Control and Cybernetics

vol. **38** (2009) No. 4A

Revisiting the analysis of optimal control problems with several state constraints*†

by

J. Frédéric Bonnans[1] and Audrey Hermant[2]

[1] INRIA-Saclay and CMAP, Ecole Polytechnique
91128 Palaiseau, France

[2] Centre d'Analyse de Défense, DGA
16 bis avenue Prieur de la Côte d'Or, 94114 Arcueil, France

Abstract: This paper improves the results of and gives shorter proofs for the analysis of state constrained optimal control problems than presented by the authors in Bonnans and Hermant (2009b), concerning second order optimality conditions and the well-posedness of the shooting algorithm. The hypothesis for the second order necessary conditions is weaker, and the main results are obtained without reduction to the normal form used in that reference, and without analysis of high order regularity results for the control. In addition, we provide some numerical illustration. The essential tool is the use of the "alternative optimality system".

Keywords: optimal control, state constraints, shooting algorithm, second order optimality conditions.

1. Introduction

In these notes we give an account of some recent results on the analysis of state-constrained optimal control problems (see the classical review by Hartl, Sethi and Vickson, 1995) using in an essential way the approach called "alternative optimality system". This approach was introduced in an informal way by Bryson, Denham and Dreyfus (1963), Jacobson, Lele and Speyer (1971). The basic idea is, when a state constraint is active over an interval of time, to replace it with its time derivative of the smallest order such that the control appears. Adding in a proper way the junction conditions, it is then possible to state a shooting algorithm. Maurer (1979) gave a sound mathematical basis to this approach by defining in a precise way the alternative costate, and giving the alternative formulation of the optimality conditions, but restricting the analysis mainly to (important) special cases such as a single control and constraint. So, many important questions remained open for a long time, such as the junction analysis

*The authors thank the two referees for their useful comments.
†Submitted: January 2009; Accepted: September 2009.

in the general case, second order optimality conditions, and the well-posedness of numerical algorithms, among them the shooting algorithm.

There are a number of recent paper devoted to state-constrained problems, see Malanowski (2007) and references therein, and Malanowski and Maurer (2001). An application to the computation of constrained spline functions is discussed in Opfer and Oberle (1988). Let us highlight Malanowski and Maurer (1998) where a complete theory is obtained for the first order state constraints. High order constraints were studied in Bonnans and Hermant (2007-2009) and Hermant (2008,2009a,b). In this paper we revisit the problem studied in Bonnans and Hermant (2009b), by improving some results and simplifying some of the proofs. Part of the simplification is due to the fact that we are able to obtain the main results without using the reduction to the normal form (see Bonnans and Hermant, 2009b) and also without analysing high order regularity results for the control. Other important points are second-order necessary optimality conditions with weaker hypotheses than in Bonnans and Hermant (2007,2009b), the simplification of the presentation of junction condition for the alternative costate and linearized costate (and its relation to the jump of the standard costate) and a detailed analysis of the difference of costs of the associated quadratic subproblems. This last point is related to the fact that a junction point for a given constraint causes jumps in the multipliers of the other active constraints. This is why the case of several state constraints is essentially more difficult than the case of a single one, discussed in many papers. We provide also a more compact and self-contained presentation of the results. The analysis allows for state constraints of order higher than two. However, the analysis of the shooting algorithm excludes boundary arcs with such constraints, as expected (see Remark 3).

The paper is organized as follows. Section 2 discusses some consequences of Pontryagin's principle and provides some results on the continuity of the control. For the latter we rely on the fact that the control minimizes the Hamiltonian. In this section we establish high order regularity on arcs with constant active set of constraints. The shooting algorithm (and related questions on second-order optimality conditions) is presented in Section 3, using reduction of isolated contact points and the alternative optimality system. Second-order optimality conditions are presented in Section 4. The well-posedness of the shooting algorithm is established in Section 5, assuming in particular that no boundary arc has state constraint of order higher than two. In Section 6, we present a numerical application of the shooting algorithm on two academic problems involving three state constraints of order 1 and 2, respectively.

Notation. The set of integers from i to j is denoted $\{i{:}j\}$. The operator ":=" means a definition of the l.h.s. The cardinal of a set I is denoted by $|I|$. The open (respectively closed) Euclidean ball of center x and radius R is denoted by $B(x,R)$ (respectively $\bar{B}(x,R)$).

For the value of functions of time only, as say y, we denote y_t the value at time t and the one of its ith component is denoted $y_{i,t}$. If u and y depend on time, and h is a function of (t, u, y), we denote by $h_{i,y}(t, u_t, y_t)$ the partial derivative w.r.t y of its ith component. When necessary for clarity we denote partial derivatives say like $D_t h(t, u_t, y_t)$.

The Sobolev space $W^{m,s}(0, T, I\!R^n)$, where m is a positive integer and $s \in [1, \infty]$, is the set of functions in $L^s(0, T, I\!R^n)$, whose weak time derivatives also belong to $L^s(0, T, I\!R^n)$. Elements of $W^{1,s}(0, T, I\!R^n)$ are Hölder (respectively Lipschitz) functions of time for $s \in [1, \infty)$ (respectively $s = \infty$).

The space of functions of $[0, T] \to I\!R^n$ with bounded variations is denoted $BV(0, T, I\!R^n)$. The measure associated with $\eta \in BV(0, T)$ is denoted $\mathrm{d}\eta$. Elements of $BV(0, T, I\!R^n)$ have for all time $t \in [0, T]$, left and right limits (right limit for $t = 0$, left limit for $t = T$). If a function of time, say η, has left or right limits at time t, the latter are denoted η_t^- and η_t^+, respectively. The convex combinations of the latter are denoted $\eta_t^\sigma := \sigma\eta_t^+ + (1-\sigma)\eta_t^-$, for $\sigma \in [0, 1]$, and the *jump* at time t of η is $[\eta_t] := \eta_t^+ - \eta_t^-$.

By $I\!R^{n*}$ we denote the dual of $I\!R^n$, identified with the space of n dimensional horizontal vectors.

2. Pontryagin's principle

2.1. Statement

In this section we study optimal control problems with state constraints, of the following type:

$$\begin{cases} \text{Min} \displaystyle\int_0^T \ell(u_t, y_t)\mathrm{d}t + \phi(y_T); \\ \dot{y}_t = f(u_t, y_t); \quad t \in (0, T); \quad y_0 = y^0; \\ g(y_t) \leq 0; \quad t \in [0, T], \end{cases} \tag{1}$$

with $\ell : I\!R^m \times I\!R^n \to I\!R$, $f : I\!R^m \times I\!R^n \to I\!R^n$, $g : I\!R^n \to I\!R^r$, $r \geq 1$, $y^0 \in I\!R^n$ given. All data f, g, ℓ, ϕ are of class C^∞, and f is Lipschitz. Denote the control and state spaces by

$$\mathcal{U} := L^\infty(0, T, I\!R^m); \quad \mathcal{Y} := W^{1,\infty}(0, T, I\!R^n). \tag{2}$$

For a given $u \in \mathcal{U}$, the state equation (i.e., the differential equation in the second row of (1)) has a unique solution $y(u) \in \mathcal{Y}$. The *generalized Hamiltonian function* $H : I\!R \times I\!R^m \times I\!R^n \times I\!R^{n*} \to I\!R$ is

$$H(\alpha, u, y, p) := \alpha\,\ell(u, y) + pf(u, y). \tag{3}$$

DEFINITION 1 *We say that* $(\bar{u}, \bar{y}) \in \mathcal{U} \times \mathcal{Y}$ *is a* generalized Pontryagin extremal *if there exists* $\bar{\alpha} \geq 0$ *and* $\bar{\eta} \in BV(0, T, I\!R^r)$, *with* $(\bar{\alpha}, \mathrm{d}\bar{\eta}) \neq 0$, *and a costate*

