Structural stability of bang--bang trajectories with a double switching time in the minimum time problem

STRUCTURAL STABILITY OF BANG–BANG TRAJECTORIES WITH A DOUBLE SWITCHING TIME IN THE MINIMUM TIME PROBLEM*

LAURA POGGIOLINI†

Abstract. In this paper we consider the problem of structural stability of strong local optimisers for the minimum time problem in the case when the nominal problem has a bang-bang strongly local optimal control which exhibits a double switch.

Key words. Hamiltonian methods, bang-bang controls, sufficient second order conditions

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

1. Introduction.

An important issue in optimal control problem where the optimal control is bang-bang is the stability of the switching structure under small perturbations of the data. This problem may arise either in practical optimal control problems, where the data are not known exactly, or in numerical approximation techniques of the optimal control.

In this paper we address such issue for the minimum time problem between two submanifolds of a finite dimensional manifold M in the case when the dynamics is affine with respect to the control and the control takes values in a box of $\mathbb{R}^m$. Namely, the following optimal control is studied:

$$T \to \min, \tag{1a}$$

$$\dot{\xi}(t) = f_0(\xi(t)) + \sum_{s=1}^{m} u_s(t) f_s(\xi(t)) \quad \text{a.e. } t \in [0,T], \tag{1b}$$

$$\xi(0) \in N_0, \qquad \xi(T) \in N_f, \tag{1c}$$

$$u(t) = (u_1(t), \dots, u_m(t)) \in [-1,1]^m \text{ a.e. } t \in [0,T]. \tag{1d}$$

For such a problem, the triple (T, ξ, u) is said to be an *admissible triple* for problem (1) if $T > 0$ and the couple $(\xi, u) \in W^{1,\infty}([0,T], M) \times L^\infty([0,T], \mathbb{R}^m)$ satisfies (1b), (1c) and (1d). We are going to consider a triple $(\widehat{T}, \widehat{\xi}, \widehat{u})$, that satisfies the necessary conditions for optimality (i.e. Pontryagin Maximum Principle), where the control is bang-bang with multiple switching.

We recall that for a control affine system where the control set is a compact convex polyhedron Δ, a control function $\widehat{u}$ is said to be bang-bang if it is piecewise constant with values in the vertexes of Δ, i.e. there exist $0 = \widehat{\tau}_0 < \widehat{\tau}_1 < \widehat{\tau}_2 \ldots < \widehat{\tau}_r < \widehat{\tau}_{r+1} = T$ such that $\widehat{u}|_{(\widehat{\tau}_{i-1}, \widehat{\tau}_i)}$ is a vertex of Δ for any $i = 1, \dots, r+1$ and $\widehat{u}|_{(\widehat{\tau}_{i-1}, \widehat{\tau}_i)} \neq \widehat{u}|_{(\widehat{\tau}_i, \widehat{\tau}_{i+1})}$ for any $i = 1, \dots, r$. The times $\widehat{\tau}_1, \dots, \widehat{\tau}_r$ are called *switching times* of $\widehat{u}$. For each switching time $\widehat{\tau}_i$, the multiplicity of $\widehat{\tau}_i$ is defined as the dimension of the face of Δ containing the vertexes $\widehat{u}|_{(\widehat{\tau}_{i-1}, \widehat{\tau}_i)}$ and $\widehat{u}|_{(\widehat{\tau}_i, \widehat{\tau}_{i+1})}$. Thus, if $\Delta = [-1,1]^m$ is the standard box of $\mathbb{R}^m$, the multiplicity of the switching time $\widehat{\tau}_i$ is given by the number of components that change sign at time $\widehat{\tau}_i$.

The literature on bang-bang controls is by now very rich for the case of simple switches, i.e. for the case when all the switching times have multiplicity one. See, for

*The author was partially supported by Gruppo Nazionale per lAnalisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM)

†Dipartimento di Matematica e Informatica "Ulisse Dini", Università degli Studi di Firenze, Italy (laura.poggiolini@unifi.it, http://www.dma.unifi.it/~poggiolini/).

example, [1, 10, 9, 15] for local optimality results and [5, 2] for structural stability results. On the other hand, at least to the author's knowledge, the literature on bang-bang controls with multiple switches is much more scarce than the one with simple switches only. L^1-local optimality results for bang-bang controls with multiple switches in the minimum time problem between two fixed end points were given in [18].

The problems of strong local optimality and structural stability of bang-bang extremals with a double switch, and a finite number of simple switches, in a Mayer problem were addressed in [13] and [3], respectively. The minimum time problem was studied in [14] where the authors consider the case when a double switch occurs and all the other switches are simple. They prove that under suitable regularity conditions, and assuming the coercivity of the second order approximation of a certain finite-dimensional subproblem of the given one, the triple $(\widehat{T}, \widehat{\xi}, \widehat{u})$ is in fact a state-local minimiser of the problem. See Definition 1 for a precise definition of this kind of strong local optimality.

Here we consider the same case as in [14] and we study the structural stability of the locally optimal control $\widehat{u}$ under smooth perturbations of the data of the problem: the drift f_0, the controlled vector fields $f_1, f_2, \ldots, f_m$ and the submanifolds of the initial and final constraints.

In particular we are interested in understanding how the existence of the double switch and the bang-bang structure of the locally optimal control are affected by small perturbations of the data. Such a situation is in fact not generic and we are going to show that under the same assumptions that ensure state-local optimality of the reference triple plus a full rank condition, the bang-bang structure of the locally optimal control is stable under small perturbations, even though the double switching time may decouple into two simple switching times, i.e. the number of bang arcs may increase of one unit as the double switch may decouple in two simple switches but the number of switches of each control component and the sequence of values it takes are stable under small perturbations of the data.

The proof is carried out by Hamiltonian methods, which were also used in [14] to prove the state local optimality result for the nominal problem. The same methods were also used in [16] and [17] to prove state local optimality and structural stability of a bang-singular-bang extremal in the minimum time problem between two fixed end points.

As in [14], for the sake of notational simplicity we shall confine ourselves to the case when $M = \mathbb{R}^n$, $m = 2$ and only the double switch occurs. However, as all the results are invariant under a change of coordinates, they can be easily generalised to the case when the state space is a smooth finite dimensional manifold. Moreover, the presence of a finite number of simple switches occuring either before and/or after the double one can be treated at the expenses of a much heavier notation, see for example [12, 13]. Thus the nominal problem (1) simplifies to

$$
\begin{aligned}
(\mathbf{P_0}) \qquad & T \to \min, \\
& \dot{\xi}(t) = f_0(\xi(t)) + u_1(t) f_1(\xi(t)) + u_2(t) f_2(\xi(t)) \quad \text{a.e. } t \in [0, T], \\
& \xi(0) \in N_0, \qquad \xi(T) \in N_f, \\
& |u_s(t)| \le 1 \quad s = 1, 2 \qquad \text{a.e. } t \in [0, T].
\end{aligned}
$$

Without loss of generality we can assume that $\widehat{u}$ is given by

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

We assume that ($\mathbf{P_0}$) is the problem we obtain when $r = 0$ in the following parameter dependent problem ($\mathbf{P_r}$):

$$\begin{aligned} (\mathbf{P_r}) \qquad & T \to \min, \\ & \dot{\xi}(t) = f_0^r(\xi(t)) + u_1(t) f_1^r(\xi(t)) + u_2(t) f_2^r(\xi(t)) \quad \text{a.e. } t \in [0,T], \\ & \xi(0) \in N_0^r, \qquad \xi(T) \in N_f^r, \\ & |u_s(t)| \le 1 \quad s = 1,2 \qquad \text{a.e. } t \in [0,T]. \end{aligned}$$

The parameter r belongs to some ball B_R centered at the origin of $\mathbb{R}^k$ and radius $R > 0$. For notational simplicity we choose $\mathbb{R}^n$ as state–space; all the data are assumed to be smooth, more precisely the maps

$$(r,x) \in B_R \times \mathbb{R}^n \mapsto f_i^r(x) \in \mathbb{R}^n, \qquad i = 0,1,2$$

are assumed to be C^2 and the submanifolds of the initial and final constraints are given as regular intersections of zero-level sets of C^2 maps from $B_R \times \mathbb{R}^n$ to $\mathbb{R}$, i.e.

$$\begin{aligned} & N_0^r = \bigcap_{i=1}^{n_0} \left(\Phi_i^{0,r}\right)^{-1}(0), \\ & \mathrm{D}\Phi_i^{0,r}(x) \text{ are linearly independent at } x \quad \forall (r,x) \in B_R \times \mathbb{R}^n. \end{aligned}$$

and

$$\begin{aligned} & N_f^r = \bigcap_{j=1}^{n_f} \left(\Phi_j^{f,r}\right)^{-1}(0), \\ & \mathrm{D}\Phi_j^{f,r}(x) \text{ are linearly independent at } x \quad \forall (r,x) \in B_R \times \mathbb{R}^n. \end{aligned}$$

We are interested in state-local optimisers according to the following definition:

DEFINITION 1 (state-local optimality). *The trajectory ξ of an admissible triple (T,ξ,u) for problem ($\mathbf{P_r}$) is a state-local minimiser of such problem if there are neighbourhoods $\mathcal{U}$ of its range $\xi([0,T])$, $\mathcal{U}_0$ of $\xi(0)$ and $\mathcal{U}_f$ of $\xi(T)$ such that ξ is a minimum time trajectory among the admissible trajectories of ($\mathbf{P_r}$) whose range is in $\mathcal{U}$, whose initial point is in $N_0^r \cap \mathcal{U}_0$ and whose final point is in $N_f^r \cap \mathcal{U}_f$.*

REMARK 1.1. *Notice that state-local optimality is a kind of strong local optimality, in the sense that there is no localisation with respect to the control, but only with respect to the trajectories. Moreover state-local optimality is stronger than the classical notion of strong-local optimality where one considers the C^0 distance between trajectories, i.e. one considers only triples (T,ξ,u) where the graph of the trajectory ξ is close to the graph of the reference trajectory $\widehat{\xi}$, and T is close to $\widehat{T}$, see e.g. [10].*

Assuming that $(\widehat{T},\widehat{\xi},\widehat{u})$ satisfies normal PMP with adjoint covector $\widehat{\lambda}$, the sufficient conditions for state-local optimality as stated in [14] plus a full rank condition which ensures the uniqueness of the adjoint covector, we prove that for small R each problem

($\mathbf{P_r}$), $r \in B_R$ has a state-local optimal trajectory (T^r, ξ^r, u^r) (with adjoint covector λ^r) where u^r componentwise preserves the bang-bang structure of $\widehat{u}$ and T^r is close to $\widehat{T}$. Moreover the switching times and the final time depend smoothly on the parameter r and λ^r is the only Pontryagin extremal of ($\mathbf{P_r}$) whose graph is close to the graph of $\widehat{\lambda}$.

We would like to recall that this set of assumptions (which concern the nominal problem only) is the same set of assumptions that is required in the classical regular case for the stability of weak local optimisers, see [6, 7, 8].

2. Notation. We are going to use some basic notions from symplectic geometry. For any manifold $N \subset \mathbb{R}^n$ and any $x \in N$, the tangent space and the cotangent space to N in x are denoted as T_xN and T_x^*N, respectively. We recall that the cotangent bundle $T^*\mathbb{R}^n$ to $\mathbb{R}^n$ can be identified with the Cartesian product $(\mathbb{R}^n)^* \times \mathbb{R}^n = T_x^*\mathbb{R}^n \times T_x\mathbb{R}^n$ for any $x \in \mathbb{R}^n$. The projection from $T^*\mathbb{R}^n$ onto $\mathbb{R}^n$ is denoted as $\pi \colon \ell \in T^*\mathbb{R}^n \mapsto \pi\ell \in \mathbb{R}^n$. We shall write $T_x\mathbb{R}^n$ instead of $\mathbb{R}^n$, to emphasize the fact that we are dealing with tangent vectors.

