



HAL
open science

Strong local optimality for generalized L1 optimal control problems

Francesca C Chittaro, Laura Poggiolini

► **To cite this version:**

Francesca C Chittaro, Laura Poggiolini. Strong local optimality for generalized L1 optimal control problems. *Journal of Optimization Theory and Applications*, In press. hal-01585634v1

HAL Id: hal-01585634

<https://hal.science/hal-01585634v1>

Submitted on 11 Sep 2017 (v1), last revised 2 Jul 2018 (v2)

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.

Strong local optimality for generalized L^1 optimal control problems

Francesca C. Chittaro¹ and Laura Poggiolini^{2‡}

September 11, 2017

Abstract

In this paper, we analyse control affine optimal control problems with a cost functional involving the absolute value of the control. The Pontryagin extremals associated with such systems are given by (possible) concatenations of bang arcs with singular arcs and with *inactivated arcs*, that is, arcs where the control is identically zero. Here we consider Pontryagin extremals given by a bang-inactive-bang concatenation. We establish sufficient optimality conditions for such extremals, in terms of some regularity conditions and of the coercivity of a suitable finite-dimensional second variation.

1 Introduction

In recent years, optimal control problems aiming at minimising the L^1 -norm of the control have received an increasing attention: for instance, they model problems coming from neurobiology [1], mechanics [2, 3], and fuel-consumption [4, 5, 6]. As noticed in these papers, a peculiarity of optimal control problems involving the L^1 -norm is the fact that the optimal control vanishes along nontrivial time intervals; this property is referred to as *sparsity*, and the piece of optimal trajectories corresponding to the zero control are known as *inactivated arcs*, *cost arcs* or *zero arcs*.

The arising of sparse controls for L^1 minimisation problems is well known also in the frame of infinite-dimensional optimal control, see for instance [7, 8, 9], where usually the L^1 -norm of the control is added to the integral cost in order to induce sparse solutions.

In this paper we are concerned with a generalisation of the L^1 minimisation problem of the following form:

$$\text{minimise } \int_0^T |u(t)\psi(\xi(t))| dt \text{ subject to} \quad (1a)$$

$$\dot{\xi}(t) = f_0(\xi(t)) + u(t)f_1(\xi(t)), \quad (1b)$$

$$\xi(0) = \hat{x}_0, \quad \xi(T) = \hat{x}_f, \quad (1c)$$

$$|u(t)| \leq 1 \quad \text{a.e. } t \in [0, T], \quad (1d)$$

where the state ξ belongs to a smooth n -dimensional manifold M , f_0 and f_1 are smooth vector fields on M , ψ is a smooth real-valued function on M and $T > 0$ is fixed. For such optimal control problems, a well known necessary condition is the Pontryagin Maximum Principle (PMP),

**This work was supported by the projects “Dinamiche non autonome, sistemi Hamiltoniani e teoria del controllo” (GNAMPA 2015), “Dinamica topologica, sistemi Hamiltoniani e teoria del controllo” (GNAMPA 2016) and by UTLN - Appel à projet “Chercheurs invités”.

[†]Francesca C. Chittaro is with Aix Marseille Université, CNRS, ENSAM, LSIS UMR 7296, 13397 Marseille, France, and Université de Toulon, CNRS, LSIS UMR 7296, 83957 La Garde, France francesca.chittaro@univ-tln.fr

[‡]Laura Poggiolini is with Dipartimento di Matematica e Informatica “Ulisse Dini”, Università degli Studi di Firenze, via di Santa Marta 3, 50139 Firenze laura.poggiolini@unifi.it

possibly in its non-smooth version, see [10]. More precisely, PMP implies that for every optimal trajectory $\xi(\cdot)$ of (1) there exist a constant $\nu \leq 0$ and a suitable curve $\lambda(t) \in T_{\xi(t)}^*M$ such that

$$u(t)F_1(\lambda(t)) + \nu|u(t)\psi(\xi(t))| = \max_{v \in [-1,1]} \{vF_1(\lambda(t)) + \nu|v\psi(\xi(t))|\}, \text{ a.e. } t, \quad (2)$$

where $F_1(\lambda(t)) = \langle \lambda(t), f_1(\xi(t)) \rangle$. In particular, in the normal case ($\nu = -1$), equation (2) implies that the optimal control may have three possible behaviours:

- If there exists an interval $I \subset [0, T]$ such that $|F_1(\lambda(t))| > |\psi(\xi(t))| \quad \forall t \in I$, then $u(t) = \text{sgn}(F_1(\lambda(t)))$. In this case we say that the interval I is a *regular bang arc*, and we call u a *bang control*.
- If $|F_1(\lambda(t))| < |\psi(\xi(t))| \quad \forall t \in I$, then the Hamiltonian (2) attains its maximum only if $u \equiv 0$ in I . In this case, following the terminology of [1], we say that I is an *inactivated arc*.
- If $|F_1(\lambda(t))| = |\psi(\xi(t))| \quad \forall t \in I$, then the maximising control is not uniquely determined by (2). Indeed, if $F_1(\lambda(t)) = |\psi(\xi(t))|$, then the maximum in (2) is attained for every $u \in [0, 1]$, and, analogously, if $F_1(\lambda(t)) = -|\psi(\xi(t))|$, then the maximum is attained for every $u \in [-1, 0]$. We say that I is a *singular arc*.

In this paper we state sufficient optimality conditions for solutions of (1b)-(1c)-(1d) given by a concatenation of bang and inactivated arcs. More precisely, we assume that we are given an admissible trajectory $\hat{\xi}: [0, T] \rightarrow M$ satisfying

$$\dot{\hat{\xi}}(t) = \begin{cases} f_0(\hat{\xi}(t)) + u_1 f_1(\hat{\xi}(t)) & t \in [0, \hat{\tau}_1), \\ f_0(\hat{\xi}(t)) & t \in (\hat{\tau}_1, \hat{\tau}_2), \\ f_0(\hat{\xi}(t)) + u_3 f_1(\hat{\xi}(t)) & t \in (\hat{\tau}_2, T] \end{cases}, \quad \hat{\xi}(0) = \hat{x}_0, \quad \hat{\xi}(T) = \hat{x}_f,$$

for some $\hat{\tau}_1, \hat{\tau}_2 \in (0, T)$, $\hat{\tau}_1 < \hat{\tau}_2$ and some $u_1, u_3 \in \{-1, 1\}$. We provide some first and second order sufficient conditions that guarantee the *strong-local* optimality of $\hat{\xi}(\cdot)$, accordingly to the following definition:

Definition 1 *A curve $\hat{\xi}: [0, T] \rightarrow M$, solution of (1b)-(1c) with associated control $\hat{u}(\cdot)$, is a strong-local minimiser of (1) if there exists a neighbourhood \mathcal{U} in $[0, T] \times M$ of the graph of $\hat{\xi}$ such that $\int_0^T |\hat{u}(t)\psi(\hat{\xi}(t))| dt \leq \int_0^T |u(t)\psi(\xi(t))| dt$ for every admissible trajectory $\xi: [0, T] \rightarrow M$ of (1b)-(1c)-(1d), with associated control $u(\cdot)$ and whose graph is in \mathcal{U} .*

If $\int_0^T |\hat{u}(t)\psi(\hat{\xi}(t))| dt < \int_0^T |u(t)\psi(\xi(t))| dt$ for any admissible trajectory other than $\hat{\xi}$ and whose graph is in \mathcal{U} , we say that $\hat{\xi}$ is a strict strong-local minimiser.

In [3], the authors investigate necessary and sufficient optimality conditions for control problems with cost functional affine in the L^1 -norm of the control: the original optimal control problem is extended in such a way that the candidate optimal trajectory happens to be a bang-bang extremal of the extended problem, so that sufficient conditions are stated in terms of optimality conditions of finite-dimension optimisation problems.

Our approach relies on Pontryagin Maximum Principle and on Hamiltonian methods ([11, 12]). This technique consists in covering a neighbourhood of the graph of the reference trajectory so that the admissible trajectories can be lifted to the cotangent bundle, where the costs can be compared. Notice that it is possible to lift all the trajectories whose graph is close to the one of $\hat{\xi}(\cdot)$, independently on the associated controls: this permits us to obtain an optimality result in the *strong* norm. The main point in this procedure is to be able to define a suitable Lagrangian manifold Λ of the cotangent bundle and a suitable diffeomorphism from $[0, T] \times \Lambda_0$ onto a neighbourhood of the graph of the reference trajectory. In the standard theory (that is, for smooth costs and open sets of control, see [11]), this diffeomorphism is the composition of the canonical projection with the flow of the maximised Hamiltonian, restricted to some suitable Lagrangian submanifold; in particular, the reference extremal is locally optimal as long as the projection of the Hamiltonian flow, restricted to Λ_0 , is invertible with smooth inverse. In order to apply similar methods to the problem under consideration, two main ingredients are needed:

- *the regularity conditions*, that guarantee that the maximised Hamiltonian of PMP is well defined, continuous and piecewise smooth in a neighbourhood of the range of the reference extremal $\widehat{\lambda}: [0, T] \rightarrow T^*M$ associated with $\widehat{\xi}$ in PMP;
- *the coercivity of the second variation*: indeed, following the approach initiated in [13, 14], we show that if the second variation associated to a finite dimensional sub-problem of (1) is coercive, then the projection of the flow of the maximised Hamiltonian emanating from is invertible and we can lift the admissible trajectories to the cotangent bundle.

We would like to stress the fact that the regularity conditions are very close to the necessary ones, in the sense that they sum up to requiring strict inequalities where PMP yields mild ones, plus a non tangency condition which is generically satisfied.

Finally, we remark that the conditions that we require are quite easy to check on a given extremal: indeed, the regularity assumptions are just sign conditions along the extremal, while the second variation reduces to a quadratic function of one real variable.

2 Preliminaries

2.1 Notations

Given a vector field f on M , the Lie derivative at a point $q \in M$ of a smooth function $\varphi: M \rightarrow \mathbb{R}$ with respect to f is denoted with $L_f\varphi(q) = \langle d\varphi(q), f(q) \rangle$, and $L_f^2\varphi(q) = L_f(L_f\varphi)(q)$.

The Lie bracket of two vector fields f, g is denoted as commonly with $[f, g]$.

Let T^*M be the cotangent bundle and $\pi: T^*M \rightarrow M$ its projection on M . For every vector field on M , we denote with the corresponding capital letter the associated Hamiltonian function, that is $F(\ell) = \langle \ell, f(\pi\ell) \rangle$, $\ell \in T^*M$.

Denote with σ 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

$$\langle dF(\ell), \cdot \rangle = \sigma_\ell(\cdot, \vec{F}(\ell)).$$

The Hamiltonian flow associated with the Hamiltonian vector field \vec{F} , and starting at time $t = 0$, is denoted with the cursive \mathcal{F}_t .

Finally, we define the function $\Psi: T^*M \rightarrow \mathbb{R}$ as $\Psi := \psi \circ \pi$. As we shall extensively use it, we recall that $\vec{\Psi}(\ell) = (-D\psi(\pi\ell), 0)$.

2.2 Assumptions

Assume that we are given an admissible trajectory $\widehat{\xi}$ for our control system, that is, a solution of (1b), corresponding to some measurable control $\widehat{u}(\cdot)$ with values in $[-1, 1]$, that satisfies the boundary conditions (1c). We assume moreover that the control is of the form bang-inactivated-bang, i.e. there exist $\widehat{\tau}_1 < \widehat{\tau}_2 \in (0, T)$, and $u_1, u_3 \in \{-1, 1\}$ such that

$$\widehat{u}(t) = \begin{cases} u_1 & t \in [0, \widehat{\tau}_1), \\ 0 & t \in (\widehat{\tau}_1, \widehat{\tau}_2), \\ u_3 & t \in (\widehat{\tau}_2, T]. \end{cases} \quad (3)$$

The trajectory $\widehat{\xi}$ and its associated control \widehat{u} are respectively called *reference trajectory* and *reference control*.

