \widehat{f}_t \circ \xi(t), \\ \xi(\widehat{t}_2) = q. \end{cases}$$

Analogously, $\widehat{\mathscr{F}}_t$ denotes the Hamiltonian flow associated with the Hamiltonian $\widehat{F}_t(\ell) = \langle \ell\,,\ \widehat{f}_t(\pi\ell) \rangle$, from time $\widehat{t}_2$ to time t. Finally, we denote

$$\widehat{c}_t(x) := c \circ \widehat{S}_t(x).$$

As already discussed in the introduction, for generic state-constrained optimal control problems, Pontryagin Maximum Principle involves the existence of a non-negative regular measure μ [1], supported only on the boundary interval $[\widehat{t}_1, \widehat{t}_2]$, which plays the role of the multiplier of the state constraint (see [15] for a survey). The measure μ appears in the adjoint equation for the covector. However, since the problem is autonomous and Assumptions 1-2 hold true, then the measure μ is absolutely continuous with respect to the Lebesgue measure; as a consequence, one can show that the adjoint covector is absolutely continuous and, in the Maximum Principle, we can deal directly with the density $-\eta$ of the measure μ (see [15, Proposition 4.1, Proposition 4.2] and [8, 21, 22]).

In order to state the PMP for the problem under concern, let us first define the pre-Hamiltonian associated with the optimal control problem

$$h(\ell, \eta, u) = F_0(\ell) + uF_1(\ell) + \eta c(\pi\ell), \quad \ell \in T^*M,\ \eta \in \mathbb{R},\ u \in [-1, 1]$$

and the maximized Hamiltonian

$$H_{\max}(\ell, \eta) = \max_{u \in [-1,1]} \big(F_0(\ell) + uF_1(\ell) + \eta c(\pi\ell) \big), \quad \ell \in T^*M,\ \eta \in \mathbb{R}.$$

Theorem 2.3 (PMP). *Assume that* $(\widehat{\xi}, \widehat{u})$ *is an optimal pair for* (2.1)-(**M**)*, that* (2.2)-(2.3) *hold true and that Assumptions* 1-2 *are satisfied by the pair. Then,*

[1] see, e.g., Chapter 1 in [10] for the regularity of measures defined via monotone right-continuous functions

there exist an absolutely continuous curve $\widehat{\lambda}\colon [0,T] \to T^*M$, *a continuous curve* $\widehat{\eta}\colon [0,T] \to (-\infty, 0]$ *and a constant* $p_0 \in \{0,1\}$ *such that*

$$(\widehat{\lambda}(t), \widehat{\eta}(t), p_0) \neq 0 \qquad \forall t \in [0,T], \tag{2.5a}$$

$$\pi\widehat{\lambda}(t) = \widehat{\xi}(t) \qquad \forall t \in [0,T], \tag{2.5b}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\lambda}(t) = \vec{h}(\widehat{\lambda}(t), \widehat{\eta}(t), \widehat{u}(t)) \qquad a.e.\ t \in [0,T], \tag{2.5c}$$

$$h(\widehat{\lambda}(t), \widehat{\eta}(t), \widehat{u}(t)) = H_{\max}(\widehat{\lambda}(t), \widehat{\eta}(t)) \qquad a.e.\ t \in [0,T], \tag{2.5d}$$

$$\widehat{\eta}(t)c(\widehat{\xi}(t)) = 0 \qquad \forall t \in [0,T], \tag{2.5e}$$

and the following transversality condition holds true

$$\widehat{\lambda}(T)|_{T_{\widehat{\xi}(T)}\mathcal{N}_f} = -p_0\,\mathrm{d}\psi(\widehat{\xi}(T)).$$

Let $(\widehat{T}, \widehat{\xi}, \widehat{u})$ *be an optimal triple for* (2.1)-(**T**) *such* (2.2)-(2.3) *and Assumptions* 1-2 *hold true. Then there exist an absolutely continuous curve* $\widehat{\lambda}\colon [0,\widehat{T}] \to T^*M$, *a continuous curve* $\widehat{\eta}\colon [0,\widehat{T}] \to (-\infty,0]$ *and a constant* $p_0 \in \{0,1\}$ *such that equations* (2.5) *hold on* $[0,\widehat{T}]$ *and moreover*

$$\widehat{\lambda}(\widehat{T}) \in T^{\perp}_{\widehat{\xi}(\widehat{T})}\mathcal{N}_f, \tag{2.6}$$

$$h(\widehat{\lambda}(t), \widehat{\eta}(t), \widehat{u}(t)) \equiv p_0 \quad \forall t \in [0,\widehat{T}].$$

A curve λ satisfying equations (2.5) is called an *extremal*; if $p_0 = 1$, then it is a *normal* extremal, otherwise it is said to be *abnormal.*

In the following, we assume that $\widehat{\lambda}$ is an extremal associated with the reference pair $(\widehat{\xi}, \widehat{u})$ (triple $(\widehat{T}, \widehat{\xi}, \widehat{u})$). In analogy with the notations for the entry/exit and switching points, we define

$$\widehat{\ell}_0 = \widehat{\lambda}(0), \qquad \widehat{\ell}_i = \widehat{\lambda}(\widehat{t}_i),\ i = 1, \ldots, N, \qquad \widehat{\ell}_T = \widehat{\lambda}(T).$$

PMP, in particular equation (2.5d), imposes that $\widehat{u}(t)F_1(\widehat{\lambda}(t)) \geq 0$ for a.e. t. Together with Assumption 1, this gives $F_1(\widehat{\lambda}(t)) \equiv 0$ for $t \in [\widehat{t}_1, \widehat{t}_2]$. Differentiating this equality with respect to time, we also obtain that

$$F_{01}(\widehat{\lambda}(t)) - \widehat{\eta}(t)L_{f_1}c(\widehat{\xi}(t)) = 0 \qquad \forall t \in (\widehat{t}_1, \widehat{t}_2),$$

where $F_{01}(\ell) := \langle \ell\,,\, [f_0, f_1](\pi\ell)\rangle$ is the Hamiltonian associated with the field $[f_0, f_1]$; thanks to Assumption 2, we can recover the expression of $\widehat{\eta}(t)$:

$$\widehat{\eta}(t) = \frac{F_{01}(\widehat{\lambda}(t))}{L_{f_1}c(\widehat{\xi}(t))} \quad \forall t \in (\widehat{t}_1, \widehat{t}_2). \tag{2.7}$$

PMP also states that, on each interval $(\widehat{t}_{i-1}, \widehat{t}_i)$, $i = 1, 3, \ldots, N+1$, the Hamiltonian

$$H_i := F_0 + u_i F_1$$

coincides with $H_{\max}$ along the reference extremal. As a consequence, at each $\widehat{t}_i$, $i = 1, 3, \ldots, N$, we have $\frac{\mathrm{d}}{\mathrm{d}t}(H_{i+1} - H_i)(\widehat{\lambda}(t)) \geq 0$.

Our next assumptions are in fact a strengthened version of the weak inequalities implied by PMP. The first one implies that, on each interval

$(\widehat{t}_{i-1}, \widehat{t}_i)$, $i = 1, 3, \dots, N+1$, u_i is the only value of the control maximizing the pre-Hamiltonian.

Assumption 3 (Regularity of the bang arcs).

$$u_1 F_1(\widehat{\lambda}(t)) > 0 \qquad t \in [0, \widehat{t}_1),$$
$$u_i F_1(\widehat{\lambda}(t)) > 0 \qquad t \in (\widehat{t}_{i-1}, \widehat{t}_i), \ i = 3, \dots, N,$$
$$u_{N+1} F_1(\widehat{\lambda}(t)) > 0 \qquad t \in (\widehat{t}_N, T].$$

The following assumption assures that, at each switching time, the reference extremal crosses transversely the zero-level set of the Hamiltonian F_1. As this property is stable for perturbations of the crossing point, it will be crucial for the construction of the maximized flow.

Assumption 4 (Regularity of the switching points).

$$\frac{\mathrm{d}}{\mathrm{d}t}(H_{i+1} - H_i)(\widehat{\lambda}(t))|_{t=\widehat{t}_i} > 0, \quad i = 3, \dots, N.$$

Equivalently, $\boldsymbol{\sigma}_{\widehat{\ell}_i}\left(\vec{H}_i, \vec{H}_{i+1}\right) = (\widehat{u}_{i+1} - \widehat{u}_i) F_{01}(\widehat{\ell}_i) > 0$ *for* $i = 3, \dots, N$.

Finally, recalling that $\widehat{\eta}$ is non-positive and supported on $[\widehat{t}_1, \widehat{t}_2]$, we impose the following regularity assumption on $\widehat{\eta}$.

Assumption 5 (regularity of $\widehat{\eta}$). *We assume that* $\sup_{t \in (\widehat{t}_1, \widehat{t}_2)} \widehat{\eta}(t) < 0.$

We remark that, under the last assumption, $\widehat{\eta}$ is discontinuous at the junction times $\widehat{t}_1, \widehat{t}_2$. Moreover, as a consequence of Assumptions 2-5 and equation (2.7) we get the following

Proposition 2.4. *There exists a neighborhood* $\mathcal{V}$ *of* $\widehat{\lambda}([\widehat{t}_1, \widehat{t}_2])$ *in* T^*M *such that*

$$u_1 F_{01}(\ell) < 0, \quad u_3 F_{01}(\ell) > 0 \qquad \forall \ell \in \mathcal{V}.$$

3. Over-maximized Hamiltonian

For classical regular cases (e.g., when the maximized Hamiltonian is C^2), Hamiltonian methods provide sufficient optimality conditions by directly estimating the difference between the cost associated with the reference trajectory and the one associated with any other admissible trajectory, by means of the integral of a one-form depending on the maximized Hamiltonian (see [3, Chapter 17]).

With some wariness, this approach can be extended to the regular bang-bang case (see [5, 25]). Indeed, since the internal arcs are regular bang-bang, the maximized Hamiltonian is piecewise smooth in a neighborhood of the range of the restriction $\widehat{\lambda}|_{[0, \widehat{t}_1 - \epsilon) \cup (\widehat{t}_2 + \epsilon, T]}$, in the following sense: fix $\epsilon > 0$, and consider a neighborhood $\mathcal{O}_1$ of $\widehat{\lambda}([0, \widehat{t}_1 - \epsilon))$ and a neighborhood $\mathcal{O}_2$ of $\widehat{\lambda}((\widehat{t}_2 + \epsilon, T])$; as $H_{\max} = F_0 + u_1 F_1$ along the first bang arc, we can always choose $\mathcal{O}_1$ such that the two Hamiltonians coincide on $\mathcal{O}_1$, and conclude that $H_{\max}|_{\mathcal{O}_1}$ is smooth. Analogously, it is always possible to choose $\mathcal{O}_2$ such that $H_{\max}$ is smooth on the set $\{\ell \in \mathcal{O}_2 : F_1(\ell) \neq 0\}$, i.e., $H_{\max}$ is continuous in $\mathcal{O}_2$ and it is not differentiable only at the points where F_1 vanishes.

By contrast, along boundary arcs, the maximized Hamiltonian coincides with F_0, but it is not associated with a unique value of the control, as all Hamiltonians of the form $F_0 + vF_1$, $v \in [-1, 1]$, coincide and realize the maximum of $h(\ell, \eta, u)$ over the admissible values of the control. Anyway, no selection of the associated multi-valued Hamiltonian vector field is suitable to construct the field of state-extremals used to compare the costs associated with the candidate trajectories (see for instance [26] for an explanation of the phenomenon, in the case of unconstrained singular arcs; this fact is indeed typical of singular arcs, in both constrained and unconstrained control problems). In [30], the author overcame this issue by substituting the maximized Hamiltonian with a smooth *overmaximized* one, (locally) satisfying some crucial properties: its flow must be tangent to the zero-level set of F_1 and the reference extremal must be an integral curve of its. Due to analogies between boundary and singular arcs, we get inspiration from the methods used for the latter ones to treat our case.

First of all, let us define the set

$$\Sigma = \{\ell \in T^*M \colon F_1(\ell) = 0\}.$$

By Assumption 2 and Remark 2.1, the set $\mathcal{U} := \Sigma \cap \pi^{-1}(U_{\mathcal{C}})$ is a codimension 1 embedded submanifold of T^*M.

We are now going to construct a new Hamiltonian K_2 whose flow has nice invariant and over-maximality properties. Although the construction may look similar to the one in [30], here the result is quite different, as we will point out in Remark 3.4.

Lemma 3.1. *Possibly shrinking $U_{\mathcal{C}}$, there exists a smooth function $\tau \colon U_{\mathcal{C}} \to \mathbb{R}$ such that $\tau(x) = 0$ if and only if $c(x) = 0$, and*

$$c(\exp(\tau(x)u_1 f_1)(x)) = 0 \quad \forall x \in U_{\mathcal{C}}.$$

Moreover,

$$c(x) < 0 \iff \tau(x) > 0,$$

and

$$\mathrm{d}\tau(x) = -\frac{\mathrm{d}c(y)\exp(\tau(x)u_1 f_1)_*}{u_1 L_{f_1} c(y)}, \qquad \textit{where} \;\; y := \exp(\tau(x)u_1 f_1)(x). \quad (3.1)$$

Proof. Notice that

$$\frac{\partial}{\partial s} c\left(\exp(s u_1 f_1)(x)\right)\Big|_{s=0} = u_1 L_{f_1} c(x) > 0 \qquad \forall x \in U_{\mathcal{C}},$$

by Assumption 2. The thesis follows from the implicit function theorem. □

Define the Hamiltonian functions H_0, $K_2 \colon \mathcal{U} \to \mathbb{R}$ as

$$H_0(\ell) := F_0 \circ \exp(\tau(\pi\ell)u_1\vec{F}_1)(\ell) \qquad K_2(\ell) := H_0(\ell) + v_s(\pi\ell)F_1(\ell).$$

Proposition 3.2. *Possibly shrinking $\mathcal{U}$, the following properties hold:*

1. $H_0(\ell) \geq F_0(\ell)$ *for every* $\ell \in \mathcal{U}$ *such that* $c(\pi\ell) \leq 0$. $H_0(\ell) = F_0(\ell)$ *if and only if* $c(\pi\ell) = 0$.
2. *Let* $\vec{c}$ *denote the Hamiltonian vector field associated with* $c \circ \pi$. *Then*
$$\vec{H}_0(\ell) = \exp(-su_1\vec{F}_1)_*\Big(\vec{F}_0 + \frac{F_{01}}{L_{f_1}c \circ \pi}\vec{c}\Big) \circ \exp(su_1\vec{F}_1)|_{s=\tau(\pi\ell)}(\ell).$$
3. $\mathcal{U}$ *is invariant under the Hamiltonian flow of* H_0 *and* K_2.
4. $\Sigma_0 := \{\ell \in \mathcal{U} \colon c(\pi\ell) = 0\} = \{\ell \in T^*M \colon F_1(\ell) = 0,\ c(\pi\ell) = 0,\ \pi\ell \in U_C\}$ *is a submanifold of* Σ *and it is invariant under the flow associated with the Hamiltonian* K_2.
5. $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\lambda}(t) = \vec{K}_2(\widehat{\lambda}(t))$ *for every* $t \in [\widehat{t}_1, \widehat{t}_2]$.