$\bar{p} \in BV(0,T,\mathbb{R}^{n*})$, *such that a.e.* $t \in (0,T)$*:*

$$\dot{\bar{y}}_t = f(\bar{u}_t, \bar{y}_t) \quad \text{a.e. in } [0,T], \tag{4}$$

$$-\mathrm{d}\bar{p}_t = H_y(\bar{\alpha}, \bar{u}_t, \bar{y}_t, \bar{p}_t)\mathrm{d}t + \sum_{i=1}^{r} g_i'(\bar{y}_t)\mathrm{d}\bar{\eta}_{i,t} \quad \text{on } [0,T], \tag{5}$$

$$\bar{u}_t \in \operatorname*{argmin}_w H(\bar{\alpha}, w, \bar{y}_t, \bar{p}_t), \quad \text{a.a. on } \;]0,T[, \tag{6}$$

and in addition

$$g_i(\bar{y}_t) \le 0;\ \mathrm{d}\bar{\eta}_{i,t} \ge 0;\ t \in [0,T];\ \int_0^T g_i(\bar{y}_t)\mathrm{d}\bar{\eta}_{i,t} = 0,\ i \in \{1:r\}, \tag{7}$$

$$\bar{y}(0) = y^0; \quad \bar{p}_T = \bar{\alpha}\phi'(\bar{y}_T). \tag{8}$$

We can rewrite the costate equation (5), with final condition in (8), in integral form:

$$\bar{p}_t = \bar{\alpha}\phi'(\bar{y}_t) + \int_t^T H_y(\bar{\alpha}, \bar{u}_s, \bar{y}_s, \bar{p}_s)\mathrm{d}s + \sum_{i=1}^{r} \int_t^T g_i'(\bar{y}_s)\mathrm{d}\bar{\eta}_{i,s}, \text{ for all } t \in [0,T]. \tag{9}$$

The following is well-known, see Section 5.2 in Ioffe (1979).

THEOREM 1 *Let* $\bar{u} \in \mathcal{U}$ *be an optimal control and* $\bar{y}$ *be the associated state for problem* (1)*. Then* $(\bar{u}, \bar{y})$ *is a generalized Pontryagin extremal.*

2.2. Continuity of the control

The *total derivative* of a function of the state, say $g(y)$, is by the definition the function $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^r$ whose expression is

$$g^{(1)}(u,y) := g'(y)f(u,y). \tag{10}$$

Along a *trajectory* (u,y) (i.e., a solution of the state equation), $g^{(1)}(u_t, y_t)$ is equal to $\frac{\mathrm{d}}{\mathrm{d}t} g(y_t)$. In a similar way we can define higher order derivatives. These formal expressions are the sum of all partial derivatives multiplied by the corresponding derivative of the variable, understood as formal variables (and not true time derivatives) except for y whose derivative is replaced by $f(u,y)$. Denoting the partial derivatives by subscripts, we obtain for instance the mapping $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^r$

$$g^{(2)}(u,\dot{u},y) = g_u^{(1)}(u,y)\dot{u} + g_y^{(1)}(u,y)f(u,y). \tag{11}$$

As long as the total derivatives do no depend on u (respectively the derivatives of u), we may denote them as $g^{(i)}(y)$ (respectively $g^{(i)}(u,y)$).

DEFINITION 2 (i) *For $i \in \{1:r\}$, the order of the state constraint $g_i(y)$ is the smallest positive integer q_i such that*

$$g_{i,u}^{(k)}(u,y) = 0, \quad \text{for all } (u,y) \in I\!R^m \times I\!R^n \text{ and } 0 \leq k < q_i. \tag{12}$$

Then $g_i^{(k)}(u,y)$ does not depend on the derivatives of u for $k \leq q_i$.
(ii) *We say that the state constraint i is* regular along a trajectory $(\bar{u},\bar{y}) \in \mathcal{U} \times \mathcal{Y}$, *such that $\bar{u}$ is continuous, if $g_{i,u}^{(q_i)}(\bar{u}_t,\bar{y}_t) \neq 0$, for all $t \in [0,T]$.*

For a state constraint g_i of order q_i, and $k \in \{1:(q_i-1)\}$, we may write $g_i^{(k)}(y)$ instead of $g_i^{(k)}(u,y)$, and we have

$$g_i^{(k+1)}(u,y) = g_{i,y}^{(k)}(y)f(u,y); \quad g_{i,u}^{(k+1)}(u,y) = g_{i,y}^{(k)}(y)f_u(u,y). \tag{13}$$

For instance, if $q_i \geq 2$, then skipping arguments of g and f:

$$\begin{cases} g_i^{(2)}(u,y) & = g_{i,y}^{(1)}f = g_i''(f,f) + g_i'f_yf; \\ g_{i,u}^{(2)}(u,y) & = g_{i,y}^{(1)}f_u = (g_i''f + g_i'f_y)\,f_u. \end{cases} \tag{14}$$

The set $I(t)$ of active state constraints at time $t \in [0,T]$ is defined by

$$I(t) := \{i \in \{1:r\};\ g_i(\bar{y}(t)) = 0\}. \tag{15}$$

Define the set of state constraints of order κ, and those active at time t along the trajectory $(\bar{u},\bar{y})$ by:

$$I_\kappa := \{1 \leq i \leq r;\ q_i = \kappa\}; \quad I_\kappa(t) := \{i \in I_\kappa;\ g_i(\bar{y}_t) = 0\}. \tag{16}$$

For control variables with left and right limits at every time, a strong Legendre-Clebsch type condition, along the direction of jump of the control, is as follows:

$$\begin{cases} \text{For some } \alpha > 0\text{:}\ \alpha|[\bar{u}_t]|^2 \leq H_{uu}(\bar{u}_t^\sigma,\bar{y}_t,\bar{p}_t^\sigma)([\bar{u}_t],[\bar{u}_t]), \\ \qquad\qquad\qquad\qquad \text{for all } \sigma \in [0,1],\ t \in [0,T]. \end{cases} \tag{17}$$

If the control is continuous, the hypothesis of linear independence w.r.t. the control of *first-order* state constraints is as follows:

$$\{g_{i,u}^{(1)}(\bar{u}_t,\bar{y}_t)\}_{i \in I_1(t)} \ \text{ is of rank } |I_1(t)|, \text{ for all } t \in [0,T]. \tag{18}$$

We recall that, being of bounded variation, $\bar{p}$ has left and right limits.

THEOREM 2 *Let $(\bar{u},\bar{y})$ be a Pontryagin extremal for problem (P).*
(i) *Let $R > \|\bar{u}\|_\infty$. If $H(\cdot,\bar{y}_t,\bar{p}_t^{\pm})$ has, for all $t \in [0,T]$, a unique minimum over $\bar{B}(0,R)$, and if* (17) *holds, then $\bar{u}$ is continuous.*
(ii) *If $\bar{u}$ is continuous and* (18) *hold, then the components of $\bar{\eta}$ associated with first order state constraint are continuous, and $H_{uu}(\bar{u}_t,\bar{y}_t,\bar{p}_{t^\pm})$ is a continuous function of time.*

Proof. (i) By assumption, $H(u, \bar{y}_t, \bar{p}_t^{\pm})$ has a unique point of minimum at time t at the point $\tilde{u}_t \in \bar{B}(0, R)$. Then $\tilde{u} = \bar{u}$ a.e. and we may take $\bar{u} = \tilde{u}$. When, say, $t \uparrow \tau$, $\tilde{u}_t$ has at least one cluster point a. It is easy to check that $H(\cdot, \bar{y}_t, \bar{p}_t^-)$ attains its minimum on $\bar{B}(0, R)$ at the point a, implying the existence of left and right limits for $\bar{u}$ at time τ. By the costate equation (5), $\bar{p}$ has at most countably many jumps, of type

$$[\bar{p}_t] = \bar{p}_t^+ - \bar{p}_t^- = -\sum_{i=1}^{r} \nu_i g_i'(\bar{y}_t), \qquad \text{with } \nu_i := [\bar{\eta}_{i,t}] \geq 0. \tag{19}$$