The canonical Liouville one–form $\boldsymbol{s}$ on $T^*\mathbb{R}^n$ and the associated canonical symplectic two-form $\boldsymbol{\sigma} = \mathrm{d}\boldsymbol{s}$ allow to associate to any, possibly time-dependent, smooth Hamiltonian $F_t \colon T^*\mathbb{R}^n \to \mathbb{R}$, the unique Hamiltonian vector field $\overrightarrow{F_t}$ such that

$$\boldsymbol{\sigma}(v, \overrightarrow{F_t}(\ell)) = \langle \mathrm{d}F_t(\ell)\, ,\, v\rangle, \quad \forall v \in T_\ell T^*\mathbb{R}^n.$$

Choosing coordinates $\ell = (p, x) \in (\mathbb{R}^n)^* \times \mathbb{R}^n$, we have

$$\overrightarrow{F_t}(p, x) = \left(\frac{-\partial F_t}{\partial x}, \frac{\partial F_t}{\partial p} \right)(p, x).$$

To any vector field $f \colon \mathbb{R}^n \to \mathbb{R}^n$ we associate the Hamiltonian function F

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

so that $\overrightarrow{F}(p, x) = \big(- p\, \mathrm{d}f(x), f(x) \big)$.

We denote by $\widehat{f_t}$ the piecewisely time-dependent vector field associated to the reference control:

$$\widehat{f_t} := f_0 + \widehat{u}_1(t) f_1 + \widehat{u}_2(t) f_2$$

and by h_1, h_2 its restrictions to the time intervals $[0, \widehat{\tau})$ and $(\widehat{\tau}, \widehat{T}]$, respectively:

$$h_1 := \widehat{f_t}\Big|_{[0,\widehat{\tau})} = f_0 - f_1 - f_2, \qquad h_2 := \widehat{f_t}\Big|_{(\widehat{\tau},\widehat{T}]} = f_0 + f_1 + f_2.$$

In what follows we shall also need the vector fields

$$\begin{aligned} k_1 &:= f_0 + f_1 - f_2 = h_1 + 2f_1 = h_2 - 2f_2, \\ k_2 &:= f_0 - f_1 + f_2 = h_1 + 2f_2 = h_2 - 2f_1. \end{aligned}$$

The associated Hamiltonian functions are denoted by the same letter, but capitalized. Namely

$$\begin{aligned} H_1(\ell) &:= \langle \ell\, ,\, h_1(\pi\ell)\rangle, \qquad H_2(\ell) := \langle \ell\, ,\, h_2(\pi\ell)\rangle, \\ K_1(\ell) &:= \langle \ell\, ,\, k_1(\pi\ell)\rangle, \qquad K_2(\ell) := \langle \ell\, ,\, k_2(\pi\ell)\rangle. \end{aligned}$$

Analougously we define the parameter dependent vector fields

$$h_1^r := f_0^r - f_1^r - f_2^r, \qquad h_2^r := f_0^r + f_1^r + f_2^r,$$
$$k_1^r := f_0^r + f_1^r - f_2^r, \qquad k_2^r := f_0^r - f_1^r + f_2^r,$$

and the associated parameter dependent Hamiltonians

$$H_1^r := F_0^r - F_1^r - F_2^r, \qquad H_2^r := F_0^r + F_1^r + F_2^r,$$
$$K_1^r := F_0^r + F_1^r - F_2^r, \qquad K_2^r := F_0^r - F_1^r + F_2^r.$$

The maximised Hamiltonian of the nominal control system (**P₀**) is well defined in the whole cotangent bundle $T^*\mathbb{R}^n$ and is denoted by $H^{\max}$:

$$\begin{aligned} H^{\max}(\ell) :=& \max\left\{F_0(\ell) + u_1 F_1(\ell) + u_2 F_2(\ell) \colon (u_1, u_2) \in [-1,1]^2\right\} \\ =& F_0(\ell) + |F_1(\ell)| + |F_2(\ell)|\,. \end{aligned}$$

Throughout the paper, the symbol $\mathcal{O}(x)$ denotes a neighborhood of x in its ambient space. The flow starting at time $t = 0$ of the time-dependent vector field $\widehat{f}_t$ is defined in a neighborhood $\mathcal{O}(\widehat{x}_0)$ for any $t \in [0, \widehat{T}]$ and is denoted by $\widehat{S}_t \colon \mathcal{O}(\widehat{x}_0) \to \mathbb{R}^n$, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{S}_t(x) = \widehat{f}_t \circ \widehat{S}_t(x) \quad \text{a.e. } t \in [0, \widehat{T}], \qquad \widehat{S}_0(x) = x.$$

We denote by $\widehat{x}_0 := \widehat{\xi}(0)$ and by $\widehat{x}_f := \widehat{\xi}(\widehat{T}) = \widehat{S}_T(\widehat{x}_0)$ the end points of the reference trajectory and by $\widehat{x}_d := \widehat{\xi}(\widehat{\tau}) = \widehat{S}_{\widehat{\tau}}(\widehat{x}_0)$ the point corresponding to the switching time.

Analougously, the flow starting at time $t = 0$ of the time-dependent Hamiltonian vector field associated to $\widehat{F}_t(\ell) := \langle \ell\,, \widehat{f}_t(\pi\ell)\rangle$ is defined in a neighborhood $\mathcal{O}(\widehat{\ell}_0)$ of $\widehat{\ell}_0 := \widehat{\lambda}(0)$ for any $t \in [0, \widehat{T}]$ and is denoted by $\widehat{\mathcal{F}}_t \colon \mathcal{O}(\widehat{\ell}_0) \to T^*\mathbb{R}^n$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathcal{F}}_t(\ell) = \overrightarrow{\widehat{F}_t} \circ \widehat{\mathcal{F}}_t(\ell) \quad \text{a.e. } t \in [0, \widehat{T}], \qquad \widehat{\mathcal{F}}_0(\ell) = \ell.$$

Given a smooth function $\gamma \colon \mathcal{O}(x) \subset \mathbb{R}^n \to \mathbb{R}$ and a vector $\delta x \in T_x\mathbb{R}^n$, the Lie derivative of γ with respect to the vector δx at the point x is denoted by $\delta x \cdot \gamma\,(x)$, i.e. $\delta x \cdot \gamma\,(x) = \langle \mathrm{D}\gamma(x)\,, \delta x\rangle$. If $f \colon \mathcal{O}(x) \to \mathbb{R}^n$ is a smooth vector field, then $f \cdot \gamma\,(x)$ is the Lie derivative of γ at x with respect to the vector $f(x)$, i.e. $f \cdot \gamma\,(x) := \langle \mathrm{D}\gamma(x)\,, f(x)\rangle$. Given two smooth vector fields $f,\ g \colon \mathbb{R}^n \to \mathbb{R}^n$, then the Lie bracket $[f, g]$ is given by the vector field $(\mathrm{D}g)f - (\mathrm{D}f)g$.

We identify any bilinear form Q on a vector space V with a linear form $Q \colon V \to V^*$: $Q(v, w) = \langle Qv\,, w\rangle$. Given W, linear subspace of V we thus say that a vector $v \in V$ is in $W^{\perp_Q}$ if $Q(v, w) = 0$ for any $w \in W$. We denote the associate quadratic form with the same letter but calligraphic: $\mathcal{Q}[v]^2 = Q(v, v)$.

Finally, given an interval $[t_1, t_2]$ and a function $\varphi \colon (t_1, t_2) \to \mathbb{R}^n$ we use the symbol $⨍_{t_1}^{t_2} \varphi(s)\,\mathrm{d}s$ to denote the mean value of φ in (t_1, t_2):

$$⨍_{t_1}^{t_2} \varphi(s)\,\mathrm{d}s := \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \varphi(s)\,\mathrm{d}s.$$

3. Assumptions. We now state the assumptions on the nominal extremal triple $\big(\widehat{T}, \widehat{\xi}, \widehat{u}\big)$ of (**P₀**). Besides the necessary conditions for optimality, namely Pontryagin Maximum Principle (PMP) –which we assume to hold in its normal form– we require that the triple $\big(\widehat{T}, \widehat{\xi}, \widehat{u}\big)$ satisfies the conditions that ensure state-local optimality, as stated in [14]: regularity along the bang arcs, regularity at the switching time and the coercivity of the second order variation associated to some finite-dimensional subproblem of the given one. Moreover we assume the uniqueness of the adjoint covector associated to the reference triple $\big(\widehat{T}, \widehat{\xi}, \widehat{u}\big)$ in Pontryagin Maximum Principle for ($\mathbf{P_0}$).

Assumption 1 (Normal PMP). *There exists an absolutely continuous curve* $\widehat{\lambda}\colon [0, \widehat{T}] \to T^*\mathbb{R}^n$ *satisfying the following properties*

$$\pi\widehat{\lambda}(t) = \widehat{\xi}(t), \qquad \forall t \in [0, \widehat{T}], \tag{2a}$$

$$\dot{\widehat{\lambda}}(t) = \overrightarrow{\widehat{F}}_t(\widehat{\lambda}(t)), \qquad a.e.\ t \in [0, \widehat{T}], \tag{2b}$$

$$\widehat{F}_t(\widehat{\lambda}(t)) = H^{\max}(\widehat{\lambda}(t)) = 1, \qquad a.e.\ t \in [0, \widehat{T}], \tag{2c}$$

$$\widehat{\lambda}(0)\Big|_{T_{\widehat{x}_0}N_0} = 0, \quad \widehat{\lambda}(\widehat{T})\Big|_{T_{\widehat{x}_f}N_f} = 0. \tag{2d}$$

In coordinates we put $\widehat{\lambda}(t) := \Big(\widehat{\mu}(t), \widehat{\xi}(t)\Big)$ where $\widehat{\mu}(t) \in T^*_{\widehat{\xi}(t)}\mathbb{R}^n \quad \forall t \in [0, \widehat{T}]$.
Here and in what follows we shall use the following notation:

$$\widehat{\ell}_0 := \widehat{\lambda}(0),\ \widehat{\ell}_d := \widehat{\lambda}(\widehat{\tau}),\ \widehat{\ell}_f := \widehat{\lambda}(\widehat{T}),\ \widehat{p}_0 := \widehat{\mu}(0),\ \widehat{p}_d := \widehat{\mu}(\widehat{\tau}),\ \widehat{p}_f := \widehat{\mu}(\widehat{T}).$$

Remark 3.1. *The adjoint covector* $\widehat{\mu}$ *is a solution to the ODE*

$$\dot{\mu}(t) = \frac{-\partial F_t}{\partial x}\Big(\mu(t), \widehat{\xi}(t)\Big) = -\left\langle \mu(t)\,,\, \mathrm{d}\widehat{f}_t(\widehat{\xi}(t))\right\rangle$$

so that $\widehat{\mu}(t) = \widehat{p}_0 \widehat{S}_{t*}^{-1} \quad \forall t \in [0, \widehat{T}]$ *and the transversality conditions* (2d) *read*

$$\widehat{p}_0 = \sum_{i=1}^{n-n_0} \widehat{a}_i\, \mathrm{D}\Phi_i^{0,0}(\widehat{x}_0), \qquad \widehat{p}_f = \sum_{j=1}^{n-n_f} \widehat{b}_j\, \mathrm{D}\Phi_i^{f,0}(\widehat{x}_f),$$

for some $\widehat{a} = \big(\widehat{a}_1, \ldots, \widehat{a}_{n-n_0}\big) \in \mathbb{R}^{n-n_0}$, $\widehat{b} = \Big(\widehat{b}_1, \ldots, \widehat{b}_{n-n_f}\Big) \in \mathbb{R}^{n-n_f}$.

Remark 3.2. *As* $\widehat{\lambda}$ *is a normal extremal then the transversality conditions* (2d) *together with the maximality condition* (2c) *yield* $h_1(\widehat{x}_0) \notin T_{\widehat{x}_0}N_0$ *and* $h_2(\widehat{x}_f) \notin T_{\widehat{x}_f}N_f$.