Proof. Claim 1. By definition, $H_0(\ell) = F_0(\ell)$ whenever $c(\pi\ell) = 0$.
Pick some $\ell \in \mathcal{U}$ such that $c(\pi\ell) = 0$. Then

$$\langle \mathrm{d}(H_0 - F_0)(\ell)\,,\,\delta\ell\rangle = -u_1F_{01}(\ell)\langle \mathrm{d}\tau(\pi\ell)\,,\,\pi_*\delta\ell\rangle = u_1F_{01}(\ell)\frac{\langle \mathrm{d}c(\pi\ell)\,,\,\pi_*\delta\ell\rangle}{u_1L_{f_1}c(\pi\ell)}.$$

Thanks to the fact that $u_1L_{f_1}c(\widehat{\xi}(t)) > 0$ for $t \in [\widehat{t}_1, \widehat{t}_2]$, to Proposition 2.4 and to Lemma 3.1, we get the claim.

Claim 2. Let $\ell \in \mathcal{U}$ and set $\ell' := \exp(\tau(\pi\ell)u_1\vec{F}_1)(\ell)$. Then, for every $X \in T_\ell\Sigma$, we get

$$\begin{aligned}
&\boldsymbol{\sigma}_\ell(X, \vec{H}_0) = \langle \mathrm{d}H_0(\ell)\,,\,X\rangle\\
&\quad = \langle \mathrm{d}F_0(\ell')\exp(su_1\vec{F}_1)_*\,,\,X\rangle + \langle \mathrm{d}F_0(\ell')\,,\,u_1\vec{F}_1(\ell')\rangle\langle \mathrm{d}\tau(\pi\ell)\pi_*\,,\,X\rangle\\
&\quad = \boldsymbol{\sigma}_{\ell'}\big(\exp(su_1\vec{F}_1)_*X, \vec{F}_0\big) + u_1F_{01}(\ell')\frac{\boldsymbol{\sigma}_{\ell'}\big(\exp(su_1\vec{F}_1)_*X, \vec{c}\big)}{u_1L_{f_1}c(\pi\ell')} =\\
&\quad = \boldsymbol{\sigma}_\ell\Big(X, \exp(-su_1\vec{F}_1)_*\Big(\vec{F}_0 + \frac{F_{01}\,\vec{c}}{L_{f_1}c \circ \pi}\Big) \circ \exp(su_1\vec{F}_1)\Big), \qquad s = \tau(\pi\ell),
\end{aligned}$$

where here above we used equation (3.1).

Claim 3. We prove the invariance of Σ with respect to the flow of $\vec{H}_0$. The second part of the claim then follows trivially. Let $\ell \in \Sigma$ set, as above, $\ell' := \exp(\tau(\pi\ell)u_1\vec{F}_1)(\ell)$. We get

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s}F_1 \circ \exp(s\vec{H}_0)(\ell)\Big|_{s=0} = \boldsymbol{\sigma}_\ell\big(\vec{H}_0, \vec{F}_1\big)\\
&= \boldsymbol{\sigma}_\ell\Big(\Big(\exp(-\tau(\pi\ell)u_1\vec{F}_1)_*\Big(\vec{F}_0 + \frac{F_{01}}{L_{f_1}(c \circ \pi)}\vec{c}\Big), \vec{F}_1\Big)\\
&= \boldsymbol{\sigma}_{\ell'}\Big(\Big(\vec{F}_0 + \frac{F_{01}}{L_{f_1}(c \circ \pi)}\vec{c}\Big), \vec{F}_1\Big) = F_{01}(\ell') + \frac{F_{01}(\ell')}{L_{f_1}c(\pi\ell')}(-L_{f_1}c(\pi\ell')) = 0.
\end{aligned}$$

Claim 4. Let $\ell \in \Sigma_0$. Then $\vec{F}_1(\ell)$ is tangent to Σ but, by Assumption 2, it is not tangent to $\{\ell \in \mathcal{U} \colon c(\pi\ell) = 0\}$. This proves that Σ_0 is a codimension-one submanifold of Σ.

Moreover, for any $\ell \in \Sigma_0$ we have

$$\frac{\mathrm{d}}{\mathrm{d}s} c(\pi \circ \exp(s\vec{K}_2)(\ell))|_{s=0} = \langle \mathrm{d}c(\pi\ell)\,,\, \pi_*\Big(\vec{F}_0 + \frac{F_{01}}{L_{f_1}(c\circ\pi)}\vec{c} + (\upsilon_s \circ \pi)\vec{F}_1\Big)(\ell)\rangle$$
$$= \langle \mathrm{d}c(\pi\ell)\,,\, f_0(\pi\ell) + \upsilon_s(\pi\ell) f_1(\pi\ell)\rangle = 0.$$

Claim 5. We finally notice that, thanks to (2.4) and (2.7) and the fact that $\widehat{\lambda}(t) \in \Sigma_0$, for every $t \in [\widehat{t}_1, \widehat{t}_2]$ we have that

$$\vec{K}_2(\widehat{\lambda}(t)) = \vec{F}_0(\widehat{\lambda}(t)) + \frac{F_{01}(\widehat{\lambda}(t))}{L_{f_1}c(\widehat{\xi}(t))}\vec{c}(\widehat{\lambda}(t)) + \upsilon_s(\widehat{\xi}(t))\vec{F}_1(\widehat{\lambda}(t))$$
$$= \vec{F}_0(\widehat{\lambda}(t)) + \widehat{\eta}(t)(-\mathrm{d}c(\widehat{\xi}(t)), 0) + \widehat{u}_s(t)\vec{F}_1(\widehat{\lambda}(t)) = \vec{h}(\widehat{\lambda}(t), \widehat{\eta}(t), \widehat{u}(t)).$$

□

Remark 3.3. *Notice that the projections*

$$\pi_*\vec{H}_0(\ell) = \exp(-\tau(\pi\ell)u_1 f_1)_* f_0 \circ \exp(\tau(\pi\ell)u_1 f_1)(\pi\ell),$$
$$\pi_*\vec{K}_2(\ell) = \exp(-\tau(\pi\ell)u_1 f_1)_* f_0 \circ \exp(\tau(\pi\ell)u_1 f_1)(\pi\ell) + \upsilon_s(\pi\ell) f_1(\pi\ell)$$

are well defined vector fields on M. Thus, in the following, we set

$$h_0(x) := \exp(-\tau(x)u_1 f_1)_* f_0 \circ \exp(\tau(x)u_1 f_1)(x),$$
$$k_2(x) := h_0(x) + \upsilon_s(x) f_1(x),$$

so that $H_0(\ell) = \langle \ell\,,\, h_0(\pi\ell)\rangle$ and $K_2(\ell) = \langle \ell\,,\, k_2(\pi\ell)\rangle$.

Remark 3.4 (On the difference between constrained and unconstrained singular arcs). *We already remarked that, if Assumption* 1 *holds, then the boundary arc resembles a singular arc of an unconstrained problem: indeed, Assumption* 1 *together with equation* (2.5d) *forces the Hamiltonian F_1 to vanish along the boundary arc. However, the two kinds of arc have an intrinsically different nature: on the one hand, for boundary arcs of the first order, the control function can be recovered as a true feedback control, that is, it is a function of $\widehat{\xi}(t)$ only; on the other hand, for unconstrained singular extremals, the control may be recovered as a feedback of the whole extremal $\widehat{\lambda}(t)$ (under some regularity conditions).*

Another notable difference concerns the value of the Hamiltonian F_{01} along the arc. In fact, PMP imposes F_{01} to be null along any unconstrained singular arc; this fact is crucial in order to define an overmaximized Hamiltonian in a neighbourhood of the singular arc (see [30]*). For boundary arcs of the first order, equation* (2.7) *just imposes that F_{01} must not change sign; if, in addition, Assumption* 5 *is satisfied, then F_{01} cannot vanish along the arc.*

This fact, in particular, implies that there is no overmaximized Hamiltonian in a full neighbourhood of the arc. Assume indeed, by contradiction, that there exists some Hamiltonian function $H_+ \geq F_0$, defined in neighborhood of the range of the boundary arc, such that

- *$H_+ = F_0$ on $\Sigma \cap \pi^{-1}(\mathcal{C})$,*
- *the Hamiltonian flow associated with H_+ preserves Σ, $\boldsymbol{\sigma}(\vec{F}_1, \vec{H}_+) \equiv 0$.*

Let $\ell \in \Sigma \cap \pi^{-1}(\mathcal{C})$ and consider the curve $\gamma(s) = \exp(u_1 s\vec{F}_1)(\ell)$, $s \in (-\epsilon, \epsilon)$. By Assumption 2, the projection of γ on M is transversal to $\mathcal{C}$. Moreover,

$$\frac{d}{ds}(H_+ - F_0)(\gamma(s))|_{s=0} = \boldsymbol{\sigma}_\ell\big(u_1\vec{F}_1, \vec{H}_+ - \vec{F}_0\big) = u_1 F_{01}(\ell) < 0,$$

which contradicts the fact that $H_+ \geq F_0$ in a neighborhood of ℓ.

Anyway, this fact does not constitute an issue for our construction: indeed, as all admissible trajectories do satisfy the state constraint, we just need the overmaximized Hamiltonian to be defined on a "half tubular neighbourhood" of the boundary arc.

To construct a well-defined over-maximized Hamiltonian flow, we must wisely concatenate the flow of $H_{\max}$ (around the bang arcs) with the one of K_2 (in a "half neighborhood" of the boundary arc). The following lemma states that the junction can be done with no particular issues.

Lemma 3.5. *Let $\epsilon > 0$. Possibly shrinking $U_{\mathcal{C}}$, for every $\ell \in \mathcal{U}$ the following inequalities hold:*

$$u_1 F_1 \circ \exp((t - \widehat{t}_1)\vec{H}_1)(\ell) > 0 \qquad \forall t \in (-\epsilon, \widehat{t}_1),$$
$$u_3 F_1 \circ \exp((t - \widehat{t}_2)\vec{H}_3)(\ell) > 0 \qquad \forall t \in (\widehat{t}_2, \widehat{t}_3 - \epsilon).$$

Proof. A Taylor expansion gives

$$u_1 F_1 \circ \exp((t - \widehat{t}_1)\vec{H}_1)(\ell) = u_1 F_1(\ell) + (t - \widehat{t}_1) u_1 F_{01}(\ell) + O((t - \widehat{t}_1)^2).$$

Recalling that $\ell \in \Sigma$, Proposition 2.4 yields the claim in a left-hand neighborhood $(\widehat{t}_1 - \delta, \widehat{t}_1)$ of $\widehat{t}_1$. For $t \in (-\epsilon, \widehat{t}_1 - \delta)$, the claim holds true by continuity (possibly shrinking the neighborhood $\mathcal{U}$).
The proof on the other bang interval is analogous. □

We still have to define the overmaximised flow for $t \geq \widehat{t}_3$. As already pointed out, the maximized Hamiltonian is well defined, continuous, and piecewise smooth in a neighborhood $\mathcal{O}_2$ of $\{\widehat{\lambda}(t) : t \in [\widehat{t}_2 + \epsilon, T]\}$, but its associated Hamiltonian vector field is not well defined on the switching surfaces (that is, on the set $\{\ell : F_1(\ell) = 0\}$). Nevertheless, here below we are proving that the flow associated with $H_{\max}$ is well defined, at least in $\mathcal{O}_2$ (eventually shrunk). Indeed, we can define, locally around $\widehat{\ell}_2$, some smooth functions $\boldsymbol{t_k}$, $k = 3, \dots, N$, that describe the times when the maximized flow hits a switching surface.

Lemma 3.6. *There exist a neighborhood $\mathcal{U}_2$ of $\widehat{\ell}_2$, $\mathcal{U}_2 \subset \mathcal{U}$, and a smooth function $\boldsymbol{t_3} \colon \mathcal{U}_2 \to \mathbb{R}$ such that*

$$\big(H_4 - H_3\big) \circ \exp((\boldsymbol{t_3}(\ell) - \widehat{t}_2)\vec{H}_3)(\ell) = 0, \quad \boldsymbol{t_3}(\widehat{\ell}_2) = \widehat{t}_3.$$

Moreover, for every $\delta\ell \in T_{\widehat{\ell}_2}(T^*M)$ it holds

$$\langle \mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\,\delta\ell\rangle = -\frac{\boldsymbol{\sigma}\big(\exp((\widehat{t}_3 - \widehat{t}_2)\vec{H}_3)_*\delta\ell, \vec{H}_4 - \vec{H}_3\big)(\widehat{\ell}_3)}{\boldsymbol{\sigma}\big(\vec{H}_3, \vec{H}_4\big)(\widehat{\ell}_3)}.$$

The proof is a straightforward application of the implicit function theorem, and relies on Assumption 4.

Analogously, we can define $\boldsymbol{t_k}$ for $k \geq 4$. First of all, we set

$$\phi_3(\ell) = \exp((\boldsymbol{t_3}(\ell) - \widehat{t}_2)\vec{H}_3)(\ell),$$

and, for $k = 4, \ldots, N$ and $\ell \in \mathcal{U}_2$ we recursively define

$$\begin{aligned} \psi_k(t,\ell) &= (H_{k+1} - H_k) \circ \exp((t - \boldsymbol{t_{k-1}}(\ell))\vec{H}_k) \circ \phi_{k-1}(\ell) \\ \boldsymbol{t_k}(\ell) &= \min_{t > \boldsymbol{t_{k-1}}(\ell)} \{t \colon \psi_k(t,\ell) = 0\} \\ \phi_k(\ell) &= \exp((\boldsymbol{t_k}(\ell) - \boldsymbol{t_{k-1}}(\ell))\vec{H}_k) \circ \phi_{k-1}(\ell). \end{aligned}$$

The existence and smoothness of the functions $\boldsymbol{t_k}$ are granted by the implicit function theorem and Assumptions 3-4. Finally, we set $\boldsymbol{t_{N+1}} \equiv T$.

Given $\epsilon > 0$ small enough, we define the flow

$$\mathcal{H} \colon (t,\ell) \in (-\epsilon, T+\epsilon) \times \mathcal{U}_2 \mapsto \mathcal{H}_t(\ell) \in T^*M$$

as

$$\mathcal{H}_t(\ell) = \begin{cases} \exp(t - \widehat{t}_1)\vec{H}_1 \circ \exp(\widehat{t}_1 - \widehat{t}_2)\vec{K}_2(\ell) & \text{if } t \in (-\epsilon, \widehat{t}_1], \\ \exp(t - \widehat{t}_2)\vec{K}_2(\ell) & \text{if } t \in [\widehat{t}_1, \widehat{t}_2], \\ \exp(t - \widehat{t}_2)\vec{H}_3(\ell) & \text{if } t \in (\widehat{t}_2, \boldsymbol{t_3}(\ell)], \\ \exp(t - \boldsymbol{t_{k-1}})\vec{H}_k \circ \phi_{k-1}(\ell) & \begin{array}{l}\text{if } t \in (\boldsymbol{t_{k-1}}(\ell), \boldsymbol{t_k}(\ell)], \\ \quad k = 4, \ldots, N+1. \end{array} \end{cases} \tag{3.2}$$

$\mathcal{H}_t(\ell)$ is the Hamiltonian flow associated with the piecewise time-dependent Hamiltonian

$$H(t,\ell) = \begin{cases} H_{\max}(\ell) & \text{if } t \in (-\epsilon, \widehat{t}_1), \\ K_2(\ell) & \text{if } t \in [\widehat{t}_1, \widehat{t}_2], \\ H_{\max}(\ell) & \text{if } t \in (\widehat{t}_2, T], \end{cases}$$

defined in a neighborhood of the graph of the reference extremal. To lighten the notation, sometimes we write $H_t(\cdot) = H(t,\cdot)$.