In analogy to the classical bang-bang case, we say that $\widehat{\tau}_1$ and $\widehat{\tau}_2$ are the *switching times* of the reference control \widehat{u} and we call $\widehat{\xi}(\widehat{\tau}_1)$, $\widehat{\xi}(\widehat{\tau}_2)$ the *switching points* of $\widehat{\xi}$. For the sake of future notation we set $\widehat{\tau}_0 := 0$, $\widehat{\tau}_3 := T$ and $\widehat{I}_j := (\widehat{\tau}_{j-1}, \widehat{\tau}_j)$, $j = 1, 2, 3$. Moreover, we call h_j the vector field that defines the reference trajectory in the interval \widehat{I}_j , that is, $h_1 = f_0 + u_1 f_1$, $h_2 = f_0$ and

$h_3 = f_0 + u_3 f_1$, and the *reference time-dependent vector field* $\widehat{h}_t = f_0 + \widehat{u}(t)f_1$, so that $\widehat{h}_t \equiv h_j$ for $t \in \widehat{I}_j$. We denote with $\widehat{S}_t(x)$ the solution at time t of the Cauchy problem

$$\dot{\xi}(t) = \widehat{h}_t(\xi(t)), \quad \xi(0) = x. \quad (4)$$

Our first assumption ensures that the cost function is smooth along the reference trajectory but for a finite number of points. It also guarantees that the Hamiltonian vector field associated with the maximised Hamiltonian (defined in (7) here below) is well defined and C^1 but in a codimension 1 subset, as explained in Remark 1 below. It is also crucial in defining the second variation of an appropriate subproblem of (1).

Assumption 1 *The cost ψ does not annihilate at the switching points of the reference trajectory, that is $\psi(\widehat{\xi}(\widehat{\tau}_i)) \neq 0$, $i = 1, 2$. Moreover, for every $t \in [0, T]$, the vector field \widehat{h}_t is never tangent to the set $\{\psi = 0\}$ along $\widehat{\xi}(t)$, that is $L_{\widehat{h}_t} \psi(\widehat{\xi}(t)) \neq 0$.*

Thanks to Assumption 1, $\psi \circ \widehat{\xi}$ annihilates at most a finite number of times. In particular, we assume that $\psi(\widehat{\xi}(t))$ annihilates n_1 times in the interval \widehat{I}_1 and n_3 times in the interval \widehat{I}_3 (where we admit the cases $n_1 = 0$ and $n_3 = 0$). We denote as \widehat{s}_{1i} , $i = 0, \dots, n_1$, the zeroes of $\psi \circ \widehat{\xi}$ occurring in \widehat{I}_1 and \widehat{s}_{3i} , $i = 0, \dots, n_3$, the zeroes of $\psi \circ \widehat{\xi}$ occurring in \widehat{I}_3 . Set $a_0 = \text{sgn}(\psi(\widehat{x}_0))$ and $a_2 = \text{sgn}(\psi(\widehat{\xi}(\widehat{\tau}_2)))$. In particular

$$\begin{aligned} \text{sgn}(\psi \circ \widehat{\xi}|_{(\widehat{s}_{1\ i-1}, \widehat{s}_{1i})}) &= a_0(-1)^{i-1}, \quad i \in \{1, \dots, n_1 + 1\}, \\ \text{sgn}(\psi \circ \widehat{\xi}|_{(\widehat{s}_{3\ i-1}, \widehat{s}_{3i})}) &= a_2(-1)^{i-1}, \quad i \in \{1, \dots, n_3 + 1\}, \end{aligned}$$

(here we set $\widehat{s}_{10} = 0$, $\widehat{s}_{1\ n_1+1} = \widehat{\tau}_1$, $\widehat{s}_{30} = \widehat{\tau}_2$, $\widehat{s}_{3\ n_3+1} = T$).

Remark 1 *By continuity and Assumption 1, there exist two neighbourhoods $\mathcal{V}_1, \mathcal{V}_3$ of the range of $\widehat{\xi}|_{\widehat{I}_1}$, $\widehat{\xi}|_{\widehat{I}_3}$ respectively, such that the zero level sets of $\psi|_{\mathcal{V}_1}$ and $\psi|_{\mathcal{V}_3}$ are hypersurfaces of M , transverse to h_1 and h_3 , respectively. Moreover the zero level set of $\psi|_{\mathcal{V}_j}$, $j = 1, 3$, has n_j connected components.*

Our next assumption is that the reference pair satisfies PMP in its normal form; in principle, non-smooth versions of PMP (see e.g. [10, Theorem 22.26]) are required. Indeed our problem can be seen as a hybrid control problem as defined in [10, Section 22.5] with the switching surface S given by $\{(t, x, y) : \psi(x) = 0, y = x\}$. Nevertheless, thanks to Assumption 1, in the case under study [10, Theorem 22.26] reduces to the standard smooth version of PMP (as stated, for instance, in [11]).

In order to apply it, we define the (normal) *control-dependent Hamiltonian* and *reference Hamiltonian* respectively as

$$\mathfrak{h}(u, \ell) := F_0 + uF_1(\ell) - |u\psi(\pi\ell)|, \quad (u, \ell) \in [-1, 1] \times T^*M \quad (5)$$

$$\widehat{H}_t(\ell) := F_0 + \widehat{u}(t)F_1(\ell) - |\widehat{u}(t)\psi(\pi\ell)|, \quad \ell \in T^*M \quad (6)$$

and we assume the following.

Assumption 2 (PMP) *There exists a Lipschitzian curve $\widehat{\lambda}: [0, T] \rightarrow T^*M$ such that*

$$\begin{aligned} \dot{\widehat{\lambda}}(t) &= \overrightarrow{\widehat{H}_t}(\widehat{\lambda}(t)) & \text{a.e. } t \in [0, T], \\ \pi\widehat{\lambda}(t) &= \widehat{\xi}(t), & \forall t \in [0, T]. \\ \widehat{H}_t(\widehat{\lambda}(t)) &= \max_{u \in [-1, 1]} \mathfrak{h}(u, \ell), & \text{a.e. } t \in [0, T]. \end{aligned} \quad (7)$$

The curve $\widehat{\lambda}(\cdot)$ is called the *reference extremal*. In the following, we set

$$\widehat{\ell}_0 = \widehat{\lambda}(0), \quad \widehat{\ell}_1 = \widehat{\lambda}(\widehat{\tau}_1), \quad \widehat{\ell}_2 = \widehat{\lambda}(\widehat{\tau}_2), \quad \widehat{\ell}_T = \widehat{\lambda}(T).$$

Remark 2 Assumption 1 ensures that in a neighbourhood of the range of $\widehat{\lambda}(\cdot)$ the switching surface is composed by two non-intersecting connected components, $\{F_1 = |\Psi|\}$ and $\{F_1 = -|\Psi|\}$. Moreover, it guarantees that the Hamiltonian vector field associated with \widehat{H}_t is well defined along the reference trajectory, except at the times $t = \widehat{s}_{1i}$, $i = 1, \dots, n_1$, \widehat{s}_{3i} , $i = 1, \dots, n_3$, $\widehat{\tau}_1, \widehat{\tau}_2$, where it has possibly different left-sided and right-sided limits.

Remark 3 Let $\xi(\cdot)$ be an admissible trajectory of (1b) that satisfies the Pontryagin Maximum Principle (PMP from now on) with associated extremal $\lambda(\cdot)$. If the costate $\lambda(t)$ does not belong to the set

$$\{\ell \in T^*M : |F_1(\ell)| = |\Psi(\ell)|\}, \quad (8)$$

then the control associated with ξ can be recovered uniquely by equation (7). Analogously to the smooth case, we call such set the switching surface of problem (1).

It is easy to see that the reference Hamiltonian (6) is piecewise-constant with respect to time, and it takes the values

$$\widehat{H}_t(\ell) := \begin{cases} H_1^{\sigma_{0i}} := F_0 + u_1 F_1 + \sigma_{0i} \psi \circ \pi & t \in [\widehat{s}_{1\ i-1}, \widehat{s}_{1i}] \ i = 1, \dots, n_1 + 1, \\ H_2 := F_0 & t \in [\widehat{\tau}_1, \widehat{\tau}_2], \\ H_3^{\sigma_{2i}} := F_0 + u_3 F_1 + \sigma_{2i} \psi \circ \pi & t \in [\widehat{s}_{3\ i-1}, \widehat{s}_{3i}] \ i = 1, \dots, n_3 + 1, \end{cases} \quad (9)$$

where we used the following symbols:

$$\sigma_{0i} = a_0(-1)^i, \ i = 1, \dots, n_1, \quad \sigma_{2i} = a_2(-1)^i, \ i = 1, \dots, n_3.$$

Because of the maximality condition of PMP (equation (7)), the following inequalities hold along the reference extremal:

$$u_1 F_1(\widehat{\lambda}(t)) \geq |\psi(\widehat{\xi}(t))|, \ t \in \widehat{I}_1, \quad |F_1(\widehat{\lambda}(t))| \leq |\psi(\widehat{\xi}(t))|, \ t \in \widehat{I}_2, \quad (10)$$

$$u_3 F_1(\widehat{\lambda}(t)) \geq |\psi(\widehat{\xi}(t))|, \ t \in \widehat{I}_3. \quad (11)$$

In addition, the following relations must hold:

$$\frac{d}{dt}(H_2 - H_1^{\sigma_{0n_1}})(\widehat{\lambda}(t))|_{t=\widehat{\tau}_1} \geq 0, \quad \frac{d}{dt}(H_3^{\sigma_{20}} - H_2)(\widehat{\lambda}(t))|_{t=\widehat{\tau}_2} \geq 0. \quad (12)$$

Even though the derivative $\widehat{\lambda}(t)$ is discontinuous at $t = \widehat{\tau}_i$, $i = 1, 2$, the two derivatives in (12) exist and are well defined. In particular, they can be expressed as follows:

$$\frac{d}{dt}(H_2 - H_1^{\sigma_{0n_1}})(\widehat{\lambda}(t))\Big|_{t=\widehat{\tau}_1} = \sigma \left(\overrightarrow{H_1^{\sigma_{0n_1}}}, \overrightarrow{H_2} \right) (\widehat{\ell}_1), \quad (13)$$

$$\frac{d}{dt}(H_3^{\sigma_{20}} - H_2)(\widehat{\lambda}(t))\Big|_{t=\widehat{\tau}_2} = \sigma \left(\overrightarrow{H_2}, \overrightarrow{H_3^{\sigma_{20}}} \right) (\widehat{\ell}_2). \quad (14)$$

Remark 4 Due to the nonlinearity introduced by the absolute value of the control in the integral cost, two consecutive bang arcs can exist only if ψ annihilates at the switching point between such arcs (this is a straightforward application of the PMP and of previous remarks). In particular, we relate this fact to [3, Theorem 3.1], that states the impossibility of the existence of bang-bang junctions: indeed, the results of [3] hold for the case $\psi \equiv 1$ (plus some other control-independent terms in the cost).

As already anticipated in the Introduction, in order to apply the Hamiltonian methods we need the maximised Hamiltonian (right hand side of equation (7)) to be well defined and sufficiently regular in a neighbourhood of the range of the reference extremal; this is guaranteed by the strengthened version of the necessary conditions (10) and (12).

Assumption 3 *Along each arc the reference control is the only one that maximises over $u \in [-1, 1]$ the control-dependent Hamiltonian evaluated along $\widehat{\lambda}$. Equivalently, in equations (10) the strict inequalities hold true.*

The following assumptions, Assumptions 4 and 5, are the strict form of the necessary conditions (13)-(14). They basically state that the change of value in the control \widehat{u} can be detected by the first order approximation of $(H_2 - H_1^{\sigma_{0n_1}}) \circ \widehat{\lambda}$ at $\widehat{\tau}_1$ and $(H_3^{\sigma_{20}} - H_2) \circ \widehat{\lambda}$ at $\widehat{\tau}_2$, respectively. They ensure the transversality of the intersection between the reference extremal $\widehat{\lambda}$ and the switching surface (8), thus avoiding bang-singular junctions.

