Strong local optimality for a bang-bang-singular extremal: the fixed-free case

STRONG LOCAL OPTIMALITY FOR A BANG-BANG-SINGULAR EXTREMAL: THE FIXED-FREE CASE*

LAURA POGGIOLINI† AND GIANNA STEFANI†

Abstract. In this paper we give sufficient conditions for a Pontryagin extremal trajectory, consisting of two bang arcs followed by a partially or totally singular one, to be a strong local minimizer for a Mayer problem. The problem is defined on $\mathbb{R}^n$ and the end-points constraints are of fixed-free type. We use a Hamiltonian approach and its connection with the second order conditions in the form of a LQ accessory problem. An example is proposed. All the results are coordinate free so they also hold on a manifold.

Key words. Sufficient optimality conditions, singular control, second variation, Hamiltonian methods

AMS subject classifications. 49K15, 49J15, 93C10

1. Introduction.

In this paper we consider a Mayer problem on a fixed time interval $[0, T]$ and governed by a control affine dynamics. We study the strong local optimality of a trajectory consisting of two bang arcs followed by a singular one.

In Optimal Control literature two different kinds of local optimality are usually considered: *weak local optimality*, i.e. with respect to the $C^0 \times L^\infty$-distance of the couples (trajectory, associated control); *strong local optimality*, i.e. with respect to the C^0-distance of admissible trajectories, without any localization on the controls and which is defined below. An intermediate kind of local optimality, called *Pontryagin local optimality*, is also studied in the literature, see for example [10]. In our case Pontryagin local optimality reduces to local optimality with respect to the $C^0 \times L^1$-distance of the couples (trajectory, associated control).

Here we give sufficient optimality conditions for the reference trajectory to be a strong local minimizer in the case when the end-point constraints are of fixed-free type. We also recall that since a Bolza problem can always be reduced to a Mayer one, sufficient optimality conditions can be also derived for a Bolza problem.

Control affine systems can be modeled in different ways; since we want to consider both bang and partially singular arcs (see Definition 1.2), we model the system as follows: let $X_1, \ldots, X_m$ be smooth vector fields on $\mathbb{R}^n$ and let $\mathcal{X}$ be their convex hull, i.e.

$$\mathcal{X}(x) = \Big\{ \sum_{i=1}^{m} u_i X_i(x) \colon u = (u_1, \ldots, u_m) \in \Delta \Big\},$$

where $\Delta := \{u \in \mathbb{R}^m \colon u_i \geq 0,\ i = 1, \ldots, m,\ \sum_{i=1}^m u_i = 1\}$ is the simplex of $\mathbb{R}^m$.

Given a smooth function $c \colon \mathbb{R}^n \to \mathbb{R}$ and a point $x_0 \in \mathbb{R}^n$, we consider a Mayer problem of the following kind

$$\text{minimize } \ c(\xi(T)) \text{ subject to} \tag{1a}$$

$$\dot\xi(t) \in \mathcal{X}(\xi(t)) \quad \text{a.e. } t \in [0, T], \tag{1b}$$

$$\xi(0) = x_0. \tag{1c}$$

*Submitted to the editors 8th March 2018.

Funding: The authors are partially supported by Gruppo Nazionale per l'Analisi Matematica, la Probabilitè le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

†DIMAI, Università degli Studi di Firenze, Italy (laura.poggiolini@unifi.it, gianna.stefani@unifi.it, http://www.dma.unifi.it/~poggiolini/).

Equivalently, by Filippov's theorem, equation (1b) can also be written as

$$\dot{\xi}(t) = \sum_{i=1}^{m} \upsilon_i(t) X_i(\xi(t)), \quad \text{a.e. } t \in [0,T], \quad \upsilon \in L^\infty([0,T],\Delta).$$

Our aim is to give sufficient conditions for a reference trajectory $\widehat{\xi}\colon [0,T] \to \mathbb{R}^n$ of (1b)-(1c) to be indeed a *strong local minimizer* of the problem according to the following definition

DEFINITION 1.1. *An admissible trajectory $\widehat{\xi}\colon [0,T] \to \mathbb{R}^n$ for an optimal control problem is a* strong local minimizer *if it is a minimizer among the admissible trajectories which are in a neighborhood of $\widehat{\xi}$ with respect to the C^0 topology.*

We consider the case when $\widehat{\xi}$ is the concatenation of bang and singular arcs.

DEFINITION 1.2. *Given an admissible trajectory ξ and a time interval $(t_1, t_2) \subset [0,T]$, we say that $\xi|_{(t_1,t_2)}$ is*

- *a* bang arc *of ξ if $\dot{\xi}(t)$ is the same vertex of $\mathcal{X}(\xi(t))$ for any $t \in (t_1, t_2)$,*
- *a* singular arc *of ξ if $\dot{\xi}(t)$ is in the relative interior of a face of $\mathcal{X}(\xi(t))$ for any $t \in (t_1,t_2)$. A singular arc is called* totally singular *if the dimension of the face is maximal, otherwise it is called* partially singular.

Notice that in the single input case, $m = 2$, singular means totally singular. This case was considered in [21] where a proof was only sketched.

Here we assume there exist times $\widehat{\tau}_1, \widehat{\tau}_2$ such that $0 < \widehat{\tau}_1 < \widehat{\tau}_2 < T$, vector fields h_1, h_2, $h_3 \in \{X_1, \ldots X_m\}$, (where h_1 and h_3 might be the same vector field) and a measurable function $\widehat{\upsilon} \in L^\infty([\widehat{\tau}_2, T], (0,1))$ such that $\widehat{\xi}$ is the absolutely continuous solution to the following Cauchy problem

$$\begin{aligned}
&\dot{\xi}(t) = h_1(\xi(t)) && t \in [0, \widehat{\tau}_1),\\
&\dot{\xi}(t) = h_2(\xi(t)) && t \in [\widehat{\tau}_1, \widehat{\tau}_2),\\
&\dot{\xi}(t) = \widehat{\upsilon}(t) h_3(\xi(t)) + (1 - \widehat{\upsilon}(t))\, h_2(\xi(t)) && \text{a.e. } t \in [\widehat{\tau}_2, T],\\
&\xi(0) = x_0.
\end{aligned} \tag{2}$$

Thus, if $m = 2$, then $(\widehat{\tau}_2, T)$ is a totally singular arc, else it is a partially singular one. Denoting by $f_{\rm d} := h_3 - h_2$ we can write the dynamics on the singular arc as

$$\dot{\xi}(t) = h_2(\xi(t)) + \widehat{\upsilon}(t) f_{\rm d}(\xi(t)), \qquad t \in (\widehat{\tau}_2, T).$$

We also define the time-dependent reference vector field $\widehat{f}_t$ as

$$\widehat{f}_t := \begin{cases} h_1 & t \in [0, \widehat{\tau}_1),\\ h_2 & t \in (\widehat{\tau}_1, \widehat{\tau}_2),\\ h_2 + \widehat{\upsilon}(t) f_{\rm d} & \text{a.e. } t \in (\widehat{\tau}_2, T]. \end{cases} \tag{3}$$

REMARK 1.1. *In this paper we consider a case study for our Hamiltonian approach, i.e. when the final point is not constrained and the initial one is fixed. Indeed in this case the second order conditions give the possibility of constructing a field of non intersecting* almost extremals *and this is sufficient to obtain the result. The extension to a problem with constrained final point requires adding a penalty term by taking advantage of a classical result on quadratic forms due to Hestenes, see [8].*

In a future paper, [17], we shall extend the result to the case when the end points are constrained to smooth sub-manifolds of $\mathbb{R}^n$ and the cost depends on both the end points.

Since the main necessary condition for strong local optimality, Pontryagin Maximum Principle (PMP), is naturally set in the cotangent bundle $(\mathbb{R}^n)^* \times \mathbb{R}^n$, we give our sufficient conditions in such framework. We then use a Hamiltonian approach and its connection with the second order conditions.

The main idea is to use the symplectic properties of the cotangent bundle to compare the costs of neighboring admissible trajectories by lifting them to such bundle. To do so we define a suitable Hamiltonian flow $\mathcal{H}$, emanating from a horizontal Lagrangian sub-manifold Λ, $\mathcal{H}\colon (t,\ell) \in [0,T] \times \Lambda \mapsto \mathcal{H}_t(\ell) \in T^*\mathbb{R}^n$. Since the final point is free, we consider flows backwards in time, with T as a starting time, Section 4.1. The existence of this flow will be ensured by the regularity assumptions on the extremal $\widehat{\lambda}$ given by PMP, see Assumptions 1–5.

From a Hamiltonian point of view the sufficient conditions sum up to proving the existence of a tubular neighborhood of $[0,T] \times \{\widehat{\lambda}(T)\}$ in $[0,T] \times \Lambda$ where the map $\mathrm{id} \times \pi\mathcal{H}$ is locally invertible, see Theorem 4.2. Thanks to the compactness of the time interval $[0,T]$ it suffices to prove that $\pi\mathcal{H}_t$ is invertible for any t, see Theorem 4.3. The connection with a suitable second order approximation (2^{nd} *variation*) is obtained as shown in the following lines.

1. The 2^{nd} variation J''_{ext} is in the form of a coordinate-free linear-quadratic (LQ) problem on the interval $[\widehat{\tau}_2, T]$ and it is obtained applying an intrinsic version of Goh transformation. Indeed we can obtain our sufficient conditions either when J''_{ext} is coercive or when $L_{f_{\mathrm{d}}} c \equiv 0$ (a fact which prevents the coercivity), provided a suitable restriction of J''_{ext} is coercive, see Section 3.2.
2. We show that the derivative of $\mathcal{H}_t$ along $\widehat{\lambda}$ is, up to an isomorphism, the linear Hamiltonian flow associated to the LQ problem, see Section 5.2.
3. From the coercivity of the 2^{nd} variation we deduce that $\pi\mathcal{H}_t$ is locally invertible (Sections 5.3 and 5.4), so that we can compare the costs of neighboring admissible trajectories by lifting them to the cotangent bundle. In our case proving the local invertibility is equivalent to requiring the invertibility of $\pi\mathcal{H}_t$ for any $t \in [\widehat{\tau}_2, T]$ and the sufficient conditions for the optimality of a bang-bang trajectory of a suitable Mayer problem on $[0, \widehat{\tau}_2]$, see Remark 4.3.

REMARK 1.2. *The Hamiltonian approach allows to prove strong local optimality of the reference trajectory in the case of a partially singular arc, by giving regularity conditions on the reference control, while in the second order conditions only the singular component of the control is considered.*

We also point out out that our result applies also to the case $L_{f_{\mathrm{d}}} c \equiv 0$, a case which, up to the authors knowledge, has not been considered so far. In Section 6.1 we provide an example where this condition holds and to which our theory applies.

REMARK 1.3. *In [19] we considered the bang-singular-bang case for the minimum time problem in a single input control system with fixed end points. For technical reasons the construction of almost-extremals provided in [19] works for Mayer problems only if the singular arc is the first or the last one. The technique can be applied to the concatenation of an arbitrary number of bang arcs and a singular one, provided the singular arc is the initial or the final one. The possibility of modifying the technique in order to consider any concatenation of bang and singular arcs is currently being studied.*

In the case of bang-bang extremals for either a Mayer or a Bolza problem, Hamiltonian methods have been successfully exploited in [3, 12, 14, 15], while bang-bang extremals in the minimum time problem have been studied in [18, 16].

Bang-bang extremals are extensively studied in the literature, also with other methods, see for example [9, 11, 5] and the references therein.

Hamiltonian methods have also been applied to singular extremals, see [22, 6] and to concatenations of bang and singular arcs, see [19] and the references therein.

The literature is rich of results that involve some localization of the control and with different approaches, see e.g. [4] and the references therein.

We should also like to mention that Hamiltonian methods can also be successfully employed to obtain sufficient conditions to structure stability of minimizers, see e.g. [7, 20, 13].

2. Preliminaries.

2.1. Notation. We start by recalling some basic facts and by introducing some specific notations. We identify any bi-linear form Q on a vector space W with a linear form $Q\colon W \to W^*$, we write $Q(v,w) = \langle Qv\,,\, w\rangle$, and we denote the associate quadratic form as $Q(v,v) = Q[v]^2$.