Maximality condition (2c) implies, for any $i = 1, 2$ and for almost every $t \in [0, \widehat{T}]$,

$$\widehat{u}_i(t)F_i(\widehat{\lambda}(t)) = \widehat{u}_i(t)\langle\widehat{\lambda}(t)\,,\, f_i(\widehat{\xi}(t))\rangle \geq 0.$$

We assume that the bang arcs of $\widehat{\lambda}$ are regular, i.e., we assume that at each point $\widehat{\lambda}(t)$, $t \neq \widehat{\tau}$, the maximum of the Hamiltonian is achieved only by $u = \widehat{u}(t) = (\widehat{u}_1(t), \widehat{u}_2(t))$, i.e.,

$$F_0(\widehat{\lambda}(t)) + u_1F_1(\widehat{\lambda}(t)) + u_2F_2(\widehat{\lambda}(t)) < H^{\max}(\widehat{\lambda}(t)) = 1 \qquad \forall (u_1, u_2) \in [-1, 1]^2 \setminus \{(\widehat{u}_1(t), \widehat{u}_2(t))\}.$$

In terms of the controlled Hamiltonians F_1 and F_2 this can be stated as follows:

ASSUMPTION 2 (Regularity along the bang arcs). *Let $i = 1, 2$. If $t \neq \widehat{\tau}$, then*

$$\widehat{u}_i(t)F_i(\widehat{\lambda}(t)) = \widehat{u}_i(t)\langle\widehat{\lambda}(t)\,,\, f_i(\widehat{\xi}(t))\rangle > 0. \tag{3}$$

REMARK 3.3. *Because of the normality condition in PMP, it holds $F_0(\widehat{\lambda}(t)) - 1 = -\widehat{u}_1(t)F_1(\widehat{\lambda}(t)) - \widehat{u}_2(t)F_2(\widehat{\lambda}(t))$ for all $t \in [0, \widehat{T}]$. By continuity, from (3), we get $F_1(\widehat{\ell}_d) = F_2(\widehat{\ell}_d) = 0$ so that $F_0(\widehat{\ell}_d) = 1$. Therefore $f_0(\widehat{x}_d) \notin \operatorname{span}\{f_1(\widehat{x}_d),\ f_2(\widehat{x}_d)\}$.*

From the necessary maximality condition (2c) we get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}2\,F_i\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}-} &= \frac{\mathrm{d}}{\mathrm{d}t}\,(K_i - H_1)\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}-} \geq 0,\\
\frac{\mathrm{d}}{\mathrm{d}t}2\,F_i\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}+} &= \frac{\mathrm{d}}{\mathrm{d}t}\,(H_2 - K_j)\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}+} \geq 0,
\end{aligned}\qquad i, j \in \{1, 2\},\ i \neq j.$$

We assume that the above inequalities are strict:

ASSUMPTION 3 (Regularity at the double switching time).

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(K_\nu - H_1)\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}-} > 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\,(H_2 - K_\nu)\circ\widehat{\lambda}(t)\Big|_{t=\widehat{\tau}+} > 0, \quad \nu = 1, 2.$$

Assumption 3 is called the STRONG BANG-BANG LEGENDRE CONDITION FOR THE DOUBLE SWITCHING TIME. Equivalently, this assumption can be expressed in terms of the Lie brackets of vector fields or in terms of the canonical symplectic structure $\boldsymbol{\sigma}(\cdot,\cdot)$ on $T^*\mathbb{R}^n$.

PROPOSITION 2. *Assumption 3 is equivalent to*

$$\begin{aligned}
\langle\widehat{\ell}_d\,,\,[h_1, k_\nu]\,(\widehat{x}_d)\rangle &= \boldsymbol{\sigma}\left(\overrightarrow{H_1}, \overrightarrow{K_\nu}\right)(\widehat{\ell}_d) > 0,\\
\langle\widehat{\ell}_d\,,\,[k_\nu, h_2]\,(\widehat{x}_d)\rangle &= \boldsymbol{\sigma}\left(\overrightarrow{K_\nu}, \overrightarrow{H_2}\right)(\widehat{\ell}_d) > 0,
\end{aligned}\qquad \nu = 1, 2.$$

An easy computation proves the following equivalent condition

PROPOSITION 3. *Assumption 3 is equivalent to*

$$\langle\widehat{\ell}_d\,,\,[f_0, f_i]\,(\widehat{x}_d)\rangle > \left|\langle\widehat{\ell}_d\,,\,[f_1, f_2]\,(\widehat{x}_d)\rangle\right|, \qquad i = 1, 2, \tag{4}$$

i.e.

$$\boldsymbol{\sigma}\left(\overrightarrow{F_0}, \overrightarrow{F_i}\right)(\widehat{\ell}_d) > \left|\boldsymbol{\sigma}\left(\overrightarrow{F_1}, \overrightarrow{F_2}\right)(\widehat{\ell}_d)\right|, \qquad i = 1, 2.$$

In what follows we shall also need to reformulate Assumption 3 in terms of the pull-backs of the vector fields h_ν and k_ν along the reference flow $\widehat{S}_t$. Define

$$g_\nu(x) := \widehat{S}_{\widehat{\tau}\,*}^{-1} h_\nu \circ \widehat{S}_{\widehat{\tau}}(x), \quad j_\nu(x) := \widehat{S}_{\widehat{\tau}\,*}^{-1} k_\nu \circ \widehat{S}_{\widehat{\tau}}(x), \qquad \nu = 1, 2 \tag{5}$$

and let G_ν, J_ν be the associated Hamiltonians. Then a straightforward computation yields.

PROPOSITION 4. *Assumption 3 is equivalent to*

$$\begin{aligned}
\langle\widehat{\ell}_0\,,\,[g_1, j_\nu]\,(\widehat{x}_0)\rangle &= \boldsymbol{\sigma}\left(\overrightarrow{G}_1, \overrightarrow{J_\nu}\right)(\widehat{\ell}_0) > 0,\\
\langle\widehat{\ell}_0\,,\,[j_\nu, g_2]\,(\widehat{x}_0)\rangle &= \boldsymbol{\sigma}\left(\overrightarrow{J_\nu}, \overrightarrow{G}_2\right)(\widehat{\ell}_0) > 0,
\end{aligned}\qquad \nu = 1, 2. \tag{6}$$

Also, we assume that $\widehat{\xi}$ has no self-intersection:

ASSUMPTION 4. *The reference trajectory $\widehat{\xi}\colon [0, \widehat{T}] \to \mathbb{R}^n$ is injective.*

4. The second order variation. The second order variation is the second order approximation of a finite-dimensional subproblem of ($\mathbf{P_0}$) obtained by keeping the same endpoint constraints and restricting the set of admissible controls. Namely, we allow for independent variations of the switching times of each of the two reference control components $\widehat{u}_1$ and $\widehat{u}_2$. This subproblem is then extended by allowing for variations of the initial points of trajectories on a neighborhood of $\widehat{x}_0$ in $\mathbb{R}^n$. We penalise the latter variations with a smooth cost α that vanishes on N_0.

We allow for perturbations of the final time, of the initial point of trajectories on N_0, of the final point on N_f and of the switching time of either component of the reference control: let $\tau_1 := \widehat{\tau} + \varepsilon_1$ and $\tau_2 := \widehat{\tau} + \varepsilon_2$ be the perturbed switching times of the first and of the second component of $\widehat{u}$, respectively, and let $\tau_3 := \widehat{T} + \varepsilon_3$ be the perturbation of the final time $\widehat{T}$.

Let $\alpha\colon \mathbb{R}^n \to \mathbb{R}$ be a smooth nonnegative function vanishing on N_0. We remove the constraint on the initial point $\xi(0)$ introducing the penalty cost α on such point. We thus obtain the following problem in the unknowns x, ε_1, ε_2, ε_3:

$$\alpha(x) + \widehat{T} + \delta_3 \to \min, \tag{7a}$$

$$\dot{\xi} = \begin{cases} h_1(\xi(t)) & t \in (0, \widehat{\tau} + \delta_1), \\ k_\nu(\xi(t)) & t \in (\widehat{\tau} + \delta_1, \widehat{\tau} + \delta_2), \\ h_2(\xi(t)) & t \in (\widehat{\tau} + \delta_2, \widehat{T} + \delta_3), \end{cases} \tag{7b}$$

$$\xi(0) = x \in \mathbb{R}^n, \qquad \xi(\widehat{T} + \delta_3) \in N_f, \tag{7c}$$

$$\delta_1 := \min\{\varepsilon_1, \varepsilon_2\}, \quad \delta_2 := \max\{\varepsilon_1, \varepsilon_2\}, \quad \delta_3 := \varepsilon_3, \tag{7d}$$

$$\nu = \begin{cases} 1 & \text{if } \varepsilon_1 \le \varepsilon_2, \\ 2 & \text{if } \varepsilon_1 > \varepsilon_2. \end{cases} \tag{7e}$$

Let g_ν, j_ν, $\nu = 1, 2$ be the pullbacks along the reference flow of the vector fields h_ν and k_ν, as defined in equation (5). Let $\widehat{N}_f$ be the pullback of N_f to time $t = 0$ along the reference flow:

$$\widehat{N}_f := \widehat{S}_{\widehat{T}}^{-1}(N_f)$$

and let $T_{\widehat{x}_0}\widehat{N}_f = \widehat{S}_{\widehat{T}\,*}^{-1}(T_{\widehat{x}_f}N_f)$ be its tangent space at $\widehat{x}_0$.
By the transversality condition (2d) at the reference final time $\widehat{T}$, there exists a smooth function $\beta\colon \mathbb{R}^n \to \mathbb{R}$ that vanishes on N_f and such that $\mathrm{d}\beta(\widehat{x}_f) = -\widehat{\ell}_f$. Also let $\widehat{\beta}$ be the pull-back of β along the reference flow, $\widehat{\beta} := \beta \circ \widehat{S}_{\widehat{T}}$ so that, by Remark 3.1,

$$\widehat{\beta}\colon \mathcal{O}(\widehat{x}_0) \to \mathbb{R}, \quad \widehat{\beta}\Big|_{\mathcal{O}(\widehat{x}_0)\cap\widehat{N}_f} \equiv 0, \qquad \mathrm{d}\widehat{\beta}(\widehat{x}_0) = -\widehat{p}_0.$$

Let us set

$$a_1 := \delta_1, \qquad b := \delta_2 - \delta_1 = |\varepsilon_2 - \varepsilon_1|, \qquad a_2 := \delta_3 - \delta_2;$$

then the second order approximations of problem (7), for $\nu = 1, 2$, are defined on the closed half-spaces

$$V_\nu^+ := \big\{(\delta x, a_1, b, a_2) \in T_{\widehat{x}_0}\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}\colon \\ \delta x + a_1 g_1(\widehat{x}_0) + b\, j_\nu(\widehat{x}_0) + a_2 g_2(\widehat{x}_0) \in T_{\widehat{x}_0}\widehat{N}_f\big\}$$

and are given by

$$
\begin{aligned}
\mathcal{Q}_\nu[\delta x, a_1, b, a_2]^2 &= \mathrm{D}^2(\alpha + \widehat{\beta})(\widehat{x}_0)[\delta x]^2 + 2\,\delta x \cdot (a_1 g_1 + b\, j_\nu + a_2 g_2) \cdot \widehat{\beta}(\widehat{x}_0) \\
&\quad + (a_1 g_1 + b\, j_\nu + a_2 g_2)^2 \cdot \widehat{\beta}(\widehat{x}_0) \\
&\quad + a_1 b\, [g_1, j_\nu] \cdot \widehat{\beta}(\widehat{x}_0) + a_1 a_2\, [g_1, g_2] \cdot \widehat{\beta}(\widehat{x}_0) + b\, a_2\, [j_\nu, g_2] \cdot \widehat{\beta}(\widehat{x}_0),
\end{aligned} \tag{8}
$$

see [13] for the construction. The restrictions of $\mathcal{Q}_\nu$ to the sets

$$
\begin{aligned}
V^+_{0,\nu} := \big\{&(\delta x, a_1, b, a_2) \in T_{\widehat{x}_0} N_0 \times \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \colon \\
&\delta x + a_1 g_1(\widehat{x}_0) + b\, j_\nu(\widehat{x}_0) + a_2 g_2(\widehat{x}_0) \in T_{\widehat{x}_0} \widehat{N}_f \big\}, \qquad \nu = 1, 2,
\end{aligned}
$$

are indeed the second order approximation of (**$\mathbf{P_0}$**).

We are now in a position to state our assumption on the second order approximation of subproblem (7).