Assumption 4 $\frac{d}{dt}(H_2 - H_1^{\sigma_{0n_1}})(\widehat{\lambda}(t))|_{t=\widehat{\tau}_1} > 0$

Assumption 5 $\frac{d}{dt}(H_3^{\sigma_{20}} - H_2)(\widehat{\lambda}(t))|_{t=\widehat{\tau}_2} > 0$

2.3 The second variation

This section is devoted to the construction of the second variation associated with the problem (1), in the spirit of [15, 13, 16, 14].

In order to compute the second variation of the cost functional, one should take into account all the admissible variations of the reference trajectory $\widehat{\xi}$. Actually, as it will be proved, it suffices to consider only the trajectories corresponding to controls having the same bang-inactive-bang structure of the reference one, that is, to allow only variations of the switching times $\widehat{\tau}_1$ and $\widehat{\tau}_2$. We thus obtain the following two dimensional sub-problem of (1):

$$\min_{0 < \tau_1 < \tau_2 < T} \int_{I_1 \cup I_3} |\psi(\xi(t))| dt \quad (15)$$

subject to

$$\begin{cases} \dot{\xi} = h_i \circ \xi(t) & t \in I_i, \quad i = 1, 2, 3, \\ \xi(0) = \widehat{\mathbf{x}}_0, \quad \xi(T) = \widehat{\mathbf{x}}_f, \\ I_1 = (0, \tau_1), \quad I_2 = (\tau_1, \tau_2), \quad I_3 = (\tau_2, T), \end{cases} \quad (16)$$

It is well known that, on a smooth manifold, the second derivatives are highly coordinate-dependent, unless they are computed with respect to variations contained in the kernel of the differential of the function to be derived. To overcome this problem and obtain an intrinsic expression of the second variation of the cost (15), we introduce two smooth functions $\alpha, \beta: M \rightarrow \mathbb{R}$ satisfying

$$d\alpha(\widehat{\mathbf{x}}_0) = \widehat{\ell}_0, \quad d\beta(\widehat{\mathbf{x}}_f) = -\widehat{\ell}_T, \quad (17)$$

and remove the constraint on the initial point of the trajectories, thus obtaining the extended sub-problem:

$$\min_{\tau_1, \tau_2} \left(\alpha(\xi(0)) + \beta(\xi(T)) + \int_{I_1 \cup I_3} |\psi(\xi(t))| dt \right) \quad (18)$$

subject to

$$\begin{cases} \dot{\xi} = h_i \circ \xi(t) & t \in I_i, \quad i = 1, 2, 3, \\ \xi(0) \in M, \quad \xi(T) = \widehat{\mathbf{x}}_f \\ I_1 = (0, \tau_1), \quad I_2 = (\tau_1, \tau_2), \quad I_3 = (\tau_2, T) \end{cases} \quad (19)$$

It is easy to see, by PMP, that the differential of the new cost (18) is zero at $\widehat{\mathbf{x}}_0$.

The necessity of summing variations belonging to tangents spaces based at different points of the reference trajectory is a typical issue in geometric control. We get around this problem by pulling-back the solutions of (19) to the initial point $\widehat{\mathbf{x}}_0$ by means of the flow of the reference vector field; in this way, the variations of the trajectory will evolve on the tangent space $T_{\widehat{\mathbf{x}}_0}M$.

To simplify the expressions of the pull-backs, we first reparametrise time in such a way that each interval I_j is mapped into the corresponding reference interval \widehat{I}_j ; to do that, we consider the *variations of the lengths of the intervals*

$$\theta_i := (\tau_i - \tau_{i-1}) - (\widehat{\tau}_i - \widehat{\tau}_{i-1}),$$

we set $\boldsymbol{\theta} := (\theta_1, \theta_2, \theta_3)$, and define the piecewise-affine reparametrisation $\varphi_{\boldsymbol{\theta}}: [0, T] \rightarrow [0, T]$ by means of the Cauchy problem

$$\varphi_{\boldsymbol{\theta}}(0) = 0, \quad \dot{\varphi}_{\boldsymbol{\theta}}(t) = \frac{\tau_i - \tau_{i-1}}{\widehat{\tau}_i - \widehat{\tau}_{i-1}} \quad \forall t \in [\widehat{\tau}_{i-1}, \widehat{\tau}_i], \quad i = 1, 2, 3.$$

Denoting the *pull-back fields* as

$$g_i := \widehat{S}_{t*}^{-1} h_i \circ \widehat{S}_t, \quad t \in [\widehat{\tau}_{i-1}, \widehat{\tau}_i], \quad (20)$$

the *pullback system* of (19) has the form

$$\begin{cases} \dot{\zeta}_t(x, \boldsymbol{\theta}) = \frac{\theta_i}{\widehat{\tau}_i - \widehat{\tau}_{i-1}} g_i(\zeta_t(x, \boldsymbol{\theta})) & t \in [\widehat{\tau}_{i-1}, \widehat{\tau}_i], \quad i = 1, 2, 3, \\ \zeta_0(x, \boldsymbol{\theta}) = x, \end{cases}$$

where we notice that $\dot{\varphi}_{\boldsymbol{\theta}}(t) = 1 + \frac{\theta_i}{\widehat{\tau}_i - \widehat{\tau}_{i-1}}$ for every $t \in \widehat{I}_i$. Finally, setting

$$\begin{aligned} \widehat{\beta} &:= \beta \circ \widehat{S}_T, \quad \widehat{\psi}_t := \psi \circ \widehat{S}_t, \\ \widehat{\gamma}(x) &:= \alpha(x) + \widehat{\beta}(x) + \int_{\widehat{I}_1 \cup \widehat{I}_3} |\widehat{\psi}_s(x)| ds, \end{aligned} \quad (21)$$

the cost can be written as

$$J(x, \boldsymbol{\theta}) = \alpha(x) + \widehat{\beta}(\zeta_T(x, \boldsymbol{\theta})) + \int_{\widehat{I}_1 \cup \widehat{I}_3} \dot{\varphi}_{\boldsymbol{\theta}}(t) |\widehat{\psi}_t(\zeta_t(x, \boldsymbol{\theta}))| dt.$$

Since we need to compute the first and second variations of J , it is necessary to get rid of the absolute value inside the integral, that is, to locate the zeroes of the function $\widehat{\psi}_t(\zeta_t(x, \boldsymbol{\theta}))$ for $t \in \widehat{I}_1 \cup \widehat{I}_3$. They turn out to be smooth functions of the initial state x and of the switching time variations $\boldsymbol{\theta}$, as the following Lemma states.

Lemma 1 *Locally around $(\widehat{\mathbf{x}}_0, \mathbf{0})$, there exist n_1 smooth functions \mathfrak{s}_{1i} of $(x, \boldsymbol{\theta})$ such that $\mathfrak{s}_{1i}(\widehat{\mathbf{x}}_0, \mathbf{0}) = \widehat{\mathfrak{s}}_{1i}$ and*

$$\widehat{\psi}_{\mathfrak{s}_{1i}(x, \boldsymbol{\theta})} \circ \exp\left(\mathfrak{s}_{1i}(x, \boldsymbol{\theta}) \frac{\theta_1}{\widehat{\tau}_1} g_1\right)(x) = 0, \quad i = 1, \dots, n_1. \quad (22)$$

Analogously, there exist n_3 smooth functions \mathfrak{s}_{3i} of $(x, \boldsymbol{\theta})$ such that $\mathfrak{s}_{3i}(\widehat{\mathbf{x}}_0, \mathbf{0}) = \widehat{\mathfrak{s}}_{3i}$ and

$$\widehat{\psi}_{\mathfrak{s}_{3i}(x, \boldsymbol{\theta})} \circ \exp\left((\mathfrak{s}_{3i}(x, \boldsymbol{\theta}) - \widehat{\tau}_2) \frac{\theta_3}{T - \widehat{\tau}_2} g_3\right) \circ \exp(\theta_2 g_2) \circ \exp(\theta_1 g_1)(x) = 0, \quad i = 1, \dots, n_3. \quad (23)$$

Moreover, the following relations hold at the point $(\widehat{\mathbf{x}}_0, \mathbf{0})$:

$$\begin{aligned} \frac{\partial \mathfrak{s}_{1i}}{\partial x} &= -\frac{L_{\delta x} \widehat{\psi}_{\widehat{\mathfrak{s}}_{1i}}(\widehat{\mathbf{x}}_0)}{L_{g_1} \widehat{\psi}_{\widehat{\mathfrak{s}}_{1i}}(\widehat{\mathbf{x}}_0)}, & \frac{\partial \mathfrak{s}_{1i}}{\partial \theta_1} &= -\frac{\widehat{\mathfrak{s}}_{1i}}{\widehat{\tau}_1}, & \frac{\partial \mathfrak{s}_{1i}}{\partial \theta_2} &= \frac{\partial \mathfrak{s}_{1i}}{\partial \theta_3} = 0, \\ \frac{\partial \mathfrak{s}_{3i}}{\partial x} &= -\frac{L_{\delta x} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)}, & \frac{\partial \mathfrak{s}_{3i}}{\partial \theta_1} &= -\frac{L_{g_1} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)}, \\ \frac{\partial \mathfrak{s}_{3i}}{\partial \theta_2} &= -\frac{L_{g_2} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)}, & \frac{\partial \mathfrak{s}_{3i}}{\partial \theta_3} &= -\frac{\widehat{\mathfrak{s}}_{3i} - \widehat{\tau}_2}{T - \widehat{\tau}_2}. \end{aligned}$$

The Lemma is proved applying the implicit function theorem to the functions defined in (22) and (23); this can be done thanks to Assumption 1. As usual, we set $\mathfrak{s}_{10} = 0$, $\mathfrak{s}_{1n_1+1} = \widehat{\tau}_1$, $\mathfrak{s}_{30} = \widehat{\tau}_2$, $\mathfrak{s}_{3n_3+1} = T$.

Taking advantage of the functions \mathfrak{s}_{1i} , \mathfrak{s}_{3i} , we can rewrite the cost $J(x, \theta)$ as

$$\begin{aligned} J(x, \theta) &= \alpha(x) + \widehat{\beta}(\zeta_T(x, \theta)) + a_0 \left(1 + \frac{\theta_1}{\widehat{\tau}_1}\right) \sum_{i=1}^{n_1+1} (-1)^{i-1} \int_{\mathfrak{s}_{1i-1}(x, \theta)}^{\mathfrak{s}_{1i}(x, \theta)} \widehat{\psi}_t \circ \exp\left(\frac{t\theta_1}{\widehat{\tau}_1} g_1\right)(x) dt \\ &\quad + a_2 \left(1 + \frac{\theta_3}{T - \widehat{\tau}_2}\right) \times \\ &\quad \times \sum_{i=1}^{n_3+1} (-1)^{i-1} \int_{\mathfrak{s}_{3i-1}(x, \theta)}^{\mathfrak{s}_{3i}(x, \theta)} \widehat{\psi}_t \circ \exp\left(\frac{(t - \widehat{\tau}_2)\theta_3}{T - \widehat{\tau}_2} g_3\right) \circ \exp(\theta_2 g_2) \circ \exp(\theta_1 g_1)(x) dt, \end{aligned}$$

and compute its second variation at the point $(\widehat{\mathbf{x}}_0, \mathbf{0})$ along all variations $(\delta x, \varepsilon) \in T_{\widehat{\mathbf{x}}_0}M \times \mathbb{R}^3$ compatible with the constraints (19), that is, along the variations in the space

$$V = \{(\delta x, \varepsilon) : \delta x + \sum_{i=1}^3 \varepsilon_i g_i = 0, \sum_{i=1}^3 \varepsilon_i = 0\}. \quad (24)$$