In this paper we use notation from differential geometry and some basic element of the theory of symplectic manifolds referred to the trivial cotangent bundle $T^*\mathbb{R}^n = (\mathbb{R}^n)^* \times \mathbb{R}^n$, see for example [1]. We take advantage of the intrinsic notation from differential geometry as it is more compact and clear. In particular we distinguish between points in $\mathbb{R}^n$, usually denoted as x and tangent vectors to $\mathbb{R}^n$, denoted as δx.

Given a C^1 vector field f on $\mathbb{R}^n$, we denote as $\exp tf(x)$ the flow at time t emanating from a point x at time 0, i.e. $\exp tf(x)$ is the solution to

$$\dot\xi(t) = f(\xi(t)), \quad \xi(0) = x.$$

If g is another C^1 vector field, then the Lie bracket between f and g is denoted as $[f,g]$, i.e. $[f,g](x) := \mathrm{D}g(x)f(x) - \mathrm{D}f(x)g(x)$.

If $\alpha\colon \mathbb{R}^n \to \mathbb{R}$ is a C^2 function, $\mathrm{d}\alpha$ is its differential, while $\mathrm{D}^2\alpha$ is the second derivative of α. Moreover $L_f\alpha\,(x) := \langle \mathrm{d}\alpha(x)\,,\, f(x)\rangle$ is the Lie derivative of α with respect to the vector field f at the point x.

Finally, if G is a C^1 map from a manifold X in a manifold Y, its tangent map at a point $x \in X$ is denoted as T_xG, or simply as G_* if the point x is clear from the context. In particular, if $\alpha\colon \mathbb{R}^n \to \mathbb{R}$ is a C^2 function and $\delta x \in \mathbb{R}^n$, we denote by the symbol $\mathrm{d}\alpha_*\delta x$ the couple $\big(\mathrm{D}^2\alpha(x)(\delta x, \cdot), \delta x\big) \in (\mathbb{R}^n)^* \times \mathbb{R}^n$ whenever the point x is clear from the context.

We denote by $\pi\colon \ell = (p,x) \in T^*\mathbb{R}^n \mapsto x \in \mathbb{R}^n$ the projection on the base space. The symbol $\boldsymbol{s}$ denotes the canonical Liouville one–form on $T^*\mathbb{R}^n$: $\boldsymbol{s} := \sum_{i=1}^n p^i\,\mathrm{d}x_i$. The associated canonical symplectic two–form $\boldsymbol{\sigma} = \mathrm{d}\boldsymbol{s} = \sum_{i=1}^n \mathrm{d}p^i \wedge \mathrm{d}x_i$ allows one to associate to any, possibly time-dependent, smooth Hamiltonian $H_t\colon T^*\mathbb{R}^n \to \mathbb{R}$, a Hamiltonian vector field $\overrightarrow{H_t}$, by

$$\boldsymbol{\sigma}\left(V, \overrightarrow{H_t}(\ell)\right) = \langle \mathrm{d}H_t(\ell)\,,\, V\rangle, \quad \forall\, V \in T_\ell T^*\mathbb{R}^n, \tag{4}$$

i.e. $\overrightarrow{H_t}(\ell) = \left(-\dfrac{\partial H_t}{\partial x}(\ell), \dfrac{\partial H_t}{\partial p}(\ell)\right), \quad \forall \ell = (p,x) \in T^*\mathbb{R}^n.$

We keep this notation throughout the paper, namely the overhead arrow denotes the vector field associated to a Hamiltonian, moreover the script letter denotes its flow from time T, unless otherwise stated.

Finally we recall that any vector field f on $\mathbb{R}^n$ defines, by lifting to the cotangent bundle, a Hamiltonian

$$F\colon \ell = (p, x) \in T^*\mathbb{R}^n \mapsto \langle p\,,\, f(x)\rangle \in \mathbb{R}.$$

In particular we denote by H_1, H_2, H_3 the Hamiltonians associated with h_1, h_2, h_3, respectively and by $H_{i_1 i_2 \dots i_k}$, $i_1, \dots, i_k \in \{1,\, 2\; 3\}$, the Hamiltonian associated with the vector field $h_{i_1 i_2 \dots i_k} := [h_{i_1}, [\dots [h_{i_{k-1}}, h_{i_k}] \dots]$.

The flow from time T of the reference vector field $\widehat{f}_t$ defined in (3), is a local diffeomorphism defined in a neighborhood of the point $\widehat{x}_T := \widehat{\xi}(T)$. For each $t \in [0, T]$ we denote such flow as $\widehat{S}_t \colon \mathbb{R}^n \to \mathbb{R}^n$, while

$$\widehat{F}_t = \begin{cases} H_1 & \text{if } \; t \in [0, \widehat{\tau}_1), \\ H_2 & \text{if } \; t \in (\widehat{\tau}_1, \widehat{\tau}_2), \\ H_2 + \widehat{v}(t) H_3 & \text{if } \; t \in (\widehat{\tau}_2, T], \end{cases}$$

denotes the time–dependent reference Hamiltonian and $\mathcal{F}_t$ denotes its flow.

2.2. The necessary conditions. We start by stating the main necessary condition of optimality, i.e. Pontryagin Maximum Principle (PMP). Since there is no constraint on the final point, then PMP must hold in its normal form:

Assumption 1 (Normal PMP). *There exists an absolutely continuous mapping* $\widehat{\mu}\colon [0, T] \to (\mathbb{R}^n)^*$ *such that a.e.* $t \in [0, T]$

$$\dot{\widehat{\mu}}(t) = -\widehat{\mu}(t)\, \mathrm{D}\widehat{f}_t(\widehat{\xi}(t)), \qquad \widehat{\mu}(T) = -\,\mathrm{dc}(\widehat{x}_T),$$
$$\widehat{F}_t(\widehat{\mu}(t), \widehat{\xi}(t)) = \max\left\{ \langle \widehat{\mu}(t)\,,\, X\rangle \colon X \in \mathcal{X}(\widehat{\xi}(t)) \right\}.$$

$\widehat{\mu}(t)$ is called adjoint covector and the trajectory $\widehat{\xi}$ of the system is called a *state extremal* of problem (1) while the couple $\widehat{\lambda}(t) := \big(\widehat{\mu}(t), \widehat{\xi}(t)\big)$ is called an *extremal* of problem (1). We use the following notation for the end points and for the switching points of $\widehat{\lambda}(t) \in T^*\mathbb{R}^n$:

$$\widehat{\ell}_f := \widehat{\lambda}(T), \quad \widehat{\ell}_2 := \widehat{\lambda}(\widehat{\tau}_2) = \widehat{\mathcal{F}}_{\widehat{\tau}_2}(\widehat{\ell}_f), \quad \widehat{\ell}_1 := \widehat{\lambda}(\widehat{\tau}_1) = \widehat{\mathcal{F}}_{\widehat{\tau}_1}(\widehat{\ell}_f), \quad \widehat{\ell}_0 := \widehat{\lambda}(0) = \widehat{\mathcal{F}}_0(\widehat{\ell}_f).$$

We call $\widehat{\lambda}\Big|_{[0,\widehat{\tau}_1)}$ and $\widehat{\lambda}\Big|_{(\widehat{\tau}_1,\widehat{\tau}_2)}$ bang arcs, while $\widehat{\lambda}\Big|_{(\widehat{\tau}_2,T]}$ is a singular arc.

Thanks to the structure of the reference trajectory, PMP gives the following necessary conditions:

1. On the first bang arc, $t \in [0, \widehat{\tau}_1]$, we get $H_1(\widehat{\lambda}(t)) \geq \langle \widehat{\mu}(t)\,,\, X\rangle$, $\forall X \in \mathcal{X}(\widehat{\xi}(t))$.
2. On the second bang arc, $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$, we get $H_2(\widehat{\lambda}(t)) \geq \langle \widehat{\mu}(t)\,,\, X\rangle$, $\forall X \in \mathcal{X}(\widehat{\xi}(t))$, in particular $H_1(\widehat{\ell}_2) = H_2(\widehat{\ell}_2)$.
3. On the singular arc, $t \in [\widehat{\tau}_2, T]$, we get

$$(H_2 + \widehat{v}(t) F_{\mathrm{d}})\,(\widehat{\lambda}(t)) \geq \langle \widehat{\mu}(t)\,,\, X\rangle, \quad \forall X \in \mathcal{X}(\widehat{\xi}(t)),$$

which implies $F_{\mathrm{d}}(\widehat{\lambda}(t)) \equiv 0$ and, by differentiation,

$$(5)\quad \frac{\mathrm{d}}{\mathrm{d}t} F_{\mathrm{d}}(\widehat{\lambda}(t)) = H_{23}(\widehat{\lambda}(t)) = 0, \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2} F_{\mathrm{d}}(\widehat{\lambda}(t)) = -H_{232}(\widehat{\lambda}(t)) + \widehat{v}(t)\mathbb{L}(\widehat{\lambda}(t)) = 0,$$

$$\mathbb{L}(\ell) := (H_{323} + H_{232})(\ell) = \langle p\,,\, [f_{\mathrm{d}}, [h_2, f_{\mathrm{d}}]]\,(x)\rangle, \quad \ell = (p, x) \in T^*\mathbb{R}^n.$$

4. At the first switching time we get $H_{12}(\widehat{\ell}_1) = \dfrac{\mathrm{d}}{\mathrm{d}t}\,(H_2 - H_1)\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}_1} \geq 0$, see [3].
5. At the second switching time we get $H_{232}(\widehat{\ell}_2) = -\dfrac{\mathrm{d}^2}{\mathrm{d}t^2} F_{\mathrm{d}}\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}_2^-} \geq 0$, see [19].

Moreover, other necessary conditions are known to hold along singular arcs, namely the Goh condition (which in this case is automatically satisfied) and the *generalized Legendre condition* (GLC), see e.g. [1], applied to the sub-problem where the controlled vector field is constrained on the edge whose extrema are h_2 and h_3

$$\mathbb{L}(\widehat{\lambda}(t)) = \langle\widehat{\mu}(t)\,,\,(h_{323}+h_{232})\,(\widehat{\xi}(t))\rangle \geq 0 \quad \forall t\in[\widehat{\tau}_2, T]. \tag{6}$$

We recall that the generalized Legendre condition (GLC) takes this form because we deal with Pontryagin Maximum Principle. If one considers the minimum principle, as in [4] and in [25], then GLC is given by the reverse inequality.

3. Assumptions and main result.

3.1. Regularity conditions. We now state regularity conditions by requiring strict inequalities to hold whenever necessary conditions yield mild inequalities.

Assumption 2 (Regularity along the bang arcs).

$$H_1(\widehat{\lambda}(t)) > \langle\widehat{\mu}(t)\,,\,X\rangle, \qquad \forall X\in\mathcal{X}(\widehat{\xi}(t))\setminus\{h_1(\widehat{\xi}(t))\}, \quad \forall t\in[0,\widehat{\tau}_1),$$
$$H_2(\widehat{\lambda}(t)) > \langle\widehat{\mu}(t)\,,\,X\rangle, \qquad \forall X\in\mathcal{X}(\widehat{\xi}(t))\setminus\{h_2(\widehat{\xi}(t))\}, \quad \forall t\in(\widehat{\tau}_1,\widehat{\tau}_2),$$

namely we require that the reference control is the only maximizing control along each bang arc.

Assumption 3 (Regularity along the singular arc). *For any* $a\in[0,1]$ *and any* $t\in[\widehat{\tau}_2, T]$

$$H_2(\widehat{\lambda}(t)) + \widehat{v}(t) f_{\mathrm{d}}(\widehat{\lambda}(t)) > \langle\widehat{\mu}(t)\,,\,X\rangle, \quad \forall X\in\mathcal{X}(\widehat{\xi}(t)),\ X(\widehat{\xi}(t)) \neq (h_2 + a f_{\mathrm{d}})\,(\widehat{\xi}(t)),$$

i.e. we require that the set of maximizers along the singular arc is only the edge defined by h_2 and h_3.