We have that

$$\begin{aligned} 0 &= H_u(\bar{u}_t^+, \bar{y}_t, \bar{p}_t^+) - H_u(\bar{u}_t^-, \bar{y}_t, \bar{p}_t^-) \\ &= \int_0^1 (H_{uu}(\bar{u}_t^\sigma, \bar{y}_t, \bar{p}_t^\sigma)[\bar{u}_t] + [\bar{p}_t] f_u(\bar{u}_t^\sigma, \bar{y}_t))\, \mathrm{d}\sigma. \end{aligned} \tag{20}$$

Using (19) and observing that $g_i' f_u = g_{i,u}^{(1)} = 0$ if $q_i > 1$, we obtain that

$$\int_0^1 H_{uu}(\bar{u}_t^\sigma, \bar{y}_t, \bar{p}_t^\sigma)[\bar{u}_t]\mathrm{d}\sigma \; = \; \sum_{i \in I_1} \int_0^1 \nu_i g_{i,u}^{(1)}(\bar{u}_t^\sigma, \bar{y}_t)\mathrm{d}\sigma. \tag{21}$$

Taking the scalar product of both sides of (21) by $[\bar{u}_t]$, we get using hypothesis (17) that

$$\alpha |[\bar{u}_t]|^2 \;\; \leq \;\; \sum_{i \in I_1} \int_0^1 \nu_i g_{i,u}^{(1)}(\bar{u}^\sigma, \bar{y}_t)[\bar{u}_t]\mathrm{d}\sigma = \sum_{i \in I_1} \nu_i \left[g_i^{(1)}(\bar{u}_t, \bar{y}_t) \right]. \tag{22}$$

If $\nu_i > 0$, then $g_i(\bar{y}_t) = 0$, and hence $[g_i^{(1)}(\bar{u}_t, \bar{y}_t)] \leq 0$, since t is a local maximum of $g_i(\bar{y}_t)$. Therefore, the right-hand side in (22) is nonpositive, implying $[\bar{u}_t] = 0$. Point (i) follows.
(ii) Since $[\bar{u}_t] = 0$, the right-hand side of (21) (with $\bar{u}_t^\sigma = \bar{u}_t$) is zero. We conclude with (18) that the components of $\bar{\eta}$ associated with first order state constraint are continuous. In addition, as $g_i'(y) f_u(y, u)$ is identically zero whenever $q_i > 1$, it follows that $H_{uu}(\bar{u}_t, \bar{y}_t, \bar{p}_{t^\pm})$ is a continuous function of time. ∎

Remark 1 We note that Theorem 2 improves the related statements in Bonnans and Hermant (2009b) and Maurer (1979) by using a weak hypothesis (17). In the sequel, if $(\bar{u}, \bar{y})$ is a Pontryagin extremal and $\bar{u}$ is continuous, we will say that $(\bar{u}, \bar{y})$ is a continuous Pontryagin extremal (this, of course, does not imply the continuity of the multiplier $\bar{\eta}$ or of the costate $\bar{p}$).

2.3. Smoothness on each arc

If $0 \leq a < b \leq T$, we say that (a, b) is an *arc* of the Pontryagin extremal $(\bar{u}, \bar{y}, \bar{p}, \bar{\eta})$ if (a, b) is a maximal interval of $[0, T]$, over which the set of active constraints $I(t)$ is constant. Let q_i be the order of the ith state constraint, set $q := (q_1, \dots, q_r)$, and define $G(u,y) : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^r$ by

$$G_i(u,y) := g_i^{(q_i)}(u,y), \quad i \in \{1:r\}. \tag{23}$$

The hypothesis of *linear independence of gradients of active constraints w.r.t. the control* is

$$\{G_{i,u}(\bar{u}_t, \bar{y}_t)\}_{i\in I(t)} \text{ is of full rank, for all } t \in [0,T]. \tag{24}$$

We also need a *strong Legendre-Clebsch condition along the kernel of active constraints*:

$$\left\{ \begin{array}{l} \text{For some } \alpha > 0, \text{ for all } t \in [0,T],\, v \in \mathbb{R}^m : \\ \alpha |v|^2 \le H_{uu}(\bar{u}_t, \bar{y}_t, \bar{p}_t)(v,v), \;\text{ if } G_{i,u}(\bar{u}_t, \bar{y}_t)v = 0, \;\text{ for all }\; i \in I(t). \end{array} \right. \tag{25}$$

We introduce, like in Maurer (1979), the *alternative multipliers* $\eta^{i,k}$, where $i \in \{1:r\}$ and $k \in \{1:q_i\}$, and $\bar{\eta}^q$:

$$\eta_t^{i,1} := -\bar{\eta}_{i,t}; \quad \eta_t^{i,k} := \int_t^T \eta_s^{i,k-1} \mathrm{d}s; \quad \bar{\eta}_i^q := \eta^{i,q_i}. \tag{26}$$

The *alternative costate* (of order q) is defined as

$$\bar{p}_t^q := \bar{p}_t - \sum_{i=1}^r \sum_{j=1}^{q_i} \eta_t^{i,j} g_{i,y}^{(j-1)}(\bar{y}_t) \tag{27}$$

and the corresponding *alternative Hamiltonian* $H^q : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n*} \times \mathbb{R}^{r*}$ is

$$H^q(u,y,p^q,\eta^q) := \ell(u,y) + p^q f(u,y) + \eta^q G(u,y). \tag{28}$$

This derivation of the alternative multipliers and the following proposition are due to Maurer (1979):

$$-\dot{\bar{p}}_t^q = H_y^q(\bar{u}_t, \bar{y}_t, \bar{p}^q, \bar{\eta}^q), \quad t \in (0,T). \tag{29}$$

$$H(u, \bar{y}_t, \bar{p}_t) = H^q(u, \bar{y}_t, \bar{p}^q, \bar{\eta}^q), \quad \text{for all } u \in \mathbb{R}^m. \tag{30}$$

Proposition 1 *Let $(\bar{u}, \bar{y}, \bar{p}, \bar{\eta})$ be a continuous Pontryagin extremal satisfying* (24)-(25). *Then $\bar{u}$ and $\bar{\eta}^q$ are of class C^∞ over any arc (and therefore so are $\bar{p}_t$ and $\bar{\eta}_t$).*

Proof. Let us denote by I^* the set of active constraints over an arc (a,b), and let $G_{I^*}(u,y) := (G_i(u,y))_{i\in I^*}$. The two algebraic equations

$$\left\{ \begin{array}{ll} H_u^q(\bar{u}_t, \bar{y}_t, \bar{p}_t^q, \bar{\eta}_t^q) & = 0, \\ G_{I^*}(\bar{u}_t, \bar{y}_t) & = 0, \end{array} \right. \tag{31}$$

hold over (a, b), and their Jacobian w.r.t. the algebraic variables $(u, \bar{\eta}^q)$ is

$$\mathrm{Jac}_{I(t)} := \begin{pmatrix} H_{uu}(\bar{u}_t, \bar{y}_t, \bar{p}_t) & (G_{I^*,u}(\bar{u}_t, \bar{y}_t))^\top \\ G_{I^*,u}(\bar{u}_t, \bar{y}_t) & 0 \end{pmatrix}, \tag{32}$$

which, by (24)-(25), is invertible. By hypothesis, $\bar{u}$ is continuous, and since

$$0 = H_u^q(\bar{u}_t, \bar{y}_t, \bar{p}_t^q, \bar{\eta}_t^q) = H_u(\bar{u}_t, \bar{y}_t, \bar{p}_t^q) + \bar{\eta}_t^q G_{I^*,u}(\bar{u}_t, \bar{y}_t), \tag{33}$$

(24) implies that $\bar{\eta}^q$ is a continuous function of $(\bar{u}_t, \bar{y}_t, \bar{p}_t^q)$. Since the algebraic variables $\bar{u}$ and $\bar{\eta}^q$ are continuous, the implicit function theorem implies that they are (locally in time) functions of class C^∞ of $(\bar{y}_t, \bar{p}_t^q)$, so that $(\bar{y}, \bar{p}^q)$ is on (a, b) solution of a differential equation with C^∞ data. The conclusion follows. ∎

3. The shooting algorithm

3.1. Formulation

We say that $\tau \in [0, T]$ is a *junction point* if $I(t)$ is not constant for t close to τ. The set of junction points is closed, and therefore, has a finite cardinal iff each junction point is an *isolated junction point*. We note that τ is an isolated junction point (i.e., is not a limit point of the set of junction points) iff there are two arcs of the form (a, τ) and (τ, b). All junction points are isolated iff there are finitely many junction points, and iff there are finitely many arcs.