After some easy but tedious manipulations we end up with the following expression:

$$\begin{aligned} J''[\delta x, \varepsilon]^2 &= \frac{1}{2} D^2 \widehat{\gamma}(\widehat{\mathbf{x}}_0)[\delta x]^2 + \\ &\quad + a_2 \sum_{i=1}^{n_3} (-1)^i \frac{L_{\delta x} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)} \left(2L_{\delta x + \varepsilon_1 g_1 + \varepsilon_2 g_2} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0) - L_{\delta x} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)\right) \\ &\quad + a_0 \sum_{i=1}^{n_1} (-1)^i \frac{\left(L_{\delta x} \widehat{\psi}_{\widehat{\mathfrak{s}}_{1i}}(\widehat{\mathbf{x}}_0)\right)^2}{L_{g_1} \widehat{\psi}_{\widehat{\mathfrak{s}}_{1i}}(\widehat{\mathbf{x}}_0)} + \int_{\widehat{I}_3} L_{\delta x} L_{\varepsilon_1 g_1 + \varepsilon_2 g_2} \left|\widehat{\psi}_t\right|(\widehat{\mathbf{x}}_0) dt + L_{\delta x} L_{\sum_{i=1}^3 \varepsilon_i g_i} \widehat{\beta}(\widehat{\mathbf{x}}_0) \\ &\quad + a_0 \frac{\varepsilon_1}{2} (-1)^{n_1} \left(L_{\delta x} \widehat{\psi}_{\widehat{\tau}_1}(\widehat{\mathbf{x}}_0) + L_{\delta x + \varepsilon_1 g_1} \widehat{\psi}_{\widehat{\tau}_1}(\widehat{\mathbf{x}}_0)\right) + a_2 (-1)^{n_3} \varepsilon_3 L_{\delta x + \sum_{i=1}^3 \varepsilon_i g_i} \widehat{\psi}_T(\widehat{\mathbf{x}}_0) \\ &\quad - \frac{\varepsilon_3^2}{2} L_{g_3} \widehat{\psi}_T(\widehat{\mathbf{x}}_0) + a_2 \sum_{i=1}^{n_3} (-1)^i \frac{\left(L_{\varepsilon_1 g_1 + \varepsilon_2 g_2} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)\right)^2}{L_{g_3} \widehat{\psi}_{\widehat{\mathfrak{s}}_{3i}}(\widehat{\mathbf{x}}_0)} + \frac{1}{2} \int_{\widehat{I}_3} L_{\varepsilon_1 g_1 + \varepsilon_2 g_2}^2 \left|\widehat{\psi}_t\right|(\widehat{\mathbf{x}}_0) dt \\ &\quad + \frac{\varepsilon_1 \varepsilon_2}{2} \int_{\widehat{I}_3} L_{[g_1, g_2]} \left|\widehat{\psi}_t\right|(\widehat{\mathbf{x}}_0) dt + L_{\sum_{i=1}^3 \varepsilon_i g_i}^2 \widehat{\beta}(\widehat{\mathbf{x}}_0) + \frac{1}{2} \sum_{1 \leq i < j \leq 3} \varepsilon_i \varepsilon_j L_{[g_i, g_j]} \widehat{\beta}(\widehat{\mathbf{x}}_0). \end{aligned} \quad (25)$$

Remark 5 *The terms containing the Lie derivative of $|\widehat{\psi}_t|$ are defined everywhere, except for a finite number of values of t , therefore the integrals are well defined.*

We recall that the space of admissible variations for the original subproblem (15)-(16) is

$$V_0 = \{(\delta x, \varepsilon) : \delta x + \sum_{i=1}^3 \varepsilon_i g_i = 0, \sum_{i=1}^3 \varepsilon_i = 0, \delta x = 0\}.$$

Indeed, only the restriction to V_0 of (25) is independent of the choice of α and β , while its value on the whole V depends on the choice of these two functions.

Moreover we notice that we can write the second variation as the sum $J'' = \frac{1}{2} D^2 \widehat{\gamma}(\widehat{\mathbf{x}}_0)[\delta x]^2 + J''_0$, where J''_0 does not depend on α . Since $D^2 \widehat{\gamma}(\widehat{\mathbf{x}}_0)[\delta x]^2|_{V_0}$ is null, then, if $J''|_{V_0}$ is coercive, we can always choose α in such a way that J'' is coercive on the whole V . In Section 4 we will prove that the coercivity of J'' on V implies the existence of a suitable Lagrangian submanifold such that the projection of the maximised Hamiltonian flow is an invertible map between this manifold and its image on M . In view of this, our last assumption is the following:

Assumption 6 J'' is coercive on V_0 .

Remark 6 We stress that, in order to apply the Hamiltonian methods, we need to know the expression of the second variation, equation (25), of the extended sub-problem. Nevertheless, the invertibility is guaranteed under Assumption 6, i.e. it suffices to check the coercivity of the quadratic form

$$\begin{aligned} J''[0, \varepsilon]^2 &= a_0 \frac{\varepsilon_1}{2} (-1)^{n_1} L_{\varepsilon_1 g_1} \widehat{\psi}_{\widehat{\tau}_1}(\widehat{\mathbf{x}}_0) - \frac{\varepsilon_3^2}{2} L_{g_3} \widehat{\psi}_T(\widehat{\mathbf{x}}_0) + a_2 \varepsilon_3^2 \sum_{i=1}^{n_3} (-1)^i L_{g_3} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0) \\ &+ \frac{1}{2} \int_{\widehat{I}_3} L_{\varepsilon_1 g_1 + \varepsilon_2 g_2}^2 \left| \widehat{\psi}_t \right|(\widehat{\mathbf{x}}_0) dt + \frac{\varepsilon_1 \varepsilon_2}{2} \int_{\widehat{I}_3} L_{[g_1, g_2]} \left| \widehat{\psi}_t \right|(\widehat{\mathbf{x}}_0) dt + \frac{1}{2} \sum_{1 \leq i < j \leq 3} \varepsilon_i \varepsilon_j L_{[g_i, g_j]} \widehat{\beta}(\widehat{\mathbf{x}}_0). \end{aligned} \quad (26)$$

on the triples $\varepsilon \in \mathbb{R}^3$ satisfying

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0 \quad \varepsilon_1 g_1(\widehat{\mathbf{x}}_0) + \varepsilon_2 g_2(\widehat{\mathbf{x}}_0) + \varepsilon_3 g_3(\widehat{\mathbf{x}}_0) = 0.$$

In particular, this means that, if the two vectors $(g_3 - g_2)(\widehat{\mathbf{x}}_0)$ and $(g_2 - g_1)(\widehat{\mathbf{x}}_0)$ are linearly independent, then the space V_0 is trivial, so that the second variation is coercive by definition. Else, V_0 is a one-dimensional linear space.

3 Hamiltonian formulation

3.1 Construction of the maximised Hamiltonian flow

The maximised Hamiltonian is well defined, continuous and piecewise smooth in T^*M . On the contrary, its associated Hamiltonian vector field is not well defined on the switching surfaces and on the zero level set of Ψ . The scope of this section is to prove that, nevertheless, the flow associated with the maximised Hamiltonian is well defined, at least in a tubular neighbourhood of the graph of the reference extremal. This property relies on the fact that the reference extremal crosses transversally the hypersurfaces of discontinuity of the Hamiltonian vector field.

Fix $\varepsilon > 0$; then on a suitable neighbourhood of the range of $\widehat{\lambda}|_{[0, \widehat{\tau}_1 - \varepsilon]}$ the maximised Hamiltonian vector field is ill-defined on the zero level set of Ψ , else it coincides with $\vec{F}_0 + u_1 \vec{F}_1 - \text{sgn}(\psi) \vec{\Psi}$. By Assumption 1, in this neighbourhood the set $\{\Psi = 0\}$ has n_1 connected components, transversal to $\widehat{\lambda}(\cdot)$. By means of the implicit function theorem, we can express the hitting times of the integral curves of $\vec{F}_0 + u_1 \vec{F}_1 - \text{sgn}(\psi) \vec{\Psi}$ with each connected component as a smooth function of their starting point.

Proposition 1 *There exists a neighbourhood \mathcal{U} of $\widehat{\ell}_0$ in T^*M such that for any $\ell \in \mathcal{U}$ there exists a unique $s_{11} = s_{11}(\ell)$ such that $s_{11}(\widehat{\ell}_0) = \widehat{s}_{11}$ and*

$$\Psi \circ \exp(s_{11}(\ell) \vec{H}_1^{\sigma_{01}})(\ell) = 0 \quad \forall \ell \in \mathcal{U}.$$

Proof. Since $\Psi \circ \exp(t \vec{H}_1^{\sigma_{01}})(\ell) = \psi \circ \exp(th_1)(\pi \ell)$, it suffices to notice that

$$\frac{\partial}{\partial t} \psi \circ \exp(th_1)(\pi \ell) \Big|_{(\widehat{s}_{11}, \widehat{\ell}_0)} = L_{h_1} \psi \left(\widehat{\xi}(\widehat{s}_{11}) \right).$$

Thus, Assumption 1 yields the result. \square

The curves $\exp(t \vec{H}_1^{\sigma_{01}})(\ell)$ cross transversally the zero level set of Ψ (thanks to the fact that they project on the integral curves of $f_0 + u_1 f_1$ and to Assumption 1), so that for t in a sufficiently small right neighbourhood of $s_{11}(\ell)$ the maximised Hamiltonian on the points $\exp(t \vec{H}_1^{\sigma_{01}})(\ell)$ is

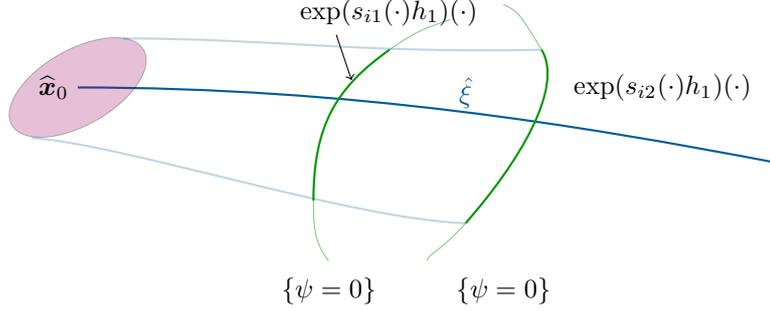


Figure 1: The times s_{1i} , $i = 1, 2$

$H_1^{\sigma_{02}}$, that is, $\overrightarrow{H_1^{\sigma_{01}}}$ is no more the vector field associated with the maximised Hamiltonian. On the other hand, $H_1^{\sigma_{02}}$ is the maximised Hamiltonian also at the points

$$\exp((t - s_{11}(\ell))\overrightarrow{H_1^{\sigma_{02}}}) \circ \exp(s_{11}(\ell)\overrightarrow{H_1^{\sigma_{01}}})(\ell), \quad \ell \in \mathcal{U}. \quad (27)$$

Then, for t in a right neighbourhood of $s_{11}(\ell)$, the maximised Hamiltonian flow is (27).

Iterating the same argument as above, we can prove the following

Proposition 2 *Possibly shrinking the neighbourhood \mathcal{U} , for every $i = 1, \dots, n_1$ there exists a unique smooth function $s_{1i}: \mathcal{U} \rightarrow \mathbb{R}$ such that $s_{1i}(\widehat{\ell}_0) = \widehat{s}_{1i}$ and*

$$\Psi \circ \exp(s_{1i}(\ell) - s_{1i-1}(\ell))\overrightarrow{H_1^{\sigma_{0i}}} \circ \dots \circ \exp(s_{11}(\ell)\overrightarrow{H_1^{\sigma_{01}}})(\ell) = 0 \quad \forall \ell \in \mathcal{U}.$$

Moreover, the differential of s_{1i} at $\widehat{\ell}_0$ is given by

$$\langle ds_{1i}(\widehat{\ell}_0), \delta\ell \rangle = - \frac{L_{\pi_* \delta\ell} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)}{L_{g_1} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)}.$$

Taking advantage of the functions \widehat{s}_{1i} defined here above, on the set $\{(t, \ell) \in \mathbb{R} \times \mathcal{U} : 0 \leq t \leq s_{1n_1}(\ell)\}$ we can write the maximised Hamiltonian flow as the following concatenation:

$$\mathcal{H}_t(\ell) = \exp((t - s_{1i-1}(\ell))\overrightarrow{H_1^{\sigma_{0i}}}) \circ \mathcal{H}_{s_{1i-1}(\ell)}(\ell) \quad t \in [s_{1i-1}(\ell), s_{1i}(\ell)], \quad i = 1, \dots, n_1, \quad (28)$$

where we defined $s_{10}(\ell) \equiv \widehat{s}_{10} = 0$.