Assumption 4 (Regularity at the switching points).

$$H_{12}(\widehat{\ell}_1) > 0, \qquad H_{232}(\widehat{\ell}_2) > 0. \tag{7}$$

Assumption 5 (Strong generalised Legendre condition).

$$R(t) := \mathbb{L}(\widehat{\lambda}(t)) = \langle\widehat{\mu}(t)\,,\,[f_{\mathrm{d}},[h_2,f_{\mathrm{d}}]]\,(\widehat{\xi}(t))\rangle > 0 \qquad t\in[\widehat{\tau}_2, T]. \tag{SGLC}$$

Thanks to (SGLC) from (5) we can recover the control along the singular arc:

$$\widehat{v}(t) = \frac{H_{232}}{\mathbb{L}}(\widehat{\lambda}(t)) \qquad \forall t\in(\widehat{\tau}_2, T],$$

so that, by recurrence, one can easily prove that $\widehat{v}\in C^\infty([\widehat{\tau}_2,T],(0,1))$.

The condition $\widehat{v}(t)\in(0,1)$ reads

$$H_{232}(\widehat{\lambda}(t)) > 0, \quad H_{323}(\widehat{\lambda}(t)) > 0 \qquad \forall t\in[\widehat{\tau}_2, T]. \tag{8}$$

Thus the reference vector field is discontinuous at $\widehat{\xi}(\widehat{\tau}_2)$ if and only if $\widehat{v}(\widehat{\tau}_2^+) > 0$, i.e. if and only if the regularity condition at $\widehat{\tau}_2$, equation (7), holds.

3.2. The extended second variation. The second order conditions will be derived studying a sub-problem of the given one. Namely we consider the reference vector field $\widehat{f}_t$ and allow only for perturbations of $\widehat{\upsilon}$ on the singular interval $(\widehat{\tau}_2, T)$ and for perturbations of the switching time $\widehat{\tau}_1$. Following the ideas of [19], we represent the perturbations of the first switching time $\widehat{\tau}_1$ by a new positive control υ_0 which is a reparametrization of time. The sub-problem can be written as

$$\text{(9a)} \qquad \text{Minimize } c(\xi(T)) \;\; \text{subject to}$$

$$\text{(9b)} \qquad \dot{\xi}(t) = \begin{cases} \upsilon_0(t) h_1(\xi(t)) & t \in (0, \widehat{\tau}_1), \\ \upsilon_0(t) h_2(\xi(t)) & t \in (\widehat{\tau}_1, \widehat{\tau}_2), \\ h_2(\xi(t)) + \upsilon(t) f_{\mathrm{d}}(\xi(t)) & t \in (\widehat{\tau}_2, T), \end{cases}$$

$$\text{(9c)} \qquad \xi(0) = x_0, \quad \upsilon_0(t) > 0, \; \int_0^{\widehat{\tau}_2} \upsilon_0(t)\, \mathrm{d}t = \widehat{\tau}_2, \quad \upsilon(t) \in (0, 1).$$

Problem (9) gives rise to a linear quadratic problem on the singular arc $[\widehat{\tau}_2, T]$ where the variation of the first switching time τ_1 produces a cost at time $\widehat{\tau}_2$. Set

$$\text{(10)} \qquad g_t := \widehat{S}_{t*}^{-1} f_{\mathrm{d}} \circ \widehat{S}_t \,, \; t \in [\widehat{\tau}_2, T], \quad k_i := \widehat{S}_{\widehat{\tau}_1 *}^{-1} h_i \circ \widehat{S}_{\widehat{\tau}_1}, \; i = 1, 2, \quad k := k_1 - k_2,$$

i.e. g_t is the push-forward of f_{d} from time $t \in [\widehat{\tau}_2, T]$ to time T while k_i is the push-forwards of h_i, $i = 1, 2$, from the first switching time $\widehat{\tau}_1$ to T. With this notation the second variation of (9) as defined in [2] and written in terms of the push-forwards to time T instead of pullbacks to time 0, is given by

$$\text{(11)} \qquad J''[(\delta x, \delta\upsilon_0, \delta\upsilon)]^2 = \int_{\widehat{\tau}_2}^T \delta\upsilon(t) L_{\delta\eta(t)} L_{g_t} c\,(\widehat{x}_T)\; dt + \frac{\varepsilon_0^2}{2} \left(L_k^2 c\,(\widehat{x}_T) + H_{12}(\widehat{\ell}_1) \right)$$

subject to

$$\dot{\delta\eta}(t) = \delta\upsilon(t) g_t(\widehat{x}_T), \qquad \delta\eta(T) = \delta x \in \mathbb{R}^n, \qquad \delta\eta(\widehat{\tau}_2) = \varepsilon_0 k(\widehat{x}_T),$$
$$\varepsilon_0 = \int_0^{\widehat{\tau}_1} \delta\upsilon_0(t)\, dt = -\int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta\upsilon_0(t)\, dt.$$

We then extend the second variation to a new quadratic form called *extended second variation*. Following the same lines as in the appendix of [19] and setting

$$w(t) := \int_t^{\widehat{\tau}_2} \delta\upsilon(s)\, \mathrm{d}s, \qquad \varepsilon_1 := w(T),$$

the extended second variation of (9) is given by the following LQ problem on $[\widehat{\tau}_2, T]$.

$$\text{(12)} \qquad \begin{aligned} &J''_{\mathrm{ext}}[(\delta x, \varepsilon_0, \varepsilon_1, w)]^2 = -\varepsilon_1 L_{\delta x} L_{f_{\mathrm{d}}} c\,(\widehat{x}_T) - \frac{\varepsilon_1^2}{2} L_{f_{\mathrm{d}}}^2 c\,(\widehat{x}_T) + \\ &+ \frac{\varepsilon_0^2}{2} \left(L_k^2 c\,(\widehat{x}_T) + H_{12}(\widehat{\ell}_1) \right) + \frac{1}{2} \int_{\widehat{\tau}_2}^T \left(2\, w(t) L_{\zeta(t)} L_{\dot{g}_t} c\,(\widehat{x}_T) + w(t)^2 R(t) \right) \mathrm{d}t \end{aligned}$$

subject to

$$\text{(13)} \qquad \dot{\zeta}(t) = w(t) \dot{g}_t(\widehat{x}_T), \qquad \zeta(\widehat{\tau}_2) = \varepsilon_0\, k(\widehat{x}_T), \qquad \zeta(T) = \delta x + \varepsilon_1 f_{\mathrm{d}}(\widehat{x}_T).$$

This means that we consider the quadratic form $J''_{\rm ext}$ defined by (12) on the linear space, called *space of admissible variations*, given by

$$\mathcal{W}_{\rm ext} := \{(\delta x, \varepsilon_0, \varepsilon_1, w) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times L^2([\widehat{\tau}_2, T]) \colon (13) \text{ admits a solution}\}.$$

Notice that

$$\dot{g}_t = \widehat{S}_{t*}^{-1} h_{23} \circ \widehat{S}_t, \quad t \in [\widehat{\tau}_2, T]. \tag{14}$$

We consider two cases:

• $L_{f_{\rm d}} c \equiv 0$. $J''_{\rm ext}$ cannot be coercive and we only require its coercivity on the subspace of $\mathcal{W}_{\rm ext}$ given by $\varepsilon_1 = 0$. Notice that $f_{\rm d}(\widehat{x}_T) \neq 0$. Indeed, if $f_{\rm d}(\widehat{x}_T) = 0$, then $\mathbb{L}(\widehat{\ell}_f) = \langle \widehat{\mu}(T), [f_{\rm d}, h_{23}](\widehat{x}_T)\rangle = \langle {\rm d}c, {\rm D}f_{\rm d} h_{23}\rangle(\widehat{x}_T) = L_{h_{23}} L_{f_{\rm d}} c(\widehat{x}_T) = 0$, a contradiction to (SGLC).

• $L_{f_{\rm d}} c$ is not identically zero. Choosing the variation $\delta e = (-f_{\rm d}(\widehat{x}_T), 0, 1, 0)$ in (12) we get $J''_{\rm ext}[\delta e]^2 = L^2_{f_{\rm d}} c(\widehat{x}_T)$ and we require $L^2_{f_{\rm d}} c(\widehat{x}_T) > 0$.

In this case the set, locally defined near $\widehat{x}_T$ in $\mathbb{R}^n$,

$$\widetilde{M} := \{x \in \mathbb{R}^n \colon L_{f_{\rm d}} c(x) = 0\},$$

is a hyper-surface whose tangent space at $\widehat{x}_T$ is

$$T_{\widehat{x}_T} \widetilde{M} = \{\delta z \in \mathbb{R}^n \colon L_{\delta z} L_{f_{\rm d}} c(\widehat{x}_T) = 0\}.$$

For $x = \exp(r f_{\rm d})(z)$, $z \in \widetilde{M}$ set $\widetilde{c}(x) := c(z)$, i.e. we extend $c|_{\widetilde{M}}$ as a constant function along the integral lines of $f_{\rm d}$. In a sufficiently small neighborhood $\mathcal{O}$ of $\widehat{x}_T$, the function $\widetilde{c} \colon \mathcal{O} \to \mathbb{R}$ is smooth and it enjoys the following properties

$$\begin{aligned} &\widetilde{c}(\widehat{x}_T) = c(\widehat{x}_T), \quad {\rm d}\widetilde{c}(\widehat{x}_T) = {\rm d}c(\widehat{x}_T), \\ &\widetilde{c}(x) \le c(x), \qquad L_{f_{\rm d}} \widetilde{c}(x) = 0 \quad \forall x \in \mathcal{O}. \end{aligned} \tag{15}$$

Following [22] it can be shown that the coercivity of (12) on $\mathcal{W}_{\rm ext}$ is equivalent to $L^2_{f_{\rm d}} c(\widehat{x}_T) > 0$ plus the coercivity of

$$\begin{aligned} \widetilde{J}[(\delta x, \varepsilon_0, w)]^2 = &\frac{\varepsilon_0^2}{2} \left(L_k^2 \widetilde{c}(\widehat{x}_T) + H_{12}(\widehat{\ell}_1) \right) + \\ &+ \frac{1}{2} \int_{\widehat{\tau}_2}^T \left(2\, w(t) L_{\zeta(t)} L_{\dot{g}_t} \widetilde{c}(\widehat{x}_T) + R(t) w(t)^2 \right) {\rm d}t \end{aligned} \tag{16}$$

subject to

$$\dot{\zeta}(t) = w(t) \dot{g}_t(\widehat{x}_T), \quad \zeta(\widehat{\tau}_2) = \varepsilon_0\, k(\widehat{x}_T), \quad \zeta(T) = \delta x \in \mathbb{R}^n. \tag{17}$$

This is exactly the same formula we obtain in the case $L_{f_{\rm d}} c \equiv 0$ setting $\widetilde{c} := c$.

We can now state our final assumption, concerning the second variation $\widetilde{J}$.

Assumption 6. *We assume the following conditions hold:*

a) The quadratic form $\widetilde{J}$, (16), is coercive on

$$\widetilde{\mathcal{W}} := \{(\delta x, \varepsilon_0, w) \in \mathbb{R}^n \times \mathbb{R} \times L^2([\widehat{\tau}_2, T], \mathbb{R}) \colon (17) \textit{ admits a solution}\}.$$

b) Either $L^2_{f_{\rm d}} c(\widehat{x}_T) > 0$ or $L_{f_{\rm d}} c \equiv 0$ in a neighborhood $\mathcal{O}$ of $\widehat{x}_T$ in $\mathbb{R}^n$.

3.3. The main result. We can now state the main result of this paper

THEOREM 3.1. *Let $\widehat{\xi}$ be the admissible trajectory defined in* (2). *Assume that $\widehat{\xi}$ is a state extremal (Assumption 1) satisfying the regularity Assumptions 2–5. If Assumption 6 is satisfied, then $\widehat{\xi}$ is a strict strong local optimal trajectory of* (1).

Indeed, in Section 5 we prove that Assumptions 1-5 plus a) of Assumption 6 imply that $\widehat{\xi}$ is a strict strong locally optimal trajectory for the cost $\widetilde{c}(\xi(T))$.

This concludes the proof in the case $L_{f_{\rm d}}c \equiv 0$. When $L^2_{f_{\rm d}}c\,(\widehat{x}_T) > 0$, the claim is proved by (15).