We finally introduce the function $\boldsymbol{u} : [\widehat{t}_2, T] \times \mathcal{U}_2 \to \{-1, 1\}$, assigning to (t,ℓ) the value of the maximizing control, as follows:

$$\boldsymbol{u}(t,\ell) = \widehat{u}_j \qquad \text{if } t \in [\boldsymbol{t_{j-1}}(\ell), \boldsymbol{t_j}(\ell)), \qquad j = 3, \ldots, N+1.$$

Lemma 3.7. *Let $\epsilon > 0$. For every $\ell \in \mathcal{U}_2$ the following inequalities hold:*

$$\boldsymbol{u}(t,\ell) F_1 \circ \mathcal{H}_t(\ell) > 0 \qquad \forall t \in (\boldsymbol{t_{j-1}}(\ell), \boldsymbol{t_j}(\ell)), \qquad j = 3, \ldots, N+1.$$

The proof is a straightforward consequence of Assumptions 3-4 and the definition of the times $\boldsymbol{t_j}$'s.

4. The second variation

We start by computing the second variation for the problem (2.1)-(**T**). First of all, we recast it in the following Mayer problem in $\mathbb{R} \times M$ on the fixed time interval $[0, \widehat{T}]$:

$$\begin{cases} \dot{\xi}^0(t) = u_0(t), \\ \dot{\xi}(t) = u_0(t)(f_0 + u(t)f_1) \circ \xi(t), \\ \xi^0(0) = 0, \quad \xi(0) = \widehat{x}_0, \\ \xi^0(\widehat{T}) \in \mathbb{R}, \quad \xi(\widehat{T}) \in \mathcal{N}_f, \\ u_0(t) > 0, \quad |u(t)| \leq 1 \quad \text{a.e. } t \in [0, \widehat{T}]. \end{cases} \tag{4.1}$$

We set $\boldsymbol{\xi}(t) = \begin{pmatrix} \xi^0(t) \\ \xi(t) \end{pmatrix}$, and we notice that $\widehat{\boldsymbol{\xi}}(t) = \begin{pmatrix} t \\ \widehat{\xi}(t) \end{pmatrix}$ satisfies the PMP with adjoint covector $\widehat{\boldsymbol{\mu}}(t) = (-\widehat{p}_0, \widehat{\mu}(t))$.

Following [4], we associate with the control system (4.1) the penalized cost

$$J(\boldsymbol{\xi}) = \alpha(\xi(0)) - p_0\xi^0(0) + \beta(\xi(\widehat{T})) + p_0\xi^0(\widehat{T}) - \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s)c(\xi(s))\mathrm{d}s,$$

where α is a smooth function defined in a neighborhood of $\widehat{x}_0$, such that $\mathrm{d}\alpha(\widehat{x}_0) = \widehat{\ell}_0$, and β is a smooth function defined in a neighborhood of $\widehat{x}_T$, such that $\mathrm{d}\beta(\widehat{x}_T) = -\widehat{\ell}_T$.

Let us stress that the "true" intrinsic extension of the cost, as defined in [4] and typically used to compute the intrinsic second variation, would be just $\alpha(\xi(0)) - p_0\xi^0(0) + \beta(\xi(T)) + p_0\xi^0(T)$. Though, from what follows, it will be clear that the first variation of such functional is not zero, so that, its second variation is not well-defined.

We consider a sub-problem of (4.1) by allowing only control functions of the following form:

$$u_0(t) \equiv 1, \quad t \in [0, \widehat{t}_2), \quad \text{and} \quad \int_{\widehat{t}_2}^{\widehat{T}} u_0(s)\mathrm{d}s = \widehat{T} - \widehat{t}_2, \tag{4.2}$$

$$u(t) \equiv \widehat{u}(t), \quad t \notin [\widehat{t}_1, \widehat{t}_2]. \tag{4.3}$$

This is equivalent to considering variations of the control $u(t)$ along the boundary arc only, and variations of the switching times t_k, $k = 3, \ldots N$, without allowing for variations of the final cost $\xi^0(\widehat{T})$ of the reference triple $(\widehat{T}, \widehat{\xi}, \widehat{u})$. Let $\boldsymbol{\xi}(t)$ be any solution of (4.1)-(4.2)-(4.3) and set

$$\boldsymbol{\varkappa}_t := \begin{pmatrix} \varkappa_t^0 \\ \varkappa_t \end{pmatrix} = \widehat{\boldsymbol{S}}_{\boldsymbol{t}}^{-1}(\boldsymbol{\xi}(t)).$$

For the sake of conciseness, we introduce the following pull-backs (to the exit time $\widehat{t}_2$):

$$\widehat{\mathcal{N}_f} := \widehat{S}_T^{-1}(\mathcal{N}_f), \quad k_t := \widehat{S}_{t*}^{-1} f_1 \circ \widehat{S}_t, \quad g_j := \widehat{S}_{\widehat{t}_j *}^{-1} h_j \circ \widehat{S}_{\widehat{t}_j}, \ j = 3, \ldots, N+1.$$

We point out that, for $j = 3, \dots, N+1$, $g_j = \widehat{S}_{t*}^{-1} h_j \circ \widehat{S}_t$ for every $t \in [\widehat{t}_{j-1}, \widehat{t}_j]$.

It is easy to check that $\boldsymbol{\varkappa}$ satisfies the following system

$$\begin{cases} \dot{\varkappa}_t^0 = u_0(t) - 1, \\ \dot{\varkappa}_t = (u_0(t) - 1)\widehat{S}_{t*}^{-1} \widehat{f}_t \circ \widehat{S}_t(\varkappa_t) + u_0(t)(u(t) - \widehat{u}(t))k_t(\varkappa_t), \\ \varkappa_0^0 = \widehat{t}_2, \quad \varkappa_0 = \widehat{x}_2, \\ \varkappa_{\widehat{T}}^0 \in \mathbb{R}, \quad \varkappa_{\widehat{T}} \in \widehat{\mathcal{N}_f}. \end{cases} \tag{4.4}$$

We also introduce the pullbacks of the extended costs α and β, i.e. we define $\widehat{\alpha} = \alpha \circ \widehat{S}_0$ and $\widehat{\beta} = \beta \circ \widehat{S}_{\widehat{T}}$. Then, the cost functional J can be written in terms of the pull-back trajectory $\boldsymbol{\varkappa}$ as

$$\begin{aligned} J(\boldsymbol{\varkappa}) = p_0\Big(\varkappa_{\widehat{T}}^0 + \int_{\widehat{t}_2}^{\widehat{T}} u_0(s)\mathrm{d}s\Big) - p_0\Big(\varkappa_0^0 + \int_0^{\widehat{t}_2} u_0(s)\mathrm{d}s\Big) \\ + \widehat{\alpha}(\varkappa_0) + \widehat{\beta}(\varkappa_{\widehat{T}}) - \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s)\widehat{c}_s(\varkappa_s)\mathrm{d}s \\ = p_0\big(\varkappa_{\widehat{T}}^0 - \varkappa_0^0 + \widehat{T} - 2\,\widehat{t}_2\big) + \widehat{\alpha}(\varkappa_0) + \widehat{\beta}(\varkappa_{\widehat{T}}) - \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s)\widehat{c}_s(\varkappa_s)\mathrm{d}s. \end{aligned} \tag{4.5}$$

We remark that, as $\boldsymbol{\varkappa}$ is completely determined by its initial condition and the controls (u_0, u), then J is actually a function of $\varkappa_0^0, \varkappa_0, u_0, u$. We point out that, although in (4.4) $\boldsymbol{\varkappa}_0$ is fixed, for the sake of future computations, we treat $\varkappa_0$ as a free point in M. On the other hand, it would be pointless to consider the whole $\boldsymbol{\varkappa}_0$ as a free point in $\mathbb{R} \times M$: indeed, as $\varkappa^0$ appears only linearly in J, we could not obtain a coercive second variation with respect to $\big(\varkappa^0, \varkappa, u_0, u\big)$. Hence, we fix $\varkappa_0^0 = \widehat{t}_2$ and allow only for variations of $\varkappa_0$ and for variations of the controls satisfying (4.2)-(4.3). In Appendix A we show how PMP grants that the first variation $\mathrm{d}J$ at $(\widehat{x}_2, u_0 \equiv 1, \widehat{u})$ is null.

Let us now compute the second variation of the cost. First of all we simplify the notations, by introducing

$$\mathcal{A}(t, x) := \widehat{\alpha}(x) - \int_{\widehat{t}_1}^{t} \widehat{\eta}(s)\widehat{c}_s(x)\mathrm{d}s.$$

We notice that, thanks to PMP, $L_{f_1}\mathcal{A}(\widehat{t}_2, \widehat{x}_2) = 0$. As explained in [4], the second variation does not depend on the particular choice of α, provided that $\mathrm{d}\alpha(\widehat{x}_0) = \widehat{\ell}_0$; therefore, we can choose α in such a way that $L_{f_1}\mathcal{A}(\widehat{t}_2, x) = 0$ for every x in a neighborhood of $\widehat{x}_2$. Introducing the control variations

$$\delta u(t) = u(t) - \widehat{u}(t),\ t \in [\widehat{t}_1, \widehat{t}_2], \qquad \delta u_0(t) = u_0(t) - 1,\ t \in [\widehat{t}_2, \widehat{T}],$$

and setting

$$\delta_k = \int_{\widehat{t}_{k-1}}^{\widehat{t}_k} \delta u_0(s)\mathrm{d}s, \quad k = 3, \dots, N+1, \tag{4.6}$$

the linearization of system (4.4) for $\varkappa_t$ reads

$$\delta\dot{\varkappa}_t = \begin{cases} 0 & t \in [0, \widehat{t}_1), \\ \delta u(t) k_t(\widehat{x}_2) & t \in [\widehat{t}_1, \widehat{t}_2), \\ \delta u_0(t) g_k(\widehat{x}_2) & t \in [\widehat{t}_{k-1}, \widehat{t}_k),\ k = 3, \ldots, N+1, \end{cases} \tag{4.7}$$

together with the boundary conditions

$$\delta\varkappa_{\widehat{t}_1} = 0, \qquad \delta\varkappa_{\widehat{T}} = \delta y \in T_{\widehat{x}_2}\widehat{\mathcal{N}}_f, \tag{4.8}$$

and the control constraint

$$\sum_{k=3}^{N+1} \delta_k = \int_{\widehat{t}_2}^{\widehat{T}} \delta u_0(t) \mathrm{d}t = 0, \tag{4.9}$$

so that

$$\delta\varkappa_{\widehat{T}} = \delta\varkappa_{\widehat{t}_2} + \sum_{k=3}^{N+1} \delta_k g_k(\widehat{x}_2).$$

The system (4.7)-(4.8) and the control constraint (4.9) define the space of admissible variations

$$W = \Big\{ (\delta y, \delta u, \boldsymbol{\delta}) \in T_{\widehat{x}_2}\widehat{\mathcal{N}}_f \times L^2([\widehat{t}_1, \widehat{t}_2]) \times \mathbb{R}^{N-2} \colon \sum_{k=3}^{N+1} \delta_k = 0, \\ \text{and system (4.7)} - \text{(4.8) is satisfied} \Big\}.$$

After some computations (see Appendix A), we are able to provide the expression of the second variation associated with the cost J along $\widehat{\boldsymbol{\lambda}}$:

$$\begin{aligned} J''[\delta y, \delta u, \boldsymbol{\delta}]^2 = {} & \frac{1}{2} \mathrm{D}^2 \Big(\mathcal{A}(\widehat{t}_2, \cdot) + \widehat{\beta} \Big)(\widehat{x}_2)[\delta\varkappa_{\widehat{t}_2}]^2 \\ & + L_{\delta\varkappa_{\widehat{t}_2}} L_{\sum_{i=3}^{N+1} \delta_i g_i} \widehat{\beta}(\widehat{x}_2) + \frac{1}{2} L^2_{\sum_{i=3}^{N+1} \delta_i g_i} \widehat{\beta}(\widehat{x}_2) \\ & + \frac{1}{2} \sum_{3 \le i < j \le N+1} \delta_i \delta_j L_{[g_i, g_j]} \widehat{\beta}(\widehat{x}_2) \\ & - \int_{\widehat{t}_1}^{\widehat{t}_2} \delta u(r) L_{\delta\varkappa_{\widehat{t}_2}} L_{k_r} \mathcal{A}(r, \widehat{x}_2) \mathrm{d}r \\ & + \int_{\widehat{t}_1}^{\widehat{t}_2} \Big(\int_{\widehat{t}_1}^{r} \delta u(r) \delta u(s) L_{k_r} L_{k_s} \mathcal{A}(s, \widehat{x}_2) \mathrm{d}s \Big) \mathrm{d}r. \end{aligned} \tag{4.10}$$

We apply to J'' a standard procedure called Goh's transformation, that is, we set

$$w(t) = -\int_{\widehat{t}_1}^{t} \delta u(s) \mathrm{d}s, \quad \varepsilon_2 = w(\widehat{t}_2),$$

and we consider the state

$$\zeta(t) = \delta\varkappa_t + w(t) k_t(\widehat{x}_2) - \varepsilon_2 f_1(\widehat{x}_2), \quad t \in [\widehat{t}_1, \widehat{t}_2].$$

It is immediate to check that ζ solves the following boundary value problem

$$\begin{cases}\dot\zeta(t) = w(t)\dot k_t(\widehat{x}_2), \\ \zeta(\widehat{t}_1) = -\varepsilon_2 f_1(\widehat{x}_2), \quad \zeta(\widehat{t}_2) = \delta x,\end{cases} \tag{4.11}$$

where we set $\delta x := \delta \varkappa_{\widehat{t}_2}$. The *extended second variation* thus reads

$$\begin{aligned} J''_{ext}[\delta x, w, \varepsilon_2, \boldsymbol{\delta}]^2 = {} & \frac{1}{2}\mathrm{D}^2\big(\mathcal{A}(\widehat{t}_2,\cdot) + \widehat{\beta}\big)(\widehat{x}_2)[\delta x]^2 \\ & + L_{\delta x} L_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2) + \frac{1}{2}L^2_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2) \\ & + \frac{1}{2}\sum_{3\le i<j\le N+1}\delta_i\delta_j L_{[g_i,g_j]}\widehat{\beta}(\widehat{x}_2) \\ & - \varepsilon_2 L_{f_1}\int_{\widehat{t}_1}^{\widehat{t}_2} w(s)\Big(L_{\dot k_s}\mathcal{A}(s,\widehat{x}_2) + L_{k_s}\dot{\mathcal{A}}(s,\widehat{x}_2)\Big)\mathrm{d}s \\ & - \int_{\widehat{t}_1}^{\widehat{t}_2} w(s)L_{\zeta(s)}\Big(L_{\dot k_s}\mathcal{A}(s,\widehat{x}_2) + L_{k_s}\dot{\mathcal{A}}(s,\widehat{x}_2)\Big)\mathrm{d}s \\ & + \frac{1}{2}\int_{\widehat{t}_1}^{\widehat{t}_2} w(s)^2\Big(L_{[k_s,\dot k_s]}\mathcal{A}(s,\widehat{x}_2) + L^2_{k_s}\dot{\mathcal{A}}(s,\widehat{x}_2)\Big)\mathrm{d}s, \end{aligned} \tag{4.12}$$

and is defined over the set of variations

$$\begin{aligned} \widetilde{W} = \Big\{ & (\delta x, w, \varepsilon_2, \boldsymbol{\delta}) \in T_{\widehat{x}_2}M \times L^2([\widehat{t}_1,\widehat{t}_2]) \times \mathbb{R}^{N-1} \colon \sum_{k=3}^{N+1}\delta_k = 0, \\ & \delta x + \sum_{k=3}^{N+1}\delta_k g_k(\widehat{x}_2) \in T_{\widehat{x}_2}\widehat{\mathcal{N}_f} \ \text{ and the system (4.11) is satisfied}\Big\}. \end{aligned}$$

Remark 4.1. *The "true" extension of W would be the set*

$$\begin{aligned} \Big\{ & (\delta y, w, \varepsilon_2, \boldsymbol{\delta}) \in T_{\widehat{x}_2}\widehat{\mathcal{N}_f} \times L^2([\widehat{t}_1,\widehat{t}_2]) \times \mathbb{R}^{N-1} \colon \sum_{k=3}^{N+1}\delta_k = 0, \\ & \delta x = \delta y - \sum_{k=3}^{N+1}\delta_k g_k(\widehat{x}_2) \ \textit{ and the system (4.11) is satisfied}\Big\}. \end{aligned}$$

As δy does not appear explicitly in (4.10), nor in (4.12), we prefer to choose δx in the pool of variations.