Remark 7 *Actually, the functions s_{1i} depend only on the projection $\pi\ell$. We write them as functions of ℓ for symmetry with other functions s_{3i} that will be defined below.*

Reasoning as above, we see that, for t belonging on a right neighbourhood of $s_{1n_1}(\ell)$, the maximised Hamiltonian at the points $\exp(t - s_{1n_1}(\ell))\overrightarrow{H_1^{\sigma_{0n_1+1}}} \circ \mathcal{H}_{s_{1n_1}(\ell)}(\ell)$ is $H_1^{\sigma_{0n_1+1}}$; in particular, $H_1^{\sigma_{0n_1+1}}$ is the maximised Hamiltonian along its integral curves until such curves intersect the hypersurface $\{H_2 - H_1^{\sigma_{0n_1+1}} = 0\}$. As above, thanks to the regularity assumptions (in this case Assumption 4), we can characterise the intersection time as a smooth function of the initial point ℓ .

Proposition 3 *Possibly shrinking \mathcal{U} , for any $\ell \in \mathcal{U}$ there exists a unique $\tau_1 = \tau_1(\ell)$ such that $\tau_1(\widehat{\ell}_0) = \widehat{\tau}_1$ and*

$$(H_2 - H_1^{\sigma_{0n_1+1}}) \circ \exp((\tau_1(\ell) - s_{1n_1}(\ell))\overrightarrow{H_1^{\sigma_{0n_1+1}}}) \circ \mathcal{H}_{s_{1n_1}(\ell)}(\ell) = 0$$

Moreover, $\tau_1(\ell)$ is smooth and its differential at $\widehat{\ell}_0$ is

$$\begin{aligned} \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle &= \frac{1}{\sigma(\overrightarrow{H_1^{\sigma_0 n_1+1}}, \overrightarrow{H_2})(\widehat{\ell}_1)} \\ &\times \left\{ -\sigma(\widehat{\mathcal{H}}_{\widehat{\tau}_1^*} \delta\ell, (\overrightarrow{H_2} - \overrightarrow{H_1^{\sigma_0 n_1+1}})(\widehat{\ell}_1)) + 2 \sum_{i=1}^{n_1} \sigma_{0i} \frac{L_{\pi^* \delta\ell} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0) L_{g_2-g_1} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)}{L_{g_1} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)} \right\} \end{aligned} \quad (29)$$

Proof. The proof is analogous to that of Proposition 1. It suffices to notice that

$$\frac{\partial}{\partial t} (H_2 - H_1^{\sigma_0 n_1+1}) \circ \exp((t - s_{1n_1}(\ell)) \overrightarrow{H_1^{\sigma_0 n_1+1}}) \circ \mathcal{H}_{s_{1n_1}(\ell)}(\ell) \Big|_{(\widehat{\tau}_1, \widehat{\ell}_0)} = \sigma(\overrightarrow{H_1^{\sigma_0 n_1+1}}, \overrightarrow{H_2})(\widehat{\ell}_1),$$

which is positive by Assumption 4. \square

We extend the maximised Hamiltonian flow \mathcal{H}_t to the whole interval $[0, \tau_1(\ell)]$ setting

$$\mathcal{H}_t(\ell) = \exp((t - s_{1n_1}(\ell)) \overrightarrow{H_1^{\sigma_0 n_1+1}}) \circ \mathcal{H}_{s_{1n_1}(\ell)}(\ell), \quad t \in [s_{1n_1}(\ell), \tau_1(\ell)].$$

Analogously, Assumption 4 implies that the function

$$(F_1(\ell) - |\psi(\pi\ell)|) \circ \exp((t - s_{1n_1}(\ell)) \overrightarrow{H_1^{\sigma_0 n_1+1}}) \circ \mathcal{H}_{s_{1n_1}(\ell)}(\ell)$$

is strictly decreasing with respect to t , for t belonging to a neighbourhood of $\tau_1(\ell)$, so that $H_1^{\sigma_0 n_1+1}$ cannot be the maximised Hamiltonian on $\exp((t - s_{1n_1}(\ell)) \overrightarrow{H_1^{\sigma_0 n_1+1}}) \circ \mathcal{H}_{s_{1n_1}(\ell)}(\ell)$ if $t > \tau_1(\ell)$. Indeed, the maximised Hamiltonian flow is

$$\exp((t - \tau_1(\ell)) \overrightarrow{H_2}) \circ \mathcal{H}_{\tau_1(\ell)}(\ell).$$

Along these curves, for $t > \tau_1(\ell)$, H_2 is the maximised Hamiltonian until the hypersurface $\{H_2 - H_3^{\sigma_{21}} = 0\}$ is reached. The time for this to happen is, as above, a smooth function of ℓ , thanks to Assumption 5. The proof is completely analogous to that of Proposition 3.

Proposition 4 *Possibly shrinking \mathcal{U} , for every $\ell \in \mathcal{U}$ there exists a unique $\tau_2 = \tau_2(\ell)$ such that $\tau_2(\widehat{\ell}_0) = \widehat{\tau}_2$ and*

$$(H_3^{\sigma_{21}} - H_2) \circ \exp(\tau_2(\ell) - \tau_1(\ell)) \overrightarrow{H_2} \circ \mathcal{H}_{\tau_1(\ell)}(\ell) = 0.$$

Moreover, the differential of $\tau_2(\ell)$ at $\widehat{\ell}_0$ is given by

$$\begin{aligned} \langle d\tau_2(\widehat{\ell}_0), \delta\ell \rangle &= \frac{1}{\sigma(\overrightarrow{H_2}, \overrightarrow{H_3^{\sigma_{21}}})(\widehat{\ell}_2)} \left\{ -\sigma(\widehat{\mathcal{H}}_{\widehat{\tau}_2^*} \delta\ell, (\overrightarrow{H_3^{\sigma_{21}}} - \overrightarrow{H_2})(\widehat{\ell}_2)) \right. \\ &+ \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle \sigma(\widehat{\mathcal{H}}_{\widehat{\tau}_2^*} \widehat{\mathcal{H}}_{\widehat{\tau}_1^*}^{-1}(\overrightarrow{H_2} - \overrightarrow{H_1^{\sigma_0 n_1+1}})(\widehat{\ell}_1), \overrightarrow{H_3^{\sigma_{21}}} - \overrightarrow{H_2})(\widehat{\ell}_2) \\ &\left. + 2 \sum_{i=1}^{n_1} \sigma_{0i} \frac{L_{\pi^* \delta\ell} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0) L_{g_3-g_2} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)}{L_{g_1} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)} \right\}. \end{aligned} \quad (30)$$

We thus extend the maximised Hamiltonian flow to the interval $[0, \tau_2(\ell)]$ setting

$$\mathcal{H}_t(\ell) = \exp((t - \tau_1(\ell)) \overrightarrow{H_2}) \circ \mathcal{H}_{\tau_1(\ell)}(\ell), \quad t \in [\tau_1(\ell), \tau_2(\ell)].$$

The construction of the maximised Hamiltonian flow on the whole interval $[0, T]$ follows the same lines: we characterise the discontinuities of the vector field as smooth functions of the initial state, and then we concatenate the corresponding Hamiltonian flows. Regularity assumptions are, as usual, crucial for this.

Proposition 5 *Possibly shrinking \mathcal{U} , for every $i = 1, \dots, n_3$ there exist a unique smooth function $s_{3i}: \mathcal{U} \rightarrow \mathbb{R}$ such that $s_{3i}(\widehat{\ell}_0) = \widehat{s}_{3i}$ and*

$$\Psi \circ \exp(s_{3i}(\ell) - s_{3i}(\widehat{\ell}_0)) \overrightarrow{H_3^{\sigma_{2i}}} \circ \dots \circ \exp((s_{31}(\ell) - \tau_2(\ell)) \overrightarrow{H_3^{\sigma_{21}}}) \circ \mathcal{H}_{\tau_2(\ell)}(\ell) = 0 \quad \forall \ell \in \mathcal{U}.$$

Moreover, the differential of s_{3i} at $\widehat{\ell}_0$ is given by

$$\begin{aligned} \langle ds_{3i}(\widehat{\ell}_0), \delta \ell \rangle &= \frac{-1}{L_{g_3} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0)} \left\{ L_{\pi_* \delta \ell} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0) - \langle d\tau_1(\widehat{\ell}_0), \delta \ell \rangle L_{g_2 - g_1} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0) \right. \\ &\quad \left. - \langle d\tau_2(\widehat{\ell}_0), \delta \ell \rangle L_{g_3 - g_2} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0) \right\}, \quad i = 1, \dots, n_3. \end{aligned}$$

We finally define the maximised flow, for initial conditions in \mathcal{U} , on the whole $[0, T]$:

$$\mathcal{H}_i(\ell) = \exp((t - s_{3i-1}(\ell)) \overrightarrow{H_3^{\sigma_{2i}}}) \circ \mathcal{H}_{s_{3i-1}(\ell)}(\ell) \quad t \in [s_{3i-1}(\ell), s_{3i}(\ell)], \quad i = 1, \dots, n_3 + 1, \quad (31)$$

where we put $s_{30}(\ell) := \tau_2(\ell)$ and $s_{3n_3+1}(\ell) \equiv T$.

3.2 Hamiltonian form of the second variation

In this section we propose an alternative representation of the second variation, more compact and easier to handle with. To do that, we establish an isomorphism between $T_{\widehat{\mathbf{x}}_0}^* M \times T_{\widehat{\mathbf{x}}_0} M$ and $T_{\widehat{\ell}_0}(T^* M)$, and we map the Hamiltonians defined in (9) to some Hamiltonian functions G_i'' defined on $T_{\widehat{\mathbf{x}}_0}^* M \times T_{\widehat{\mathbf{x}}_0} M$; we then express the second variation J'' in terms of these Hamiltonians. The new expression of the second variation highlights its links with the Hamiltonian vector fields and with the maximised Hamiltonian flow.

First, we define the following anti-symplectic isomorphism¹ between $T_{\widehat{\mathbf{x}}_0}^* M \times T_{\widehat{\mathbf{x}}_0} M$ and $T_{\widehat{\ell}_0}(T^* M)$:

$$\iota(\delta p, \delta x) = (-\delta p + A[\delta x, \cdot], \delta x),$$

where A is the symmetric bilinear form on $T_{\widehat{\mathbf{x}}_0} M$ defined by

$$\begin{aligned} A[\delta x, \delta y] &:= D^2 \left(-\widehat{\beta} - \int_{\widehat{I}_1 \cup \widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) [\delta x, \delta y] - 2a_0 \sum_{i=1}^{n_1} (-1)^i \frac{L_{\delta x} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0) L_{\delta y} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)}{L_{g_1} \widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)} \\ &\quad - 2a_2 \sum_{i=1}^{n_3} (-1)^i \frac{L_{\delta x} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0) L_{\delta y} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0)}. \end{aligned}$$

We then set

$$\overrightarrow{G_1''} := \iota^{-1} \overrightarrow{H_1^{\sigma_{01}}}(\widehat{\ell}_0), \quad \overrightarrow{G_2''} = \iota^{-1} \widehat{\mathcal{H}}_{\tau_1}^{-1} \overrightarrow{H_2}(\widehat{\ell}_1), \quad \overrightarrow{G_3''} := \iota^{-1} \widehat{\mathcal{H}}_{\tau_2}^{-1} \overrightarrow{H_3^{\sigma_{21}}}(\widehat{\ell}_2). \quad (32)$$

By computation (see Appendix A), one can see that each $\overrightarrow{G_i''}$ is the constant Hamiltonian vector field associated with the following linear Hamiltonian functions:

$$\begin{aligned} G_1''(\delta p, \delta x) &= \langle \delta p, g_1(\widehat{\mathbf{x}}_0) \rangle + L_{\delta x} L_{g_1} \left(\widehat{\beta} + \int_{\widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) + a_0 (-1)^{n_1} L_{\delta x} \widehat{\psi}_{\widehat{\tau}_1}(\widehat{\mathbf{x}}_0) \\ &\quad + 2a_2 \sum_{i=1}^{n_3} (-1)^i \frac{L_{g_1} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0) L_{\delta x} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0)}, \end{aligned} \quad (33)$$

$$\begin{aligned} G_2''(\delta p, \delta x) &= \langle \delta p, g_2(\widehat{\mathbf{x}}_0) \rangle + L_{\delta x} L_{g_2} \left(\widehat{\beta} + \int_{\widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) \\ &\quad + 2a_2 \sum_{i=1}^{n_3} (-1)^i \frac{L_{g_2} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0) L_{\delta x} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{s}_{3i}}(\widehat{\mathbf{x}}_0)}, \end{aligned} \quad (34)$$

$$G_3''(\delta p, \delta x) = \langle \delta p, g_3(\widehat{\mathbf{x}}_0) \rangle + L_{\delta x} L_{g_3} \widehat{\beta}(\widehat{\mathbf{x}}_0) + a_2 (-1)^{n_3} L_{\delta x} \widehat{\psi}_T(\widehat{\mathbf{x}}_0). \quad (35)$$

¹in particular, if ς is the standard symplectic form on the product $T_{\widehat{\mathbf{x}}_0}^* M \times T_{\widehat{\mathbf{x}}_0} M$, then $\varsigma = -\sigma \circ (\iota, \iota)$.