4. Hamiltonian approach.

The first step in applying the Hamiltonian approach described in the Introduction, is the construction of an over-maximized Hamiltonian flow. Indeed the presence of a singular arc prevents us from using the maximized Hamiltonian (see [19]) which can be used in the classical case, i.e. when it is C^2, see [1]. The over-maximized Hamiltonian was introduced in [22] and then used in [19, 20, 6] . In [24] the authors give a systematic extension of the classical techniques to the case of an over-maximized Hamiltonian whose flow is only Lipschitz continuous.

4.1. The over-maximized flow. In this section we describe how the regularity conditions allow to define in a neighborhood $\mathcal{U}$ of the graph of $\widehat{\lambda}$ in $[0,T] \times T^*\mathbb{R}^n$, a time-dependent Hamiltonian function $H : \mathcal{U} \to \mathbb{R}$ whose flow satisfies the assumptions stated in [24]. We consider the flow of the over-maximized Hamiltonian emanating from the following Lagrangian manifold:

$$\Lambda := \{(\mathrm{d}(-\widetilde{c})(x), x) : x \in \mathcal{O}\}\,. \tag{18}$$

In our case the assumptions of [24] read as follows.

1. The flow $(t,\ell) \in [0,T] \times \Lambda \mapsto \mathcal{H}_t(\ell) := (\mu_t(\ell), \xi_t(\ell)) \in T^*\mathbb{R}^n$ is Lipschitz continuous.
2. The function $\Phi\colon (t,\ell) \in [0,T] \times \Lambda \mapsto \langle \mu_t(\ell)\,,\, \pi\overrightarrow{H_t} \circ \mathcal{H}_t(\ell)\rangle - H_t \circ \mathcal{H}_t(\ell) \in \mathbb{R}$ is Lipschitz-Caratheodory i.e.
   - For almost every $t \in [0,T]$ the map $\ell \in \Lambda \mapsto \Phi(t,\ell) \in \mathbb{R}$ is locally Lipschitz.
   - For each $\ell \in \Lambda$ the map $t \in [0,T] \mapsto \Phi(t,\ell) \in \mathbb{R}$ is bounded measurable.
   - For any compact set $K \subset \Lambda$ there is an essentially bounded measurable function $m\colon [0,T] \to \mathbb{R}$ such that
   $$\|\Phi(t,\ell_1) - \Phi(t,\ell_2)\| \le m(t)\,\|\ell_1 - \ell_2\|\,, \qquad \forall \ell_1,\ell_2 \in \Lambda.$$
3. The following over-maximality conditions hold:
   (a) $H_t \circ \mathcal{H}_t(\ell) \ge H^{\max} \circ \mathcal{H}_t(\ell)$, for any $(t,\ell) \in [0,T] \times \Lambda$,
   (b) $H_t \circ \widehat{\lambda}(t) = \widehat{F}_t \circ \widehat{\lambda}(t) = H^{\max} \circ \widehat{\lambda}(t)$, for a.e. $t \in [0,T]$,
   (c) $\overrightarrow{H_t} \circ \widehat{\lambda}(t) = \overrightarrow{\widehat{F}_t} \circ \widehat{\lambda}(t)$, for a.e. $t \in [0,T]$.

The coercivity of the second variation will then guarantee the invertibility of the projected over-maximized flow of such Hamiltonian.

In order to describe H_t, $t \in [\widehat{\tau}_2, T]$ notice that the SGLC condition (Assumption 5) implies that there exists a neighborhood $\mathcal{O}_{\rm s}$ of the range of the singular arc $\widehat{\lambda}([\widehat{\tau}_2, T])$ in $T^*\mathbb{R}^n$ such that the sets

$$\Sigma := \{\ell \in \mathcal{O}_{\rm s} \colon F_{\rm d}(\ell) = 0\} = \{\ell \in \mathcal{O}_{\rm s} \colon H_2(\ell) = H_3(\ell)\}\,,$$
$$\mathcal{S} := \{\ell \in \Sigma \colon H_{23}(\ell) = 0\} = \{\ell \in \mathcal{O}_{\rm s} \colon H_2(\ell) = H_3(\ell),\ H_{23}(\ell) = 0\}$$

are smooth simply connected manifolds of codimension 1 and 2, respectively. More precisely $\overrightarrow{H_{23}}$ is transverse to Σ in $\mathcal{O}_{\rm s}$, while $\overrightarrow{F_{\rm d}}$ is tangent to Σ and transverse to $\mathcal{S}$ in Σ, see [19]. Moreover we can define a smooth function $u_{\rm s} \colon \mathcal{O}_{\rm s} \to \mathbb{R}$ and the feedback Hamiltonian as follows:

$$u_{\rm s}(\ell) = \frac{H_{232}}{\mathbb{L}}(\ell), \qquad H^{\mathcal{S}}(\ell) = H_2(\ell) + u_{\rm s}(\ell) F_{\rm d}(\ell).$$

REMARK 4.1. *On the singular interval* $[\widehat{\tau}_2, T]$, $\widehat{v}(t) = u_{\rm s}(\widehat{\lambda}(t))$ *and* $\widehat{\lambda}$ *is the solution to the Cauchy problem*

$$\dot{\lambda}(t) = \overrightarrow{H^{\mathcal{S}}}(\lambda(t)), \quad \lambda(\widehat{\tau}_2) = \widehat{\ell}_2. \tag{19}$$

In [19] the authors prove that possibly restricting $\mathcal{O}_{\rm s}$, the following implicit function problem has a solution $\theta \colon \mathcal{O}_{\rm s} \to \mathbb{R}$:

$$\begin{cases} H_{23} \circ \exp(\theta(\ell)\overrightarrow{F_{\rm d}})(\ell) = 0, \\ \theta(\ell) = 0 \quad \text{if } H_{23}(\ell) = 0, \end{cases}$$

and

$$\langle {\rm d}\theta(\ell_{\mathcal{S}}) \,,\, \delta\ell \rangle = \frac{-\,\boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{H_{23}}(\ell_{\mathcal{S}})\right)}{\mathbb{L}(\ell_{\mathcal{S}})}, \qquad \forall \ell_{\mathcal{S}} \colon H_{23}(\ell_{\mathcal{S}}) = 0.$$

The over-maximized Hamiltonian is defined starting from

$$\widetilde{H}_2(\ell) := H_2 \circ \exp(\theta(\ell)\overrightarrow{F_{\rm d}})(\ell).$$

LEMMA 4.1. *Possibly restricting* $\mathcal{O}_{\rm s}$ *the following properties hold*

1. $\widetilde{H}_2(\ell) \geq H_2(\ell)$ *for any* $\ell \in \Sigma$ *and equality holds if and only if* $\ell \in \mathcal{S}$.
2. *For any* $\ell_{\mathcal{S}} \in \mathcal{S}$ *and* $\delta\ell \in T_{\ell_{\mathcal{S}}}\Sigma$

$$ {\rm d}\Big(\widetilde{H}_2 - H_2\Big)(\ell_{\mathcal{S}}) = 0, \qquad {\rm D}^2\left(\widetilde{H}_2 - H_2\right)(\ell_{\mathcal{S}})[\delta\ell]^2 = \frac{\left(\boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{H_{23}}(\ell_{\mathcal{S}})\right)\right)^2}{\mathbb{L}(\ell_{\mathcal{S}})}. \tag{20}$$

3. $\overrightarrow{\widetilde{H}_2}$ *is tangent to* Σ *and, setting,*

$$H_t(\ell) := \widetilde{H}_2(\ell) + \widehat{v}(t) F_{\rm d}(\ell), \quad \forall (t, \ell) \in [\widehat{\tau}_2, T] \times \mathcal{O}_{\rm s} \tag{21}$$

we easily get that $\overrightarrow{H_t}$ *is tangent to* Σ.

4. $\widehat{\lambda}\Big|_{[\widehat{\tau}_2, T]}$ *solves the Cauchy problem* $\dot{\lambda}(t) = \overrightarrow{H_t}(\lambda(t)), \quad \lambda(T) = \widehat{\ell}_f$.
5. $\left[\overrightarrow{F_{\rm d}}, \overrightarrow{H_t}\right] \equiv 0$ *on* Σ *hence* $\overrightarrow{F_{\rm d}}$ *is invariant on* Σ *with respect to the flow of* $\overrightarrow{H_t}$*:*

$$\overrightarrow{F_{\rm d}} \circ \mathcal{H}_t(\ell) = \mathcal{H}_{t*}\overrightarrow{F_{\rm d}}(\ell), \qquad \forall (t, \ell) \in [\widehat{\tau}_2, T] \times \Sigma. \tag{22}$$

The proof of Lemma 4.1 can be done adapting the results in [19] and completes the analysis of the singular arc.

The bang arcs present problems of a different kind. Namely we need to define the *switching times* near the reference switching points $\widehat{\ell}_1$ and $\widehat{\ell}_2$.

For (t,ℓ) near the graph of $\widehat{\lambda}$ we need $H_t(\ell) \geq H^{\max}(\ell)$, but in [19] it is shown that the backwards flow of $\overrightarrow{H_2}$ from time $\widehat{\tau}_2$ is the maximized one *if and only if* $H_{23}(\ell) \geq 0$. In order to overcome this problem we introduce a *correction* of such a flow from time $\widehat{\tau}_2$ by keeping it on Σ whenever $H_{23}(\ell) < 0$.

In [19] it is shown that, thanks to the second inequality in Assumption 4, the implicit function theorem applied to the problem:

$$\begin{cases} H_{23} \circ \exp(t_2 - \widehat{\tau}_2)\overrightarrow{\widetilde{H}}_2\,(\ell) = 0, \\ t_2(\ell) = \widehat{\tau}_2 \quad \text{ if } H_{23}(\ell) = 0. \end{cases}$$

defines, in a neighborhood $\mathcal{O}_2$ of $\widehat{\ell}_2$, a function $t_2 \colon \mathcal{O}_2 \to \mathbb{R}$ such that if $\ell \in \Sigma$, then $t_2(\ell) = \widehat{\tau}_2$ if and only if $\ell \in \mathcal{S}$; moreover

$$\langle \mathrm{d}t_2(\widehat{\ell}_2)\,,\, \delta\ell\rangle = \frac{-\,\boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{H_{23}}(\widehat{\ell}_2)\right)}{H_{223}(\widehat{\ell}_2)} \quad \forall \delta\ell \in T^*\mathbb{R}^n.$$

We set

$$\tau_2(\ell) := \min\left\{t_2(\ell), \widehat{\tau}_2\right\} = \begin{cases} t_2(\ell) & \text{if } H_{23}(\ell) < 0, \\ \widehat{\tau}_2 & \text{if } H_{23}(\ell) \geq 0. \end{cases}$$

The next step is the definition of the switching time $\tau_1 \colon \mathcal{O}_2 \to \mathbb{R}$, possibly shrinking $\mathcal{O}_2$. Actually, thanks to the first inequality in Assumption 4, the implicit function theorem applies also to

$$\begin{cases} (H_2 - H_1) \circ \exp\left(\tau_1 - \tau_2(\ell)\right)\overrightarrow{H_2} \circ \exp\left(\tau_2(\ell) - \widehat{\tau}_2\right)\overrightarrow{\widetilde{H}_2}(\ell) = 0, \\ \tau_1(\widehat{\ell}_2) = \widehat{\tau}_1, \end{cases}$$

see e.g. [3]. Setting

$$\widetilde{k}(x) := \widehat{S}_{\widehat{\tau}_2 *} k \circ \widehat{S}_{\widehat{\tau}_2}^{-1}(x), \quad \widetilde{K}(p,x) := \langle p\,,\, \widetilde{k}(x)\rangle, \tag{23}$$

the linearization of τ_1 at $\widehat{\ell}_2$ is given by

$$\langle \mathrm{d}\tau_1(\widehat{\ell}_2)\,,\, \delta\ell\rangle = \frac{\boldsymbol{\sigma}\left(\exp\left(\widehat{\tau}_1 - \widehat{\tau}_2\right)\overrightarrow{H_2}_*\delta\ell, \left(\overrightarrow{H_1} - \overrightarrow{H_2}\right)(\widehat{\ell}_1)\right)}{H_{12}(\widehat{\ell}_1)} = \frac{\boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{\widetilde{K}}(\widehat{\ell}_2)\right)}{H_{12}(\widehat{\ell}_1)}. \tag{24}$$

We can now define the flow $(t,\ell) \mapsto \mathcal{H}_t(\ell)$ backwards in time emanating from a neighborhood $\mathcal{O}_f$ of $\widehat{\ell}_f$ in $T^*\mathbb{R}^n$ at time T.