We say that $\tau \in [0, T]$ is a *contact point* for constraint $i \in \{1 : r\}$ if $i \in I(\tau)$. If, in addition, $i \notin I(t)$ for $t \neq \tau$, close to τ, then we say that τ is an *isolated contact point* or a *touch point*. If the measure $\bar{\eta}$ has a nonzero (zero) jump at the junction time τ, we say that τ is an *essential* (*non essential*) junction point. The alternative optimality system allows also for proving the following important result.

Lemma 1 *Let* $(\bar{u}, \bar{y}, \bar{p}, \bar{\eta})$ *be a continuous Pontryagin extremal satisfying* (24)-(25). *Let* $\tau \in (0, T)$ *be an isolated touch point associated with just a first order state constraint. Then* τ *is a non essential touch point.*

Proof. Let i_0 be the index of the first order state constraint. Since (24) implies (18), we already know by theorem 2(ii) that $[\bar{\eta}_{i_0,\tau}] = 0$, and hence, $[\bar{\eta}^q_{i_0,\tau}] = 0$. Remember that $(\bar{y}, \bar{p}^q)$ is the solution of the state equation and (29). For t close to, and different from τ, the index set $I(t)$ is a constant I^*. By the arguments of the proof of proposition 1, it follows that $(\bar{u}_t, \bar{\eta}_t^q)$ is a smooth function of $(\bar{y}, \bar{p}^q)$. Therefore, $(\bar{u}, \bar{y}, \bar{p}^q, \bar{\eta}^q)$ is of class C^∞ for t close to τ, implying $[\bar{\eta}_\tau] = 0$, as was to be proved. ∎

For $y \in I\!R^n$, consider the vector

$$\Gamma(y) := \begin{pmatrix} g_1(y) & \cdots & g_1^{(q_1-1)}(y) & \cdots & g_r(y) & \cdots & g_r^{(q_r-1)}(y) \end{pmatrix}^\top. \tag{34}$$

PROPOSITION 2 *Let* $(\bar{u}, \bar{y})$ *be a trajectory satisfying* (24). *Then, for all* $t \in [0, T]$, *the restriction of* $\Gamma'(\bar{y}_t)$ *to the active constraints at time* t *has full rank, equal to* $\sum_{i \in I(t)} q_i$.

The proof is based on the following lemma, due to Maurer (1979). Set $q_{max} := \max_i q_i$.

LEMMA 2 *Let the trajectory* (u, y) *be such that* u *is of class* C^∞ *over* $(a, b) \subset [0, T]$. *For* $k \in \{1 : (q_{max} - 1)\}$, *define the mappings* $A_k : (a, b) \to \mathbb{R}^{n \times m}$ *by:*

$$\begin{cases} A_0(t) := f_u(u_t, y_t) \\ A_k(t) := f_y(u_t, y_t) A_{k-1}(t) - \dot{A}_{k-1}(t), \qquad 1 \le k \le q_{max} - 1. \end{cases} \tag{35}$$

Then, for all $t \in (a, b)$ *and* $i = 1, \ldots, r$, *we have:*

$$\begin{cases} g_{i,y}^{(j)}(y_t) A_k(t) = 0 \qquad \text{for } k, j \ge 0, \quad k + j \le q_i - 2, \\ g_{i,y}^{(j)}(y_t) A_{q_i - j - 1}(t) = g_{i,u}^{(q_i)}(u_t, y_t) \qquad \text{for } 0 \le j \le q_i - 1. \end{cases} \tag{36}$$

Proof. We first show that for all $j = 0, \ldots, q_i - 1$, the following assertion

$$g_{i,y}^{(j)}(y_t) A_k(t) = 0 \qquad \forall\, t \in (a, b) \tag{37}$$

implies that

$$g_{i,y}^{(j+1)}(u_t, y_t) A_k(t) = g_{i,y}^{(j)}(y_t) A_{k+1}(t) \qquad \forall\, t \in (a, b). \tag{38}$$

Indeed, since for $j \le q_i$

$$g_{i,y}^{(j)}(u, y) \;=\; g_{i,yy}^{(j-1)}(y) f(u, y) + g_{i,y}^{(j-1)}(y) f_y(u, y), \tag{39}$$

by derivation of (37) w.r.t. time, we get

$$\begin{aligned} 0 &= g_{i,yy}^{(j)}(y_t) f(u_t, y_t) A_k(t) + g_{i,y}^{(j)}(y_t) \dot{A}_k(t) \\ &= g_{i,yy}^{(j)}(y_t) f(u_t, y_t) A_k(t) + g_{i,y}^{(j)}(f_y(u_t, y_t) A_k(t) - A_{k+1}(t)) \\ &= g_{i,y}^{(j+1)}(u_t, y_t) A_k(t) - g_{i,y}^{(j)}(y_t) A_{k+1}(t). \end{aligned}$$

This gives (38). Also, $g_{i,u}^{(j)}(u_t, y_t) = g_{i,y}^{(j-1)}(y_t) f_u(u_t, y_t) = g_{i,y}^{(j-1)}(y_t) A_0(t)$ for $j = 1, \ldots, q_i$. Since $g_{i,u}^{(j)} = 0$ for $j \le q_i - 1$, it follows that $g_{i,y}^{(j)}(y_t) A_0(t) = 0$, for $j = 0, \ldots, q_i - 2$. By (38), we deduce that $g_{i,y}^{(j)}(y_t) A_1(t) = 0$ for $j = 0, \ldots, q_i - 3$. By induction, this proves the first equation in (36). Since $g_{i,y}^{(q_i-2)}(y_t) A_0(t) = 0 = g_{i,y}^{(q_i-3)}(y_t) A_1(t) = \cdots = g_{i,y}(y_t) A_{q_i-2}(t)$, by (38) we obtain $g_{i,u}^{(q_i)}(y_t) = g_{i,y}^{(q_i-1)}(y_t) A_0(t) = g_{i,y}^{(q_i-2)}(y_t) A_1(t) = \cdots = g_{i,y}(y_t) A_{q_i-1}(t)$, which proves the second equation in (36). ∎

Proof (Proof of Proposition 2). Given $M(\lambda) := \sum_{i\in I(t)} \sum_{j=1}^{q_i-1} \lambda_{i,j} g_{i,y}^{(j)}(\bar{y}_t)$ such that $M(\lambda) = 0$, we have to prove that $\lambda = 0$. By the definition of the state order, $M(\lambda) f_u(\bar{u}_t, \bar{y}_t) = \sum_{i\in I(t)} \lambda_{i,q_i-1} g_{i,u}^{(q_i)}(\bar{y}_t)$, which, in view of (24), implies $\lambda_{i,q_i-1} = 0$, for $i \in I(t)$. So, for $k = 1$ the following relation holds:

$$\lambda_{i,j} = 0, \quad \text{when } \max(0, q_i - k) \le j \le q_i - 1, \text{ for } i \in I(t). \tag{40}$$