Assumption 5. *For each $\nu = 1, 2$, $\mathcal{Q}_\nu$ is coercive on $V^+_{0,\nu}$.*

Since both $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are quadratic forms, we may as well remove the constraint $b \geq 0$ and let them be defined and coercive on the linear spaces

$$
\begin{aligned}
V_{0,\nu} := \big\{&(\delta x, a_1, b, a_2) \in T_{\widehat{x}_0} N_0 \times \mathbb{R}^3 \colon \\
&\delta x + a_1 g_1(\widehat{x}_0) + b\, j_\nu(\widehat{x}_0) + a_2 g_2(\widehat{x}_0) \in T_{\widehat{x}_0} \widehat{N}_f \big\}, \qquad \nu = 1, 2.
\end{aligned} \tag{9}
$$

Also let

$$
\begin{aligned}
V_\nu := \big\{&(\delta x, a_1, b, a_2) \in T_{\widehat{x}_0} \mathbb{R}^n \times \mathbb{R}^3 \colon \\
&\delta x + a_1 g_1(\widehat{x}_0) + b\, j_\nu(\widehat{x}_0) + a_2 g_2(\widehat{x}_0) \in T_{\widehat{x}_0} \widehat{N}_f \big\}, \qquad \nu = 1, 2.
\end{aligned} \tag{10}
$$

By [4] we obtain the following:

Theorem 5. *If the second order approximations $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are coercive on $V_{0,1}$ and $V_{0,2}$ respectively, then there exists a smooth function $\alpha \colon \mathbb{R}^n \to \mathbb{R}$ such that $\alpha|_{N_0} \equiv 0$, $\mathrm{d}\alpha(\widehat{x}_0) = \widehat{\ell}_0$ and both $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are coercive quadratic forms on V_1 and V_2, respectively.*

The main result of [14] is the following:

Theorem 6. *Assume $\big(\widehat{T}, \widehat{\xi}, \widehat{u}\big)$ is an admissible triple for the minimum time problem (1). Assume the triple is bang-bang with only one switching time which is a double switching time. Assume the triple satisfies Assumptions 1-5, i.e. PMP, the regularity assumptions along the bang arcs and at the double switching time, insjectivity of the trajectory, and the coercivity assumption. Then, $\widehat{\xi}$ is a strict state-locally optimal trajectory.*

5. The uniqueness of the adjoint covector.

In order to prove our structural stability result we need one further assumption which was not required in [14], i.e. the uniqueness of the adjoint covector associated to the reference triple $\big(\widehat{T}, \widehat{\xi}, \widehat{u}\big)$ in PMP for problem (**$\mathbf{P_0}$**).

Assumption 6. *$\widehat{\mu}$ is the only adjoint covector associated to $\widehat{\xi}$.*

The uniqueness assumption can in fact be stated in terms of the data of the nominal problem (**P_0**). For any $i = 0, 1, 2$, let $\widetilde{f}_i$ be the pull-back of f_i along the reference flow from the double switching time $\widehat{\tau}$ to time 0:

$$\widetilde{f}_i(x) := \widehat{S}_{\widehat{\tau}*}^{-1} f_i \circ \widehat{S}_{\widehat{\tau}}(x) = \exp(-\widehat{\tau}h_1)_* f_i \circ \exp \widehat{\tau}h_1(x).$$

Lemma 7. *Assumption 6 holds if and only if*

$$\operatorname{span}\left\{T_{\widehat{x}_0}N_0,\ T_{\widehat{x}_0}\widehat{N}_f, \widetilde{f}_0(\widehat{x}_0), \widetilde{f}_1(\widehat{x}_0), \widetilde{f}_2(\widehat{x}_0)\right\} = \mathbb{R}^n.$$

Proof. For ease of notation set $C := \operatorname{span}\left\{T_{\widehat{x}_0}N_0,\ T_{\widehat{x}_0}\widehat{N}_f, \widetilde{f}_0(\widehat{x}_0), \widetilde{f}_1(\widehat{x}_0), \widetilde{f}_2(\widehat{x}_0)\right\}$.

1. Let Assumption 6 hold and assume, by contradiction, that $C \neq \mathbb{R}^n$. Then there exists $p \in C^{\perp}$, $p \neq 0$:

$$\langle p\,,\, \delta x\rangle = 0 \quad \forall \delta x \in T_{\widehat{x}_0}N_0 + T_{\widehat{x}_0}\widehat{N}_f, \quad \langle p\,,\, \widetilde{f}_i(\widehat{x}_0)\rangle = 0 \quad \forall i = 0, 1, 2.$$

Let $\mu(t) := (\widehat{p}_0 + p)\,\widehat{S}_{t*}^{-1} = \widehat{\mu}(t) + p\widehat{S}_{t*}^{-1}$.
If $t \in [0, \widehat{\tau}]$ then $\langle p\widehat{S}_{t*}^{-1}\,,\, h_1(\widehat{\xi}(t))\rangle = \langle p\,,\, \left(\widetilde{f}_0 - \widetilde{f}_1 - \widetilde{f}_2\right)(\widehat{x}_0)\rangle = 0$. If $t \in (\widehat{\tau}, T]$, then $\langle p\widehat{S}_{t*}^{-1}\,,\, h_2(\widehat{\xi}(t))\rangle = \langle p\widehat{S}_{\widehat{\tau}*}^{-1}\,,\, h_2(\widehat{\xi}(\widehat{\tau}))\rangle = \langle p\,,\, \left(\widetilde{f}_0 + \widetilde{f}_1 + \widetilde{f}_2\right)(\widehat{x}_0)\rangle = 0$.
As $\mu(0)|_{T_{\widehat{x}_0}N_0 + T_{\widehat{x}_0}\widehat{N}_f} = \widehat{p}_0|_{T_{\widehat{x}_0}N_0 + T_{\widehat{x}_0}\widehat{N}_f}$, it is easily checked that $\lambda(t) := \left(\mu(t), \widehat{\xi}(t)\right)$ satisfies PMP, a contradiction.

2. Assume $C = \mathbb{R}^n$ and suppose, by contradiction, there exists a pair $(\mu(t), p_0)$ in $T^*_{\widehat{\xi}(t)}\mathbb{R}^n \times \{0, 1\}$ which, together with the reference triple $\left(\widehat{T}, \widehat{\xi}, \widehat{u}\right)$ satisfies PMP. Thus the following conditions hold:

$$\begin{aligned} &\langle \mu(t)\,,\, f_0(\widehat{\xi}(t))\rangle + \widehat{u}_1(t)\langle \mu(t)\,,\, f_1(\widehat{\xi}(t))\rangle + \widehat{u}_2(t)\langle \mu(t)\,,\, f_2\widehat{\xi}(t))\rangle = \\ &= F_0(\mu(t), \widehat{\xi}(t)) + \left|F_1(\mu(t), \widehat{\xi}(t))\right| + \left|F_2(\mu(t), \widehat{\xi}(t))\right| = p_0 \in \{0, 1\}; \end{aligned} \tag{11}$$

$$\exists p \in (T_{\widehat{x}_0}N_0)^{\perp} \cap \left(T_{\widehat{x}_0}\widehat{N}_f\right)^{\perp} : \mu(t) = p\widehat{S}_{t*}^{-1}. \tag{12}$$

As in $t = \widehat{\tau}$ the double switch of $\widehat{u}$ occurs, we have

$$\langle \mu(\widehat{\tau})\,,\, f_1(\widehat{\xi}(\widehat{\tau}))\rangle = \langle \mu(\widehat{\tau})\,,\, f_2(\widehat{\xi}(\widehat{\tau}))\rangle = 0,$$

so that $\langle \mu(\widehat{\tau})\,,\, f_0(\widehat{\xi}(\widehat{\tau}))\rangle = p_0$, that is:

$$\langle p\,,\, \widetilde{f}_1(\widehat{x}_0)\rangle = \langle p\,,\, \widetilde{f}_2(\widehat{x}_0)\rangle = 0, \quad \langle p\,,\, \widetilde{f}_0(\widehat{x}_0)\rangle = p_0. \tag{13}$$

We now distinguish between two cases:

1. if $\left(\mu(t), \widehat{\xi}(t)\right)$ is an abnormal extremal ($p_0 = 0$) then, by (12) and (13), $p \in C^{\perp}$. As $C = \mathbb{R}^n$ this means that $p = 0$, so that $\mu(t) \equiv 0$, a contradiction in PMP.

2. if $\left(\mu(t), \widehat{\xi}(t)\right)$ is a normal extremal ($p_0 = 1$) then, by (12) and (13), p acts on $C = \mathbb{R}^n$ in the same way as $\widehat{p}_0$, so that $p = \widehat{p}_0$ and $\mu(t) = \widehat{\mu}(t)$, i.e. $\widehat{\mu}$ is the only adjoint covector associated to $\widehat{\xi}$. ∎

6. The main result. We are now in a position to state the main results of this paper, which will be proved in the following sections.

THEOREM 8. *Under Assumptions 1-6 there exists* $\widetilde{R} \in (0, R)$ *such that for any* $r \in B_{\widetilde{R}}$*, problem* ($\mathbf{P_r}$) *has a bang-bang state-local minimiser* (T^r, u^r, ξ^r)*. Each control component of* u^r *has exactly one switching time. Let* τ_i^r *be the switching time of* u_i^r*,* $i = 1, 2$*. At time* τ_i^r *the control component* u_i^r *switches from the value* -1 *to the value* 1*. The final time* T^r *and the switching times* τ_1^r*,* τ_2^r *depend smoothly on* r*.*

REMARK 6.1. *Notice that the switching times* τ_1^r*,* τ_2^r *in Theorem 8 may either coincide or be different, i.e. we may either have a double switching time or two simple switching times, still the bang-bang structure is preserved and singular arcs cannot occur.*

THEOREM 9. *Under Assumptions 1-6 there exists* $\widetilde{R} \in (0, R)$*,* $\varepsilon > 0$ *and a neighborhood* $\mathcal{V}$ *of the graph of* $\widehat{\lambda}$ *in* $\mathbb{R} \times T^*\mathbb{R}^n$ *such that for any* $r \in B_{\widetilde{R}}$*, the extremal pair* λ^r *associated to the local minimum time triple* (T^r, u^r, ξ^r) *of Theorem 8 is the only extremal pair whose final time is in* $[\widehat{T} - \varepsilon, \widehat{T} + \varepsilon]$ *and whose graph is in* $\mathcal{V}$*.*

6.1. The coercivity of the second order variations. In [14], in order to prove the strong local optimality result, the authors consider the bilinear form Q_ν associated to $\mathcal{Q}_\nu$, $\nu = 1, 2$, i.e. if $\delta e = (\delta x, a_1, b, a_2)$, $\delta f = (\delta y, c_1, d, c_2) \in V_{0,\nu}$ then

$$\begin{aligned} Q_\nu[\delta e, \delta f] = & \, \mathrm{D}^2(\alpha + \widehat{\beta})(\widehat{x}_0)(\delta x, \delta y) + \delta y \cdot (a_1 g_1 + b\, j_\nu + a_2 g_2) \cdot \widehat{\beta}\,(\widehat{x}_0) \\ & + \delta x \cdot (c_1 g_1 + d\, j_\nu + c_2 g_2) \cdot \widehat{\beta}\,(\widehat{x}_0) \\ & + (c_1 g_1 + d\, j_\nu + c_2 g_2) \cdot (a_1 g_1 + b\, j_\nu + a_2 g_2) \cdot \widehat{\beta}\,(\widehat{x}_0) \\ & + d a_1 [g_1, j_\nu] \cdot \widehat{\beta}\,(\widehat{x}_0) + c_2 a_1 [g_1, g_2] \cdot \widehat{\beta}\,(\widehat{x}_0) + c_2 b [j_\nu, g_2] \cdot \widehat{\beta}\,(\widehat{x}_0) \,. \end{aligned}$$

The bilinear forms Q_ν can be written in a more compact way by introducing the linear Hamiltonians

$$\begin{aligned} G_i'' &\colon (\delta p, \delta x) \in (\mathbb{R}^n)^* \times \mathbb{R}^n \mapsto \langle \delta p\,,\, g_i(\widehat{x}_0) \rangle + \delta x \cdot g_i \cdot \widehat{\beta}\,(\widehat{x}_0) \in \mathbb{R}, \\ J_\nu'' &\colon (\delta p, \delta x) \in (\mathbb{R}^n)^* \times \mathbb{R}^n \mapsto \langle \delta p\,,\, j_\nu(\widehat{x}_0) \rangle + \delta x \cdot j_\nu \cdot \widehat{\beta}\,(\widehat{x}_0) \in \mathbb{R}, \end{aligned}$$

and the associated constant Hamiltonian vector fields $\overrightarrow{G}_i''$ and $\overrightarrow{J_\nu}''$. An easy computation shows that

$$\boldsymbol{\sigma}\left((\delta p, \delta x), \overrightarrow{G}_1'' \right) = G_i''(\delta p, \delta x), \qquad \boldsymbol{\sigma}\left((\delta p, \delta x), \overrightarrow{J_\nu}'' \right) = J_\nu''(\delta p, \delta x),$$

$$G_i''(\overrightarrow{G}_j'') = [g_j, g_i] \cdot \widehat{\beta}\,(\widehat{x}_0)\,, \qquad G_i''(\overrightarrow{J_\nu}'') = [j_\nu, g_i] \cdot \widehat{\beta}\,(\widehat{x}_0) = -J_\nu''(\overrightarrow{G}_i'').$$

With these equalities at hand it is just a straightforward computation to prove the following proposition.