For any $t \in [\widehat{\tau}_2, T]$ we choose as $\mathcal{H}_t(\ell)$ the flow of $\overrightarrow{H_t}$ defined in (21).

For $t < \widehat{\tau}_2$, setting $\widetilde{\ell} := \mathcal{H}_{\widehat{\tau}_2}(\ell)$, we define

$$\mathcal{H}_t(\ell) := \begin{cases} \exp(t - \widehat{\tau}_2)\overrightarrow{\widetilde{H}_2}(\widetilde{\ell}) & t \in [\tau_2(\widetilde{\ell}), \widehat{\tau}_2], \\ \exp(t - \tau_2(\widetilde{\ell}))\overrightarrow{H_2} \circ \mathcal{H}_{\tau_2(\widetilde{\ell})}(\widetilde{\ell}) & t \in [\tau_1(\widetilde{\ell}), \tau_2(\widetilde{\ell})), \\ \exp(t - \tau_1(\widetilde{\ell}))\overrightarrow{H_1} \circ \mathcal{H}_{\tau_1(\widetilde{\ell})}(\widetilde{\ell}) & t \in [0, \tau_1(\widetilde{\ell})), \end{cases} \tag{25}$$

see Figure 1.

REMARK 4.2. *Notice that $\mathcal{H}$ is C^∞ on $[\widehat{\tau}_2^+, T] \times \mathcal{O}_f$ and it is Lipschitz continuous on $[0, \widehat{\tau}_2^-] \times \mathcal{O}_f$. Actually it is C^1 but on $\{(t, \mathcal{H}_t(\ell)) \colon t = \tau_1(\widetilde{\ell})\}$. Indeed on the set $\{(t, \mathcal{H}_t(\ell)) \colon t = \tau_2(\widetilde{\ell})\}$ it is C^1 since $\mathcal{H}_{\tau_2(\widetilde{\ell})}(\widetilde{\ell}) \in \mathcal{S}$, so that $\overrightarrow{\widetilde{H}}_2(\mathcal{H}_{\tau_2(\widetilde{\ell})}(\widetilde{\ell})) = \overrightarrow{H_2}(\mathcal{H}_{\tau_2(\widetilde{\ell})}(\widetilde{\ell}))$, by (20).*

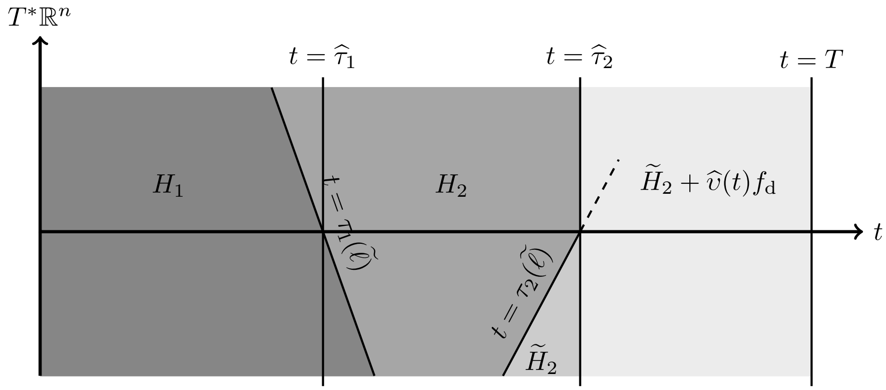

Fig. 1. *The over-maximized Hamiltonian along its flow emanating from Λ at time T. The picture shows when the transition from one smooth piece to another is defined and where the over-maximized Hamiltonian is actually greater than the maximized Hamiltonian of the control system.*

4.2. Hamiltonian sufficient conditions. In this section we state and prove the sufficient conditions for strong local optimality of $\widehat{\xi}$ in terms of Hamiltonian flow.

First we prove that if the projected over-maximized flow emanating from Λ is locally Lipschitz invertible then $\widehat{\xi}$ is a strong local minimizer, Theorem 3.1. Afterwards we give the second order conditions that ensure this invertibility property, Theorem 4.2.

Theorem 4.2. *Let Λ be defined in (18). Assume that*

$$\mathrm{id}\times\pi\mathcal{H}_:\ (t,\ell)\in[0,T]\times\Lambda\mapsto(t,\pi\mathcal{H}_t(\ell))\in\mathcal{U}. \tag{26}$$

is locally Lipschitz invertible onto a neighborhood $\mathcal{U}$ of the graph of $\widehat{\xi}$ in $[0,T]\times\mathbb{R}^n$. Then $\widehat{\xi}$ is a strict strong locally optimal trajectory for the cost $\widetilde{c}(\xi(T))$ subject to (1b)-(1c).

Proof. Clearly $(\mathrm{id}\times\pi\mathcal{H})^{-1}(t,\widehat{\xi}(t))=(t,\widehat{\ell}_f)$ for any $t\in[0,T]$. Let $\xi\colon[0,T]\to\mathbb{R}^n$ be an admissible trajectory for (1) whose graph is in $\mathcal{U}$ and let

$$(t,\ell(t)):=(\mathrm{id}\times\pi\mathcal{H})^{-1}(t,\xi(t)),\qquad\lambda(t):=\mathcal{H}_t(\ell(t))=(\mu(t),\xi(t))\,,\qquad t\in[0,T].$$

If $\varphi\colon[0,1]\to\Lambda$ is a smooth curve such that $\varphi(0)=\ell(T)$, $\varphi(1)=\widehat{\ell}_f$ then we can consider the closed path in $[0,T]\times\Lambda$ obtained by the concatenation of the curves

$$t\mapsto(t,\ell(t)),\qquad s\mapsto(T,\varphi(s)),\qquad t\mapsto(T-t,\widehat{\ell}_f).$$

Integrating the one-form $\boldsymbol{\omega}:=\mathcal{H}^*\left(\boldsymbol{s}-H_t\,\mathrm{d}t\right)$ (which is exact on $[0,T]\times\Lambda$, see [24]), we obtain

$$\begin{aligned}0=\oint\boldsymbol{\omega}=\int_{\mathrm{id}\times\ell}\Big(\langle\mu(t)\,,\,\dot{\xi}(t)\rangle-H_t(\lambda(t))\Big)\,\mathrm{d}t+\int_{\varphi}\mathcal{H}_t^*\boldsymbol{s}+&\\-\int_{\mathrm{id}\times\widehat{\ell}_f}\Big(\langle\widehat{\mu}(t)\,,\,\dot{\widehat{\xi}}(t)\rangle-H_t(\widehat{\lambda}(t))\Big)\,\mathrm{d}t.&\end{aligned}\tag{27}$$

By construction of the over-maximized Hamiltonian H_t, the integrand is non positive along the curve $\mathrm{id}\times\ell$ and is identically zero along the curve $\mathrm{id}\times\widehat{\ell}_f$. Therefore

$$
\begin{aligned}
(28) \qquad 0 \leq \int_\varphi \mathcal{H}_t^* \boldsymbol{s} &= \int_0^1 \langle \varphi(s)\,,\, \frac{\mathrm{d}}{\mathrm{d}s}(\pi\varphi)(s)\rangle \,\mathrm{d}s \\
&= \int_0^1 \langle \mathrm{d}(-\widetilde{c})(\pi\varphi(s))\,,\, \frac{\mathrm{d}}{\mathrm{d}s}(\pi\varphi)(s)\rangle \,\mathrm{d}s = \widetilde{c}(\xi(T)) - \widetilde{c}(\widehat{x}_T).
\end{aligned}
$$

Thus $\widetilde{c}(\xi(T)) \geq \widetilde{c}(\widehat{x}_T)$, i.e. $\widehat{\xi}$ is a strong local minimizer for the cost $\widetilde{c}$. Let us show that in fact it is a strict one.

If $\widetilde{c}(\xi(T)) = \widetilde{c}(\widehat{x}_T)$, then (27)-(28) imply that

$$
(29) \qquad \langle \mu(t)\,,\, \dot{\xi}(t)\rangle - H_t(\lambda(t)) = 0 \qquad \text{a.e. } t \in [0, T]
$$

so that

$$
(30) \qquad \langle \mu(t)\,,\, \dot{\xi}(t)\rangle = H^{\max}(\lambda(t)) = H_t(\lambda(t)) \qquad \text{a.e. } t \in [0, T].
$$

Since $\xi(0) = \widehat{\xi}(0)$, we also have $\lambda(0) = \widehat{\ell}_0$. On the interval $[0, \widehat{\tau}_2)$, equation (30) and the maximality condition imply that $\mathrm{d}(H_t - H^{\max})|_{\lambda(t)} = 0$ a.e. $t \in [0, \widehat{\tau}_2]$, so that $\overrightarrow{H}^{\max}(\lambda(t)) = \overrightarrow{H_t}(\lambda(t))$ a.e. $t \in [0, \widehat{\tau}_2]$. Thus $\lambda(t) = \widehat{\lambda}(t)$ for any $t \in [0, \widehat{\tau}_2]$, in particular $\lambda(\widehat{\tau}_2) = \widehat{\ell}_2$.

For $t \in [\widehat{\tau}_2, T]$, equation (29) yields $\widetilde{H}_2(\lambda(t)) = H_2(\lambda(t))$, i.e. $\lambda(t) \in \mathcal{S}$. Define

$$
\Sigma_{\xi(t)} := \left\{ p \in (\mathbb{R}^n)^* : \langle p\,,\, f_{\mathrm{d}}(\xi(t))\rangle = 0 \right\}
$$

and consider the function

$$
\Omega \colon p \in \Sigma_{\xi(t)} \mapsto \langle p\,,\, \dot{\xi}(t)\rangle - H_t(p, \xi(t)) \in \mathbb{R}.
$$

By PMP the function Ω is non positive and by (29) it is null in $\mu(t)$, therefore

$$
(31) \qquad \langle \delta p\,,\, \dot{\xi}(t) - \pi\overrightarrow{H_t}(\lambda(t))\rangle = 0, \quad \forall \delta p \in (\mathbb{R}^n)^*, \text{ such that } \langle \delta p\,,\, f_{\mathrm{d}}(\xi(t))\rangle = 0.
$$

Hence there exists $b(t) \in \mathbb{R}$ such that

$$
\dot{\xi}(t) = \pi\overrightarrow{H_t}(\lambda(t)) + b(t) f_{\mathrm{d}}(\xi(t)) \qquad \forall t \in [\widehat{\tau}_2, T].
$$

By (22), $(\pi\mathcal{H}_t)_*^{-1} f_{\mathrm{d}} \circ (\pi\mathcal{H}_t)(\ell(t)) = \overrightarrow{F_{\mathrm{d}}}(\ell(t))$ so that

$$
\begin{aligned}
&\dot{\ell}(t) = (\pi\mathcal{H}_t)_*^{-1}\left(\dot{\xi}(t) - \pi\overrightarrow{H_t}(\lambda(t))\right) = b(t)\,(\pi\mathcal{H}_t)_*^{-1} f_{\mathrm{d}}(\xi(t)) = b(t)\overrightarrow{F_{\mathrm{d}}}(\ell(t)), \\
&\dot{\lambda}(t) = \overrightarrow{H_t}(\lambda(t)) + \mathcal{H}_{t*}\dot{\ell}(t) = \overrightarrow{H_2}(\lambda(t)) + (\widehat{v}(t) + b(t))\,\overrightarrow{F_{\mathrm{d}}}(\lambda(t)).
\end{aligned}
$$

Finally, since $\lambda(t) \in \mathcal{S}$, we get

$$
(32) \qquad 0 = \boldsymbol{\sigma}\left(\dot{\lambda}(t), \overrightarrow{H_{23}}(\lambda(t))\right) = -H_{232}(\lambda(t)) + (\widehat{v}(t) + b(t))\,\mathbb{L}(\lambda(t)),
$$

so that $\widehat{v}(t) + b(t) = u_{\mathrm{s}}(\lambda(t))$. Therefore $\lambda(t)$ and $\widehat{\lambda}(t)$ solve the same Cauchy problem (19). This proves that $\lambda \equiv \widehat{\lambda}$ and hence the strict strong local optimality of $\widehat{\xi}$. ◻

As we want to obtain second order sufficient conditions, we take a Hamiltonian approach based on the linearization of the flow from

$$
L_T := T_{\widehat{\ell}_f}\Lambda = \left\{ \mathrm{d}(-\widetilde{c})_*\delta x \colon \delta x \in \mathbb{R}^n \right\}.
$$

Our construction is naturally split in two parts by the time $t = \widehat{\tau}_2$. In particular we point out that if $(\pi\mathcal{H}_{\widehat{\tau}_2})_* : L_T \to \mathbb{R}^n$ is an isomorphism, then there exists at least a smooth function $\alpha_2 \colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathrm{d}\alpha_2(\widehat{x}_2) = \widehat{\mu}(\widehat{\tau}_2), \qquad \mathcal{H}_{\widehat{\tau}_2 *} L_T = \{\mathrm{d}\alpha_{2*}\delta x \colon \delta x \in \mathbb{R}^n\}. \tag{33}$$

We explicitly point out that α_2 is not uniquely determined but only its first and second order derivatives at $\widehat{x}_2$ are determined by (33).