Let us write the extended second variation as

$$\Phi(\delta x, w, \varepsilon_2, \boldsymbol{\delta}) + \frac{1}{2}\int_{\widehat{t}_1}^{\widehat{t}_2} w(s)^2\big(L_{[k_s,\dot k_s]}\mathcal{A}(s,\widehat{x}_2) + L^2_{k_s}\dot{\mathcal{A}}(s,\widehat{x}_2)\big)\mathrm{d}s,$$

and notice that $\Phi(\delta x, w, \varepsilon_2, \boldsymbol{\delta})$ is weakly continuous [2]. Then, [17, Theorem 4.3] states that J''_{ext} is weakly lower-semicontinuous on $\widetilde{W}$ if and only if the coefficient of $w(s)^2$ in the integral is non-negative. A reasonable assumption for the well-posedness of the LQ problem (4.11)-(4.12) would be to ask the expression

$$L_{[k_t, \dot{k}_s]}\mathcal{A}(t, \widehat{x}_2) + L^2_{k_t}\dot{\mathcal{A}}(t, \widehat{x}_2) = F_{101}(\widehat{\lambda}(t)) - \widehat{\eta}(t)L^2_{f_1}c(\widehat{\xi}(t))$$

to be positive for every $t \in [\widehat{t}_1, \widehat{t}_2]$; however, this does not seem justified by PMP and seems to be artificial. Nonetheless, it is common, in the literature on optimality conditions for state-constrained problems (see e.g. [6]), not to consider control variations along the boundary arc. For these reasons, we reduce ourselves to the variations $(\delta x, w, \varepsilon_2, \boldsymbol{\delta})$ in $\widetilde{W}$ such that $w \equiv 0$, i.e. we make the following assumption

Assumption 6. *Let*

$$\mathcal{W}_0 = \Big\{(\delta x, \boldsymbol{\delta}) \in T_{\widehat{x}_2}M \times \mathbb{R}^{N-2} \colon \delta x + \sum_{k=3}^{N+1} \delta_k g_k(\widehat{x}_2) \in T_{\widehat{x}_2}\widehat{\mathcal{N}_f}, \quad \sum_{k=3}^{N+1} \delta_k = 0,\ \ \delta x \in \mathbb{R}f_1(\widehat{x}_2)\Big\}.$$

We assume that the quadratic form

$$\mathscr{J}''[(\delta x, \boldsymbol{\delta})]^2 = \frac{1}{2}\mathrm{D}^2\big(\mathcal{A}(\widehat{t}_2, \cdot) + \widehat{\beta}\big)(\widehat{x}_2)[\delta x]^2 + L_{\delta x}L_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2) + \frac{1}{2}L^2_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2) + \frac{1}{2}\sum_{3 \le i < j \le N+1}\delta_i\delta_j L_{[g_i, g_j]}\widehat{\beta}(\widehat{x}_2)$$

is coercive on $\mathcal{W}_0$.

Remark 4.2. *The procedure of considering control variations along the boundary arc, extend them through the Goh transformation, and finally setting their infinite-dimensional part to zero, could appear unjustified and useless. It is however a standard procedure to compute, at a first attempt, the second variation on a very large set (on which, usually, the quadratic form cannot be coercive), and then to reduce it in a suitable way. The aim is indeed to have enough variations, in order to ensure the invertibility of the flow, but not too many as to prevent coercivity. The particular choice of the space* $\mathcal{W}_0$ *will be discussed in Remark* 5.3.

$\mathcal{W}_0$ is a subspace of the linear space

$$\mathcal{W} = \Big\{(\delta x, \boldsymbol{\delta}) \in T_{\widehat{x}_2}M \times \mathbb{R}^{N-2} \colon \delta x + \sum_{k=3}^{N+1} \delta_k g_k(\widehat{x}_2) \in T_{\widehat{x}_2}\widehat{\mathcal{N}_f},\ \sum_{k=3}^{N+1} \delta_k = 0\Big\}.$$

[2]this is a consequence of the fact that weak convergence of the controls implies strong convergence of the trajectories, see [20].

Indeed, consider the projection $\Pi_{\mathcal{C}}\colon T_{\widehat{x}_2}M \to T_{\widehat{x}_2}\mathcal{C}$ with $\ker \Pi_{\mathcal{C}} = \mathbb{R}f_1(\widehat{x}_2)$: we notice that $\mathcal{W}_0 = \big\{(\delta x, \boldsymbol{\delta})\colon (\delta x, \boldsymbol{\delta}) \in \mathcal{W},\ \Pi_{\mathcal{C}}\delta x = 0\big\}$. We define on $\mathcal{W}$ the following quadratic form

$$\mathcal{K}[(\delta x, \boldsymbol{\delta})]^2 = \|\Pi_{\mathcal{C}}\delta x\|^2.$$

Since $\ker \mathcal{K} = \mathcal{W}_0$, and under Assumption 6, we can apply [17, Theorem 13.2] to $\mathcal{J}''$ and $\mathcal{K}$ on $\mathcal{W}$ and get that there exists $\rho > 0$ large enough such that the quadratic form $\mathcal{J}''_\rho = \mathcal{J}'' + \rho\mathcal{K}$ is coercive on $\mathcal{W}$.

We can always construct a coordinate chart $\{z_1, \ldots, z_n\}$ around $\widehat{x}_2$ such that $f_1 \equiv \frac{\partial}{\partial z_1}$ and $\mathcal{C} = \{z_1 = 0\}$. Defining, in these coordinates, the function $K(z_1, \ldots, z_n) = \frac{1}{2}\sum_{i=2}^n z_i^2$, we obtain that $\mathcal{K}$ is the Hessian of K at $\widehat{x}_2$. Then, setting $\mathcal{A}_\rho(x) := \mathcal{A}(\widehat{t}_2, x) + \rho K(x)$, we can write

$$\begin{aligned}\mathcal{J}''_\rho[(\delta x, \boldsymbol{\delta})]^2 = {}& \frac{1}{2}\mathrm{D}^2(\mathcal{A}_\rho + \widehat{\beta})(\widehat{x}_2)[\delta x]^2 + L_{\delta x} L_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2) \\ & + \frac{1}{2}L^2_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2) + \frac{1}{2}\sum_{3\le i<j\le N+1}\delta_i\delta_j L_{[g_i,g_j]}\widehat{\beta}(\widehat{x}_2).\end{aligned}$$

We have proved the following proposition.

Proposition 4.3. *The quadratic form $\mathcal{J}''_\rho$ is coercive on $\mathcal{W}$.*

We end this section by providing an alternative expression for the quadratic form $\mathcal{J}''_\rho[(\delta x, \boldsymbol{\delta})]^2$. First of all, we recall that

$$\boldsymbol{\sigma}\big(\vec{G}_j, \vec{G}_k\big)(\widehat{\ell}_2) = \langle \widehat{\ell}_2, [g_j, g_k](\widehat{x}_2)\rangle = -L_{[g_j,g_k]}\widehat{\beta}(\widehat{x}_2), \quad i, j = 3, \ldots, N+1.$$

Then, we notice that

$$\boldsymbol{\sigma}\big(\mathrm{d}\mathcal{A}_{\rho*}\delta x, \vec{G}_i\big) = \mathrm{D}^2\big(\mathcal{A}_\rho + \widehat{\beta}\big)(\widehat{x}_2)[\delta x, g_i] - L_{\delta x}L_{g_i}\widehat{\beta}(\widehat{x}_2), \quad i = 3, \ldots, N+1.$$

Finally, we can provide the following equivalent expressions for $\mathcal{J}''_\rho$:

$$\begin{aligned}\mathcal{J}''_\rho[(\delta x, \boldsymbol{\delta})]^2 = {}& \frac{1}{2}\mathrm{D}^2\big(\mathcal{A}_\rho + \widehat{\beta}\big)(\widehat{x}_2)\Big[\delta x, \delta x + \sum_{i=3}^{N+1}\delta_i g_i\Big] \\ & + \frac{1}{2}L_{\delta x + \sum_{i=3}^{N+1}\delta_i g_i} L_{\sum_{j=3}^{N+1}\delta_j g_j}\widehat{\beta}(\widehat{x}_2) \\ & + \frac{1}{2}\sum_{i=3}^{N+1}\boldsymbol{\sigma}\Big(\delta_i\vec{G}_i, \mathrm{d}\mathcal{A}_{\rho*}\delta x + \sum_{3\le j<i}\delta_j\vec{G}_j\Big)(\widehat{\ell}_2) \\ = {}& \frac{1}{2}\boldsymbol{\sigma}\Big(\mathrm{d}\mathcal{A}_{\rho*}\delta x + \sum_{i=3}^{N+1}\delta_i\vec{G}_i, \mathrm{d}(-\widehat{\beta})_*\big(\delta x + \sum_{i=3}^{N+1}\delta_i g_i\big)\Big)(\widehat{\ell}_2) \\ & + \frac{1}{2}\sum_{i=3}^{N+1}\boldsymbol{\sigma}\Big(\delta_i\vec{G}_i, \mathrm{d}\mathcal{A}_{\rho*}\delta x + \sum_{3\le j<i}\delta_j\vec{G}_j\Big)(\widehat{\ell}_2).\end{aligned} \tag{4.13}$$

4.1. The second variation of the cost (M)

The second variation for the problem (2.1)-(**M**) can be computed following the same procedure, with some minor changes. As the problem is already in Mayer form, we do not need to extend the control system; though, in order to take into account the variations of the switching times, we still perform a time reparameterization, by introducing the control u_0. The problem (2.1)-(**M**) thus becomes

$$\begin{cases} \dot{\xi}(t) = u_0(t)(f_0 + u(t)f_1) \circ \xi(t), \\ \xi(0) = \widehat{x}_0, \quad \xi(T) \in N_f, \\ u_0(t) > 0, \quad |u(t)| \leq 1, \quad \text{a.e. } t \in [0, T]. \end{cases} \tag{4.14}$$

We can associate to (4.14) the penalized cost

$$J(\xi) = \alpha(\xi(0)) + \beta(\xi(T)) - \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s) c(\xi(s)) \mathrm{d}s,$$

where α is a smooth function defined in a neighborhood of $\widehat{x}_0$ such that $\mathrm{d}\alpha(\widehat{x}_0) = \widehat{\ell}_0$ and β is a smooth function defined in a neighborhood of $\widehat{x}_T$ such that

$$\beta|_{\mathcal{N}_f} \equiv p_0 \psi \quad \text{and} \quad \mathrm{d}\beta(\widehat{x}_T) = -\widehat{\ell}_T.$$

We allow for control functions satisfying the constraints (4.2)-(4.3) and we repeat the same procedure carried above. The extended second variation and the subspace $\mathcal{W}_0$ have the same expression as in the previous section, and Assumption 6 assumes the same form.

5. Invertibility of the flow

Let $U_{\widehat{x}_2}$ be a sufficiently small neighborhood of $\widehat{x}_2$ in M. We define the Lagrangian submanifold

$$\Lambda_\rho = \left\{ \mathrm{d}\mathcal{A}_\rho(x) \colon x \in U_{\widehat{x}_2} \right\}.$$

By definition, $\widehat{\ell}_2 \in \Lambda_\rho$, Λ_ρ is contained in Σ and $\pi \colon \Lambda_\rho \to U_{\widehat{x}_2}$ is surjective. We have moreover the following fact.

Proposition 5.1. *Let $\epsilon > 0$. Possibly shrinking $U_{\widehat{x}_2}$, there exists a tubular neighborhood $\mathcal{V}$ of the graph of $\widehat{\xi}$ in $(-\epsilon, \widehat{t}_2 + \epsilon) \times M$ such that the map*

$$\mathrm{id} \times \pi\mathcal{H} \colon (t, \ell) \in (-\epsilon, \widehat{t}_2 + \epsilon) \times \Lambda_\rho \mapsto (t, \pi\mathcal{H}_t(\ell)) \in \mathcal{V}$$

is a diffeomorphism.

Proof. It suffices to show that, for every $t \in (-\epsilon, \widehat{t}_2 + \epsilon)$, the linear map $(\pi\mathcal{H}_t)_* \colon T_{\widehat{\ell}_2}\Lambda_\rho \to T_{\widehat{\xi}(t)}M$ has maximal rank. This is indeed true because, for any fixed t, this map is just the linearisation of the flow, at time t, of a vector field in M. More precisely

$$(\pi\mathcal{H}_t)_*\delta\ell = \begin{cases} \exp\left((t-\widehat{t}_1)h_1\right)_* \exp\left(\widehat{t}_1-\widehat{t}_2)k_2\right)_* \pi_*\delta\ell & \text{if } t\in(-\epsilon,\widehat{t}_1), \\ \exp\left((t-\widehat{t}_2)k_2\right)_*\pi_*\delta\ell & \text{if } t\in[\widehat{t}_1,\widehat{t}_2], \\ \exp\left((t-\widehat{t}_2)h_3\right)_*\pi_*\delta\ell & \text{if } t\in(\widehat{t}_2,\widehat{t}_2+\epsilon). \end{cases}$$

The result follows from the uniqueness of the solutions of ODEs. □

We stress that, on $(-\epsilon,\widehat{t}_2+\epsilon)\times\Lambda_\rho$, the invertibility of $\mathrm{id}\times\pi\mathcal{H}$ does not require Assumption 6 to hold, as it relies on the properties of the flow $\mathcal{H}_t$ only. Instead, to prove its invertibility on the whole $(-\epsilon,T+\epsilon)\times\Lambda_\rho$, we need Assumption 6, as the following Proposition shows.