PROPOSITION 10. *For any admissible variation* $\delta e = (\delta x, a_1, b, a_2) \in V_\nu$ *and any* $\delta p \in (\mathbb{R}^n)^*$ *let*

$$(\delta p_T, \delta x_T) := (\delta p, \delta x) + a_1 \overrightarrow{G}_1'' + b\overrightarrow{J_\nu}'' + a_2 \overrightarrow{G}_2''.$$

Then

$$\begin{aligned} Q_\nu[\delta e, \delta f] = & \, \mathrm{D}^2(\alpha + \widehat{\beta})(\widehat{x}_0)(\delta x, \delta y) + \langle \delta p\,,\, \delta y \rangle - \langle \delta p_T\,,\, \delta y + c_1 g_1 + d j_\nu + c_2 g_2 \rangle \\ & + c_1 G_1''\,(\delta p, \delta x) + d J_\nu'' \left((\delta p, \delta x) + a_1 \overrightarrow{G}_1'' \right) + c_2 G_2'' \left((\delta p, \delta x) + a_1 \overrightarrow{G}_1'' + d\overrightarrow{J_\nu}'' \right). \end{aligned}$$

Proposition 11. *Let* $\delta e = (\delta x, a_1, b, a_2)$ *be an admissible variation such that* $\delta e \in V_{0,\nu}$.

1. $\delta e \in V_{0,\nu}^{\perp_{Q_\nu}}$ *if and only if there exists* $\delta p \in (\mathbb{R}^n)^*$ *such that*

$$\text{(14a)} \qquad \delta p = -\mathrm{D}^2(\alpha + \widehat{\beta})(\widehat{x}_0)(\delta x, \cdot) + \omega_0, \qquad \omega_0 \in \left(T_{\widehat{x}_0} N_0\right)^{\perp},$$

$$\text{(14b)} \qquad G_1''(\delta p, \delta x) = \boldsymbol{\sigma}\left((\delta p, \delta x), \overrightarrow{G}_1''\right) = 0,$$

$$\text{(14c)} \qquad J_\nu''\left((\delta p, \delta x) + a_1 \overrightarrow{G}_1''\right) = \boldsymbol{\sigma}\left((\delta p, \delta x) + a_1 \overrightarrow{G}_1'', \overrightarrow{J_\nu}''\right) = 0,$$

$$\text{(14d)} \qquad G_2''\left((\delta p, \delta x) + a_1 \overrightarrow{G}_1'' + b\overrightarrow{J_\nu}''\right) = \boldsymbol{\sigma}\left((\delta p, \delta x) + a_1 \overrightarrow{G}_1'' + b\overrightarrow{J_\nu}'', \overrightarrow{G}_2''\right) = 0,$$

$$\text{(14e)} \qquad \delta p_T \in \left(T_{\widehat{x}_0} \widehat{N}_f\right)^{\perp}.$$

2. Assume the coercivity assumption, Assumption 5, holds. If there exists $\delta p \in (\mathbb{R}^n)^*$ *such that equations (14) hold, then* δe *is the trivial variation* $(0, 0, 0, 0)$.

Consider the Lagrangian manifold of the initial transversality conditions

$$\Lambda_0 := \left\{\ell = \mathrm{d}\alpha(x) + \omega \colon x \in N_0,\ \omega \in (T_x N_0)^{\perp},\ H_1(\ell) = 1\right\}$$

so that

$$T_{\widehat{\ell}_0} \Lambda_0 := \left\{\delta\ell = \mathrm{d}\alpha_* \delta x + \omega \colon \delta x \in T_{\widehat{x}_0} N_0,\ \omega \in (T_{\widehat{x}_0} N_0)^{\perp},\ \boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{H_1}(\widehat{\ell}_0)\right) = 0\right\}.$$

Let $i\colon (\delta p, \delta x) \in (\mathbb{R}^n)^* \times \mathbb{R}^n \mapsto \delta\ell := -\delta p + \mathrm{d}(-\widehat{\beta})\delta x \in T^*\mathbb{R}^n$. The map i is an antisymplectic isomorphism. Moreover

$$\begin{aligned}
i\overrightarrow{G}_1'' &= \overrightarrow{H_1}(\widehat{\ell}_0) = \overrightarrow{G}_1(\widehat{\ell}_0) = \widehat{\mathcal{F}}_{\widehat{\tau}\,*}^{-1} \overrightarrow{H_1} \circ \widehat{\mathcal{F}}_{\widehat{\tau}}(\widehat{\ell}_0), \\
i\overrightarrow{G}_2'' &= \overrightarrow{G}_2(\widehat{\ell}_0) = \widehat{\mathcal{F}}_{\widehat{\tau}\,*}^{-1} \overrightarrow{H_2} \circ \widehat{\mathcal{F}}_{\widehat{\tau}}(\widehat{\ell}_0) = \widehat{\mathcal{F}}_{\widehat{T}\,*}^{-1} \overrightarrow{H_2} \circ \widehat{\mathcal{F}}_{\widehat{T}}(\widehat{\ell}_0), \\
i\overrightarrow{J_\nu}'' &= \overrightarrow{J_\nu}(\widehat{\ell}_0) = \widehat{\mathcal{F}}_{\widehat{\tau}\,*}^{-1} \overrightarrow{K_\nu} \circ \widehat{\mathcal{F}}_{\widehat{\tau}}(\widehat{\ell}_0) \qquad \nu = 1, 2,
\end{aligned}$$

and $T_{\widehat{\ell}_0} \Lambda_0 = iL_0''$ where

$$L_0'' := \left\{(\delta p, \delta x) : \delta x \in T_{\widehat{x}_0} N_0,\ \delta p = -\mathrm{D}^2(\alpha + \widehat{\beta})(\widehat{x}_0)(\delta x, \cdot) + \omega,\right.$$
$$\left. \omega \in (T_{\widehat{x}_0} N_0)^{\perp},\ \boldsymbol{\sigma}\left((\delta p, \delta x), \overrightarrow{G}_1''\right) = 0\right\}.$$

Lemma 12. *Under Assumptions 1 to 6 there exist* $\widetilde{R} \in (0, R)$, $\varepsilon > 0$ *and a neighborhood* $\mathcal{O}(\widehat{\ell}_0)$ *of* $\widehat{\ell}_0$ *in* $T^*\mathbb{R}^n$ *such that for any* $r \in B_{\widetilde{R}}$, *there exists a unique bang-bang extremal pair* $\lambda^r = (\mu^r, \xi^r)$ *of* (**P**$_{\mathbf{r}}$) *having the following properties:*

1. λ^r *is a normal extremal and* $\lambda^r(0) \in \mathcal{O}(\widehat{\ell}_0)$*;*
2. *each component* u_i^r, $i = 1, 2$ *of the associated control* $u^r = (u_1^r, u_2^r)$ *has exactly one switching time* τ_i^r*;* $\tau_1^r, \tau_2^r \in [\widehat{\tau} - \varepsilon, \widehat{\tau} + \varepsilon]$*; at time* τ_i^r *the control component* u_i^r *switches from the value* -1 *to the value* $+1$*;*
3. $T^r \in [\widehat{T} - \varepsilon, \widehat{T} + \varepsilon]$*;*
4. τ_1^r, τ_2^r, T^r *and* $\lambda^r(0)$ *depend smoothly on* r*,*
5. *the bang arcs are regular: for* $i = 1, 2 \quad u_i^r(t) F_i^r(\lambda^r(t)) > 0 \quad \forall t \neq \tau_i^r$*,*
6. *each switching time is regular:* $\left.\dfrac{\mathrm{d}}{\mathrm{d}t} u_i^r(t) F_i^r(\lambda^r(t))\right|_{t = \tau_i^r \pm} > 0, \quad i = 1, 2.$

Proof. We prove claims 1-4 applying the implicit function theorem: for $\nu = 1, 2$ consider the following system of $2n+3$ scalar equations in the unknowns $r \in B_R$, $\ell = (p, x) \in T^*\mathbb{R}^n$, t_1, t_2, $t_3 \in \mathbb{R}$:

(15a) $\quad \ell \in \left(T_{\pi\ell} N_0^r\right)^\perp \times N_0^r,$

(15b) $\quad H_1^r(\ell) - 1 = 0,$

(15c) $\quad K_\nu^r \circ \exp t_1 \overrightarrow{H_1^r}(\ell) - 1 = 0,$

(15d) $\quad H_2^r \circ \exp(t_2 - t_1)\overrightarrow{K_\nu^r} \circ \exp t_1 \overrightarrow{H_1^r}(\ell) - 1 = 0,$

(15e)
$$\exp(t_3 - t_2)\overrightarrow{H_2^r} \circ \exp(t_2 - t_1)\overrightarrow{K_\nu^r} \circ \exp t_1 \overrightarrow{H_1^r}(\ell) \in \left(T_{\pi \exp(t_3-t_2)\overrightarrow{H_2^r} \circ \exp(t_2-t_1)\overrightarrow{K_\nu^r} \circ \exp t_1 \overrightarrow{H_1^r}(\ell)} N_f^r\right)^\perp \times N_f^r.$$

Equations (15) represent the structure of the reference extremal that we want to preserve: equation (15a) is the initial condition of problem ($\mathbf{P_r}$), together with the initial transversality condition. Equations (15b) to (15d) represent the control structure (i.e. each compoment switches from the value -1 to the value 1), while equation (15e) is the final condition of problem ($\mathbf{P_r}$), , together with the final transversality condition.