We can now state a second order sufficient condition for the strong local optimality of the reference trajectory $\widehat{\xi}$.

THEOREM 4.3. *Let $\widehat{\xi}$ be the admissible trajectory defined in (2). Assume that $\widehat{\xi}$ is a state extremal (Assumption 1) satisfying the regularity Assumptions 2–5.*

Assume moreover

1. $(\pi\mathcal{H}_t)_* : L_T \to \mathbb{R}^n$ *is an isomorphism for any $t \in [\widehat{\tau}_2, T]$, i.e. the kernel of the map is trivial;*
2. $H_{12}(\widehat{\ell}_1) - L^2_{\widetilde{k}}\alpha_2(\widehat{x}_2) > 0$ *where α_2 is any smooth function on $\mathbb{R}^n$ satisfying (33) and $\widetilde{k}$ is defined in (23).*

Then $\widehat{\xi}$ is a strict strong local minimizer for problem (1).

Proof. According to Theorem 4.2 we only need to prove that the map $\mathrm{id} \times \pi\mathcal{H}_\cdot [0, T] \times \Lambda \to [0, T] \times \mathbb{R}^n$ is one-to-one onto a neighborhood $\mathcal{U}$ of the graph of $\widehat{\xi}$. Since $[0, T]$ is a compact interval, it suffices to prove that $\mathrm{id} \times \pi\mathcal{H}_t$ is locally bi-Lipschitz in a neighborhood of $(t, \widehat{\ell}_f)$ for any $t \in [0, T]$.

For $t \neq \widehat{\tau}_1$ Remark 4.2 implies that it suffices to prove that $(\pi\mathcal{H}_t)_* : L_T \to \mathbb{R}^n$ is an isomorphism while, for $t = \widehat{\tau}_1$ we take advantage of Clarke inverse function theorem.

- Condition 1 ensures the invertibility on $[\widehat{\tau}_2, T]$.
- For $t \in (\widehat{\tau}_1, \widehat{\tau}_2)$, $(\pi\mathcal{H}_t)_* = \exp(t - \widehat{\tau}_2)h_{2*}(\pi\mathcal{H}_{\widehat{\tau}_2})_*$ which is invertible.
- If $t = \widehat{\tau}_1$, for any $\delta\ell \in L_T$, set $\widetilde{\delta\ell} = \mathcal{H}_{\widehat{\tau}_2 *}\delta\ell$. The linearization of $\pi\mathcal{H}_{\widehat{\tau}_1}$ at $\widehat{\ell}_f$ is given by

$$(\pi\mathcal{H}_{\widehat{\tau}_1})_*\delta\ell = \begin{cases} \exp(\widehat{\tau}_1 - \widehat{\tau}_2)h_{2*}\pi\widetilde{\delta\ell}, & \langle \mathrm{d}\tau_1(\widehat{\ell}_2), \widetilde{\delta\ell}\rangle < 0, \\ \exp(\widehat{\tau}_1 - \widehat{\tau}_2)h_{2*}\Big(\pi\widetilde{\delta\ell} - \langle \mathrm{d}\tau_1(\widehat{\ell}_2), \widetilde{\delta\ell}\rangle\widetilde{k}(\widehat{x}_2)\Big), & \langle \mathrm{d}\tau_1(\widehat{\ell}_2), \widetilde{\delta\ell}\rangle > 0. \end{cases}$$

If $\langle \mathrm{d}\tau_1(\widehat{\ell}_2), \widetilde{\delta\ell}\rangle\widetilde{k}(\widehat{x}_2) = 0$ for any $\widetilde{\delta\ell} \in L_T$, we are done. Otherwise, it suffices to prove that for any $a \in [0, 1]$ and any $\delta\ell \in L_T$, $\delta\ell \neq 0$

$$\pi(\widetilde{\delta\ell}) - a\langle \mathrm{d}\tau_1(\widehat{\ell}_2), \widetilde{\delta\ell}\rangle\widetilde{k}(\widehat{x}_2) \neq 0.$$

For $a = 0$ the claim is obvious thanks to assumption 1. Assume by contradiction there exists $a \in (0, 1]$ and $\delta\ell \in L_T$, $\delta\ell \neq 0$ such that

$$\pi\widetilde{\delta\ell} - a\langle \mathrm{d}\tau_1(\widehat{\ell}_2), \widetilde{\delta\ell}\rangle\widetilde{k}(\widehat{x}_2) = 0. \tag{34}$$

By (34) there exists $\rho \neq 0$ such that $\pi\widetilde{\delta\ell} = \rho\widetilde{k}(\widehat{x}_2)$ so that $\widetilde{\delta\ell} = \rho\,\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{x}_2)$ and

$$0 = \rho\widetilde{k}(\widehat{x}_2) - a\langle \mathrm{d}\tau_1(\widehat{\ell}_2), \rho\,\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{x}_2)\rangle\widetilde{k}(\widehat{x}_2).$$

Thus $1 - a\langle \mathrm{d}\tau_1(\widehat{\ell}_2), \mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{x}_2)\rangle = 0$, so that

$$0 = 1 - \frac{a}{H_{12}(\widehat{\ell}_1)}\boldsymbol{\sigma}\left(\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{x}_2), \vec{\widetilde{K}}(\widehat{\ell}_2)\right) = \frac{1}{H_{12}(\widehat{\ell}_1)}\left\{H_{12}(\widehat{\ell}_1) - aL^2_{\widetilde{k}}\alpha_2(\widehat{x}_2)\right\}.$$

Since this quantity is positive both for $a = 0$ and for $a = 1$, it is positive for any $a \in [0, 1]$. ☐

REMARK 4.3. *As already said, the switching time $\widehat{\tau}_2$ naturally splits our construction in two parts. In particular Assumption 1 of Theorem 4.3 takes into account only the problem restricted to the singular interval $[\widehat{\tau}_2, T]$. Assumption 2 coincides with the sufficient condition in [3] for a fixed-free Mayer problem on $[0, \widehat{\tau}_2]$ with cost $\alpha_2(\xi(\widehat{\tau}_2))$.*

5. Proof of the main result.

In this section we prove that the coercivity of $\widetilde{J}$ (Assumption 6 *a)*) guarantees that Assumptions 1 and 2 of Theorem 4.3 hold true. In particular Assumption 1 will be proven to hold by exploiting the coercivity of $\widetilde{J}$ on the subspace $\widetilde{\mathcal{V}}$ of the admissible variations such that $\varepsilon_0 = 0$, while Assumption 2 is proven to hold by exploiting the coercivity of $\widetilde{J}$ on the subspace of the admissible variations which are $\widetilde{J}$-orthogonal to $\widetilde{\mathcal{V}}$.

5.1. Coercivity of $\widetilde{J}$ in Hamiltonian formalism.

We start by exploiting the coercivity of $\widetilde{J}$ on $\widetilde{\mathcal{V}} := \left\{ \delta e = (\delta x, 0, w) \in \widetilde{\mathcal{W}} \right\}$, i.e.

$$\widetilde{J}[\delta e]^2 = \frac{1}{2} \int_{\widehat{\tau}_2}^{T} \left(2\, w(t) L_{\zeta(t)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) + R(t) w(t)^2 \right) \mathrm{d}t \tag{35}$$

subject to

$$\dot{\zeta}(t) = w(t) \dot{g}_t(\widehat{x}_T), \quad \zeta(\widehat{\tau}_2) = 0, \quad \zeta(T) = \delta x \in \mathbb{R}^n. \tag{36}$$

The associated Hamiltonian is given by the quadratic form

$$H''_t(\delta p, \delta x) = -\frac{1}{2R(t)} \left(\langle \delta p\, , \, \dot{g}_t(\widehat{x}_T) \rangle + L_{\delta x} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) \right)^2 \tag{37}$$

and the corresponding Hamiltonian linear system with initial conditions in the Lagrangian subspace of transversality conditions

$$L''_T := \{ (0, \delta x) \colon \delta x \in \mathbb{R}^n \}$$

is given by

$$\begin{cases} \dot{\mu}''(t) = \dfrac{1}{R(t)} \Big(\langle \mu''(t)\, , \, \dot{g}_t(\widehat{x}_T) \rangle + L_{\zeta''(t)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) \Big) L_{(\cdot)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T)\,, \\ \dot{\zeta}''(t) = \dfrac{-1}{R(t)} \Big(\langle \mu''(t)\, , \, \dot{g}_t(\widehat{x}_T) \rangle + L_{\zeta''(t)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) \Big) \dot{g}_t(\widehat{x}_T), \\ \mu''(T) = 0, \qquad \zeta''(T) = \delta x. \end{cases} \tag{38}$$

We denote the solution of (38) as $\mathcal{H}''_t(0, \delta x)$.

$\widetilde{J}$ is coercive on $\widetilde{\mathcal{V}}$ if and only if for any $t \in [\widehat{\tau}_2, T]$,

$$\delta x \neq 0 \implies \zeta''(t) \neq 0 \tag{39}$$

where ζ'' is defined in (38), see for example [23].

If $k(\widehat{x}_T) = 0$, then we get no more information since $\widetilde{\mathcal{V}} = \widetilde{\mathcal{W}}$.
Assume $k(\widehat{x}_T) \neq 0$. Since $\widetilde{J}$ is coercive on $\widetilde{\mathcal{V}}$, we just need to express its coercivity on $\widetilde{\mathcal{V}}^{\perp} := \left\{ \delta e \in \widetilde{\mathcal{W}} \colon \widetilde{J}(\delta e, \overline{\delta e}) = 0 \quad \forall \overline{\delta e} \in \widetilde{\mathcal{V}} \right\}$.