Proposition 5.2. *The map*

$$\mathrm{id}\times\pi\mathcal{H}\colon (t,\ell)\in(-\epsilon,T+\epsilon)\times\Lambda_\rho\mapsto(t,\pi\mathcal{H}_t(\ell))\in\mathcal{V}$$

is a diffeomorphism.

Proof. As a matter of fact, we just need to prove the invertibility of $\mathrm{id}\times\pi\mathcal{H}$ at the points $(t,\ell)=(\boldsymbol{t_k}(\ell),\ell)$; indeed, the map is invertible at $\widehat{\ell}_2$ for every $t<\widehat{t}_3$, since it coincides with the flow of the field h_3. In contrast, at time $t=\widehat{t}_3$ a folding phenomenon could occur, due to fact that the switching time $\boldsymbol{t_3}(\ell)$ depends on the costate too; once the invertibility at $t=\widehat{t}_3$ is proved, then we can conclude that the map is invertible up to $t<\widehat{t}_4$. We then repeat the same argument at every $\widehat{t}_k$, $k\geq 4$.

Let us then start with the third switching time. First of all, we notice that the flow $\pi\mathcal{H}_{\widehat{t}_3}$ has different analytic expressions, depending on the sign of $\boldsymbol{t_3}(\ell)-\widehat{t}_3$:

$$\pi\mathcal{H}_{\widehat{t}_3}(\ell) = \begin{cases} \exp((\widehat{t}_3-\widehat{t}_2)h_3)(\pi\ell) & \text{if } \boldsymbol{t_3}(\ell)\geq\widehat{t}_3, \\ \exp((\widehat{t}_3-\boldsymbol{t_3}(\ell))h_4)\circ\exp((\boldsymbol{t_3}(\ell)-\widehat{t}_2)h_3)(\pi\ell) & \text{if } \boldsymbol{t_3}(\ell)\leq\widehat{t}_3. \end{cases}$$

To verify the invertibility at $t=\widehat{t}_3$, we compute the generalized Clarke's Jacobian of $\pi\mathcal{H}_{\widehat{t}_3}$ at $\widehat{\ell}_2$

$$\partial\big(\pi\mathcal{H}_{\widehat{t}_3}\big)(\widehat{\ell}_2) = \overline{\mathrm{co}}\big\{\exp((\widehat{t}_3-\widehat{t}_2)h_3)_*\pi_*(\cdot), \\ \exp((\widehat{t}_3-\widehat{t}_2)h_3)_*\pi_*(\cdot)+\langle\mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\,(\cdot)\rangle(h_3-h_4)(\widehat{x}_2)\big\}.$$

Thus, in order to apply Clarke's Inverse Function Theorem (see [9]), we only need to prove that there is no $a\in[0,1]$ and no nontrivial $\delta\ell\in T_{\widehat{\ell}_2}\Lambda_\rho$ such that

$$\exp((\widehat{t}_3-\widehat{t}_2)h_3)_*\pi_*\delta\ell + a\langle\mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\,\delta\ell\rangle(h_3-h_4)(\widehat{x}_3)=0,$$

or, equivalently,

$$\pi_*\delta\ell + a\langle\mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\,\delta\ell\rangle(g_3-g_4)(\widehat{x}_2)=0. \tag{5.1}$$

We prove this fact by means of a contradiction argument: assume that we can find $a \in [0,1]$ and a nontrivial $\delta\ell \in T_{\widehat{\ell}_2}\Lambda_\rho$ that verify (5.1). Then it is easy to check that

$$(\delta x, \boldsymbol{\delta}) = (\pi_*\delta\ell, \boldsymbol{\delta}) \ \text{ such that } \ \begin{array}{l} \delta_3 = -\delta_4 = a\langle \mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\, \delta\ell\rangle, \\ \delta_j = 0 \quad \forall j = 5, \dots N+1 \end{array}$$

is a nontrivial admissible variation in $\mathcal{W}$. We evaluate $\mathcal{J}''_\rho$ on it. We obtain

$$\begin{aligned} \mathcal{J}''_\rho[(\delta x, \boldsymbol{\delta})]^2 &= \frac{1}{2}\boldsymbol{\sigma}\big(\delta_3\vec{G}_3, \mathrm{d}\mathcal{A}_{\rho*}\delta x\big) \\ &= \frac{1}{2}a\langle \mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\, \delta\ell\rangle\Big(\boldsymbol{\sigma}\big(\vec{G}_3, \delta\ell\big)(\widehat{\ell}_2) \\ &\qquad - \boldsymbol{\sigma}\big(\vec{G}_4, \delta\ell + a\langle \mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\, \delta\ell\rangle\vec{G}_3\big)(\widehat{\ell}_2)\Big) \\ &= \frac{1}{2}a(1-a)\langle \mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\, \delta\ell\rangle^2\boldsymbol{\sigma}\big(\vec{G}_4, \vec{G}_3\big)(\widehat{\ell}_2), \end{aligned}$$

which is negative, by Assumption 4. We got a contradiction with Proposition 4.3. Then, $\pi\mathcal{H}_t$ is invertible at $t = \widehat{t}_3$.

A similar argument proves the invertibility at the other switching times. First of all, we notice that, for every $\ell \in \mathcal{U}$ and $k = 4, \dots, N+1$, we have

$$\pi\mathcal{H}_{\widehat{t}_k}(\ell) = \begin{cases} \exp((\widehat{t}_k - \widehat{t}_{k-1})h_k) \circ \pi\mathcal{H}_{\widehat{t}_{k-1}}(\ell) & \text{if } \boldsymbol{t_k}(\ell) \geq \widehat{t}_k, \\ \exp((\widehat{t}_k - \boldsymbol{t_k}(\ell))h_{k+1})\circ & \\ \quad \circ \exp((\boldsymbol{t_k}(\ell) - \widehat{t}_{k-1})h_k) \circ \pi\mathcal{H}_{\widehat{t}_{k-1}}(\ell) & \text{if } \boldsymbol{t_k}(\ell) \leq \widehat{t}_k. \end{cases}$$

In order to apply Clarke's inverse function theorem at time $\widehat{t}_k$, we assume, by contradiction, that we can find $a \in [0,1]$ and a non trivial $\delta\ell \in T_{\widehat{x}_2}\Lambda_\rho$ such that

$$\pi_*\delta\ell + \sum_{j=3}^{k-1}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\, \delta\ell\rangle(g_j - g_{j+1})(\widehat{x}_2) + a\langle \mathrm{d}\boldsymbol{t_k}(\widehat{\ell}_2)\,,\, \delta\ell\rangle(g_k - g_{k+1})(\widehat{x}_2) = 0.$$

In this case, setting

$$\begin{aligned} &\delta x = \pi_*\delta\ell, \qquad \delta_3 = \langle \mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\, \delta\ell\rangle, \\ &\delta_j = \langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\, \delta\ell\rangle - \langle \mathrm{d}\boldsymbol{t_{j-1}}(\widehat{\ell}_2)\,,\, \delta\ell\rangle, \qquad j = 4, \dots, k-1, \\ &\delta_k = a\langle \mathrm{d}\boldsymbol{t_k}(\widehat{\ell}_2)\,,\, \delta\ell\rangle - \langle \mathrm{d}\boldsymbol{t_{k-1}}(\widehat{\ell}_2)\,,\, \delta\ell\rangle, \\ &\delta_{k+1} = -a\langle \mathrm{d}\boldsymbol{t_k}(\widehat{\ell}_2)\,,\, \delta\ell\rangle, \\ &\delta_i = 0 \qquad i = k+1, \dots, N+1, \end{aligned}$$

we have that $(\delta x, \boldsymbol{\delta})$ is a non trivial element of $\mathcal{W}$.

First of all, by computations we can prove that

$$\langle \mathrm{d}\boldsymbol{t_k}(\widehat{\ell}_2)\,,\, \delta\ell\rangle = -\frac{\boldsymbol{\sigma}\big(\delta\ell - \sum_{j=3}^{k-1}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\delta\ell\rangle\big(\vec{G}_{j+1} - \vec{G}_j\big), \vec{G}_{k+1} - \vec{G}_k\big)(\widehat{\ell}_2)}{\boldsymbol{\sigma}\big(\vec{G}_k, \vec{G}_{k+1}\big)(\widehat{\ell}_2)}$$

for any $k = 4, \dots, N+1$. Evaluating $\mathscr{J}''_\rho$ at $(\delta x, \boldsymbol{\delta})$, we obtain

$$\mathscr{J}''_\rho[(\delta x, \boldsymbol{\delta})]^2 = \frac{1}{2}\sum_{i=3}^{N+1} \boldsymbol{\sigma}\Big(\delta_i \vec{G}_i, \mathrm{d}\mathcal{A}_{\rho_*}\delta x + \sum_{3\le j<i} \delta_j \vec{G}_j\Big)(\widehat{\ell}_2)$$
$$= \frac{1}{2}a(1-a)\langle \mathrm{d}\boldsymbol{t_k}(\widehat{\ell}_2)\,,\,\delta\ell\rangle^2 \boldsymbol{\sigma}\big(\vec{G}_{k+1}, \vec{G}_k\big)(\widehat{\ell}_2).$$

By Assumption 4, this quantity is non-positive, which contradicts Proposition 4.3. □

Remark 5.3 (On the choice of control variations). *Hamiltonian methods work under the assumption that the submanifold Λ_ρ is Lagrangian and projects injectively in M, which is equivalent to say that it is, at least locally, the graph of some scalar function $\mathscr{A}$. The coercivity of the second variation is usually exploited to guarantee the existence of such a submanifold Λ_ρ, in the following way: one interprets the second variation as the cost of a linear-quadratic optimal control problem (the control system given by the linearization of the pullback system, as* (4.11)*); standard LQ theory (see, for instance,* [18, 31]*) indeed relates the coercivity of the quadratic cost with the invertibility of the Hamiltonian flow (associated with the LQ problem) emanating from some n-dimensional linear space, constructed in the following way: it is the direct sum of the graph of the cost at the initial point of admissible trajectories, and the orthogonal complement to the space of the constraints on the initial point of admissible trajectories. If the problem has no constraints at the initial time, then this space is the graph of the cost at the initial point.*

In particular, in the presence of constraints at the initial time, such space does not project injectively; to address this issue, we extend again the second variation (as we did by adding the quadratic form $\rho\mathcal{K}$), in order to obtain a LQ problem with no constraints at the initial point, still paying attention that the functional is coercive on the (new) space of admissible variations.

In this paper, we must moreover ensure that Λ_ρ is contained in Σ, which imposes that $f_1(\widehat{x}_2)$ must be contained in $\ker\mathcal{K}$*. In other words, the direction $\mathbb{R}f_1(\widehat{x}_2)$ cannot be added in a second time, but must be already contained in $\widetilde{W}$* [3]*. Considering the control variation δu and then performing Goh transformation is a way to "generate" this direction in the space of admissible ones. The infinite-dimensional part w is not necessary to our aims, and can therefore be set to zero.*

Remark 5.4 (On the choice of the starting time). *When Hamiltonian methods were first developed (see for instance* [1, 3, 4, 31]*), the Lagrangian submanifold Λ_ρ was defined in a neighbourhood of the initial point of the reference trajectory, and the second variation was expressed as an LQ optimal control problem, defined on the tangent space to M at the same point; as a consequence, all flows were considered to evolve forward in time from time 0. We stress however that this choice is made only to simplify notations and, in*

[3]With a little abuse of language, when we say that $\mathbb{R}f_1(\widehat{x}_2)$ is contained in $\widetilde{W}$ we mean that $f_1(\widehat{x}_2)$ is an admissible value of $\zeta(\widehat{t}_1)$

general, any point of the reference trajectory could be chosen as the privileged one. In this paper, our choice concerns a neighborhood of the point $\widehat{x}_2$ for the following reasons: first of all, defining the submanifold Λ_ρ as the graph of a function defined in a neighborhood of $\widehat{x}_2$ makes the constraint $\Lambda_\rho \subset \Sigma$ easier to write in terms of the function itself; then, computing the second variation as an LQ problem on $T_{\widehat{x}_2}M$ permits to neglect variations along the first bang arc.

6. Minimum-Time problem: optimality

We are now ready to state and prove one of our main results.

Theorem 6.1. *Let $(\widehat{T}, \widehat{\xi}, \widehat{u})$ be an admissible triple for the minimum-time problem* (2.1)-(**T**)*, satisfying PMP with associate multipliers $\widehat{\lambda}$, $\widehat{\eta}$ and $p_0 = 1$ (normal extremal). Let Assumptions* 1–6 *be satisfied. Then $(\widehat{T}, \widehat{\xi}, \widehat{u})$ is a (time, state)-local minimizer.*

If $\mathcal{N}_f = \{\widehat{x}_f\}$, the minimum is strict.

Proof. The proof follows the same lines of [3, Theorem 17.1] (see also [33]), thus we are just recalling the main arguments.

We define on $\mathbb{R} \times T^*M$ the one-form $\omega = \mathcal{H}_t^* \varsigma - H_t \circ \mathcal{H}_t \mathrm{d}t$. Notice that ω is discontinuous through $t = \widehat{t}_1$ and $t = \widehat{t}_2$. Nevertheless, as the hypotheses of [33, Lemma 3.3] are met, one can prove that ω is exact[4] on $(-\epsilon, \widehat{T} + \epsilon) \times \Lambda_\rho$.

Let $\mathcal{O}$ be a tubular neighborhood of the graph of $\widehat{\xi}$, such that $(\pi\mathcal{H})^{-1}$ is well defined in $\mathcal{O}$ and let (T, ξ, v) be an admissible triple for the control system (2.1a)-(2.1b) with graph contained in $\mathcal{O}$.

Assume, by contradiction, that $T < \widehat{T}$. Let $y \colon [T, \widehat{T}] \to \mathcal{N}_f$ be any curve such that

$$y(T) = \xi(T), \quad y(\widehat{T}) = \widehat{\xi}(\widehat{T}).$$

We define the path $\gamma \colon [0, 2\widehat{T}] \to [0, \widehat{T}] \times \Lambda_\rho$ as

$$\gamma(t) = \begin{cases} (t, \ell(t)) = (\mathrm{id} \times \pi\mathcal{H})^{-1}(t, \xi(t)) & t \in [0, T], \\ (\mathrm{id} \times \pi\mathcal{H})^{-1}(t, y(t)) & t \in [T, \widehat{T}], \\ (2\widehat{T} - t, \widehat{\ell}_2) & t \in [\widehat{T}, 2\widehat{T}]. \end{cases}$$

We recall that indeed $\widehat{\ell}_2 = (\pi\mathcal{H}_t)^{-1}(\widehat{\xi}(t))$, for every $t \in [0, \widehat{T}]$.