The linearised equations with respect to (ℓ, t_1, t_2, t_3) at the point $(r, \ell, t_1, t_2, t_3) = (0, \widehat{\ell}_0, \widehat{\tau}, \widehat{\tau}, \widehat{T})$ are given by

(16a) $\quad \delta\ell = (\delta p, \delta x) \in T_{\widehat{\ell}_0}\left(\left(T_{\widehat{x}_0} N_0\right)^\perp \times N_0\right),$

(16b) $\quad \boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{H_1}(\widehat{\ell}_0)\right) = 0,$

(16c) $\quad \boldsymbol{\sigma}\left(\exp \widehat{\tau}\overrightarrow{H_1}_* \delta\ell + \delta t_1 \overrightarrow{H_1}(\widehat{\ell}_d), (\overrightarrow{K_\nu} - \overrightarrow{H_1})(\widehat{\ell}_d)\right) = 0,$

(16d) $\quad \boldsymbol{\sigma}\left(\exp \widehat{\tau}\overrightarrow{H_1}_* \delta\ell - \delta t_1 (\overrightarrow{K_\nu} - \overrightarrow{H_1})(\widehat{\ell}_d) + \delta t_2 \overrightarrow{K_\nu}(\widehat{\ell}_d), (\overrightarrow{H_2} - \overrightarrow{K_\nu})(\widehat{\ell}_d)\right) = 0,$

(16e)
$$\widehat{\mathcal{F}}_{\widehat{T}\,*}\delta\ell + \exp(\widehat{T} - \widehat{\tau})\overrightarrow{H_2}_*\left(\delta t_1 \overrightarrow{H_1} + (\delta t_2 - \delta t_1)\overrightarrow{K_\nu} + (\delta t_3 - \delta t_2)\overrightarrow{H_2}\right)(\widehat{\ell}_d) \in T_{\widehat{\ell}_f}\left(\left(T_{\widehat{x}_f} N_f\right)^\perp \times N_f\right).$$

Notice that $\alpha|_{N_0} \equiv 0$ so that $\mathrm{d}\alpha(x) \in (T_x N_0)^\perp$ for any $x \in N_0$. Thus for every $\ell = (p, x) \in (T_{\pi\ell} N_0)^\perp \times N_0$, it holds $p - \mathrm{d}\alpha(x) \in (T_x N_0)^\perp$ so that, if $\delta\ell = (\delta p, \delta x)$ we get $\delta x \in T_{\widehat{x}_0} N_0$ $\delta p - \mathrm{D}^2\alpha(\widehat{x}_0)[\delta x, \cdot] \in (T_{\widehat{x}_0} N_0)^\perp$. Thus, taking the pull-back to time $t = 0$, the homogeneous linear system (16) admits a nontrivial solution if and only if there exists $\delta\ell = (\delta p, \delta x) \in T_{\widehat{\ell}_0} T^*\mathbb{R}^n$, δt_1, δt_2, $\delta t_3 \in \mathbb{R}$, with at least one of them being different from zero, such that

(17a) $\quad \delta\ell = \mathrm{d}\alpha_* \delta x + \omega_0, \qquad \delta x \in T_{\widehat{x}_0} N_0,\ \omega_0 \in T_{\widehat{p}_0}\left(T_{\widehat{x}_0} N_0\right)^\perp,$

(17b) $\quad \boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{G}_1(\widehat{\ell}_0)\right) = 0,$

(17c) $\quad \boldsymbol{\sigma}\left(\delta\ell + \delta t_1 \overrightarrow{G}_1(\widehat{\ell}_0), (\overrightarrow{J_\nu} - \overrightarrow{G}_1)(\widehat{\ell}_0)\right) = 0,$

(17d) $\quad \boldsymbol{\sigma}\left(\delta\ell - \delta t_1 (\overrightarrow{J_\nu} - \overrightarrow{G}_1)(\widehat{\ell}_0) + \delta t_2 \overrightarrow{J_\nu}(\widehat{\ell}_0), (\overrightarrow{G}_2 - \overrightarrow{J_\nu})(\widehat{\ell}_0)\right) = 0,$

(17e) $\quad \delta x_f := \delta x + \left(\delta t_1 g_1 + (\delta t_2 - \delta t_1) j_\nu + (\delta t_3 - \delta t_2) g_2\right)(\widehat{x}_0) \in T_{\widehat{x}_0} \widehat{N}_f,$

$$\delta\ell + \left(\delta t_1 \overrightarrow{G}_1 + (\delta t_2 - \delta t_1)\overrightarrow{J_\nu} + (\delta t_3 - \delta t_2)\overrightarrow{G}_2\right)(\widehat{\ell}_0) = \mathrm{d}(-\widehat{\beta})_* \delta x_f + \omega_f, \qquad \omega_f \in T_{\widehat{p}_0}\left(T_{\widehat{x}_0}\widehat{N}_f\right)^\perp . \tag{17f}$$

Applying the anti symplectic isomorphism i^{-1} and denoting $i^{-1}\delta\ell = (\delta p, \delta x)$, equations (17) can also be written as

$$(\delta p, \delta x) = -\mathrm{D}^2(\alpha + \widehat{\beta})(\widehat{x}_0)(\delta x, \cdot) + \omega_0, \qquad \delta x \in T_{\widehat{x}_0}N_0,\ \omega_0 \in (T_{\widehat{x}_0}N_0)^\perp , \tag{18a}$$

$$\boldsymbol{\sigma}\left((\delta p, \delta x), \overrightarrow{G}_1''\right) = 0, \tag{18b}$$

$$\boldsymbol{\sigma}\left((\delta p, \delta x) + \delta t_1 \overrightarrow{G}_1'', \overrightarrow{J_\nu}''\right) = 0, \tag{18c}$$

$$\boldsymbol{\sigma}\left((\delta p, \delta x) + \delta t_1 \overrightarrow{G}_1'' + (\delta t_2 - \delta t_1)\overrightarrow{J_\nu}'', \overrightarrow{G}_2''\right) = 0, \tag{18d}$$

$$(\delta p_T, \delta x_T) := (\delta p, \delta x) + \delta t_1 \overrightarrow{G}_1'' + (\delta t_2 - \delta t_1)\overrightarrow{J_\nu}'' + (\delta t_3 - \delta t_2)\overrightarrow{G}_2'' \in \left(T_{\widehat{x}_0}\widehat{N}_f\right)^\perp \times T_{\widehat{x}_0}\widehat{N}_f . \tag{18e}$$

Thus, by claim 1 of Proposition 11, the variation $(\delta x, \delta t_1, \delta t_2 - \delta t_1, \delta t_3 - \delta t_2)$ is in $V_{0,\nu} \cap V_{0,\nu}^{\perp_{Q_\nu}}$. As $\mathcal{Q}_\nu$ is coercive on $V_{0,\nu}$ we can apply claim 2 of Proposition 11 and we get $\delta x = 0$, $\delta t_1 = \delta t_2 = \delta t_3 = 0$, so that $\delta p = 0$ if and only if $\omega_0 = 0$. By equations (18),

$$\omega_0 \in \operatorname{span}\left\{T_{\widehat{x}_0}N_0,\ T_{\widehat{x}_0}\widehat{N}_f,\ g_1(\widehat{x}_0), j_\nu(\widehat{x}_0), g_2(\widehat{x}_0)\right\}^\perp = \\ = \operatorname{span}\left\{T_{\widehat{x}_0}N_0,\ T_{\widehat{x}_0}\widehat{N}_f,\ \widetilde{f}_0(\widehat{x}_0), \widetilde{f}_1(\widehat{x}_0),\ \widetilde{f}_2(\widehat{x}_0)\right\}^\perp ,$$

so that Assumption 6 and Lemma 7 yield the claim.

Thus we can apply the implicit function theorem to system (15). For $r \in B_{\widetilde{R}}$ let $(\ell_0^r, \tau_1^r, \tau_2^r, T^r)$, $\ell_0^r = (p_0^r, x_0^r)$ be the solution of system (15). The piecewise smooth curve $\lambda^r(t) = (\mu^r(t), \xi^r(t))$ defined by

$$\begin{cases} \left.\begin{array}{ll} \exp t\overrightarrow{H_1^r}(\ell_0^r), & t \in [0, \tau_1^r], \\ \exp(t - \tau_1^r)\overrightarrow{K_1^r} \circ \exp \tau_1^r \overrightarrow{H_1^r}(\ell_0^r), & t \in [\tau_1^r, \tau_2^r], \\ \exp(t - \tau_2^r)\overrightarrow{H_2^r} \circ \exp(\tau_2^r - \tau_1^r)\overrightarrow{K_1^r} \circ \exp \tau_1^r \overrightarrow{H_1^r}(\ell_0^r), & t \in [\tau_2^r, T^r], \end{array}\right\} & \text{if } \tau_1^r < \tau_2^r \\ \left.\begin{array}{ll} \exp t\overrightarrow{H_1^r}(\ell_0^r), & t \in [0, \tau_2^r], \\ \exp(t - \tau_2^r)\overrightarrow{K_2^r} \circ \exp \tau_2^r \overrightarrow{H_1^r}(\ell_0^r), & t \in [\tau_2^r, \tau_1^r], \\ \exp(t - \tau_1^r)\overrightarrow{H_2^r} \circ \exp(\tau_1^r - \tau_2^r)\overrightarrow{K_2^r} \circ \exp \tau_2^r \overrightarrow{H_1^r}(\ell_0^r), & t \in [\tau_1^r, T^r], \end{array}\right\} & \text{if } \tau_2^r < \tau_1^r, \\ \left.\begin{array}{ll} \exp t\overrightarrow{H_1^r}(\ell_0^r), & t \in [0, \tau_1^r], \\ \exp(t - \tau_2^r)\overrightarrow{H_2^r} \circ \exp \tau_2^r \overrightarrow{H_1^r}(\ell_0^r), & t \in [\tau_1^r, T^r], \end{array}\right\} & \text{if } \tau_2^r = \tau_1^r, \end{cases}$$

is a normal extremal of problem ($\mathbf{P_r}$) and satisfies claims 1-4

We can now complete the proof by proving claims 5-6: possibly restricting $\widetilde{R}$ and $\mathcal{O}(\widehat{\ell}_0)$ we can assume, by continuity

$$\begin{array}{ll} F_i^r(\lambda^r(t)) < 0 & \forall t \in [0, \widehat{\tau} - \varepsilon], \\ F_i^r(\lambda^r(t)) > 0 & \forall t \in [\widehat{\tau} + \varepsilon, T^r], \end{array} \qquad i = 1, 2,$$

$$\begin{aligned} &\boldsymbol{\sigma}\left(\overrightarrow{H_1^r}, \overrightarrow{K_\nu^r}\right)(\lambda^r(t)) > 0 \\ &\boldsymbol{\sigma}\left(\overrightarrow{K_\nu^r}, \overrightarrow{H_2^r}\right)(\lambda^r(t)) > 0 \end{aligned} \qquad \forall t \in [\widehat{\tau}-\varepsilon, \widehat{\tau}+\varepsilon], \qquad \nu = 1, 2, \tag{19}$$

By construction, λ^r is a normal Pontryagin extremal of ($\mathbf{P_r}$). We prove claim 5 in the case when $\tau_1^r < \tau_2^r$. The other cases are analougous. For any $t \in (\widehat{\tau}-\varepsilon, \tau_1^r)$ there exists $\theta_1 \in (t, \tau_1^r)$ such that

$$\begin{aligned} 2F_1^r(\lambda^r(t)) &= 2F_1^r(\lambda^r(\tau_1^r)) + (t-\tau_1^r)\frac{\mathrm{d}(2F_1^r \circ \lambda^r)}{\mathrm{d}t}(\theta_1) \\ &= (t-\tau_1^r)\boldsymbol{\sigma}\left(\overrightarrow{H_1^r}, 2\overrightarrow{F_1^r}\right)(\lambda^r(\theta_1)) = (t-\tau_1^r)\boldsymbol{\sigma}\left(\overrightarrow{H_1^r}, \overrightarrow{K_1^r}\right)(\lambda^r(\theta_1)) \end{aligned}$$

which is negative by (19). Analougously, for any $t \in (\tau_1^r, \tau_2^r]$ there exists $\theta_2 \in (\tau_1^r, t)$ such that

$$\begin{aligned} 2F_1^r(\lambda^r(t)) &= 2F_1^r(\lambda^r(\tau_1^r)) + (t-\tau_1^r)\frac{\mathrm{d}(2F_1^r \circ \lambda^r)}{\mathrm{d}t}(\theta_2) \\ &= (t-\tau_1^r)\boldsymbol{\sigma}\left(\overrightarrow{K_1^r}, 2\overrightarrow{F_1^r}\right)(\lambda^r(\theta_2)) = (t-\tau_1^r)\boldsymbol{\sigma}\left(\overrightarrow{H_1^r}, \overrightarrow{K_1^r}\right)(\lambda^r(\theta_2)) \end{aligned} \tag{20}$$

which is positive by (19). Finally, if $t \in (\tau_2^r, \widehat{\tau}+\varepsilon)$ there exists $\theta_3 \in (\tau_2^r, \widehat{T}+\varepsilon)$ such that

$$\begin{aligned} 2F_1^r(\lambda^r(t)) &= 2F_1^r(\lambda^r(\tau_2^r)) + (t-\tau_2^r)\frac{\mathrm{d}(2F_1^r \circ \lambda^r)}{\mathrm{d}t}(\theta_3) \\ &= 2F_1^r(\lambda^r(\tau_2^r)) + (t-\tau_2^r)\boldsymbol{\sigma}\left(\overrightarrow{H_2^r}, 2\overrightarrow{F_1^r}\right)(\lambda^r(\theta_3)) \\ &= 2F_1^r(\lambda^r(\tau_2^r)) + (t-\tau_2^r)\boldsymbol{\sigma}\left(\overrightarrow{K_2^r}, \overrightarrow{H_2^r}\right)(\lambda^r(\theta_3)) \end{aligned}$$

which is positive by (19) and (20). The proof for the sign of $F_2^r(\lambda^r(t))$ follows the same line.