Let it hold for some $k \in \{1 : q_{max}\}$. Set $I_k(t) := \{i \in I(t);\ q_i > k\}$. Then $M(\lambda) = \sum_{i\in I_k(t)} \sum_{j=1}^{q_i-1-k} \lambda_{i,j} g_{i,y}^{(j)}(\bar{y}_t)$. Let A_{k-1} be defined by (35) with $(u, y) = (\bar{u}, \bar{y})$. In view of (36), we have

$$M(\lambda) A_{k-1} = \sum_{i\in I_k(t)} \lambda_{i,q_i-1-k} g_{i,u}^{(q_i)}(\bar{y}_t) = 0, \tag{41}$$

implying in view of (24) that $\lambda_{i,q_i-1-k} = 0$ when $q_i > k$. Therefore, the result follows by induction. ■

When setting the alternative formulation, we observe that we may add an arbitrary constant to each component of η. Similarly, when defining the alternative multipliers we may add arbitrary integration constants. This will result in a difference of an arbitrary polynomial of degree $q_i - 1$ for a state constraint of order q_i. When $[0, T]$ is the union of finitely many arcs, we may choose different polynomials on each arc. By Proposition 1, $(\bar{u}_t, \bar{y}_t, \bar{p}_t^q, \bar{\eta}_t^q)$ is of class C^∞ over each arc. In the context of shooting algorithms, it is convenient to choose these constants so that the multipliers associated with nonactive constraints are equal to zero, i.e.

$$\eta_t^{i,j} = 0, \quad \text{if } i \notin I(t), \quad j \in \{1 : q_i\}. \tag{42}$$

Let us set $\nu_\tau^i := [\bar{\eta}_{i,\tau}] \ge 0$. By (19) and (27), the jumps of the original and alternative costate are related by

$$\left\{ \begin{array}{rcl} [\bar{p}_\tau^q] & = & [\bar{p}_\tau] - \displaystyle\sum_{i\in I(\tau)} \sum_{j=1}^{q_i} [\eta_\tau^{i,j}] g_{i,y}^{(j-1)}(\bar{y}_\tau) \\ & = & - \displaystyle\sum_{i\in I(\tau)} \left((\nu_\tau^i + [\eta_\tau^{i,1}]) g_i'(\bar{y}_\tau) + \sum_{j=2}^{q_i} [\eta_\tau^{i,j}] g_{i,y}^{(j-1)}(\bar{y}_\tau) \right). \end{array} \right. \tag{43}$$

If (24) holds, this uniquely determines coefficients $\nu_\tau^{i,j}$ such that

$$[\bar{p}^q(\tau)] = - \sum_{i\in I(\tau)} \sum_{j=1}^{q_i} \nu_\tau^{i,j} g_{i,y}^{(j-1)}(\bar{y}_\tau), \tag{44}$$

with

$$\left\{ \begin{array}{rcll} \nu_\tau^{i,1} & = & \nu_\tau^i + [\eta_\tau^{i,1}], & i \in I(\tau), \\ \nu_\tau^{i,j} & = & [\eta_\tau^{i,j}], & i \in I(\tau),\ j \in [2 : q_i]. \end{array} \right. \tag{45}$$

In the sequel we assume that there are finitely many arcs. Let N_b^i, N_{to}^i denote, respectively, the number of boundary arcs and touch points of the state constraint of index $i \in \{1:r\}$. Denote by $\mathcal{I}_b^i := \cup_{k=1}^{N_b^i}[\tau_{en}^{i,k}, \tau_{ex}^{i,k}]$ the closure of the union of boundary arcs of each constraint, for $i \in \{1:r\}$, and

$$\mathcal{T}_{en}^i := \{\tau_{en}^{i,1} < \cdots < \tau_{en}^{i,N_b^i}\}, \;\; \mathcal{T}_{ex}^i := \{\tau_{ex}^{i,1} < \cdots < \tau_{ex}^{i,N_b^i}\}, \tag{46}$$

and similarly denote the sets of touch and junction points of constraint i by

$$\mathcal{T}_{to}^i := \{\tau_{to}^{i,1} < \cdots < \tau_{to}^{i,N_{to}^i}\}; \quad \mathcal{T}^i := \mathcal{T}_{en}^i \cup \mathcal{T}_{ex}^i \cup \mathcal{T}_{to}^i. \tag{47}$$

The set of junction points is $\mathcal{T} := \cup_{i=1}^r \mathcal{T}^i$. The alternative formulation includes the following relations on each arc:

$$\begin{aligned}
\dot{\bar{y}}_t &= f(\bar{u}_t, \bar{y}_t) \quad \text{on } [0,T] \quad ; \quad \bar{y}_0 = y^0, & (48)\\
-\dot{\bar{p}}_t^q &= H_y^q(\bar{u}_t, \bar{y}_t, \bar{p}_t^q, \bar{\eta}_t^q) \;\; \text{on } [0,T] \setminus \mathcal{T}, & (49)\\
0 &= H_u^q(\bar{u}_t, \bar{y}_t, \bar{p}_t^q, \bar{\eta}_t^q) \;\; \text{on } [0,T] \setminus \mathcal{T}, & (50)\\
G_i(\bar{u}_t, \bar{y}_t) &= 0 \qquad \text{on } \mathcal{I}_b^i, \quad i \in \{1:r\}, & (51)\\
\bar{\eta}_{i,t}^{q_i} &= 0 \qquad \text{on } [0,T] \setminus \mathcal{I}_b^i, \quad i \in \{1:r\}. & (52)
\end{aligned}$$

Assuming, for simplicity, that the state constraints are not active at time T, we have the final condition for the costate

$$\bar{p}_T^q = \phi'(\bar{y}_T). \tag{53}$$

In view of the definition of orders of state constraints, and since a constraint reaches a maximum at a touch point, we have the following junction conditions:

$$\begin{aligned}
g_i^{(j)}(\bar{y}_\tau) &= 0 \qquad \text{if } \tau \in \mathcal{T}_{en}^i, \quad j \in \{0:(q_i-1)\}, & (54)\\
g_i(\bar{y}_\tau) &= 0 \qquad \text{if } \tau \in \mathcal{T}_{to}^i. & (55)
\end{aligned}$$

It remains to state the junction conditions for the costate. We will assume that each junction time is a junction time for a single constraint. As done in the literature (see Maurer, 1979), we fix the integration constants $[\eta^{i,j}]$ such that $\bar{p}^q$ is continuous at exit points, and at an entry (respectively touch) point has a jump involving only the derivatives (respectively the first derivative) of the entering state constraint, i.e.:

$$\begin{aligned}
[\bar{p}_\tau^q] &= 0, \quad \text{for all } \tau \in \mathcal{T}_{ex}^i, \quad i \in \{1:r\}, & (56)\\
[\bar{p}_\tau^q] &= -\sum_{j=1}^{q_i} \nu_\tau^{i,j} g_{i,y}^{(j-1)}(\bar{y}_\tau), \quad \text{for all } \tau \in \mathcal{T}_{en}^i, \quad i \in \{1:r\}, & (57)\\
[\bar{p}_\tau^q] &= -\nu_\tau^{i,1} g_i'(\bar{y}_\tau), \quad \text{for all } \tau \in \mathcal{T}_{to}^i, \quad i \in \{1:r\}. & (58)
\end{aligned}$$

Note that, for a touch point $\tau \in \mathcal{T}_{to}^i$, we have that $\bar{\eta}_{i,t}^q = 0$ for $t \neq \tau$ close to τ.