For any $\delta e = (\delta x, \varepsilon_0, w)$, $\overline{\delta e} = \left(\overline{\delta x}, \overline{\varepsilon}_0, \overline{w}\right)$ in $\widetilde{\mathcal{W}}$, let ζ and $\overline{\zeta}$ be the corresponding solutions of system (17). The bilinear form associated with $\widetilde{J}$ is given by

$$
\begin{aligned}
\widetilde{J}(\delta e, \overline{\delta e}) = & \frac{\varepsilon_0 \overline{\varepsilon}_0}{2} \left(H_{12}(\widehat{\ell}_1) + L_k^2 \widetilde{c}\,(\widehat{x}_T) \right) + \\
& + \frac{1}{2} \int_{\widehat{\tau}_2}^{T} \left(\overline{w}(t) L_{\zeta(t)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) + w(t) L_{\overline{\zeta}(t)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) + w(t)\overline{w}(t) R(t) \right) \mathrm{d}t.
\end{aligned} \tag{40}
$$

If $p(t)$ is the solution of the Cauchy problem

$$
\dot{p}(t) = -w(t) L_{(\cdot)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T)\,, \quad p(T) = 0,
$$

then an integration by parts in (40) gives

$$
\begin{aligned}
\widetilde{J}(\delta e, \overline{\delta e}) = & \frac{1}{2} \left(\varepsilon_0 \overline{\varepsilon}_0 \left(H_{12}(\widehat{\ell}_1) + L_k^2 \widetilde{c}\,(\widehat{x}_T) \right) + \langle p(\widehat{\tau}_2)\,,\, \overline{\zeta}(\widehat{\tau}_2) \rangle \right) + \\
& + \frac{1}{2} \int_{\widehat{\tau}_2}^{T} \overline{w}(t) \left(\langle p(t)\,,\, \dot{g}_t(\widehat{x}_T) \rangle + L_{\zeta(t)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) + w(t) R(t) \right) \mathrm{d}t.
\end{aligned} \tag{41}
$$

Since $\overline{\zeta}(T)$ is free, we obtain that $\overline{w}$ may be any function in $L^2([\widehat{\tau}_2, T], \mathbb{R})$. Thus, if $\delta e \in \widetilde{\mathcal{V}}^{\perp}$ then

$$
\langle p(t)\,,\, \dot{g}_t(\widehat{x}_T) \rangle + L_{\zeta(t)} L_{\dot{g}_t} \widetilde{c}\,(\widehat{x}_T) + w(t) R(t) = 0 \quad \text{a.e. } t \in [\widehat{\tau}_2, T]. \tag{42}
$$

Comparing (42) and (38) we get $(p(t), \zeta(t)) = \mathcal{H}''_t(0, \delta x) = (\mu''(t), \zeta''(t))$ so that for any $\delta e \in \widetilde{\mathcal{V}}^{\perp}$ we get $\zeta''(\widehat{\tau}_2) = \varepsilon_0 k(\widehat{x}_T)$ and

$$
\widetilde{J}[\delta e]^2 = \frac{\varepsilon_0^2}{2} \left(H_{12}(\widehat{\ell}_1) + L_k^2 \widetilde{c}\,(\widehat{x}_T) \right) + \frac{1}{2} \langle \mu''(\widehat{\tau}_2)\,,\, \zeta''(\widehat{\tau}_2) \rangle. \tag{43}
$$

Without loss of generality we can choose $\varepsilon_0 = 1$, so that the coercivity of $\widetilde{J}$ can be expressed as

$$
H_{12}(\widehat{\ell}_1) - L_k^2 (-\widetilde{c})\,(\widehat{x}_T) + \boldsymbol{\sigma} \left(\mathcal{H}''_{\widehat{\tau}_2} \left(\pi \mathcal{H}''_{\widehat{\tau}_2} \right)^{-1} k(\widehat{x}_T), (0, k(\widehat{x}_T)) \right) > 0. \tag{44}
$$

5.2. The anti-symplectic isomorphism.

In order to relate the coercivity of $\widetilde{J}$ with the properties of the flow $\mathcal{H}_t$, we define

$$
\iota \colon (\delta p, \delta x) \in (\mathbb{R}^n)^* \times \mathbb{R}^n \mapsto \delta \ell := \left(-\delta p + \mathrm{D}^2 \,(-\widetilde{c})(\widehat{x}_T)(\delta x, \cdot), \delta x \right) \in (\mathbb{R}^n)^* \times \mathbb{R}^n
$$

so that $\iota^{-1} = \iota$. The mapping ι is an anti-symplectic linear isomorphism, i.e.

$$
\boldsymbol{\sigma} \left(\iota(\delta p, \delta x), \iota(\overline{\delta p}, \overline{\delta x}) \right) = \boldsymbol{\sigma} \left((\overline{\delta p}, \overline{\delta x}), (\delta p, \delta x) \right), \quad \forall (\delta p, \delta x),\ (\overline{\delta p}, \overline{\delta x}) \in (\mathbb{R}^n)^* \times \mathbb{R}^n.
$$

The choice of the anti-symplectic isomorphism ι, instead of a symplectic one, depends on the fact that we are using PMP while for the accessory problem we are studying a minimization problem.

With this notation we get

$$
\iota L''_T = \{ \mathrm{d}(-\widetilde{c})_* \delta x \colon \delta x \in T_{\widehat{x}_T} \mathbb{R}^n \} = L_T.
$$

Following the lines of Lemma 9 in [19] one can prove the following Lemma:

LEMMA 5.1. *Let $\mathcal{H}''_t$ and $\mathcal{H}_t$ be the Hamiltonian flows associated to the quadratic Hamiltonian H''_t defined in (37) and to the over-maximized Hamiltonian H_t defined in (21), respectively. Then*

$$
\iota \mathcal{H}''_t \iota^{-1} = \widehat{\mathcal{F}}_{t*}^{-1} \mathcal{H}_{t*} \qquad \forall t \in [\widehat{\tau}_2, T]. \tag{45}
$$

5.3. Proof of Assumption 1 of Theorem 4.3. Applying (45) to $\mathcal{H}''_{t\,|L''_T}$ we get for $\delta\ell \in L_T$

$$\pi\mathcal{H}''_t\iota^{-1}\delta\ell = \pi\widehat{\mathcal{F}}_{t*}^{-1}\mathcal{H}_{t*}\delta\ell.$$

Since the Hamiltonian $\widehat{F}_t$ is the lift of the vector field $\widehat{f}_t$, we get that $\pi\widehat{\mathcal{F}}_{t*}^{-1} = \widehat{S}_{t*}^{-1}\pi$. Thus from (39) we obtain that Assumption 1 of Theorem 4.3 holds true. This implies the existence of a function α_2 as defined in (33).

5.4. Proof of Assumption 2 of Theorem 4.3. Notice that

$$\begin{aligned} &L_k^2(-\widetilde{c})\,(\widehat{x}_T) = \boldsymbol{\sigma}\left(\mathrm{d}(-\widetilde{c})_* k(\widehat{x}_T), \overrightarrow{K}(\widehat{\ell}_f)\right), \\ &\left(\pi\mathcal{H}''_{\widehat{\tau}_2}\right)^{-1} k(\widehat{x}_T) = \iota^{-1}\left(\pi\widehat{\mathcal{F}}_{\widehat{\tau}_2}^{-1}\mathcal{H}_{\widehat{\tau}_2}\right)_*^{-1} k(\widehat{x}_T) = \iota^{-1}\left(\pi\mathcal{H}_{\widehat{\tau}_2}\right)_*^{-1}\widetilde{k}(\widehat{x}_2) \end{aligned} \tag{46}$$

where the vector field $\widetilde{k}$ is defined in (23), as well as the associated Hamiltonian $\widetilde{K}$. We can compute

$$\begin{aligned} &\boldsymbol{\sigma}\left(\mathcal{H}''_{\widehat{\tau}_2}\left(\pi\mathcal{H}''_{\widehat{\tau}_2}\right)^{-1} k(\widehat{x}_T), (0, k(\widehat{x}_T))\right) = \boldsymbol{\sigma}\left(\iota^{-1}\widehat{\mathcal{F}}_{\widehat{\tau}_2*}^{-1}\mathcal{H}_{\widehat{\tau}_2*}\iota\left(\pi\mathcal{H}''_{\widehat{\tau}_2}\right)^{-1} k(\widehat{x}_T), (0, k(\widehat{x}_T))\right) \\ &= \boldsymbol{\sigma}\left(\mathrm{d}(-\widetilde{c})_* k(\widehat{x}_T), \widehat{\mathcal{F}}_{\widehat{\tau}_2*}^{-1}\mathcal{H}_{\widehat{\tau}_2*}\left(\pi\mathcal{H}_{\widehat{\tau}_2*}\right)^{-1}\widetilde{k}(\widehat{x}_2)\right) = \boldsymbol{\sigma}\left(\mathrm{d}(-\widetilde{c})_* k(\widehat{x}_T), \widehat{\mathcal{F}}_{\widehat{\tau}_2*}^{-1}\,\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{x}_2)\right). \end{aligned}$$

Thus

$$\begin{aligned} &\boldsymbol{\sigma}\left(\mathcal{H}''_{\widehat{\tau}_2}\left(\pi\mathcal{H}''_{\widehat{\tau}_2}\right)^{-1} k(\widehat{x}_T), (0, k(\widehat{x}_T))\right) - L_k^2(-\widetilde{c})\,(\widehat{x}_T) = \\ &\qquad = \boldsymbol{\sigma}\left(\mathrm{d}(-\widetilde{c})_* k(\widehat{x}_T), \widehat{\mathcal{F}}_{\widehat{\tau}_2*}^{-1}\,\mathrm{d}\,\alpha_{2*}\widetilde{k}(\widehat{x}_2) - \overrightarrow{K}(\widehat{\ell}_f)\right) = \\ &\qquad = \boldsymbol{\sigma}\left(\mathrm{d}(-\widetilde{c}\circ\widehat{S}_{\widehat{\tau}_2}^{-1})_*\,\widetilde{k}(\widehat{x}_2), \mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{x}_2) - \overrightarrow{\widetilde{K}}(\widehat{\ell}_2)\right) = \\ &\qquad = -\left(\mathrm{D}^2\alpha_2(\widehat{x}_2)[\widetilde{k}(\widehat{x}_2)]^2 + \langle\widehat{\mu}(\widehat{\tau}_2)\,,\,\mathrm{D}\widetilde{k}(\widehat{x}_2)\widetilde{k}(\widehat{x}_2)\rangle\right) = -L_{\widetilde{k}}^2\alpha_2\,(\widehat{x}_2)\,. \end{aligned} \tag{47}$$

Equations (47) and (44) complete the proof of Assumption 2 of Theorem 4.3.

6. The state-feedback single input case. The standard form for single input control affine systems is

$$\dot{\xi}(t) = f_0(\xi(t)) + u(t)f_1(\xi(t)), \quad |u(t)| \le 1. \tag{48}$$

This case was dealt with by the authors in [21]. We now consider the case when there is a state-feedback control for singular extremals, namely there exists a function $v_{\mathrm{s}}\colon \mathbb{R}^n \to \mathbb{R}$ such that $v_{\mathrm{s}}(\widehat{\xi}(t)) = \widehat{v}(t)$ for any $t \in [\widehat{\tau}_2, T]$. Indeed this is the case when the ratio $\dfrac{-F_{001}}{F_{101}}(p,x)$ does not depend on p whenever $(p,x) \in \mathcal{S}$.

Under this assumption, sufficient second order conditions have been given for optimality of trajectories containing both bang and singular arcs with respect to trajectories with the same switching structure, see [25] and the references therein.

Without any loss of generality we can assume that the dynamics driving a bang-bang-singular trajectory is given by

$$\widehat{f}_t = \begin{cases} h_1 = f_0 - f_1 & t \in [0, \widehat{\tau}_1), \\ h_2 = f_0 + f_1 & t \in [\widehat{\tau}_1, \widehat{\tau}_2), \\ f_{\mathrm{s}} := f_0 + v_{\mathrm{s}} f_1 & t \in [\widehat{\tau}_2, T], \end{cases} \tag{49}$$

hence for any $t \in [\widehat{\tau}_2, T]$, $\widehat{S}_t$ coincides with $\exp(t-T)f_{\rm s}$.

In this case we denote by F_0, $F_{\rm s}$ and F_1 the Hamiltonians associated to f_0, $f_{\rm s}$ and f_1 and by $F_{i_1 i_2 \dots i_k}$, $i_1, \dots, i_k \in \{0, {\rm s}, 1\}$ according to the rules stated in Section 2.1.

In sub-problem (9) the dynamics in the interval $[\widehat{\tau}_2, T]$ can be written as

$$\dot{\xi}(t) = f_{\rm s}(\xi(t)) + v(t) f_1(\xi(t))$$

with $v(t)$ taking values in a neighborhood of zero so that the second variation reads

$$\begin{aligned}\widetilde{J}[(\delta x, \varepsilon_0, w)]^2 = \frac{\varepsilon_0^2}{2}\left(L_k^2 \widetilde{c}(\widehat{x}_T) + H_{12}(\widehat{\ell}_1)\right) + \\ + \int_{\widehat{\tau}_2}^T \left(w(t)^2 F_{1{\rm s}1}(\widehat{\lambda}(t)) + 2w(t) L_{\zeta(t)} L_{\dot{g}_t} \widetilde{c}(\widehat{x}_T)\right) {\rm d}t\end{aligned}$$

subject to $\dot{\zeta}(t) = w(t)\dot{g}_t(\widehat{x}_T)$, $\quad \zeta(\widehat{\tau}_2) = \varepsilon_0 k(\widehat{x}_T)$, $\quad \zeta(T) = \delta x \in \mathbb{R}^n$ where

$$k = -2\widehat{S}_{\widehat{\tau}_1 *}^{-1} f_1 \circ \widehat{S}_{\widehat{\tau}_1}, \quad g_t := \widehat{S}_{t*}^{-1} f_1 \circ \widehat{S}_t, \quad \dot{g}_t := \widehat{S}_{t*}^{-1}[f_{\rm s}, f_1] \circ \widehat{S}_t.$$

REMARK 6.1. *As a consequence a necessary condition for the coercivity of $\widetilde{J}$ is*

$$\widetilde{J}_0 := L_k^2 \widetilde{c}(\widehat{x}_T) + H_{12}(\widehat{\ell}_1) > 0. \tag{50}$$

In [25], for a problem of this class, the author shows that the trajectory is optimal with respect to trajectories associated to controls with the same bang-bang-singular structure if the 2×2 matrix associated to the problem obtained by moving both the switching times is positive definite. We point out that $\widetilde{J}_0$ is the $(1,1)$-entry of such matrix.