Finally, the switching times τ_i^r are regular (claim 6) thanks to inequalities (19).□

We can now prove Theorem 8, i.e. we prove that projection ξ^r of the extremal λ^r defined in Lemma 12 is a state-local optimal trajectory for problem ($\mathbf{P_r}$).

Proof of Theorem 8. By construction and by Lemma 12, $(T^r, \xi^r = \pi\lambda^r, u^r)$ satisfies PMP in its normal form and the regularity assumptions for problem ($\mathbf{P_r}$). Thus it suffices to prove that ξ^r has no self-intersection and that the second order variation associated to ($\mathbf{P_r}$) is coercive.

Injectivity of ξ^r. We prove that, possibly restricting $\widetilde{R}$, then for any $r \in B_{\widetilde{R}}$, the trajectory ξ^r has no self intersection. The proof is carried out by showing, with a contradiction argument, that there can be no sequence $\{r_k\}_{k\in\mathbb{N}} \subset B_{\widetilde{R}}$ that converges to 0 and such that the trajectory ξ^{r_k} is not injective.

Assume by contradiction there exists a sequence $\{r_k\}_{k\in\mathbb{N}} \subset B_{\widetilde{R}}$ that converges to 0 and such that there exist $t_{1,k}$, $t_{2,k}$, $0 \le t_{1,k} < t_{2,k} \le T^{r_k}$, $\xi^{r_k}(t_{1,k}) = \xi^{r_k}(t_{2,k})$. Up to a subsequence both $t_{1,k}$ and $t_{2,k}$ converge. Let $\bar{t}_i := \lim_{k\to\infty} t_{i,k} \in [0, \widehat{T}]$, $i = 1, 2$. If $\bar{t}_1 < \bar{t}_2$, then $\widehat{\xi}(\bar{t}_1) = \widehat{\xi}(\bar{t}_2)$, that is a contradiction. Assume then $\bar{t}_1 = \bar{t}_2 =: \bar{t}$. Different cases may occur:

1. Up to a subsequence $0 \le t_{1,k} < t_{2,k} \le \tau_1^{r_k}$. In this case

$$0 = \xi^{r_k}(t_{2,k}) - \xi^{r_k}(t_{1,k}) = \int_{t_{1,k}}^{t_{2,k}} h_1^r(\xi^{r_k}(s))\,\mathrm{d}s.$$

Applying the mean value theorem componentwise we get:

$$\forall j = 1, \ldots, n \qquad \exists s_j^k \in [t_{1,k}, t_{2,k}] \colon (h_1^r)_j(\xi^{r_k}(s_j^k)) = 0. \tag{21}$$

Thus, as $k \to \infty$ in (21) we obtain $h_1(\widehat{\xi}(\bar{t})) = 0$, a contradiction since $\bar{t} \in [0, \widehat{\tau}_1]$ and $H_1(\widehat{\lambda}(t)) = 1 \quad \forall t \in [0, \widehat{\tau}_1]$.

2. Up to a subsequence $0 \le t_{1,k} \le \tau_1^{r_k} < t_{2,k} \le \tau_2^{r_k}$. In this case

$$0 = \frac{\xi^{r_k}(t_{2,k}) - \xi^{r_k}(t_{1,k})}{t_{2,k} - t_{1,k}} = \frac{\tau_1^{r_k} - t_{1,k}}{t_{2,k} - t_{1,k}} \fint_{t_{1,k}}^{\tau_1^{r_k}} h_1^r(\xi^{r_k}(s))\,\mathrm{d}s + \\ + \frac{t_{2,k} - \tau_1^{r_k}}{t_{2,k} - t_{1,k}} \fint_{\tau_1^{r_k}}^{t_{2,k}} k_1^r(\xi^{r_k}(s))\,\mathrm{d}s. \tag{22}$$

Up to a subsequence there exists $\lim_{k\to\infty} \dfrac{\tau_1^{r_k} - t_{1,k}}{t_{2,k} - t_{1,k}} = c \in [0, 1]$, so that passing to the limit in (22) we obtain

$$0 = c\, h_1(\widehat{x}_d) + (1 - c)\, k_1(\widehat{x}_d) = f_0(\widehat{x}_d) + (1 - 2c)\, f_1(\widehat{x}_d) - f_2(\widehat{x}_d).$$

A contradiction, since $f_0(\widehat{x}_d) \notin \operatorname{span}\{f_1(\widehat{x}_d), f_2(\widehat{x}_d)\}$.

3. Up to a subsequence $0 \le t_{1,k} \le \tau_1^{r_k} \le \tau_2^{r_k} \le t_{2,k}$. In this case

$$0 = \frac{\xi^{r_k}(t_{2,k}) - \xi^{r_k}(t_{1,k})}{t_{2,k} - t_{1,k}} = \frac{\tau_1^{r_k} - t_{1,k}}{t_{2,k} - t_{1,k}} \fint_{t_{1,k}}^{\tau_1^{r_k}} h_1^r(\xi^{r_k}(s))\,\mathrm{d}s + \\ + \frac{\tau_2^{r_k} - \tau_1^{r_k}}{t_{2,k} - t_{1,k}} \fint_{\tau_1^{r_k}}^{\tau_2^{r_k}} k_1^r(\xi^{r_k}(s))\,\mathrm{d}s + \frac{t_{2,k} - \tau_2^{r_k}}{t_{2,k} - t_{1,k}} \fint_{\tau_2^{r_k}}^{t_{2,k}} h_2^r(\xi^{r_k}(s))\,\mathrm{d}s. \tag{23}$$

Up to a subsequence there exist

$$\lim_{k\to\infty} \frac{\tau_1^{r_k} - t_{1,k}}{t_{2,k} - t_{1,k}} = c_1 \in [0, 1], \qquad \lim_{k\to\infty} \frac{\tau_2^{r_k} - \tau_1^{r_k}}{t_{2,k} - t_{1,k}} = c_2 \in [0, 1],$$

so that passing to the limit in (23) we obtain

$$0 = c_1\, h_1(\widehat{x}_d) + c_2\, k_1(\widehat{x}_d) + (1 - c_1 - c_2)\, h_2(\widehat{x}_d) = \\ = f_0(\widehat{x}_d) + (1 - 2c_1)\, f_1(\widehat{x}_d) - (1 - 2c_1 - 2c_2) f_2(\widehat{x}_d).$$

A contradiction, since $f_0(\widehat{x}_d) \notin \operatorname{span}\{f_1(\widehat{x}_d), f_2(\widehat{x}_d)\}$, see Remark 3.3.
In the other cases the proof follows the same line.

Coercivity of the second variation. Let (T^r, λ^r, u^r) be the extremal defined in Lemma 12, let $\xi^r := \pi\lambda^r$ and $x_0^r := \xi^r(0)$. Assume $\tau_1^r < \tau_2^r$. In this case the trajectory ξ^r is driven by the dynamics

$$\phi_t^r := \begin{cases} h_1^r, & t \in [0, \tau_1^r], \\ k_1^r, & t \in (\tau_1^r, \tau_2^r], \\ h_2^r, & t \in (\tau_2^r, T^r]. \end{cases}$$

Let S_t^r be the flow at time t associated to ϕ_t^r and consider the pull-back vector fields

$$
\begin{aligned}
g_1^r(x) &:= (S^r_{t\,*})^{-1} h_1^r \circ S_t^r(x), \quad t \in [0, \tau_1^r], \\
j_1^r(x) &:= (S^r_{t\,*})^{-1} k_1^r \circ S_t^r(x), \quad t \in [\tau_1^r, \tau_2^r], \\
g_2^r(x) &:= (S^r_{t\,*})^{-1} h_2^r \circ S_t^r(x), \quad t \in [\tau_2^r, T^r].
\end{aligned}
$$

Let α^r be a function that vanishes on N_0^r and such that $\mathrm{d}\alpha^r(\xi^r(0)) = \lambda^r(0)$. Let β^r be a smooth function that vanishes on $(S^r_{T^r})^{-1}(N_f^r)$, such that $\mathrm{d}\beta^r(\xi^r(0)) = -\lambda^r(0)$. Finally consider the linearisation of the constraints

$$
V_0^r := \left\{ \delta e = (\delta x, a_1, b, a_2) \in T_{x_0^r} N_0^r \times \mathbb{R}^3 \colon \delta x + a_1 g_1^r + b j_1^r + a_2 g_2^r \in T_{x_0^r} N_0^r \right\}.
$$

Then the second variation at the switching points, see e.g. [11], is given by

$$
\begin{aligned}
\mathcal{Q}_r[\delta e]^2 = {}& \frac{1}{2} \mathrm{D}^2 \left(\alpha^r + \beta^r \right)(x_0^r)[\delta x]^2 + \delta x \cdot (a_1 g_1^r + b j_1^r + a_2 g_2^r) \cdot \beta^r\,(x_0^r) \\
&+ \frac{1}{2}(a_1 g_1^r + b j_1^r + a_2 g_2^r)^2 \cdot \beta^r\,(x_0^r) + \frac{1}{2} a_1 b\,[g_1^r, j_1^r] \cdot \beta^r\,(x_0^r) \\
&+ \frac{1}{2} a_1 a_2\,[g_1^r, g_2^r] \cdot \beta^r\,(x_0^r) + \frac{1}{2} a_2 b\,[j_1^r, g_2^r] \cdot \beta^r\,(x_0^r).
\end{aligned}
$$

We now show, with a contradiction argument, that $\mathcal{Q}_r$ is coercive on V_0^r: assume there exists a sequence $\{r_k\}_{k\in\mathbb{N}} \subset (0, \widetilde{R})$ that converges to 0 and such that $\mathcal{Q}_{r_k}$ is not coercive on $V_0^{r_k}$, i.e. there exists $\delta e^k = \left(\delta x^k, a_1^k, b^k, a_2^k\right) \in V_0^{r_k}$ such that $\left\|\delta x^k\right\| + \left|a_1^k\right| + \left|b^k\right| + \left|a_2^k\right| = 1$ and $\mathcal{Q}_{r_k}[\delta e^k]^2 \le 0$. Up to a subsequence δe^k converges to some $\overline{\delta e} = \left(\overline{\delta x}, \overline{a}_1, \overline{b}, \overline{a}_2\right) \in V_{0,1}$ and such that $\left\|\overline{\delta x}\right\| + |\overline{a}_1| + \left|\overline{b}\right| + |\overline{a}_2| = 1$. Thus

$$
0 \ge \lim_{k\to\infty} \mathcal{Q}_{r_k}[\delta e^k]^2 = \mathcal{Q}[\overline{\delta e}]^2 > 0,
$$

a contradiction.

We have thus proved that (T^r, ξ^r, u^r), together with λ^r satisfies all the assumptions of Theorem 1 in [11], so that ξ^r is a state-locally optimal trajectory for problem ($\mathbf{P_r}$). If $\tau_2^r < \tau_1^r$ the proof follows the same lines.

Let us consider the case $\tau_1^r = \tau_2^r =: \tau^r$. In this case, as in the nominal problem ($\mathbf{P_0}$) we have to consider two different second order approximations and to prove that they are coercive on the respective half-space of linearised constraints.