By means of rather tedious computations, it is possible to prove the following equalities (details in Appendix A).

$$\begin{aligned} G_2''(\overrightarrow{G_1''}) &= L_{[g_1, g_2]} \left(\widehat{\beta} + \int_{\widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) - a_0 (-1)^{n_1} L_{g_2} \widehat{\psi}_{\widehat{\tau}_1} (\widehat{\mathbf{x}}_0) \\ &= -\sigma \left(\overrightarrow{H_1}^{\sigma_0 n_1 + 1}, \overrightarrow{H_2} \right) (\widehat{\ell}_1) \end{aligned} \quad (36)$$

$$\begin{aligned} G_3''(\overrightarrow{G_2''}) &= L_{[g_2, g_3]} \left(\widehat{\beta} + \int_{\widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) + a_2 L_{g_2} \widehat{\psi}_{\widehat{\tau}_2} (\widehat{\mathbf{x}}_0) \\ &= -\sigma \left(\overrightarrow{H_2}, \overrightarrow{H_3}^{\sigma_{21}} \right) (\widehat{\ell}_2) \end{aligned} \quad (37)$$

$$\begin{aligned} G_3''(\overrightarrow{G_1''}) &= L_{[g_1, g_3]} \left(\widehat{\beta} + \int_{\widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) - a_0 (-1)^{n_1} L_{g_3} \widehat{\psi}_{\widehat{\tau}_1} (\widehat{\mathbf{x}}_0) + a_2 L_{g_1} \widehat{\psi}_{\widehat{\tau}_2} (\widehat{\mathbf{x}}_0) \\ &= -\sigma \left(\widehat{\mathcal{H}}_{\widehat{\tau}_2} \widehat{\mathcal{H}}_{\widehat{\tau}_1}^{-1} \overrightarrow{H_1}^{\sigma_1 n_1 + 1}, \overrightarrow{H_3}^{\sigma_{21}} \right) (\widehat{\ell}_2) \end{aligned} \quad (38)$$

We also define $\omega_0(\delta x, \cdot)$ such that $\iota^{-1} d\alpha_* \delta x = (\omega_0(\delta x, \cdot), \delta x)$. By computations

$$\begin{aligned} \omega_0(\delta x, \cdot) &= -D^2 \left(\alpha + \widehat{\beta} + \int_{\widehat{I}_1 \cup \widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) [\delta x, \cdot] \\ &\quad - 2a_0 \sum_{i=1}^{n_1} (-1)^i \frac{L_{\delta x} \widehat{\psi}_{\widehat{s}_{1i}} (\widehat{\mathbf{x}}_0) L_{\widehat{\psi}_{\widehat{s}_{1i}}} (\widehat{\mathbf{x}}_0)}{L_{g_1} \widehat{\psi}_{\widehat{s}_{1i}} (\widehat{\mathbf{x}}_0)} - 2a_2 \sum_{i=1}^{n_3} (-1)^i \frac{L_{\delta x} \widehat{\psi}_{\widehat{s}_{3i}} (\widehat{\mathbf{x}}_0) L_{\widehat{\psi}_{\widehat{s}_{3i}}} (\widehat{\mathbf{x}}_0)}{L_{g_3} \widehat{\psi}_{\widehat{s}_{3i}} (\widehat{\mathbf{x}}_0)}. \end{aligned}$$

In particular we can write equations (29)-(30) and the second variation J'' in the following form:

$$\langle d\tau_1(\widehat{\ell}_0), d\alpha_* \delta x \rangle = \frac{-1}{G_2''(\overrightarrow{G_1''})} (G_2'' - G_1'') (\omega_0(\delta x), \delta x) \quad (39)$$

$$\langle d\tau_2(\widehat{\ell}_0), d\alpha_* \delta x \rangle = \frac{-1}{G_3''(\overrightarrow{G_2''})} (G_3'' - G_2'') \left((\omega_0(\delta x), \delta x) - \langle d\tau_1(\widehat{\ell}_0), d\alpha_* \delta x \rangle (\overrightarrow{G_2''} - \overrightarrow{G_1''}) \right) \quad (40)$$

$$\begin{aligned} J''[\delta e]^2 &= \frac{1}{2} \left(\varepsilon_1 G_1'' (\omega_0(\delta x), \delta x) + \varepsilon_2 G_2'' ((\omega_0(\delta x), \delta x) + \varepsilon_1 \overrightarrow{G_1''}) \right. \\ &\quad \left. + \varepsilon_3 G_3'' ((\omega_0(\delta x), \delta x) + \varepsilon_1 \overrightarrow{G_1''} + \varepsilon_2 \overrightarrow{G_2''}) \right) \end{aligned} \quad (41)$$

for every $\delta e = (\delta x, \varepsilon) \in V$. The proof is just a straightforward application of the definitions.

We end this section by writing two conditions that are equivalent to the coercivity of J'' on V . These conditions take advantage of the Hamiltonian formulation of the second variation described here above. We recall that, given any linear subspace $W \subset V$, then J'' is coercive on V if and only if it is coercive both on W and on the orthogonal complement to W with respect to the bilinear symmetric form \mathfrak{J} associated with J'' . By lengthy but standard computations, see for example [16], and setting

$$\gamma''[\delta x, \delta y] := \langle -\omega_0(\delta x), \delta y \rangle$$

it is possible to see that

$$\begin{aligned} \mathfrak{J}[\delta e, \delta f] &= \frac{1}{2} \left(\gamma''[\delta x, \delta y] + \langle \omega_0(\delta x), \delta y \rangle + \eta_1 G_1'' (\omega_0(\delta x), \delta x) + \eta_2 G_2'' ((\omega_0(\delta x), \delta x) + \varepsilon_1 \overrightarrow{G_1''}) \right. \\ &\quad \left. + \eta_3 G_3'' ((\omega_0(\delta x), \delta x) + \varepsilon_1 \overrightarrow{G_1''} + \varepsilon_2 \overrightarrow{G_2''}) \right). \end{aligned}$$

For the sake of forthcoming computations, we choose W as the linear space

$$W = \{ \delta e = (\delta x, \varepsilon) \in V : \varepsilon_3 = 0 \}. \quad (42)$$

By computations, it is easy to prove that

$$W^{\perp \mathfrak{J}} = \{ (\delta x, \varepsilon) \in V : \varepsilon_1 = \langle d\tau_1(\widehat{\ell}_0), d\alpha_* \delta x \rangle \}, \quad (43)$$

$$J''[\delta e]^2 = \frac{\varepsilon_3}{2} (G_3'' - G_2'') ((\omega_0(\delta x), \delta x) + \varepsilon_1 \overrightarrow{G_1''} + \varepsilon_2 \overrightarrow{G_2''}) \quad \forall \delta e = (\delta x, \varepsilon) \in W^{\perp \mathfrak{J}}. \quad (44)$$

4 Main result

In this section we state and prove the main result of the paper.

Theorem 1 *Let $\widehat{\xi}: [0, T] \rightarrow M$ be an admissible trajectory for the control system (1b)-(1c)-(1d) that satisfies Assumptions 1–6. Then the trajectory $\widehat{\xi}$ is a strict strong-local minimiser for the OCP (1).*

The proof is inspired by those of [12, Theorem 3.9] and [14, Theorem 4.1], and is based on the following result.

Theorem 2 *Assume that the assumptions of Theorem 1 are satisfied. Then there exist a neighbourhood U of \widehat{x}_0 such that the set*

$$\Lambda_0 = \{d\alpha(x) : x \in U\} \quad (45)$$

is a smooth simply-connected Lagrangian submanifold that contains $\widehat{\ell}_0$ and, for every $t \in [0, T]$, the map

$$\pi \circ \mathcal{H}_t|_{\Lambda_0} \quad (46)$$

is invertible onto a neighbourhood of $\widehat{\xi}(t)$ with piecewise- C^1 inverse.

Proof. (of Theorem 1) First of all, we define the following subsets of $\mathbb{R} \times T^*M$

$$\begin{aligned} \mathcal{O}_{1i} &= \{(t, \ell) : \ell \in \mathcal{U}, s_{1i-1}(\ell) \leq t \leq s_{1i}(\ell)\} \quad i = 1, \dots, n_1 + 1, \\ \mathcal{O}_2 &= \{(t, \ell) : \ell \in \mathcal{U}, \tau_1(\ell) \leq t \leq \tau_2(\ell)\}, \\ \mathcal{O}_{3i} &= \{(t, \ell) : \ell \in \mathcal{U}, s_{3i-1}(\ell) \leq t \leq s_{3i}(\ell)\} \quad i = 1, \dots, n_3 + 1, \end{aligned}$$

and the flow $\mathcal{H}: [0, T] \times \mathcal{U} \rightarrow \mathbb{R} \times T^*M$

$$\mathcal{H}(t, \ell) = (t, \mathcal{H}_t(\ell)).$$

We also define the sets $\Omega_{ij} = \mathcal{O}_{ij} \cap (\mathbb{R} \times \Lambda_0)$, $i = 1, 3$, and $\Omega_2 = \mathcal{O}_2 \cap (\mathbb{R} \times \Lambda_0)$, and we call Ω the union of all these sets. We remark that the restriction of \mathcal{H} to each of the \mathcal{O}_{ij} , \mathcal{O}_2 (as well as to each of the Ω_{ij} , Ω_2) is smooth. Moreover, thanks to Theorem 2, the map $\pi\mathcal{H}: (t, \ell) \mapsto (t, \pi \circ \mathcal{H}(t, \ell))$ is invertible with piecewise- C^1 inverse. The points of non differentiability occur when (t, ℓ) belongs to the intersection of two of the Ω_{ij} , Ω_2 . Indeed, we notice that

$$\pi \circ \mathcal{H}_t(\ell) = \begin{cases} \exp(th_1)(\ell) & t \in [0, \tau_1(\ell)], \\ \exp((t - \tau_1(\ell))h_2) \circ \exp(\tau_1(\ell)h_1)(\ell) & t \in [\tau_1(\ell), \tau_2(\ell)], \\ \exp((t - \tau_2(\ell))h_3) \circ \exp((\tau_2(\ell) - \tau_1(\ell))h_2) \circ \exp(\tau_1(\ell)h_1)(\ell) & t \in [\tau_2(\ell), T], \end{cases}$$

so that, at the switching times, the piecewise linearisation of $\pi \circ \mathcal{H}_t$ is given by

$$\pi_* \mathcal{H}_{\widehat{\tau}_1^*}(\delta\ell) = \begin{cases} \widehat{S}_{\widehat{\tau}_1^*}(\pi_* \delta\ell) & \text{for } \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle \geq 0, \\ \widehat{S}_{\widehat{\tau}_1^*}(\pi_* \delta\ell + \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle (g_1 - g_2)(\widehat{x}_0)) & \text{for } \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle \leq 0, \end{cases} \quad (47)$$

$$\pi_* \mathcal{H}_{\widehat{\tau}_2^*}(\delta\ell) = \begin{cases} \widehat{S}_{\widehat{\tau}_2^*}(\pi_* \delta\ell + \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle (g_1 - g_2)(\widehat{x}_0)) & \text{for } \langle d\tau_2(\widehat{\ell}_0), \delta\ell \rangle \geq 0, \\ \widehat{S}_{\widehat{\tau}_2^*}(\pi_* \delta\ell + \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle (g_1 - g_2)(\widehat{x}_0) \\ \quad + \langle d\tau_2(\widehat{\ell}_0), \delta\ell \rangle (g_2 - g_3)(\widehat{x}_0)) & \text{for } \langle d\tau_2(\widehat{\ell}_0), \delta\ell \rangle \leq 0. \end{cases} \quad (48)$$

Let $\omega_{(t,\ell)} = s_\ell - \max_{|u| \leq 1} \mathfrak{h}(u, \ell) dt$ be the Poincaré-Cartan one-form (in the following we omit the dependence on the basepoint (t, ℓ)). Lemma 3.3 in [12] guarantees that the one-form $\mathcal{H}^* \omega$ is exact on each of the Ω_{ij} , Ω_2 , and therefore on the whole Ω .