Relations (48)-(58) can be interpreted as the optimality conditions for the problem of minimizing the cost function under conditions (48), (51) and (54)-(55), with fixed junction times. Remember that, under standard assumptions, by Lemma 1, touch points associated with first order state constraints are non essential; therefore we can ignore them in the formulation of the shooting algorithm. The previous discussion suggests adding the equalities allowing to find these junction times:

$$
\begin{aligned}
G_i(\bar{u}_\tau^-, \bar{y}_\tau) &= 0, \quad \text{if } \tau \in \mathcal{T}_{en}^i, \quad i \in \{1:r\}, && (59)\\
G_i(\bar{u}_\tau^+, \bar{y}_\tau) &= 0, \quad \text{if } \tau \in \mathcal{T}_{ex}^i, \quad i \in \{1:r\}, && (60)\\
g_i^{(1)}(\bar{y}_\tau) &= 0, \quad \text{if } \tau \in \mathcal{T}_{to}^i \text{ and } q_i \geq 2, \quad i \in \{1:r\}. && (61)
\end{aligned}
$$

We call (48)-(61) the *shooting equations*. A solution of these equations is called a *shooting extremal*. We will establish in Section 5 that the shooting equations are, under proper assumptions, the optimality system of a well-posed quadratic problem.

Note that these equations involve *algebraic variables*, in the terminology of differential algebraic systems, i.e., functions of time whose derivative does not appear in the equations. The algebraic variables here are the control and alternative Lagrange multiplier associated with the state constraint. But these variables are to be viewed as functions of the differential variables (state and alternative costate, since the implicit function theorem applies to the "algebraic" equations (50)-(51), see the discussion in the proof of Proposition 1). In particular, there is no need for an explicit expression of the algebraic variables as the functions of the differential ones.

PROPOSITION 3 *Let $(\bar{u}, \bar{y}, \bar{p}, \bar{\eta})$ be a continuous Pontryagin extremal with finitely many junction points. If* (24) *holds, and $\bar{u}$ is continuous, then the following relations hold:*

$$
\left\{
\begin{array}{llll}
\text{(i)} & 0 = \nu_\tau^i + [\eta_\tau^{i,1}], & i \in I(\tau),\ \tau \in \mathcal{T}_{ex} \\
\text{(ii)} & 0 = [\eta_\tau^{i,j}], & i \in I(\tau),\ j \in [2:q_i],\ \tau \in \mathcal{T}_{ex}.
\end{array}
\right. \tag{62}
$$

$$
\left\{
\begin{array}{lll}
\text{(i)} & \nu_\tau^{i,1} = \nu_\tau^i + [\eta_\tau^{i,1}], & \tau \in \mathcal{T}_{en}^i, \\
\text{(ii)} & \nu_\tau^{i,j} = [\eta_\tau^{i,j}], & \tau \in \mathcal{T}_{en}^i,\ j \in [2:q_i], \\
\text{(iii)} & 0 = \nu_\tau^i + [\eta_\tau^{i,1}], & i \in I(\tau),\ \tau \in \mathcal{T}_{en} \backslash \mathcal{T}_{en}^i, \\
\text{(iv)} & 0 = [\eta_\tau^{i,j}], & i \in I(\tau),\ j \in [2:q_i],\ \tau \in \mathcal{T}_{en}.
\end{array}
\right. \tag{63}
$$

$$
\left\{
\begin{array}{lll}
\text{(i)} & \nu_\tau^{i,1} = \nu_\tau^i, & \tau \in \mathcal{T}_{to}^i, \\
\text{(ii)} & 0 = \nu_\tau^i + [\eta_\tau^{i,1}], & i \in I(\tau),\ \tau \in \mathcal{T}_{to} \backslash \mathcal{T}_{to}^i, \\
\text{(iii)} & 0 = [\eta_\tau^{i,j}], & i \in I(\tau),\ j \in [2:q_i],\ \tau \in \mathcal{T}_{to}.
\end{array}
\right. \tag{64}
$$

Proof. These relations are simple consequences of (43)-(45), (42) and (56)-(58). For (64)(i), use the fact that, if $\tau \in \mathcal{T}_{to}^i$, then $[\eta_\tau^{i,1}] = 0$ by (52). ∎

3.2. Reduction of isolated contact points

If τ is a contact point for constraint i, then $g_i(\bar{y}_t))$ attains a local maximum at τ, and hence (if these values are well-defined) $\dot{g}(\bar{y}_\tau) = 0$ and $\ddot{g}(\bar{y}_\tau) \leq 0$. We say that a touch point τ is *reducible* if $\ddot{g}(\bar{y}_\tau)$ is continuous at time τ, and $\ddot{g}(\bar{y}_\tau) < 0$.

Let $\tau \in (0, T)$ be a touch point of a feasible trajectory $(\bar{u}, \bar{y})$; set for $y \in \mathcal{Y}$:

$$\gamma(y) := \max\{g_i(y_t),\ \ t \in [\tau - \varepsilon, \tau + \varepsilon]\}, \tag{65}$$

where $\varepsilon > 0$ is so small that

$$[\tau - \varepsilon, \tau + \varepsilon] \subset [0, T] \ \text{ and } g_i(\bar{y}_t) < 0, \text{ for all } t \neq \tau,\ |t - \tau| \leq \varepsilon. \tag{66}$$

Let us see how to compute a Taylor expansion of $\gamma(\cdot)$ in the space $W^{2,\infty}(0, T)$ (the one we need for state constraints of the order of at least two), in the vicinity of a C^2 function (which will apply to optimal trajectories with continuous control).

LEMMA 3 *Let x be a C^2 function: $[a, b] \to I\!R$, having a unique maximum at some $\theta \in (a, b)$, and such that $\ddot{x}_\theta < 0$. If y is close enough to x in $X := W^{2,\infty}(a, b)$, then it has over $[a, b]$ a unique maximum $\tau(y)$, and we have*

$$\tau(y) - \tau(x) = -\dot{y}_{\tau(x)}/\ddot{x}_{\tau(x)} + o(\|y - x\|_X), \tag{67}$$

$$\max(y) = y_{\tau(y)} = y_{\tau(x)} - \tfrac{1}{2}\frac{(\dot{y}_{\tau(x)})^2}{\ddot{x}_{\tau(x)}} + o(\|y - x\|_X^2). \tag{68}$$

Proof. There exists $\varepsilon_1 > 0$ such that, for y close enough to x in X, we have

$$\ddot{y}_t < \tfrac{1}{2}\ddot{x}_\theta < 0, \quad \text{for a.a. } t \in [\theta - \varepsilon_1, \theta - \varepsilon_1], \tag{69}$$

and y has over $[a, b]$ a unique maximum $\tau(y)$ that belongs to $[\theta - \varepsilon_1, \theta - \varepsilon_1]$. When $y \to x$ in X, $\max(y)$ converges to $\max(x) = x_\theta$, and hence, $\tau(y) \to \tau(x) = \theta$. Set $\hat{\tau}(y) := \tau(y) - \tau(x)$. Since

$$-\dot{y}_{\tau(x)} = \dot{y}_{\tau(y)} - \dot{y}_{\tau(x)} = \int_{\tau(x)}^{\tau(y)} \ddot{y}_s \mathrm{d}s = \hat{\tau}(y)\ddot{x}_\theta + O(\hat{\tau}(y)\|y - x\|_X) \tag{70}$$

and $\ddot{x}_\theta \neq 0$, relation (67) follows. Since x is of class C^2 and $\dot{x}_\theta = 0$, we have

$$x_{\tau(y)} = x_{\tau(x)} + \tfrac{1}{2}\ddot{x}_{\tau(x)}(\tau(y) - \tau(x))^2 + o((\tau(y) - \tau(x))^2), \tag{71}$$

and since $|\ddot{y} - \ddot{x}| \to 0$ uniformly, by a second-order Taylor expansion, we get:

$$(y - x)_{\tau(y)} = (y - x)_{\tau(x)} + \dot{y}_{\tau(x)}(\tau(y) - \tau(x)) + o(\|y - x\|_X^2). \tag{72}$$