Since γ is a closed path contained in $[0, \widehat{T}] \times \Lambda_\rho$, then $\oint_\gamma \omega = 0$, that is

$$0 = \oint_\gamma \mathcal{H}_t^* \varsigma - H_t \circ \mathcal{H}_t \,\mathrm{d}t = \int_0^T \langle \mathcal{H}_t(\ell(t)), \dot{\xi}(t) \rangle - H_t \circ \mathcal{H}_t(\ell(t)) \,\mathrm{d}t \quad (6.1)$$

$$+ \int_T^{\widehat{T}} \langle \mathcal{H}_t(\pi\mathcal{H}_t)^{-1}(y(t)), \dot{y}(t) \rangle - H_t \circ \mathcal{H}_t(\pi\mathcal{H}_t)^{-1}(y(t)) \,\mathrm{d}t \quad (6.2)$$

$$- \int_0^{\widehat{T}} \langle \mathcal{H}_t(\widehat{\ell}_2), \dot{\widehat{\xi}}(t) \rangle - H_t \circ \mathcal{H}_t(\widehat{\ell}_2) \,\mathrm{d}t.$$

[4] in the sense of Withney flat forms, as defined in [16].

First of all, we notice that

$$\int_0^{\widehat{T}} \langle \mathcal{H}_t(\widehat{\ell}_2), \dot{\widehat{\xi}}(t)\rangle - H_t \circ \mathcal{H}_t(\widehat{\ell}_2)\,\mathrm{d}t = \int_0^{\widehat{T}} \langle \widehat{\lambda}(t), \dot{\widehat{\xi}}(t)\rangle - (F_0(\widehat{\lambda}(t)) + \widehat{u}(t)F_1(\widehat{\lambda}(t)))\,\mathrm{d}t = 0,$$

as H_t coincides with $H_{\max}$ along the reference extremal.

The term on the right hand side of (6.1) may be split into the three addenda

$$\int_0^T \langle \mathcal{H}_t(\ell(t)), \dot{\xi}(t)\rangle - H_t \circ \mathcal{H}_t(\ell(t))\,\mathrm{d}t =$$

$$= \int_0^{\widehat{t}_1} \big(F_0 + v(t)F_1\big)(\mathcal{H}_t(\ell(t))) - H_{\max}(\mathcal{H}_t(\ell(t)))\,\mathrm{d}t \tag{6.3}$$

$$+ \int_{\widehat{t}_1}^{\widehat{t}_2} \big(F_0 + v(t)F_1\big)(\mathcal{H}_t(\ell(t))) - \big(H_0 + (v_s \circ \pi)F_1\big)(\mathcal{H}_t(\ell(t)))\,\mathrm{d}t \tag{6.4}$$

$$+ \int_{\widehat{t}_2}^{T} \big(F_0 + v(t)F_1\big)(\mathcal{H}_t(\ell(t))) - H_{\max}(\mathcal{H}_t(\ell(t)))\,\mathrm{d}t. \tag{6.5}$$

Lemmas 3.5 and 3.7 assure that the integrands in (6.3) and (6.5) are non-positive. Moreover, since $\pi\mathcal{H}_t(\ell(t)) = \xi(t)$ and $c(\xi(t)) \le 0$, Claim 1 in Proposition 3.2 implies that also (6.4) is non-positive. Hence,

$$\int_0^T \langle \mathcal{H}_t(\ell(t)), \dot{\xi}(t)\rangle - H_t \circ \mathcal{H}_t(\ell(t))\,\mathrm{d}t \le 0,$$

so that the term (6.2) must be non-negative. On the other hand,

$$\int_T^{\widehat{T}} \langle \mathcal{H}_t(\pi\mathcal{H}_t)^{-1}(y(t)), \dot{y}(t)\rangle - H_t \circ \mathcal{H}_t(\pi\mathcal{H}_t)^{-1}(y(t))\,\mathrm{d}t = \int_T^{\widehat{T}} -p_0 + O(t - \widehat{T})\,\mathrm{d}t = -\widehat{p}_0(\widehat{T} - T) + o(\widehat{T} - T),$$

since $H_{\widehat{T}} \circ \mathcal{H}_{\widehat{T}}(\pi\mathcal{H}_{\widehat{T}})^{-1}(y(\widehat{T})) = H_{\widehat{T}}(\widehat{\lambda}(\widehat{T})) = p_0$ and

$$\langle \mathcal{H}_{\widehat{T}}(\pi\mathcal{H}_{\widehat{T}})^{-1}(y(\widehat{T})), \dot{y}(\widehat{T})\rangle = \langle \widehat{\lambda}(\widehat{T}), \dot{y}(\widehat{T})\rangle = 0$$

by equation (2.6). Summing up, we obtain

$$0 \le -p_0(\widehat{T} - T) + o(\widehat{T} - T) = T - \widehat{T} + o(\widehat{T} - T),$$

which contradicts the hypothesis $T < \widehat{T}$.

Assume now that $\mathcal{N}_f$ is just a point and that $T = \widehat{T}$. Then the curve γ can be replaced by the path $\gamma\colon [0, 2\widehat{T}] \to [0, \widehat{T}] \times \Lambda_\rho$ defined as

$$\gamma(t) = \begin{cases} (t, \ell(t)) & t \in [0, \widehat{T}], \\ (2\widehat{T} - t, \widehat{\ell}_2) & t \in [\widehat{T}, 2\widehat{T}]. \end{cases}$$

Thus $\oint_\gamma \omega = 0$ implies that

$$\int_0^T \langle \mathcal{H}_t(\ell(t)), \dot{\xi}(t) \rangle - H_t \circ \mathcal{H}_t(\ell(t))\, \mathrm{d}t = 0.$$

By Lemma 3.5, we have that $v(t) \equiv u_1$ for $t \in [0, \widehat{t}_1)$, so that, in this interval, $\xi(t) = \widehat{\xi}(t)$. It shall also hold $F_0(\mathcal{H}_t(\ell(t))) = H_0\big(\mathcal{H}_t(\ell(t))\big)$ for every $t \in (\widehat{t}_1, \widehat{t}_2)$; by Claim 1 in Proposition 3.2, $c(\xi(t)) = 0$ on the same interval, so that $v(t) = v_s(\xi(t))$, for every $t \in (\widehat{t}_1, \widehat{t}_2)$, then $\xi(t)$ agrees with $\widehat{\xi}(t)$ up to $\widehat{t}_2$. Finally, by Lemma 3.7, we obtain that $v(t) = \widehat{u}(t)$ for $t \in (\widehat{t}_2, T)$. Then $\xi \equiv \widehat{\xi}$. □

We recall that the coercivity of the second variation is used only to prove the invertibility of the projected flow $\pi\mathcal{H}_t$ at the switching points $\widehat{t}_k$, $k \geq 3$; indeed, Proposition 5.1 already guarantees its invertibility along the first three arcs. Therefore, if the reference trajectory is made by the concatenation of a bang, a boundary and another bang arc, Theorem 6.1 requires Assumptions 1-2-3-5 only.

We point out that, in this case, sufficient optimality conditions are of first order only.

6.1. An example

This example is inspired from [34], which models the minimum-time path planning of a drone. The constrained control system is given by

$$\begin{cases} \dot{x}(t) = y(t), \\ \dot{y}(t) = Mu(t) - \frac{1}{2}ky^2(t), \\ x(0) = 0, \quad y(0) = 1, \\ x(T) = 20, \quad y(T) = 4, \\ |y(t)| \leq V, \\ |u(t)| \leq 1, \end{cases}$$

where $M = V = 10$ and $k = 0.05$. This is indeed a problem of the form (2.1)-($\mathbf{T}$) on $\mathbb{R}^2$, with

$$f_0 = y\frac{\partial}{\partial x} - \frac{1}{2}ky^2\frac{\partial}{\partial y}, \qquad f_1 = M\frac{\partial}{\partial y}, \qquad c(x, y) = |y| - V.$$

As in the cited paper, we consider the admissible trajectory associated with the control

$$u(t) = \begin{cases} 1 & t \in [0, \widehat{t}_1), \\ \frac{1}{4} & t \in (\widehat{t}_1, \widehat{t}_2), \\ -1 & t \in (\widehat{t}_2, T], \end{cases}$$

where

$$\widehat{t}_1 = \ln\frac{19}{7}, \quad \widehat{t}_2 = \widehat{t}_1 + 2 + 2\ln\frac{416}{665}, \quad T = \widehat{t}_2 + 2\arctan\frac{1}{2} - 2\arctan\frac{1}{5}.$$

By easy but tedious computations, we can show that the trajectory satisfies the PMP with adjoint covector satisfying $p_x(t) \equiv \frac{1}{10}$ and $p_y(0) = \frac{12}{133}$.

In particular, $y(t) = 10$ for every $t \in [\widehat{t}_1, \widehat{t}_2]$, so that $[\widehat{t}_1, \widehat{t}_2]$ is a boundary arc; on this interval, the associated density is $\eta \equiv -\frac{1}{10}$.

Assumptions 1 and 5 are visibly satisfied. As $L_{f_1} c \equiv 10$, also Assumption 2 holds true.

Finally, $F_1(p_x, p_y, x, y) = 10 p_y$; again by computations, we can prove that $p_2(t) > 0$ for $t \in [0, \widehat{t}_1)$ and $p_2(t) < 0$ for $t \in (\widehat{t}_2, T]$, so that also Assumption 3 holds true. Then the trajectory is a strict strong-local time minimizer.

7. The Mayer problem: optimality

In this Section, we consider the Mayer problem (2.1)-(**M**), and we prove the following result.

Theorem 7.1. *Let $(\widehat{\xi}, \widehat{u})$ be an admissible trajectory-control pair for problem* (2.1)-(**M**) *with associate multipliers $\widehat{\lambda}$, $\widehat{\eta}$. Let Assumptions* 1–6 *be satisfied.*

If the extremal is normal, then $(\widehat{\xi}, \widehat{u})$ is a strict strong-local minimizer for problem (2.1)*. If the extremal is abnormal, then $(\widehat{\xi}, \widehat{u})$ is isolated in the C^0 topology among all admissible trajectories.*

For the proof of Theorem 7.1, we rely on the construction done in the previous sections; in particular, we recall Proposition 5.2, that assures that the map id $\times\, \pi\mathcal{H}$ is invertible from $[0, T] \times \Lambda_\rho$ onto a neighbourhood $\mathcal{O}$ of the graph of $\widehat{\xi}$, and [33, Lemma 3.3], that guarantees that the one form $\mathcal{H}_t^* \omega$, where

$$\omega := \varsigma - H_t \, \mathrm{d}t,$$

is exact on $[0, T] \times \Lambda_\rho$.

Let us consider the function

$$\begin{aligned} \theta(t, \ell) = \theta_t(\ell) &:= \mathcal{A}_\rho(\pi\ell) + \int_{\mathcal{H}_\cdot(\ell)|_{[0,t]}} \varsigma - H_s \, \mathrm{d}s \\ &= \mathcal{A}_\rho(\pi\ell) + \int_0^t \Big(\langle \mathcal{H}_s(\ell) \, , \, \pi_* \vec{H}_s \circ \mathcal{H}_s(\ell) \rangle - H_s \circ \mathcal{H}_s(\ell) \Big) \, \mathrm{d}s \end{aligned}$$

Following the same lines of [33, Section 3], we get the following result.

Proposition 7.2. *The following expressions are true*

$$\mathrm{d}\theta_t(\ell) = \mathcal{H}_t^* \varsigma, \qquad \frac{\partial}{\partial t} \theta(t, \ell) = \langle \mathcal{H}_t(\ell) \, , \, \pi_* \vec{H}_t(\mathcal{H}_t(\ell)) \rangle - H_t(\mathcal{H}_t(\ell)).$$

In particular

$$\mathrm{d}\theta(t, \ell) = \mathcal{H}^* \omega$$

and

$$\mathrm{d} \left(\theta_t \circ (\pi\mathcal{H}_t)^{-1} \right) = \mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1} \qquad \text{on } \pi\mathcal{H}_t(\Lambda_\rho). \tag{7.1}$$

Remark 7.3. *With a small abuse of notation, we denote with $\mathrm{d}\theta_t(\ell)$ the differential of $\theta(t, \cdot)$ and with $\mathrm{d}\theta(t, \ell)$ the differential with respect to both variables.*

Proof of Theorem 7.1. The proof is similar to the one of Theorem 6.1: first of all, we consider an admissible pair (ξ, v) for the control system (2.1a)-(2.1b) with graph contained in the neighborhood $\mathcal{O}$ of the graph of $\widehat{\xi}$, and we choose a curve $\mathscr{y}\colon [0,1] \to \mathcal{N}_f$ such that

$$\mathscr{y}(0) = \xi(T), \quad \mathscr{y}(1) = \widehat{x}_f.$$

We define the closed path $\gamma\colon [0, 2T+1] \to [0,T] \times \Lambda$ as

$$\gamma(t) = \begin{cases} (t, \ell(t)) = (\mathrm{id} \times \pi \circ \mathcal{H}_t)^{-1}(\xi(t)) & t \in [0,T], \\ (T, (\pi\mathcal{H}_T)^{-1}\mathscr{y}(t-T)) & t \in [T, T+1], \\ (2T+1-t, \widehat{\ell}_2) & t \in [T+1, 2T+1]. \end{cases}$$

As above, $\oint_\gamma \omega = 0$ implies that

$$0 = \oint_\gamma \mathcal{H}_t^* \varsigma - H_t \circ \mathcal{H}_t \,\mathrm{d}t \tag{7.2}$$

$$= \int_0^T \langle \mathcal{H}_t(\ell(t)), \dot{\xi}(t)\rangle - H_t \circ \mathcal{H}_t(\ell(t)) \,\mathrm{d}t \tag{7.3}$$

$$+ \int_0^1 \langle \mathcal{H}_T(\pi\mathcal{H}_T)^{-1}(\mathscr{y}(t)), \dot{\mathscr{y}}(t)\rangle \,\mathrm{d}t - \int_0^T \langle \mathcal{H}_t(\widehat{\ell}_2), \dot{\widehat{\xi}}(t)\rangle - H_t \circ \mathcal{H}_t(\widehat{\ell}_2) \,\mathrm{d}t.$$

The addendum $\int_0^T \langle \mathcal{H}_t(\widehat{\ell}_2), \dot{\widehat{\xi}}(t)\rangle - H_t \circ \mathcal{H}_t(\widehat{\ell}_2) \,\mathrm{d}t$ is null by PMP, as already seen in the proof of Theorem 6.1. Analogously, the integrand at line (7.3) is non-positive. Therefore the integral along $\mathscr{y}$ cannot be negative. Hence, thanks to (7.1), we obtain

$$0 \le \int_0^1 \langle \mathcal{H}_T(\pi\mathcal{H}_T)^{-1}(\mathscr{y}(t)), \dot{\mathscr{y}}(t)\rangle \,\mathrm{d}t = \int_0^1 \langle \mathrm{d}\big(\theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\mathscr{y}(t)), \dot{\mathscr{y}}(t)\rangle \,\mathrm{d}t$$

$$= \theta_T \circ (\pi\mathcal{H}_T)^{-1}(\mathscr{y}(1)) - \theta_T \circ (\pi\mathcal{H}_T)^{-1}(\mathscr{y}(0)),$$

i.e.

$$\theta_T \circ (\pi\mathcal{H}_T)^{-1}(\xi(T)) - \theta_T \circ (\pi\mathcal{H}_T)^{-1}(\widehat{x}_f) \le 0,$$

thus

$$p_0\psi(\xi(T)) - p_0\psi(\widehat{x}_f) \\ \ge \big(\beta + \theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\xi(T)) - \big(\beta + \theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\widehat{x}_f). \tag{7.4}$$

We notice that $\mathrm{d}\big(\beta + \theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\widehat{x}_f) = 0$, thanks to equation (7.1) and the choice of β, so that, if $p_0 = 1$, and provided that

$$\mathrm{D}^2\big(\beta + \theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\widehat{x}_f)[\delta y]^2 > 0, \qquad \forall \delta y \in T_{\widehat{x}_f}\mathcal{N}_f, \tag{7.5}$$

then $\psi(\xi(T)) \ge \psi(\widehat{x}_f)$ and $\widehat{\xi}$ is a strong-local minimum for problem (2.1)-(**M**).