The trajectory ξ^r is driven by the dynamics

$$
\phi_t^r := \begin{cases} h_1^r, & t \in [0, \tau^r], \\ h_2^r, & t \in (\tau^r, T^r]. \end{cases}
$$

Denoting again by S_t^r the flow at time t associated to ϕ^r, we consider the pullback vector fields

$$
\begin{aligned}
g_i^r(x) &:= (S^r_{\tau^r\,*})^{-1} h_1^r \circ S^r_{\tau^r}(x), \quad i = 1, 2, \\
j_\nu^r(x) &:= (S^r_{\tau^r\,*})^{-1} k_\nu^r \circ S^r_{\tau^r}(x), \quad \nu = 1, 2.
\end{aligned}
$$

Let α^r, β^r and γ^r be as before. Then the linearisation of the constraints is given by the half spaces

$$V_\nu^{+,r} := \Big\{ \delta e = (\delta x, a_1, b, a_2) \in T_{x_0^r} N_0^r \times \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \colon \\ \delta x + a_1 g_1^r(\widehat{x}_0) + b\, j_\nu^r(\widehat{x}_0) + a_2 g_2^r(\widehat{x}_0) \in T_{x_0^r} (S^r_{T^r})^{-1}(N_f^r) \Big\}, \quad \nu = 1,2$$

and the second order approximation is given by

$$\begin{aligned} \mathcal{Q}_\nu^r[\delta x, a_1, b, a_2] = {} & \mathrm{D}^2 \left(\alpha^r + \beta^r\right)(x_0^r)[\delta x]^2 + 2\delta x \cdot (a_1 g_1^r + b\, j_\nu^r + a_2 g_2^r) \cdot \beta^r\,(x_0^r) \\ & + \left(a_1 g_1^r + b\, j_\nu^r + a_2 g_2^r\right)^2 \cdot \beta^r\,(x_0^r) + a_1 b [g_1^r, j_\nu^r] \cdot \beta^r\,(x_0^r) \\ & + a_1 a_2 [g_1^r, g_2^r] \cdot \beta^r\,(x_0^r) + b\, a_2 [j_\nu^r, g_2^r] \cdot \beta^r\,(x_0^r). \end{aligned}$$

With the same contradiction argument used in the previous case it is easy to show that $\mathcal{Q}_\nu^r$ is coercive on $V_\nu^{+,r}$, $\nu = 1,2$. Thus (T^r, ξ^r, u^r), together with λ^r satisfies all the assumptions of Theorem 4.2 in [14], so that ξ^r is a state-locally optimal trajectory for problem ($\mathbf{P_r}$). ∎

7. Local uniqueness. We now prove the local uniqueness of the extremal λ^r in the cotangent bundle $T^*\mathbb{R}^n$, namely we prove Theorem 9. The proof is carried out by showing that there exists a tubular neighborhood $\mathcal{V}$ in $\mathbb{R} \times T^*\mathbb{R}^n$ of the graph of $\widehat{\lambda}$ such that, if $\widetilde{\lambda}\colon [0, \widetilde{T}] \to T^*\mathbb{R}^n$ is an extremal of ($\mathbf{P_r}$) whose graph is in $\mathcal{V}$, with $\widetilde{T}$ close to $\widehat{T}$, then the associated control $\widetilde{u} = (\widetilde{u}_1, \widetilde{u}_2)$ is bang-bang and each control component switches once and only once from the value -1 to the value 1. This implies that $\widetilde{\lambda}$ satisfies system (15) which, by the implicit function theorem, admits one and only one solution, i.e. $\widetilde{\lambda} = \lambda^r$.

By the regularity assumption at the switching time (Assumption 3) and by continuity, there exists $\overline{\delta} > 0$ such that

$$\langle \widehat{\lambda}(t)\,, [f_0, f_i]\,(\widehat{\xi}(t)) \rangle > \Big| \langle \widehat{\lambda}(t)\,, [f_1, f_2]\,(\widehat{\xi}(t)) \rangle \Big|, \quad \forall t \in [\widehat{\tau} - \overline{\delta}, \widehat{\tau} + \overline{\delta}], \quad i = 1,2.$$

For any $\delta \in (0, \overline{\delta}]$ and $i = 1,2$ define

$$\alpha_i^a(\delta) = \min \Big\{ u_i(t) F_i \circ \widehat{\lambda}(t) = -F_i \circ \widehat{\lambda}(t) \colon t \in [0, \widehat{\tau} - \delta] \Big\},$$
$$\alpha_i^p(\delta) = \min \Big\{ u_i(t) F_i \circ \widehat{\lambda}(t) = F_i \circ \widehat{\lambda}(t) \colon t \in [\widehat{\tau} + \delta, \widehat{T}] \Big\},$$

and let

$$m(\delta) := \min \Big\{ \boldsymbol{\sigma}\left(\overrightarrow{F_0}, \overrightarrow{F_1}\right)(\widehat{\lambda}(t)) - \Big| \boldsymbol{\sigma}\left(\overrightarrow{F_1}, \overrightarrow{F_2}\right)(\widehat{\lambda}(t)) \Big|,\ i = 1,2,\ t \in [\widehat{\tau} - \delta, \widehat{\tau} + \delta] \Big\}.$$

By continuity there exists $\mathcal{O}(\widehat{\ell}_0) \subset T^*\mathbb{R}^n$ such that

$$F_i \circ \widehat{\mathcal{F}}_t(\ell) < \frac{-\alpha_i^a(\delta)}{2} \qquad \forall (t,\ell) \in [0, \widehat{\tau} - \delta] \times \mathcal{O}(\widehat{\ell}_0),$$
$$F_i \circ \widehat{\mathcal{F}}_t(\ell) > \frac{\alpha_i^p(\delta)}{2} \qquad \forall (t,\ell) \in [\widehat{\tau} + \delta, \widehat{T}] \times \mathcal{O}(\widehat{\ell}_0),$$
$$\boldsymbol{\sigma}\left(\overrightarrow{F_0}, \overrightarrow{F_1}\right)(\widehat{\mathcal{F}}_t(\ell)) - \Big| \boldsymbol{\sigma}\left(\overrightarrow{F_1}, \overrightarrow{F_2}\right)(\widehat{\mathcal{F}}_t(\ell)) \Big| > \frac{m(\delta)}{2} \quad \forall (t,\ell) \in [\widehat{\tau} - \delta, \widehat{T} + \delta] \times \mathcal{O}(\widehat{\ell}_0),$$

and, again by continuity, there esists $\overline{R} > 0$ such that

$$F_i^r \circ \widehat{\mathcal{F}}_t(\ell) < \frac{-\alpha_i^a(\delta)}{4} \qquad \forall (t,\ell) \in [0, \widehat{\tau} - \delta] \times \mathcal{O}(\widehat{\ell}_0), \quad \forall r \colon |r| \le \overline{R}, \tag{24}$$

$$F_i^r \circ \widehat{\mathcal{F}}_t(\ell) > \frac{\alpha_i^p(\delta)}{4} \qquad \forall (t,\ell) \in [\widehat{\tau} + \delta, \widehat{T}] \times \mathcal{O}(\widehat{\ell}_0), \quad \forall r \colon |r| \le \overline{R}, \tag{25}$$

and

$$\boldsymbol{\sigma}\left(\overrightarrow{F_0^r}, \overrightarrow{F_1^r}\right)(\widehat{\mathcal{F}}_t(\ell)) - \left|\boldsymbol{\sigma}\left(\overrightarrow{F_1^r}, \overrightarrow{F_2^r}\right)(\widehat{\mathcal{F}}_t(\ell))\right| > \frac{m(\delta)}{4} \qquad \forall (t,\ell) \in [\widehat{\tau}-\delta, \widehat{T}+\delta] \times \mathcal{O}(\widehat{\ell}_0), \quad \forall r\colon |r| \leq \overline{R}. \tag{26}$$

Let $\widetilde{\lambda}\colon [0,\widetilde{T}] \to T^*\mathbb{R}^n$ be an extremal of ($\mathbf{P_r}$) whose graph is in the tubular set

$$\mathcal{V}_\delta = \left\{(t, \widehat{\mathcal{F}}_t(\ell))\colon t \in [0, \widehat{T}+\delta],\ \ell \in \mathcal{O}_\delta(\widehat{\ell}_0)\right\}$$

and such that $\left|\widetilde{T} - \widehat{T}\right| < \delta$.

By (24)-(25), for $i = 1, 2$,

$$F_i^r \circ \widetilde{\lambda}(t) < \frac{-\alpha_i^a}{4} \quad \forall t \in [0, \widehat{\tau}-\delta], \qquad F_i^r \circ \widetilde{\lambda}(t) > \frac{\alpha_i^p}{4} \quad \forall t \in [\widehat{\tau}+\delta, \widetilde{T}]$$

hence there exists $\widetilde{t}_i \in (\widehat{\tau}-\delta, \widehat{\tau}+\delta)$ such that $F_i^r \circ \widetilde{\lambda}(\widetilde{t}_i) = 0$. We now prove that $\widetilde{t}_i$ is the only time at which $F_i^r \circ \widetilde{\lambda}$ is zero. More precisely we show that $F_i^r \circ \widetilde{\lambda}(t)$ is strictly monotone increasing in the interval $[\widehat{\tau}-\delta, \widehat{\tau}+\delta]$. Let $\widehat{\tau}-\delta \leq s_1 < s_2 \leq \widehat{\tau}+\delta$:

$$\begin{aligned} F_i^r \circ \widetilde{\lambda}(s_2) - F_i^r \circ \widetilde{\lambda}(s_1) &= \int_{s_1}^{s_2} \frac{\mathrm{d}}{\mathrm{d}s} F_i^r \circ \widetilde{\lambda}(s)\,\mathrm{d}s = \\ &= \int_{s_1}^{s_2} \boldsymbol{\sigma}\left(\overrightarrow{F_0^r} + \widetilde{u}_1(s)\overrightarrow{F_1^r} + \widetilde{u}_2(s)\overrightarrow{F_2^r}, \overrightarrow{F_1^r}\right)(\widetilde{\lambda}(s))\,\mathrm{d}s = \\ &= \int_{s_1}^{s_2} \left(\boldsymbol{\sigma}\left(\overrightarrow{F_0^r}, \overrightarrow{F_1^r}\right) - \widetilde{u}_2(s)\boldsymbol{\sigma}\left(\overrightarrow{F_1^r}, \overrightarrow{F_2^r}\right)\right)(\widetilde{\lambda}(s))\,\mathrm{d}s > (s_2 - s_1)\frac{m(\delta)}{4}. \end{aligned}$$

Thus each component of the control $\widetilde{u}$ associated to $\widetilde{\xi} := \pi\widetilde{\lambda}$ switches once and only once from the value -1 to the value $+1$.

REFERENCES

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

[2] U. Felgenhauer, *Optimality and sensitivity for semilinear bang-bang type optimal control problems*, Int. J. Appl. Math. Comput. Sci., 14 (2004), pp. 447–454.

[3] 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.

[4] 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.

[5] J.-H. R. Kim and H. Maurer, *Sensitivity analysis of optimal control problems with bang-bang controls*, in 42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475), vol. 4, Dec 2003, pp. 3281–3286 vol.4, https://doi.org/10.1109/CDC.2003.1271649.

[6] K. Malanowski, *Two-norm approach in stability and sensitivity analisys of optimization and optimal control problems*, Advances in Math. Sciences and Applications, 2 (1993), pp. 397–443.

[7] K. Malanowski, *Regularity of solutions in stability analisys of optimization and optimal control problems*, Control and Cybernetics, 23 (1994), pp. 61–86.

[8] K. Malanowski, *Stability and sensitivity analysis for optimal control problems with control-state constraints*, in Dissertationes Mathematicae, vol. CCCXCIV, Institute of Mathematics, Polish Academy of Sciences, 2001.

[9] H. Maurer and N. P. Osmolovskii, *Second order optimality conditions for bangbang control problems*, Control and Cybernetics, 3 (2003), pp. 555–584.

[10] H. Maurer and N. P. Osmolovskii, *Second order sufficient conditions for time–optimal bang–bang control problems*, SIAM J. Control and Optimization, 42 (2004), pp. 2239–2263.

[11] 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).

[12] 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.

[13] 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.

[14] L. Poggiolini and M. Spadini, *Bang–bang trajectories with a double switching time in the minimum time problem*, ESAIM: COCV, 22 (2016), pp. 688–709, https://doi.org/10.1051/cocv/2015021.

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

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

[17] L. Poggiolini and G. Stefani, *Structural stability for bang-singular-bang extremals in the minimum time problem*, SIAM Journal on Control and Optimization, 51 (2013), pp. 3511–3531, https://doi.org/10.1137/120895421.

[18] A. V. Sarychev, *First- and second-order sufficient optimality conditions for bang-bang controls*, SIAM Journal on Control and Optimization, 35 (1997), pp. 315–340, https://doi.org/10.1137/S0363012993246191.