The regularity condition along the singular arc is trivially satisfied; the other ones read as follows:

- Regularity along the bang arcs:

$$F_1(\widehat{\lambda}(t)) < 0, \quad t \in [0, \widehat{\tau}_1), \qquad F_1(\widehat{\lambda}(t)) > 0, \quad t \in (\widehat{\tau}_1, \widehat{\tau}_2).$$

- Regularity at the switching points: $F_{01}(\widehat{\ell}_1) > 0$, $(F_{001} + F_{101})(\widehat{\ell}_2) > 0$.
- Strong generalized Legendre condition (SGLC):

$$F_{1{\rm s}1}(\widehat{\lambda}(t)) = F_{101}(\widehat{\lambda}(t)) > 0 \quad \forall t \in [\widehat{\tau}_2, T].$$

6.1. Van der Pol Oscillator. As an example consider Van der Pol Oscillator with final time $T = 4$, in the form studied in [25] where the author numerically shows the existence of a bang-bang-singular extremal and studies its optimality with respect to trajectories with the same control structure.

$$\text{minimize} \quad \xi_3(T) \quad \text{subject to} \tag{51a}$$

$$\begin{aligned}&\dot{\xi}_1(t) = \xi_2(t), \\ &\dot{\xi}_2(t) = -\xi_1(t) + \xi_2(t)\left(1 - \xi_1^2(t)\right) + u(t), \qquad \text{a.e. } t \in [0, T], \\ &\dot{\xi}_3(t) = \frac{1}{2}\left(\xi_1^2(t) + \xi_2^2(t)\right),\end{aligned} \tag{51b}$$

$$\xi(0) = (0, 1, 0), \quad u \in [-1, 1]. \tag{51c}$$

More precisely the author numerically shows that there are two bang arcs

$$\widehat{u}(t) \equiv -1 \quad \forall t \in [0, \widehat{\tau}_1), \qquad \widehat{u}(t) \equiv 1 \quad \forall t \in (\widehat{\tau}_1, \widehat{\tau}_2)$$

where $\widehat{\tau}_1 \simeq 1.3667$, $\widehat{\tau}_2 \simeq 2.4601$ and a singular arc characterized by a state-feedback control where

$$f_{\mathrm{s}} := x_2\partial_{x_1} + x_1\partial_{x_2} + \frac{x_1^2 + x_2^2}{2}\partial_{x_3}, \qquad f_{\mathrm{s}1} = -\partial_{x_1} - x_2\partial_{x_3}.$$

Let us notice that the flow of f_{s} can be computed explicitly:

$$\exp(t-T)f_{\mathrm{s}}(x) = \begin{pmatrix} x_1\cosh(t-T) + x_2\sinh(t-T) \\ x_1\sinh(t-T) + x_2\cosh(t-T) \\ x_3 + \frac{x_1x_2}{2}\left(\cosh(2(t-T)) - 1\right) + \frac{x_1^2+x_2^2}{2}\sinh(2(t-T)) \end{pmatrix}$$

so that we can also compute

$$\dot{g}_t(x) = -\cosh(t-T)\partial_{x_1} + \sinh(t-T)\partial_{x_2} - x_2\cosh(t-T)\partial_{x_3}$$

Using the numerical results in [25], in [21] the authors prove that the regularity assumptions are satisfied. Here we recall some features which are needed in the following:

$$\widehat{\mu}_3(t) \equiv 1, \quad t \in [0, T], \qquad \widehat{\mu}_1(T) = \widehat{\xi}_2(T) = 0, \qquad F_{1\mathrm{s}1}(\widehat{\lambda}(t)) \equiv 1, \quad t \in [\widehat{\tau}_2, T].$$

In order to prove that our theory applies to this example we prove that the second variation is coercive by proving (39) and (44). We thus need to write the Hamiltonian $H_t''(\delta p, \delta x)$ and the associated linear system. Since $L_T'' = \left\{(0, \delta x) \colon \delta x \in \mathbb{R}^3\right\}$ and

$$H_t''(\delta p, \delta x) = -\frac{1}{2}\left(-\delta p_1\cosh(t-T) + \delta p_2\cosh(t-T) - \delta x_2\cosh(t-T)\right)^2$$

we get $\dot{\mu}_1''(t) \equiv \dot{\mu}_3''(t) \equiv \dot{\zeta}_3''(t) \equiv 0$, $\mu_1''(T) = \mu_3''(T) = 0$, $\zeta_3(T) = \delta x_3$ so that $\mu_1''(t) \equiv \mu_3''(t) \equiv 0$, $\zeta_3''(t) \equiv \delta x_3$ and

$$\dot{\mu}_2''(t) = -\left(\mu_2''(t)\sinh(t-T) - \xi_2''(t)\cosh(t-T)\right)\cosh(t-T),$$
$$\dot{\xi}_1''(t) = \left(\mu_2''(t)\sinh(t-T) - \xi_2''(t)\cosh(t-T)\right)\cosh(t-T),$$
$$\dot{\xi}_2''(t) = -\left(\mu_2''(t)\sinh(t-T) - \xi_2''(t)\cosh(t-T)\right)\sinh(t-T).$$

Thus

$$\mathcal{H}_t''(0, \delta x) = \begin{pmatrix} \delta x_2\sinh(t-T)\,\mathrm{d}x_2 \\ \left(\delta x_1 - \delta x_2\sinh(t-T)\right)\partial_{x_1} + \delta x_2\cosh(t-T)\partial_{x_2} + \delta x_3\partial_{x_3} \end{pmatrix}$$

hence $\pi\mathcal{H}_t''(0, \delta x) = 0$ implies $\delta x = 0$, i.e. (39) is satisfied.

Inequality (44) reads $\widetilde{J}_0 + \langle \mu''(\widehat{\tau}_2), \zeta''(\widehat{\tau}_2)\rangle > 0$ when $\zeta''(\widehat{\tau}_2) = k(\widehat{x}_T)$ that is

$$\widetilde{J}_0 + k_2^2(\widehat{x}_T)\tanh(\widehat{\tau}_2 - T) > 0.$$

$\widetilde{J}_0 \simeq 215.1022$ was computed in [25]. The push-forward $k(\widehat{x}_T) = -2\widehat{S}_{\widehat{\tau}_1*}^{-1}\partial_{x_2} \circ \widehat{S}_{\widehat{\tau}_1}$ can be computed numerically, obtaining $k(\widehat{x}_T) \simeq 14.5864\partial_{x_1} - 14.6632\partial_{x_2} - 0.0005\partial_{x_3}$. Thus $k_2^2(\widehat{x}_T)\tanh(\widehat{\tau}_2 - T) \simeq -196.1122$. Hence (44) is satisfied, i.e. the second variation J'' coercive. This proves that our results apply to the Van der Pol oscillator.

Acknowledgments. The authors are grateful to Dott. Francesco Mugelli from the University of Florence for the numerical computations in the example above.

REFERENCES

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

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

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

[4] M. S. Aronna, J. F. Bonnans, A. V. Dmitruk, and P. A. Lotito, *Quadratic order conditions for bang-singular extremals*, NUMERICAL ALGEBRA, CONTROL AND OPTIMIZATION, 2 (2012), pp. 511–546. doi:10.3934/naco.2012.2.511.

[5] Z. Chen, J.-B. Caillau, and Y. Chitour, *L1-minimization for mechanical systems*, SIAM Journal on Control and Optimization, 54 (2016), pp. 1245–1265, https://doi.org/10.1137/15M1013274.

[6] F. C. Chittaro and G. Stefani, *Minimum-time strong optimality of a singular arc: The multi-input non involutive case*, ESAIM: COCV, 22 (2016), pp. 786–810, http://dx.doi.org/10.1051/cocv/2015026.

[7] U. Felgenhauer, L. Poggiolini, and G. Stefani, *Optimality and stability result for bang–bang optimal controls with simple and double switch behaviour*, CONTROL AND CYBERNETICS, 38 (2009), pp. 1305 – 1325.

[8] M. R. Hestenes, *Applications of the theory of quadratic forms in Hilbert space to calculus of variations*, Pacific J. Math, 1 (1951), pp. 525–581.

[9] U. Ledzewicz and H. Schttler, *Optimal bang-bang controls for a two-compartment model in cancer chemotherapy*, Journal of Optimization Theory and Applications, 114 (2002), pp. 609–637, https://doi.org/10.1023/A:1016027113579.

[10] A. A. Milyutin and N. P. Osmolovskii, *Calculus of Variations and Optimal Control*, vol. 180 of Translations of Mathematical Monographs, American Mathematical Society, 1998.

[11] N. Osmolovskii and H. Maurer, *Applications to Regular and Bang-Bang Control*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2012, https://doi.org/10.1137/1.9781611972368, http://epubs.siam.org/doi/abs/10.1137/1.9781611972368.

[12] L. Poggiolini, *On local state optimality of bang-bang extremals in a free horizon Bolza problem*, Rendiconti del Seminario Matematico dell'Università e del Politecnico di Torino, (2006).

[13] L. Poggiolini, *Structural stability of bang–bang trajectories with a double switching time in the minimum time problem*, SIAM JOURNAL ON CONTROL AND OPTIMIZATION, 55 (2017), pp. 3779–3798, https://doi.org/10.1137/16M1083761.

[14] L. Poggiolini and M. Spadini, *Sufficient optimality conditions for a bang-bang trajectory in a Bolza problem*, in Mathematical Control Theory and Finance, A. Sarychev, A. Shiryaev, M. Guerra, and M. d. R. Grossinho, eds., Springer Berlin Heidelberg, 2008, pp. 337–357, https://doi.org/10.1007/978-3-540-69532-5_19. 10.1007/978-3-540-69532-5_19.

[15] L. Poggiolini and M. Spadini, *Strong local optimality for a bang-bang trajectory in a Mayer problem*, SIAM Journal on Control and Optimization, 49 (2011), pp. 140–161, https://doi.org/10.1137/090771405, http://epubs.siam.org/doi/abs/10.1137/090771405.

[16] L. Poggiolini and M. Spadini, *Bang–bang trajectories with a double switching time in the minimum time problem*, ESAIM: COCV, (2015), https://doi.org/10.1051/cocv/2015021.

[17] L. Poggiolini and G. Stefani, *Strong local optimality for a bang-bang-singular extremal.* Work in progress.

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

[19] L. Poggiolini and G. Stefani, *Bang-singular-bang extremals: sufficient optimality conditions*, Journal of Dynamical and Control Systems, 17 (2011), pp. 469–514, http://dx.doi.org/10.1007/s10883-011-9127-y. 10.1007/s10883-011-9127-y.

[20] L. Poggiolini and G. Stefani, *Structural stability for bang–singular–bang extremals in the minimum time problem*, SIAM J. Control Optim., 51 (2013), pp. 3511–3531, http://dx.doi.org/10.1137/120895421.

[21] L. Poggiolini and G. Stefani, *Strong local optimality for bang-bang-singular extremals in single input control problems*, in 20th IFAC World Congress PROCEEDINGS, vol. 50 of IFAC-PAPERSONLINE, IFAC, Elsevier, 2017, pp. 6128–6133, https://doi.org/10.1016/j.ifacol.2017.08.2022.

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

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

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

[25] G. Vossen, *Switching time optimization for bang-bang and singular controls*, Journal of Optimization Theory and Applications, 144 (2010), pp. 409–429, https://doi.org/10.1007/s10957-009-9594-4.