Note that in the above expression we could neglect the second order term, which is of order $o(\|y - x\|_X^2$. Summing (71) and (71), and using (67), we get the conclusion. ∎

For reducible touch points (defined in Section 3.2), associated with state constraint i of order $q_i > 1$, by the above lemma, we can replace locally (in time) the state constraint by the corresponding (scalar) "reduced" constraint, whose maximum over time is nonpositive. Set, for $\varepsilon > 0$ small enough, and $y \in \mathcal{Y}$:

$$\mu^{i,\tau}(y) := \max_{t\in[\tau-\varepsilon,\tau+\varepsilon]} g_i(y). \tag{73}$$

If z is the solution of the linearized state equation

$$\dot z = f'(\bar u_t, \bar y_t)(v_t, z_t) \quad \text{on } [0,T]; \qquad z_0 = 0, \tag{74}$$

since $g_i'(\bar y_t) f_u(\bar u_t, \bar y_t) = 0$, we have, as $q_i > 1$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[g_i'(\bar y_t) z_t\right] = g_i''(\bar y_t)(f(\bar u_t, \bar y_t), z_t) + g_i'(\bar y_t) f_y(\bar u_t, \bar y_t) z_t = g_{i,y}^{(1)}(\bar y_t) z_t. \tag{75}$$

It follows from Lemma 3 that we have the Taylor expansion

$$\begin{aligned}\mu^{i,\tau}(\bar y + z) = \;& \mu^{i,\tau}\left[g_i(\bar y) + g_i'(\bar y) z + \tfrac{1}{2} g_i''(\bar y)(z,z)^2 + o(\|z\|_\infty^2)\right]\\ = \;& g_i(\bar y_\tau) + g_i'(\bar y_\tau) z_\tau + \\ & \tfrac{1}{2}[g_i''(\bar y_\tau)(z_\tau, z_\tau)^2 - (g_{i,y}^{(1)}(\bar y_\tau) z_\tau)^2 / \ddot g_i(\bar y_\tau)] + o(\|z\|_\infty^2).\end{aligned} \tag{76}$$

4. Second-order optimality conditions

4.1. Main result

For $s \in [2, \infty]$, set $\mathcal{V}_s := L^s(0, T, \mathbb{R}^m)$ and $\mathcal{Z}_s := W^{1,s}(0, T; \mathbb{R}^n)$. Consider the *tangent quadratic cost function* $\mathcal{J} : \mathcal{V}_2 \times \mathcal{Z}_2 \to \mathbb{R}$:

$$\begin{aligned}\mathcal{J}(v,z) \;:=& \int_0^T H_{(u,y)^2}(\bar u_t, \bar y_t, \bar p_t)(v_t, z_t)^2 \mathrm{d}t + \phi''(\bar y_T)(z_T, z_T)\\ &+ \sum_{i=1}^r \left(\int_0^T g_i''(\bar y_t)(z_t, z_t) \mathrm{d}\bar\eta_{i,t} \;-\; \sum_{\tau\in\mathcal{T}_{to}^i} [\bar\eta_{i,\tau}] \frac{(g_{i,y}^{(1)}(\bar y_\tau) z_\tau)^2}{g_i^{(2)}(\bar u_\tau, \bar y_\tau)} \right).\end{aligned} \tag{77}$$

Note that the contribution of touch points to this quadratic cost coincides with the second order term of the Taylor expansion (76). Consider also the *linearized state constraints*

$$g_i'(\bar y_t) z_t \;\le\; 0 \;\text{ on } \mathcal{I}_b^i, \;\text{ and } g_i'(\bar y_t) z_t = 0 \text{ on } \operatorname{supp}(\mathrm{d}\bar\eta_i),\; i \in \{1:r\} \tag{78}$$

$$g_i'(\bar y_\tau) z_\tau \;\le\; 0 \;\text{ for all } \tau \in \mathcal{T}_{to}^i,\; i \in \{1:r\}, \tag{79}$$

$$g_i'(\bar y_\tau) z_\tau \;=\; 0 \;\text{ if } \nu_\tau^i > 0, \text{ for all } \tau \in \mathcal{T}_{to}^i,\; i \in \{1:r\}. \tag{80}$$

Consider also the relation stronger than (78)

$$g_i'(\bar y_t) z_t = 0 \;\text{ on } \mathcal{I}_b^i,\; i \in \{1:r\}. \tag{81}$$

For $s \in [2, \infty]$, we call *critical cone* (in $\mathcal{V}_s$) the set

$$C_s(\bar{u}, \bar{y}) := \{(v, z) \in \mathcal{V}_s \times \mathcal{Z}_s;\ \ (74) \text{ and } (78)\text{-}(80) \quad \text{hold}\}, \tag{82}$$

and *strict critical cone* the set

$$C_s^S(\bar{u}, \bar{y}) := \{(v, z) \in \mathcal{V}_s \times \mathcal{Z}_s;\ \ (74) \text{ and } (79)\text{-}(81) \quad \text{hold}\}. \tag{83}$$

Obviously $C_s^S(\bar{u}, \bar{y}) \subset C_s(\bar{u}, \bar{y})$, for all $s \in [2, \infty]$. We will say that *strict complementarity holds on boundary arcs* if the support of $\mathrm{d}\bar{\eta}$ contains all boundary arcs. In that case, (78) and (81) coincide, and $C_s^S(\bar{u}, \bar{y}) = C_s(\bar{u}, \bar{y})$. We set, for $u \in \mathcal{U}$ and $y = y(u)$:

$$J(u) := \int_0^T \ell(u_t, y_t)\mathrm{d}t + \phi(y_T). \tag{84}$$

Consider the following relations:

$$\mathcal{J}(v, z) \geq 0, \quad \text{for all } (v, z) \in C_2^S(\bar{u}, \bar{y}). \tag{85}$$

$$\text{For some } \beta > 0: \quad \mathcal{J}(v, z) \geq \beta \|v\|_2^2, \quad \text{for all } (v, z) \in C_2(\bar{u}, \bar{y}). \tag{86}$$

$$\text{For some } \alpha > 0: \quad H_{uu}(\bar{u}_t, \bar{y}_t, \bar{p}_{t^\pm})(v, v) \geq \alpha |v|^2, \quad \text{for all } v \in I\!R^m. \tag{87}$$

Obviously (86) implies (85). Applying Pontryagin's priciple to the problem of minimizing $\mathcal{J}(v, z)$ over $C_2(\bar{u}, \bar{y})$, we see that (86) implies also (87).

We say that $(\bar{u}, \bar{y})$ is a local solution of (1) satisfying the (local) quadratic growth condition if, for all $\varepsilon_1 > 0$, there exists $\varepsilon_2 > 0$ such that

$$J(u) \geq J(\bar{u}) + \tfrac{1}{2}(\beta - \varepsilon_1)\|u - \bar{u}\|_2^2, \quad \text{if } \|u - \bar{u}\|_\infty \leq \varepsilon_2,\ u \text{ feasible for (1)}. \tag{88}$$

THEOREM 3 *Let* $(\bar{u}, \bar{y})$ *be a continuous Pontryagin extremal satisfying* (24), *whose all touch points for state constraints of order greater than one are reducible. Then*
(i) *(Second-order necessary condition): if* $\bar{u}$ *is a local solution of* (1), *then* (85) *holds.*
(ii) *(Second-order sufficient condition): assume, in addition, that* (86) *holds. Then,* $(\bar{u}, \bar{y})$ *satisfies the local quadratic growth condition* (88).