Let now $\xi: [0, T] \rightarrow M$ be any admissible trajectory of the control system (1b)-(1c) whose graph is contained in $\pi\mathcal{H}(\Omega)$, and call $v(t)$ its associated control function. Define moreover $\lambda_0(t)$, $\lambda(t)$

by the equalities $(t, \lambda_0(t)) := (\pi \circ \mathcal{H})^{-1}(t, \xi(t))$ and $\lambda(t) := \mathcal{H}(t, \lambda_0(t))$; in particular, $\pi\lambda(t) = \xi(t)$. Consider the two paths in Ω

$$\gamma = \{(t, \lambda_0(t)) : t \in [0, T]\} \quad \widehat{\gamma} = \{(t, \widehat{\ell}_0) : t \in [0, T]\}.$$

We notice that the concatenation of $\widehat{\gamma}(\cdot)$ with $\gamma(T - \cdot)$ is a closed path in $[0, T] \times \Lambda_0$, so that $\int_{\widehat{\gamma}} \mathcal{H}^* \omega = \int_{\gamma} \mathcal{H}^* \omega$. In particular, $\int_{\widehat{\gamma}} \mathcal{H}^* \omega = \int_0^T |\widehat{u}(t)\psi(\widehat{\xi}(t))| dt$, while

$$\int_{\gamma} \mathcal{H}^* \omega = \int_0^T s_{\lambda(t)} - \max_{|u| \leq 1} \mathfrak{h}(u, \lambda(t)) dt \leq \int_0^T |v(t)\psi(\xi(t))| dt, \quad (49)$$

and this implies that $\widehat{\xi}$ is a strong-local minimiser.

Let us now assume that $\int_0^T |\widehat{u}(t)\psi(\widehat{\xi}(t))| dt = \int_0^T |v(t)\psi(\xi(t))| dt$, that is, the equality holds in equation (49). This implies that

$$v(t)F_1(\lambda(t)) - |v(t)\psi(\xi(t))| = \max_{w \in [-1, 1]} (wF_1(\lambda(t)) - |w\psi(\xi(t))|) \quad \text{a.e. } t \in [0, T]. \quad (50)$$

By continuity, for t small enough $(t, \lambda_0(t))$ belongs to Ω_{11} , hence $wF_1(\lambda(t)) - |w\psi(\xi(t))|$ attains its maximum only for $w = u_1$; equation (50) yields that $v(t) = u_1$ a.e. and $\xi(t) = \widehat{\xi}(t)$ as long as $(t, \lambda_0(t)) \in \Omega_{11}$, that is, for $t \in [0, \widehat{s}_{11}]$. For t in a sufficiently small right neighbourhood of \widehat{s}_{11} , $(t, \xi(t))$ belongs to Ω_{11} or Ω_{12} ; in both cases, (50) implies that $v(t) = u_1$. We can proceed iteratively and obtain that $v(t) = u_1$ a.e. and $\xi(t) = \widehat{\xi}(t)$ for $t \in [0, \widehat{\tau}_1]$.

For t in a sufficiently small right neighbourhood of $\widehat{\tau}_1$, three cases are possible: $(t, \lambda_0(t))$ may belong to $\Omega_2 \setminus \Omega_{1n_1+1}$, to $\Omega_{1n_1+1} \setminus \Omega_2$, or to the intersection $\Omega_{1n_1+1} \cap \Omega_2$.

In the first case, the maximised Hamiltonian is attained for $w = 0$, so that, reasoning as above, we obtain that $v(t) = 0$ and then $\xi(t) = \widehat{\xi}(t)$ for $t \leq \widehat{\tau}_2$. If $(t, \lambda_0(t)) \in \Omega_{1n_1+1} \setminus \Omega_2$, the maximised Hamiltonian is attained for $w = u_1$ and then (50) yields that $v(t) = u_1$. This is impossible, since, by Assumption 4 and by continuity, in a neighbourhood of $\widehat{\ell}_1$ the function $F_1(\ell) - |\psi(\pi\ell)|$ is strictly decreasing along the integral lines of $\overrightarrow{H_1^{\sigma_{n_1+1}}}$.

In the last case, for t in a sufficiently small right neighbourhood of $\widehat{\tau}_1$ it holds $t = \tau_1(\lambda_0(t))$ which implies that

$$1 = \langle d\tau_1(\lambda_0(t)), \dot{\lambda}_0(t) \rangle \quad \text{for a.e. } t. \quad (51)$$

Moreover (50) implies that $v(t)$ has the same sign of u_1 , so that $f_0 + v(t)f_1$ is a convex combination of h_1 and h_2 , and there exists some $\mu(t) \in [0, 1]$ such that $\dot{\xi}(t) = \mu(t)h_1(\xi(t)) + (1 - \mu(t))h_2(\xi(t))$.

By computations $\dot{\xi}(t) = h_1(\xi(t)) + (\pi \circ \mathcal{H}_t)_* \dot{\lambda}_0(t)$ for a.e. t , so that

$$\pi_* \dot{\lambda}_0(t) = (1 - \mu(t))\pi_*(\pi \circ \mathcal{H}_t)_*^{-1}(h_2 - h_1)(\xi(t)) \quad \text{a.e. } t. \quad (52)$$

By compactness, there exists a sequence $t_n \rightarrow \widehat{\tau}_1^+$ such that $\mu(t_n) \rightarrow \bar{\mu} \in [0, 1]$. Passing to the limit in (51)-(52) we obtain that

$$\pi_* \dot{\lambda}_0(t_n) \rightarrow (1 - \bar{\mu})(g_2 - g_1)(\widehat{\mathbf{x}}_0)$$

and $1 = (1 - \bar{\mu})\langle d\tau_1(\widehat{\ell}_0), d\alpha_*(g_2 - g_1)(\widehat{\mathbf{x}}_0) \rangle$. In particular, $\bar{\mu} \in [0, 1)$ and the variation $\delta e = ((g_2 - g_1)(\widehat{\mathbf{x}}_0), 1, -1, 0)$ belongs to W . Thus, by (41) and (39) we obtain that

$$J''[\delta e]^2 = \frac{1}{2} G_2''(G_1'') \frac{\bar{\mu}}{1 - \bar{\mu}},$$

which cannot be positive, due to (13) and (36). This contradicts Assumption 6, thus this case is not possible. Therefore we must conclude that $v(t) = \widehat{u}(t)$ for a.e. $t \in [0, \widehat{\tau}_2]$.

Analogous computations show that $v(t)$ coincides with the reference control almost everywhere in the interval $[0, T]$, that is, $\xi = \widehat{\xi}$. \square

Proof. (of Theorem 2) By construction the manifold Λ_0 defined in (45) is a horizontal Lagrangian submanifold of T^*M containing $\widehat{\ell}_0$.

The map $\pi \circ \mathcal{H}_t$ is the concatenation of smooth invertible mappings (flows). To check its invertibility at $\widehat{\ell}_0$, it is sufficient to consider it at the switching times only: indeed, the map is invertible at $\widehat{\ell}_0$ for every $t < \widehat{\tau}_1$ and, if it is also invertible for $t = \widehat{\tau}_1$, then it is invertible on the whole $[0, \widehat{\tau}_2)$; analogously, if the map is invertible for $t \leq \widehat{\tau}_2$, then it is invertible for every $t \in [0, T]$.

To verify the invertibility at the switching times, we use Clarke's inverse function theorem (see [17]), that is, we prove that all convex combination of the “left” and “right” linearisations (47)–(48) have full rank.

More precisely, for the first switching time, we show by contradiction that there is no $a \in [0, 1]$ and no $\delta\ell \in T_{\widehat{\ell}_0} \Lambda_0$, $\delta\ell \neq 0$, such that

$$(1-a)\widehat{S}_{\widehat{\tau}_1*}(\pi_*\delta\ell) + a\widehat{S}_{\widehat{\tau}_1*}(\pi_*\delta\ell + (g_1 - g_2)(\widehat{\mathbf{x}}_0)\langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle) = 0. \quad (53)$$

Indeed, assume that (53) holds true for some a and some $\delta\ell \neq 0$. Then $\langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle \neq 0$, and, since $\widehat{S}_{\widehat{\tau}_1*}$ is an isomorphism, then

$$\pi_*\delta\ell + a\langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle(g_1 - g_2)(\widehat{\mathbf{x}}_0) = 0$$

i.e. $\delta e = (\pi_*\delta\ell, a\langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle, -a\langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle, 0) \in W$. As before, thanks to (39)–(36) and Assumption 4, it is possible to prove that

$$\begin{aligned} J''[\delta e]^2 &= -\frac{a}{2}\langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle (G_2'' - G_1'')((\omega_0(\delta x), \delta x) + \varepsilon_1 \overrightarrow{G_1''}) \\ &= \frac{a}{2}(1-a)\langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle^2 G_2''(\overrightarrow{G_1''}) \leq 0, \end{aligned}$$

which contradicts the coercivity of the second variation on W .

Analogously, the linearisation of the maximised flow at time $\widehat{\tau}_2$ is invertible if for every $\delta\ell \in T_{\widehat{\ell}_0} \Lambda_0$ and for every $a \in [0, 1]$ satisfying

$$\pi_*\delta\ell + \langle d\tau_1(\widehat{\ell}_0), \delta\ell \rangle(g_1 - g_2)(\widehat{\mathbf{x}}_0) + a\langle d\tau_2(\widehat{\ell}_0), \delta\ell \rangle(g_2 - g_3)(\widehat{\mathbf{x}}_0) = 0 \quad (54)$$

it must be $\delta\ell = 0$. Indeed, if (54) holds, then the variation

$$\delta e = (\delta x, \langle d\tau_1(\widehat{\ell}_0), d\alpha_*\delta x \rangle, \langle d(a\tau_2 - \tau_1)(\widehat{\ell}_0), d\alpha_*\delta x \rangle, -a\langle d\tau_2(\widehat{\ell}_0), d\alpha_*\delta x \rangle)$$

is admissible and it is contained in W^{\perp_3} . Again, using (37), we observe that

$$J''[\delta e]^2 = \frac{a}{2}(1-a)\langle d\tau_2(\widehat{\ell}_0), d\alpha_*\delta x \rangle^2 G_3''(\overrightarrow{G_2''})$$

which cannot be positive due to Assumption 5. This contradicts the coercivity of the second variation on V . The Theorem is proved. \square

5 An example

In this section we apply our result to the following OCP

$$\min_{|u(\cdot)| \leq 1} \int_0^T |u(t)x_2(t)| dt$$

subject to the control system

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u - \alpha x_2 & \alpha > 0 \\ x_1(0) = 0, x_2(0) = 0 \\ x_1(T) = X > 0, x_2(T) = 0. \end{cases}$$

It models the problem of minimising the consumption of an academic electric vehicle moving with friction along a flat road; it has been studied in details in [4].