In order to prove inequality (7.5), we start by noticing that

$$\mathrm{D}^2\big(\beta + \theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\widehat{x}_f)[\delta y]^2 = \boldsymbol{\sigma}\big(\mathrm{d}\big(\theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)_* \delta y, \mathrm{d}(-\beta)_* \delta y\big)(\widehat{\ell}_T)$$

$$= \boldsymbol{\sigma}\big(\mathcal{H}_{T*}(\pi\mathcal{H}_T)_*^{-1}\delta y, \mathrm{d}(-\beta)_*\delta y\big)(\widehat{\ell}_T). \tag{7.6}$$

Setting $\delta\ell = (\pi\mathcal{H}_T)_*^{-1}\delta y$, we notice that $\delta\ell \in T_{\widehat{x}_2}\Lambda_\rho$, so that $\delta\ell = \mathrm{d}\mathcal{A}_{\rho*}\delta x$ for some $\delta x \in T_{\widehat{x}_2}M$. Differentiating equation (3.2), we obtain that

$$\begin{aligned} \mathcal{H}_{T*}(\pi\mathcal{H}_T)_*^{-1}\delta y &= \mathcal{H}_{T*}\delta\ell \\ &= \widehat{\mathscr{F}}_{T*}\delta\ell + \sum_{j=3}^{N}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\,\delta\ell\rangle \widehat{\mathscr{F}}_{T*}\widehat{\mathscr{F}}_{\widehat{t}_j*}^{-1}\big(\vec{H}_j - \vec{H}_{j+1}\big)(\widehat{\ell}_j) \\ &= \widehat{\mathscr{F}}_{T*}\Big(\delta\ell + \sum_{j=3}^{N}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\,\delta\ell\rangle\big(\vec{G}_j - \vec{G}_{j+1}\big)(\widehat{\ell}_2)\Big). \end{aligned} \tag{7.7}$$

Substituting this expression into (7.6), we get

$$\begin{aligned} &\mathrm{D}^2\big(\beta + \theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\widehat{x}_f)[\delta y]^2 \\ &= \boldsymbol{\sigma}\Big(\widehat{\mathscr{F}}_{T*}\Big(\delta\ell + \sum_{j=3}^{N}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\,\delta\ell\rangle\big(\vec{G}_j - \vec{G}_{j+1}\big)(\widehat{\ell}_2)\Big), \mathrm{d}(-\beta)_*\delta y\Big)(\widehat{\ell}_T) \\ &= \boldsymbol{\sigma}\Big(\delta\ell + \sum_{j=3}^{N}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\,\delta\ell\rangle\big(\vec{G}_j - \vec{G}_{j+1}\big)(\widehat{\ell}_2), \mathrm{d}(-\widehat{\beta})_*\widehat{S}_{T*}^{-1}\delta y\Big)(\widehat{\ell}_2) \\ &= \boldsymbol{\sigma}\Big(\delta\ell + \sum_{j=3}^{N}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\,\delta\ell\rangle\big(\vec{G}_j - \vec{G}_{j+1}\big)(\widehat{\ell}_2), \\ &\qquad\qquad \mathrm{d}(-\widehat{\beta})_*\Big(\pi_*\delta\ell + \sum_{j=3}^{N}\langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\,\delta\ell\rangle\big(g_j - g_{j+1}\big)(\widehat{x}_2)\Big)\Big)(\widehat{\ell}_2), \end{aligned}$$

where, to compute $\widehat{S}_{T*}^{-1}\delta y$, we used the definition of $\delta\ell$, equation (7.7) and the fact that $\pi_*\widehat{\mathscr{F}}_{T*} = \widehat{S}_{T*}$.

Using the expression (4.13), it is easy to check that the above quantity is twice $\mathscr{J}_\rho''$ evaluated at the following admissible variation:

$$\begin{aligned} &\delta x = \pi_*\delta\ell, \qquad \delta_3 = \langle \mathrm{d}\boldsymbol{t_3}(\widehat{\ell}_2)\,,\,\delta\ell\rangle, \\ &\delta_j = \langle \mathrm{d}\boldsymbol{t_j}(\widehat{\ell}_2)\,,\,\delta\ell\rangle - \langle \mathrm{d}\boldsymbol{t_{j-1}}(\widehat{\ell}_2)\,,\,\delta\ell\rangle, \qquad j = 4, \dots, N \\ &\delta_{N+1} = -\langle \mathrm{d}\boldsymbol{t_N}(\widehat{\ell}_2)\,,\,\delta\ell\rangle. \end{aligned}$$

Proposition 4.3 yields the claim.

From inequality (7.5), the equality $p_0\psi(\xi(T)) = p_0\psi(\widehat{x}_f)$ can hold only if $\xi(T) = \widehat{x}_f$. Using the same arguments of Theorem 6.1, we can conclude that $\xi \equiv \widehat{\xi}$, that is, $\widehat{\xi}$ is strictly optimal.

Assume instead that $\widehat{\lambda}$ is an abnormal extremal, that is, $p_0 = 0$. Inequality (7.4), together with $\mathrm{d}\big(\beta + \theta_T \circ (\pi\mathcal{H}_T)^{-1}\big)(\widehat{x}_f) = 0$ and inequality (7.5), imply that $\xi(T) = \widehat{x}_f$ for every admissible trajectory with graph in $\mathcal{O}$.

Using the same arguments as in the proof of Theorem 6.1, based on regularity conditions, we obtain that every integral in (7.2) is zero. This in particular implies that any admissible trajectory is associated with the same controls as the reference one. Thus $\widehat{\xi}$ is isolated among all admissible trajectories. □

We end this section by remarking that, differently from the minimum-time problem, for the Mayer problem the coercivity of the second variation is needed also if $N = 2$, that is, when the boundary arc is followed by only one internal (bang) arc.

However, the problem is simple. We can distinguish two cases. If $f_1(\widehat{x}_2) \notin T_{\widehat{x}_2}\widehat{\mathcal{N}_f}$, then $\mathcal{W}_0$ is trivial, so that the second variation is, by definition, coercive, and the reference trajectory is optimal.

If instead $f_1(\widehat{x}_2) \in T_{\widehat{x}_2}\widehat{\mathcal{N}_f}$, then Assumption 6 boils down to just one inequality:

$$L^2_{f_1}\big(\mathcal{A}(\widehat{t}_2,\cdot) + \widehat{\beta}\big)(\widehat{x}_2) > 0.$$

7.1. An example

We consider the problem of minimizing the functional

$$J = \int_0^4 -(x(t) - 2)^2 \,\mathrm{d}t$$

over all the solutions of the constrained control problem

$$\begin{cases} \dot{x}(t) = u(t), \\ x(0) = x(4) = 1, \\ x(t) \geq 0, \\ u(t) \in [-1,1] \quad \text{a.e. } t \in [0,4]. \end{cases}$$

The problem here above can be written equivalently as a problem of the form (2.1)-(**M**) in $\mathbb{R}^2$, setting

$$f_0 = -(x-2)^2 \frac{\partial}{\partial y}, \quad f_1 = \frac{\partial}{\partial x}, \quad c(x,y) = -x, \quad \psi(x,y) = y,$$
$$\xi(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathcal{N}_f = \big\{(x,y) \in \mathbb{R}^2 \colon x = 1\big\}.$$

Intuition suggests to consider the admissible trajectory $\widehat{\xi}$ associated with the control

$$\widehat{u}(t) = \begin{cases} -1 & t \in [0,1), \\ 0 & t \in (1,3), \\ 1 & t \in (3,4]. \end{cases}$$

We claim that it is a local minimum. Assumptions 1-2 are trivially satisfied. Moreover, the PMP holds in normal form with $\widehat{\lambda} = (\widehat{p}_x, \widehat{p}_y, \widehat{\xi})$, where

$$\widehat{p}_x(t) = \begin{cases} -4 + (t+1)^2 & t \in [0,1], \\ 0 & t \in [1,3], \\ 4 - (t-5)^2 & t \in [3,4], \end{cases} \qquad \widehat{p}_y(t) \equiv -1, \qquad \eta \equiv -4,$$

which implies the validity of Assumption 5. As $F_1(p_x, p_y, x, y) = p_x$, also Assumption 3 is satisfied.

It is easy to check that $\widehat{\mathcal{N}_f} = \{(x, y) \in \mathbb{R}^2 \colon x = 0\}$, so that f_1 is nowhere tangent to $\widehat{\mathcal{N}_f}$. Thus $\mathcal{W}_0$ is the null subspace of $\mathbb{R}^2$, so that Assumption 6 is trivially met. Hence all the assumptions are satisfied.

8. Conclusions

The problem studied in this paper deals with extremals with one initial (bang) internal arc and a finite number of (bang) internal arcs after the boundary one. Nonetheless, in the literature (for instance in [23]), extremals involving several bang internal arcs before *and* after the boundary one can be found. We believe that the result presented in this paper can be generalized to these more involved cases, with just minor efforts.

Other cases are more challenging. As already pointed out in [32], for problems with free time, it is possible to define two notions of strong local optimality: the one given in Definition 1.1 (called (x, t_f)-local optimality, or (time, state)-local optimality), and a stronger one, in which a neighborhood of the *range* of the reference trajectory is considered (this is called x-local optimality, or state-local optimality). Hamiltonian methods have already provided sufficient optimality conditions for state-local optimality in single-input minimum time problems (see [25, 26]). It would be interesting to use this approach to find state-optimality conditions in presence of state constraints.

Further developments may include: higher order boundary arcs, in which the costate may be discontinuous at the junction points (such jumps have never been treated with Hamiltonian methods); extremals containing also singular internal arcs; the multi-input case.

Appendix A. On the computation of first and second variation

In this section, we provide more details on the constructions of the first and second variations of the functional J. As in Section 4, we focus on the minimum-time problem.

We start by recalling the coordinate expression of the reference extremal. Choosing any (local) coordinate chart on T^*M and setting $\widehat{\lambda}(t) = \big(\widehat{\mu}(t), \widehat{\xi}(t)\big)$, equation (2.5c) implies that $\widehat{\mu}(t)$ is a solution to the following linear ODE

$$\dot{\mu}(t) = -\mu(t)\mathrm{D}\widehat{f}_t(\widehat{\xi}(t)) - \widehat{\eta}(t)\,\mathrm{d}c(\widehat{\xi}(t)).$$

Then, for every $t \in [0, T]$, we have

$$\widehat{\mu}(t) = \Big(\widehat{\mu}(\widehat{t}_2) - \int_{\widehat{t}_2}^{t} \widehat{\eta}(s)\,\mathrm{d}\widehat{c}_s(\widehat{x}_2)\,\mathrm{d}s\Big)\widehat{S}_{t*}^{-1}, \tag{A.1}$$

where we recall that $\widehat{\eta}$ is supported on $[\widehat{t}_1, \widehat{t}_2]$.

Let us now prove that the first variation, with respect to $\varkappa_0, u_0$ and u, of the cost (4.5) is null. Using (A.1), we can see that

$$\mathrm{d}\widehat{\alpha}(\widehat{x}_2) = \widehat{\ell}_2 + \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s)\,\mathrm{d}\widehat{c}_s(\widehat{x}_2)\,\mathrm{d}s, \qquad \mathrm{d}\widehat{\beta}(\widehat{x}_2) = -\widehat{\ell}_T \widehat{S}_{T*} = -\widehat{\ell}_2,$$

Then we have that

$$\frac{\partial J}{\partial \varkappa_0}(\widehat{x}_2, 1, \widehat{u}) = \mathrm{d}\widehat{\alpha}(\widehat{x}_2) + \mathrm{d}\widehat{\beta}(\widehat{x}_2) - \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s)\,\mathrm{d}\widehat{c}_s(\widehat{x}_1)\,\mathrm{d}s = 0.$$

Thanks to (4.3), the first variation of J with respect to u boils down to

$$\begin{aligned}
\frac{\partial J}{\partial u}(\widehat{x}_2, 1, \widehat{u}) &= \mathrm{d}\widehat{\alpha}(\widehat{x}_2)\frac{\partial \varkappa_0}{\partial u} - \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s)\,\mathrm{d}\widehat{c}_s(\widehat{x}_1)\frac{\partial \varkappa_s}{\partial u}\,\mathrm{d}s \\
&= \langle \mathrm{d}\widehat{\alpha}(\widehat{x}_2), \int_{\widehat{t}_2}^{\widehat{t}_1} k_s(\widehat{x}_2)\mathrm{d}s\rangle - \int_{\widehat{t}_1}^{\widehat{t}_2} \widehat{\eta}(s)\langle \mathrm{d}\widehat{c}_s(\widehat{x}_1), \int_{t_2}^{s} k_r(\widehat{x}_2)\mathrm{d}r\rangle \mathrm{d}s \\
&= \int_{\widehat{t}_2}^{\widehat{t}_1} \langle \mathrm{d}\widehat{\alpha}(\widehat{x}_2) - \int_{\widehat{t}_1}^{s} \widehat{\eta}(r)\mathrm{d}\widehat{c}_r(\widehat{x}_1)\mathrm{d}r, k_s(\widehat{x}_2)\rangle \mathrm{d}s \\
&= \int_{\widehat{t}_2}^{\widehat{t}_1} \langle \widehat{\lambda}(s), f_1(\widehat{\xi}(s))\mathrm{d}s,
\end{aligned}$$

which is null by PMP. Finally

$$\begin{aligned}
\langle \frac{\partial J}{\partial u_0}(\widehat{x}_2, 1, \widehat{u})\,,\, \delta u_0\rangle &= p_0 \langle \frac{\partial z_T^0}{\partial u_0}\,,\, \delta u_0\rangle + \langle \mathrm{d}\widehat{\beta}(\widehat{x}_2)\,,\, \frac{\partial \varkappa_T}{\partial u_0}\delta u_0\rangle \\
&= \int_{\widehat{t}_2}^{\widehat{T}} \delta u_0(t)\mathrm{d}t + \langle \mathrm{d}\widehat{\beta}(\widehat{x}_2)\,,\, \int_{\widehat{t}_2}^{T} \delta u_0(t)\widehat{S}_{t*}^{-1}\widehat{f}_t \circ \widehat{S}_t(\widehat{x}_2)\mathrm{d}t\rangle \\
&= p_0 \int_{\widehat{t}_2}^{\widehat{T}} \delta u_0(t)\mathrm{d}t - \langle \widehat{\ell}_2\,,\, \sum_{k=3}^{N+1} \delta_k g_k(\widehat{x}_2)\rangle \\
&= p_0 \int_{\widehat{t}_2}^{\widehat{T}} \delta u_0(t)\mathrm{d}t - \sum_{k=3}^{N+1} \delta_k F_0(\widehat{\ell}_k),
\end{aligned}$$

where we recall that the δ_k are defined in equation (4.6). The first addendum is null because of the constraint (4.9), and the second addendum is null since $F_0(\widehat{\ell}_k) = p_0$ for any k, so that

$$\sum_{k=3}^{N+1} \delta_k F_0(\widehat{\ell}_k) = p_0 \sum_{k=3}^{N+1} \delta_k = 0.$$

Differentiating twice the cost J with respect to $\varkappa_0, u_0$ and u, we obtain an expression which depends on δu_0 only by means of the quantities $\delta_3, \ldots, \delta_{N+1}$, i.e. we get

$$J''[\delta y, \delta u, \boldsymbol{\delta}]^2 = \frac{1}{2}\mathrm{D}^2\big(\mathcal{A}(\widehat{t}_2, \cdot) + \widehat{\beta}\big)(\widehat{x}_2)[\delta\varkappa_{\widehat{t}_2}]^2$$

$$
\begin{aligned}
&+ L_{\delta x_{\widehat{t}_2}} L_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2)\\
&+ \frac{1}{2}L^2_{\sum_{i=3}^{N+1}\delta_i g_i}\widehat{\beta}(\widehat{x}_2) + \frac{1}{2}\sum_{3\le i<j\le N+1}\delta_i\delta_j L_{[g_i,g_j]}\widehat{\beta}(\widehat{x}_2)\\
&- L_{\delta x_{\widehat{t}_2}}\Big(\int_{\widehat{t}_1}^{\widehat{t}_2}\Big(\delta u(s)L_{k_s}\widehat{\alpha} - \widehat{\eta}(s)\int_s^{\widehat{t}_2}\delta u(r)L_{k_r}\widehat{c}_s\mathrm{d}r\Big)\mathrm{d}s\Big)(\widehat{x}_2) \quad \text{(A.2)}\\
&+ \int_{\widehat{t}_1}^{\widehat{t}_2}\int_{\widehat{t}_1}^{\widehat{t}_2}\delta u(r)\delta u(s)L_{k_r}L_{k_s}\widehat{\alpha}(\widehat{x}_2)\mathrm{d}r\mathrm{d}s \quad \text{(A.3)}\\
&+ \int_{\widehat{t}_1}^{\widehat{t}_2}\int_s^{\widehat{t}_2}\delta u(s)\delta u(r)L_{[k_r,k_s]}\widehat{\alpha}(\widehat{x}_2)\mathrm{d}r\mathrm{d}s \quad \text{(A.4)}\\
&- \int_{\widehat{t}_1}^{\widehat{t}_2}\widehat{\eta}(s)\int_s^{\widehat{t}_2}\Big(\int_s^{\widehat{t}_2}\delta u(r)\delta u(\tau)L_{k_r}L_{k_\tau}\widehat{c}_s(\widehat{x}_2)\mathrm{d}\tau\Big)\mathrm{d}r\mathrm{d}s \quad \text{(A.5)}\\
&- \int_{\widehat{t}_1}^{\widehat{t}_2}\widehat{\eta}(s)\int_s^{\widehat{t}_2}\Big(\int_r^{\widehat{t}_2}\delta u(r)\delta u(\tau)L_{[k_\tau,k_r]}\widehat{c}_s(\widehat{x}_2)\mathrm{d}\tau\Big)\mathrm{d}r\mathrm{d}s, \quad \text{(A.6)}
\end{aligned}
$$

with $(\delta y, \delta u, \boldsymbol{\delta}) \in W$.

Exchanging the order of integration and relabeling the variables, we remark that

$$
\int_{\widehat{t}_1}^{\widehat{t}_2}\Big(\delta u(s)L_{k_s}\widehat{\alpha}(x) - \widehat{\eta}(s)\int_s^{\widehat{t}_2}\delta u(r)L_{k_r}\widehat{c}_s(x)\mathrm{d}r\Big)\mathrm{d}s = \int_{\widehat{t}_1}^{\widehat{t}_2}\delta u(s)L_{k_s}\mathcal{A}(s,x)\mathrm{d}s,
$$

so (A.2) can be written as

$$
-L_{\delta x}\int_{\widehat{t}_1}^{\widehat{t}_2}\delta u(s)L_{k_s}\mathcal{A}(s,x)\mathrm{d}s.
$$

Also, we have that

$$
\begin{aligned}
\text{(A.5)} &= \int_{\widehat{t}_1}^{\widehat{t}_2}\left(\int_{\widehat{t}_1}^{r}\Big(\int_s^{\widehat{t}_2}\widehat{\eta}(s)\delta u(r)\delta u(\tau)L_{k_r}L_{k_\tau}\widehat{c}_s(\widehat{x}_2)\mathrm{d}\tau\Big)\mathrm{d}s\right)\mathrm{d}r\\
&= \int_{\widehat{t}_1}^{\widehat{t}_2}\left(\int_{\widehat{t}_1}^{\widehat{t}_2}\Big(\int_{\widehat{t}_1}^{\tau}\widehat{\eta}(s)\delta u(r)\delta u(\tau)L_{k_r}L_{k_\tau}\widehat{c}_s(\widehat{x}_2)\mathrm{d}s\Big)d\tau\right.\\
&\quad \left. - \int_r^{\widehat{t}_2}\Big(\int_r^{\tau}\widehat{\eta}(s)\delta u(r)\delta u(\tau)L_{k_r}L_{k_\tau}\widehat{c}_s(\widehat{x}_2)\mathrm{d}s\Big)d\tau\right)\mathrm{d}r;
\end{aligned}
$$

noticing that $\int_r^{\tau}\widehat{\eta}(s)\widehat{c}_s(\widehat{x}_2)\mathrm{d}s = \mathcal{A}(\tau,\widehat{x}_2) - \mathcal{A}(r,\widehat{x}_2)$, we obtain that

$$
\begin{aligned}
\text{(A.3)} - \text{(A.5)} &= \int_{\widehat{t}_1}^{\widehat{t}_2}\left(\int_{\widehat{t}_1}^{\widehat{t}_2}\delta u(r)\delta u(\tau)L_{k_r}L_{k_\tau}\mathcal{A}(\tau,\widehat{x}_2)\mathrm{d}\tau\right)\mathrm{d}r\\
&+ \int_{\widehat{t}_1}^{\widehat{t}_2}\left(\int_r^{\widehat{t}_2}\delta u(r)\delta u(\tau)L_{k_r}L_{k_\tau}(\mathcal{A}(r,\widehat{x}_2) - \mathcal{A}(\tau,\widehat{x}_2))\mathrm{d}\tau\right)\mathrm{d}r
\end{aligned}
$$

$$= \int_{\widehat{t}_1}^{\widehat{t}_2} \left(\int_{\widehat{t}_1}^{r} \delta u(r)\delta u(\tau) L_{k_r} L_{k_\tau} \mathcal{A}(\tau, \widehat{x}_2) \mathrm{d}\tau \right) \mathrm{d}r$$

$$+ \int_{\widehat{t}_1}^{\widehat{t}_2} \left(\int_{r}^{\widehat{t}_2} \delta u(r)\delta u(\tau) L_{k_r} L_{k_\tau} (\mathcal{A}(\tau, \widehat{x}_2) + \mathcal{A}(r, \widehat{x}_2) - \mathcal{A}(\tau, \widehat{x}_2)) \mathrm{d}\tau \right) \mathrm{d}r$$

$$= \int_{\widehat{t}_1}^{\widehat{t}_2} \left(\int_{\widehat{t}_1}^{r} \delta u(r)\delta u(\tau) \Big(L_{k_r} L_{k_\tau} + L_{k_\tau} L_{k_r} \Big) \mathcal{A}(\tau, \widehat{x}_2) \mathrm{d}\tau \Big) \mathrm{d}r.$$

Analogous computations show that

$$(\text{A.4}) - (\text{A.6}) = \int_{\widehat{t}_1}^{\widehat{t}_2} \Big(\int_{r}^{\widehat{t}_2} \delta u(r)\delta u(\tau) L_{[k_r, k_\tau]} \mathcal{A}(r, \widehat{x}_2) \mathrm{d}\tau \Big) \mathrm{d}r.$$

Summing up the addenda, we obtain the expression (4.10).

Acknowledgment

This work was completed with the support of Appel à projet "Enseignants-Chercheurs invités", Université de Toulon, and by Progetto Internazionalizzazione, Università degli Studi di Firenze.

Author contributions. F.C.C. and L.P. contributed equally to the manuscript.

Declarations.

Conflict of interests. The authors declare that they have no conflict of interests.

References

[1] A. A. Agrachev and R. V. Gamkrelidze. Symplectic geometry for optimal control. In Hector J. Sussmann, editor, *Nonlinear Controllability and Optimal Control*, volume 133 of *Pure and Applied Mathematics*. Marcel Dekker, 1990.

[2] A. A. Agrachev and R. V. Gamkrelidze. Symplectic methods for optimization and control. In B. Jacubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*, Pure and Applied Mathematics, pages 1–58, New York, 1997. Marcel Dekker.

[3] A. A. Agrachev and Yu. L. Sachkov. *Control Theory from the Geometric Viewpoint*. Springer-Verlag, 2004.

[4] A. A. Agrachev, G. Stefani, and P. Zezza. An invariant second variation in optimal control. *Internat. J. Control*, 71(5):689–715, 1998.

[5] A. A. Agrachev, G. Stefani, and P. Zezza. Strong optimality for a bang-bang trajectory. *SIAM J. Control Optimization*, 41(4):991–1014, 2002.

[6] M. S. Aronna, J. F. Bonnans, and B. S. Goh. Second order analysis of state-constrained control-affine problems. *Mathematical Programming, Series A*, 160(1):115–147, January 2016.

[7] J. F. Bonnans and A. Hermant. No-gap second-order optimality conditions for optimal control problems with a single state constraint and control. *Mathematical Programming*, 117:21–50, 2009.

[8] B. Bonnard, L. Faubourg, G. Launay, and E. Trélat. Optimal Control with state constraints and the space shuttle re-entry problem. *Journal of Dynamical and Control Systems*, 9(2):155–199, 2003.

[9] F. H. Clarke. On the inverse function theorem. *Pacific J. Mathematics*, 64(1):97–102, 1976.

[10] D. L. Cohn. *Measure Theory*. Birkhäuser Advanced Texts Basler Lehrbücher (BAT). Birkhäuser, second edition.

[11] O. Cots. Geometric and numerical methods for a state constrained minimum time control problem of an electric vehicle. *ESAIM: COCV*, 23(4):1715–1749, 2017.

[12] O. Cots, J. Gergaud, and D. Goubinat. Direct and indirect methods in optimal control with state constraints and the climbing trajectory of an aircraft. *Optimal Control Applications and Methods*, 39(1):281–301, 2018.

[13] M. do Rosario de Pinho, M. Margarida Ferreira, U. Ledzewicz, and H. Schaettler. A model for cancer chemotherapy with state-space constraints. *Nonlinear Analysis: Theory, Methods & Applications*, 63(5):e2591–e2602, 2005. Invited Talks from the Fourth World Congress of Nonlinear Analysts (WCNA 2004).

[14] M. Fehér, O. Straka, and V. Šmídl. Constrained time-optimal control of double-integrator system and its application in MPC. *Journal of Physics: Conference Series*, 783:012024, jan 2017.

[15] R. F. Hartl, S. P. Sethi, and R. G. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev*, 37(2):181–218, 1995.

[16] J. M. Heinonen. Lectures on Lipschitz analysis. 2005.

[17] M. R. Hestenes. *Calculus of Variations and Optimal Control Theory*. John Wiley & Sons, New York, New York, 1966.

[18] V. Jurdjevic. *Geometric Control Theory*, volume 52 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1997.

[19] J. B. Keller. Optimal velocity in a race. *The American Mathematical Monthly*, 81(5):474–480, 1974.

[20] W. Liu and H. J. Sussmann. Continuous dependence with respect to the input of trajectories of control-affine systems. *SIAM Journal on Control and Optimization*, 37(3):777–803, 1999.

[21] K. Malanowski. On normality of Lagrange multipliers for state constrained optimal control problems. *Optimization*, 52(1):75–91, 2003.

[22] H. Maurer. On optimal control problems with bounded state variables and control appearing linearly. *SIAM J. Control Optim.*, 15(3):345–362, 1977.

[23] H. Maurer, J.-H. R. Kim, and G. Vossen. *On A State-Constrained Control Problem in Optimal Production and Maintenance*, pages 289–308. Springer US, 2005.

[24] A. B. Pitcher. Optimal strategies for a two-runner model of middle-distance running. *SIAM Journal on Applied Mathematics*, 70(4):1032–1046, 2009.

[25] L. Poggiolini and G. Stefani. State-local optimality of a bang-bang trajectory: a Hamiltonian approach. *Systems & Control Letters*, 53:269–279, 2004.

[26] L. Poggiolini and G. Stefani. Bang-singular-bang extremals: sufficient optimality conditions. *Journal of Dynamical and Control Systems*, 17:469–514, 2011. 10.1007/s10883-011-9127-y.

[27] H. Schättler. Local fields of extremals for optimal control problems with state constraints of relative degree 1. *Journal of Dynamical and Control Systems*, 12:563–599, 01 2006.

[28] H. Seywald and E. M. Cliff. Goddard problem in presence of a dynamic pressure limit. *Journal of Guidance, Control, and Dynamics*, 16(4):776–781, 1993.

[29] G. Stefani. Minimum–time optimality of a singular arc: second order sufficient conditions. In *43th IEEE Conference on Decision and Control*, 2004.

[30] G. Stefani. Strong optimality of singular trajectories. In F. Ancona, A. Bressan, P. Cannarsa, F. Clarke, and P. Wolenski, editors, *Geometric Control and Nonsmooth Analysis*, volume 76 of *Series on Advances in Mathematics for Applied Sciences*, pages 300–326, Hackensack, NJ, 2008. World Scientific Publishing Co. Pte. Ltd. pp. 361 ISBN: 978-981-277-606-8.

[31] G. Stefani and P. Zezza. Constrained regular LQ-control problems. *SIAM J. Control Optim.*, 35(3):876–900, 1997.

[32] G. Stefani and P. Zezza. Time-optimality of a bang-bang trajectory with maple. In *2nd IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, 2003.

[33] G. Stefani and P. Zezza. *Variational Methods in Imaging and Geometric Control*, chapter A Hamiltonian approach to sufficiency in optimal control with minimal regularity conditions: Part I. De Gruyter, 2016.

[34] A. Taitler, I. Ioslovich, E. Karpas, and P.-O. Gutman. Minimum time optimal control of second order system with quadratic drag and state constraints. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 523–528, 2019.

F. C. Chittaro*
Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France
e-mail: francesca.chittaro@univ-tln.fr

L. Poggiolini
DiMaI, Università di Firenze, 50139 Firenze, Italy
e-mail: laura.poggiolini@unifi.it