Let $T_{\min}(\alpha, X)$ be the minimum time needed to reach $(X, 0)$ from $(0, 0)$; in [4] it is shown that there exists a time $T_{\lim}(\alpha, X) > T_{\min}(\alpha, X)$ such that, if $T_{\min}(\alpha, X) < T \leq T_{\lim}(\alpha, X)$, the optimal control for this problem has the form (3), with $u_1 = 1$, $u_3 = -1$ (for the analytic expressions of T_{\lim} and of the switching times as functions of α and X , we refer to [4]). In particular, the cost function $\psi(x_1, x_2) = x_2$ never vanishes along the reference trajectory (except at the endpoints), so in this case we have $a_0 = a_2 = 1$ and $n_1 = n_3 = 0$.

First of all, we compute the pull-back vector fields. Obviously, $g_1 = h_1 = f_0 + f_1$. To compute g_2 and g_3 , we consider the functions

$$\begin{aligned} \varphi_2(x, t) &= \exp(-th_1)_* h_2 \exp(th_1)(x) \\ \varphi_3(x, t, s) &= \exp(-th_1)_* \exp(-sh_2)_* h_3 \circ \circ \exp(sh_2) \circ \exp(th_1)(x), \end{aligned}$$

and notice that $g_2(x) = \varphi_2(x, \hat{\tau}_1)$ and $g_3(x) = \varphi_3(x, \hat{\tau}_1, \hat{\tau}_2)$. The pull-back vector fields can be then computed developing φ_2, φ_3 in powers of t and (t, s) , respectively; we obtain

$$g_2 = f_0 + \frac{\eta(\hat{\tau}_1)}{\alpha} f_{01} \quad g_3 = f_0 - f_1 + \frac{\eta(\hat{\tau}_1) + \eta(\hat{\tau}_2)}{\alpha} f_{01},$$

where $\eta(t) = 1 - e^{\alpha t}$ and $f_{01} = [f_0, f_1]$.

Let $(\delta x, \varepsilon)$ be an admissible variation contained in V_0 ; from $\delta x = (0, 0)$ and the linear independence of f_1 and f_{01} , we obtain that $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ must satisfy the system

$$\begin{cases} \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0 \\ \varepsilon_1 - \varepsilon_3 = 0 \\ \varepsilon_2 \eta(\hat{\tau}_1) + \varepsilon_3 (\eta(\hat{\tau}_1) + \eta(\hat{\tau}_2)) = 0, \end{cases}$$

so that, if $\hat{\tau}_1 \neq \hat{\tau}_2$, then V_0 is the trivial linear space, and the second variation is coercive by definition.

We remark that, for $\hat{\tau}_1 = \hat{\tau}_2$, the reference trajectory satisfies the PMP in the abnormal form, a case which is not covered by our theory.

A Appendix

Here we give some hints for the computations appearing in Section 3.2.

First of all let us recall that for every Hamiltonian vector field \vec{F} and every Hamiltonian flow \mathcal{K}_t , the following identity holds for every fixed t :

$$\mathcal{K}_{t*} \vec{F} = \overline{F \circ \mathcal{K}_t^{-1}}. \quad (55)$$

Proof of equations (32).

We can see that the Hamiltonian vector field associated with G_1'' is

$$\left(\begin{array}{c} -L.L_{g_1} \left(\hat{\beta} + \int_{\hat{I}_3} |\hat{\psi}_s| ds \right) (\hat{\mathbf{x}}_0) - \sigma_{0n_1} L.\hat{\psi}_{\hat{\tau}_1} (\hat{\mathbf{x}}_0) - 2 \sum_{i=1}^{n_3} \sigma_{2i} L.\hat{\psi}_{\hat{s}_{3i}} (\hat{\mathbf{x}}_0) L_{g_1} \hat{\psi}_{\hat{s}_{3i}} (\hat{\mathbf{x}}_0) L_{g_3} \hat{\psi}_{\hat{s}_{3i}} (\hat{\mathbf{x}}_0) \\ g_1(\hat{\mathbf{x}}_0) \end{array} \right).$$

Applying the antisymplectic isomorphism ι to the vector here above we obtain

$$\left(L.L_{g_1} \left(\widehat{\beta} + \int_{\widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) + \sigma_{0n_1} L.\widehat{\psi}_{\widehat{\tau}_1} (\widehat{\mathbf{x}}_0) + D^2 \left(\widehat{\beta} + \int_{\widehat{I}_1 \cup \widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0)[g_1, \cdot] - 2 \sum_{i=1}^{n_1} \sigma_{0i} L.\widehat{\psi}_{\widehat{s}_{1i}} (\widehat{\mathbf{x}}_0) \right)_{g_1(\widehat{\mathbf{x}}_0)} \quad (56)$$

Adding and subtracting $D^2\alpha(\widehat{\mathbf{x}}_0)[g_1, \cdot]$ and noticing that $d \left(\alpha + \widehat{\beta} + \int_{\widehat{I}_1 \cup \widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) = 0$ by PMP, we obtain that

$$D^2 \left(\alpha + \widehat{\beta} + \int_{\widehat{I}_1 \cup \widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0)[g_1, \cdot] = L.L_{g_1} \left(\alpha + \widehat{\beta} + \int_{\widehat{I}_1 \cup \widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0)$$

so that the upper term in (56) reduces to

$$L.L_{g_1} \left(-\alpha - \int_{\widehat{I}_1} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) + \sigma_{0n_1} L.\widehat{\psi}_{\widehat{\tau}_1} (\widehat{\mathbf{x}}_0) - 2 \sum_{i=1}^{n_1} \sigma_{0i} L.\widehat{\psi}_{\widehat{s}_{1i}} (\widehat{\mathbf{x}}_0) = \quad (57)$$

$$= -d\alpha(\widehat{\mathbf{x}}_0)Dg_1(\widehat{\mathbf{x}}_0)(\cdot) - \sigma_{01} L.\widehat{\psi}_0(\widehat{\mathbf{x}}_0) = -d\alpha(\widehat{\mathbf{x}}_0)Dh_1(\widehat{\mathbf{x}}_0)(\cdot) - \sigma_{01} L.\psi(\widehat{\mathbf{x}}_0) \quad (58)$$

This proves that the Hamiltonian vector field associated with G_1'' is $\iota^{-1}\overrightarrow{H_1^{\sigma_{01}}}$, i.e. $\overrightarrow{G_1''}$ by definition.

Analogous computations show that applying the antisymplectic isomorphism ι to the vector field associated to G_2'' we obtain

$$\left(-d\alpha(\widehat{\mathbf{x}}_0)Dg_2(\widehat{\mathbf{x}}_0) - L.L_{g_2} \int_{\widehat{I}_1} |\widehat{\psi}_s| ds(\widehat{\mathbf{x}}_0) - \sum_{i=1}^{n_1} \sigma_{0i} \frac{L.\widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)L_{g_2}\widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)}{L_{g_1}\widehat{\psi}_{\widehat{s}_{1i}}(\widehat{\mathbf{x}}_0)} \right)_{g_2(\widehat{\mathbf{x}}_0)}. \quad (59)$$

Let us choose a coordinate system (p, q) in T^*M and let (p_0, q_0) be some point. Since

$$\mathcal{H}_{\widehat{\tau}_1} \left(\begin{matrix} p_0 \\ q_0 \end{matrix} \right) = \left(\begin{matrix} \left(p_0 + \int_{\widehat{I}_1} d|\psi_s| ds \right) \widehat{S}_{\widehat{\tau}_1}^{-1} \\ \widehat{S}_{\widehat{\tau}_1}(q_0) \end{matrix} \right), \quad (60)$$

applying (55) to $\overrightarrow{H_2^{\sigma_{01}}} \circ \mathcal{H}_{\widehat{\tau}_1}$ we get the proof of (32) for $i = 2$. The proof for $i = 3$ follows the same lines.

Proof of equation (36).

The first equality in (36) is a straightforward computation. Moreover

$$\sigma \left(\overrightarrow{H_1^{\sigma_{0n_1+1}}}, \overrightarrow{H_2^{\sigma_{01}}} \right) (\widehat{\ell}_1) = -u_1 \langle \widehat{\ell}_1, [f_0, f_1] \rangle - \sigma_{0n_1} L_{g_2} \widehat{\psi}_{\widehat{\tau}_1} (\widehat{\mathbf{x}}_0).$$

Applying (60) and noticing that, by PMP, $d \left(\alpha + \int_{\widehat{I}_1} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0) = -d \left(\widehat{\beta} + \int_{\widehat{I}_3} |\widehat{\psi}_s| ds \right) (\widehat{\mathbf{x}}_0)$, also the second equality in (36) is proved.

Analogous computation prove (37)-(38).

References

- [1] B. Berret, C. Darlot, F. Jean, T. Pozzo, C. Papaxanthis, and J.-P. Gauthier. The inactivation principle: Mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements. *PLoS Comput. Biol.*, 4(10), 2008.

- [2] Z. Chen, J.-B. Caillau, and Y. Chitour. L^1 minimization for mechanical systems. *SIAM J. Control. Optim.*, 54:1245–1265, 2016.
- [3] G. Vossen and H. Maurer. On L^1 -minimization in optimal control and applications to robotics. *Optim. Control Appl. Meth.*, (27):301–321, 2006.
- [4] N. Boizot and O. Oukacha. Consumption minimisation for an academic vehicle. [hal-01384651](#), 2016.
- [5] A. J. Craig and I. Flügge-Lotz. Investigation of optimal control with a minimum-fuel consumption criterion for a fourth-order plant with two control inputs; synthesis of an efficient suboptimal control. *J. Basic Engineering*, 87:39–58, 1965.
- [6] I. M. Ross. Space trajectory optimization and L^1 -optimal control problems. In P. Gurfil, editor, *Modern Astrodynamics*, volume 1 of *Elsevier Astrodynamics Series*, pages 155 – VIII. Butterworth-Heinemann, 2006.
- [7] C. Clason and K. Kunisch. A duality-based approach to elliptic control problems in non-reflexive Banach spaces. *ESAIM: COCV*, 17:243–266, 2011.
- [8] D. Kalise, K. Kunisch, and Z. Rao. Infinite horizon sparse optimal control. *J Optim Theory Appl*, pages 1–32, 2016.
- [9] G. Stadler. Elliptic optimal control problems with L^1 -control cost and applications for the placement of control devices. *Comp. Optim. Appl.*, 44:159–181, 2009.
- [10] F. Clarke. *Functional analysis, calculus of variations and optimal control*. GTM. Springer, 2013.
- [11] A. A. Agrachev and Yu. L. Sachkov. *Control Theory from the Geometric Viewpoint*. Springer-Verlag, 2004.
- [12] 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.
- [13] A. Agrachev, G. Stefani, and P. Zezza. Strong optimality for a bang-bang trajectory. *SIAM Journal on Control and Optimization*, 41(4):991–1014, 2002.
- [14] L. Poggiolini and G. Stefani. State-local optimality of a bang-bang trajectory: a Hamiltonian approach. *System and Control Letters*, 53(2):269–279, 2004.
- [15] A. Agrachev, G. Stefani, and P. Zezza. An invariant second variation in optimal control. *Int. J. Control*, 71(5):689–715, 1998.
- [16] L. Poggiolini. On local state optimality of bang-bang extremals in a free horizon Bolza problem. *Rendiconti del seminario matematico italiano*, 64:1–23, 2006.
- [17] F. H. Clarke. On the inverse function theorem. *Pacific J. Math.*, 64(1):97–102, 1976.