We need a couple of preliminary lemmas. Let

$$|qN_b| := \sum_{i=1}^r q_i N_b^i, \quad |N_{to}| := \sum_{i=1}^r N_{to}^i. \tag{89}$$

Denote the neighborhood of the boundary arcs, for $\varepsilon > 0$, by

$$\mathcal{I}_b^{i,\varepsilon} := \cup_{k=1}^{N_b^i} [\tau_{en}^{i,k} - \varepsilon, \tau_{ex}^{i,k} + \varepsilon], \quad i \in \{1 : r\}. \tag{90}$$

Here we take $\varepsilon \geq 0$ so small that $\mathcal{I}_b^{i,\varepsilon} \subset [0,T]$, for all $i \in \{1:r\}$. By $\varphi|_{\mathcal{I}_b}$, we denote the restriction to $\mathcal{I}_b$ of function φ defined over $[0,T]$. For all $v \in \mathcal{V}$, let $z(v) \in \mathcal{Z}$ denote the solution of the linearized state equation (74).

For $s \in [2,\infty]$, set $\hat{W}_s^\varepsilon := \prod_{i=1}^r W^{q_i,s}(\mathcal{I}_b^{i,\varepsilon})$ and define the operators $\hat{\mathcal{A}}^\varepsilon : \mathcal{V}_s \to \hat{W}_s^\varepsilon$, $\mathcal{A}^\varepsilon : \mathcal{V}_s \to \hat{W}_s^\varepsilon \times I\!R^{|N_{to}|}$, and $\mathcal{A} : \mathcal{V}_s \to \hat{W}_s^0 \times I\!R^{|qN_b|} \times I\!R^{|N_{to}|}$ by

$$\hat{\mathcal{A}}_i^\varepsilon v := G_i'(\bar{u}_t, \bar{y}_t)(v_t, z_t(v)); \quad t \in \mathcal{I}_b^{i,\varepsilon}, \;\; i \in \{1:r\}, \tag{91}$$

$$\mathcal{A}^\varepsilon v := (\hat{\mathcal{A}}_1^\varepsilon v, \dots, \hat{\mathcal{A}}_r^\varepsilon v); \quad g_i'(\bar{y})z(v)(\mathcal{T}_{to}^i), \;\; i \in \{1:r\}, \tag{92}$$

$$\mathcal{A}v := \left(\hat{\mathcal{A}}_i^0(v), \;\; g_{i,y}^{\{0:(q_i-1)\}}(\bar{y})z(v)(\mathcal{T}_{en}^i), \;\; g_i'(\bar{y})z(v)(\mathcal{T}_{to}^i), \;\; i \in \{1:r\}\right). \tag{93}$$

Lemma 4 *Let $(\bar{u}, \bar{y})$ be a continuous trajectory satisfying the state constraints of* (1)*, with finitely many junction points. If* (24) *holds, then $\mathcal{A}$ and $\mathcal{A}^\varepsilon$, for $\varepsilon \geq 0$ small enough, and $s \in [2,\infty]$, are onto.*

Proof. We skip this proof whose arguments are classical, see, e.g., Lemma 4.3 in Bonnans and Hermant (2009b).

The *cone of radial critical directions* $C_s^R(\bar{u}, \bar{y})$, for $s \in [2,\infty]$, is (note that the radiality condition deals with boundary arcs only):

$$C_s^R(\bar{u}, \bar{y}) := \left\{ \begin{array}{l} (v,z) \in C_s(\bar{u}, \bar{y}); \quad \text{for some } \nu > 0 \text{ and } \varepsilon > 0 : \\ g_i(\bar{y}) + \nu g_i'(\bar{y})z \leq 0 \text{ on } \mathcal{I}_b^{i,\varepsilon}, \, i \in \{1:r\} \end{array} \right\}. \tag{94}$$

We set

$$C_s^{R,S}(\bar{u}, \bar{y}) := C_s^R(\bar{u}, \bar{y}) \cap C_s^S(\bar{u}, \bar{y}), \quad s \in [2,\infty]. \tag{95}$$

Lemma 5 *Under the assumptions of Lemma 4, the set $C_\infty^{R,S}(\bar{u}, \bar{y})$ is a dense subset, in the L^2 norm, of $C_2^S(\bar{u}, \bar{y})$.*

Proof. a) We claim that $C_\infty^S(\bar{u}, \bar{y})$ is a dense subset, in the L^2 norm, of $C_2^S(\bar{u}, \bar{y})$. Indeed, let $(\hat{v}, \hat{z}) \in C_2^S(\bar{u}, \bar{y})$. For $M > 0$, define the truncation of $\hat{v}$ as

$$v_t^M := \max(-M, \min(M, \hat{v}_t)), \quad \text{for all } t \in [0,T]. \tag{96}$$

Then $v^M \to \hat{v}$ in L^2. Denote by $\hat{v}^M$ the projection of v^M onto $C_2^S(\bar{u}, \bar{y})$. Since projections in Hilbert spaces are nonexpansive, $\hat{v}^M \to \hat{v}$ in L^2. In view of the expression (83) of the strict critical cone, $\hat{v}^M$ is solution of the problem

$$\min_{v \in \mathcal{V}_2} \tfrac{1}{2} \int_0^T |v_t - v_t^M|^2 \mathrm{d}t; \quad (74) \text{ and } (79)\text{-}(81) \quad \text{holding}. \tag{97}$$

This is a strongly convex linear optimal control problem with state constraints. Since $\mathcal{A}$ is onto, the solution $\hat{v}^M$ is characterized by the optimality condition $\hat{v}_t^M = v_t^M - p_t^M f_u(\bar{u}_t, \bar{y}_t)$, where the costate p^M is the solution of a certain adjoint equation that we do not need write, and belongs to $BV(0,T,\mathbb{R}^{n*})$. Therefore $\hat{v}^M \in C_\infty^S(\bar{u},\bar{y})$. The claim follows.

b) We claim that there exists $\hat{\alpha} > 0$ such that, for $\varepsilon > 0$ small enough, the operator $\mathcal{A}^\varepsilon : \mathcal{V}_2 \to \hat{W}_2^0 \times \mathbb{R}^{|N_{to}|}$ has a right pseudo inverse $\mathcal{A}^{\varepsilon,\dagger}$ such that

$$\mathcal{A}^\varepsilon \mathcal{A}^{\varepsilon,\dagger} w = w, \quad \|\mathcal{A}^{\varepsilon,\dagger} w\| \le \hat{\alpha}^{-1}\|w\|_\varepsilon, \text{ for all } w \in \hat{W}_2^\varepsilon \times \mathbb{R}^{|N_{to}|}. \tag{98}$$

Since, for $\varepsilon > 0$, by Lemma 4, $\mathcal{A}^\varepsilon$ is onto, there exists $\alpha_\varepsilon > 0$ such that $\mathrm{Im}(\mathcal{A}^\varepsilon)$ contains $\alpha_\varepsilon B_\varepsilon$, where B_ε denotes the unit ball in $\hat{W}_2^\varepsilon$. Fix $\varepsilon_0 > 0$. For $\varepsilon \in (0,\varepsilon_0)$ and $v \in \mathcal{U}$, $\mathcal{A}^\varepsilon v$ is the restriction of $\mathcal{A}^{\varepsilon_0} v$ to $\hat{W}_2^\varepsilon$. There is an obvious imbedding a_ε from $W^{q_i,2}(\mathcal{I}_b^{i,\varepsilon})$ into $W^{q_i,2}(\mathcal{I}_b^{i,\varepsilon_0})$, by taking the derivative of order q_i equal to zero over $\mathcal{I}_b^{i,\varepsilon_0} \setminus \mathcal{I}_b^{i,\varepsilon}$, and $\|a_\varepsilon\|$ is uniformly bounded by some constant $\hat{a}$. So, taking $\mathcal{A}^{\varepsilon,\dagger} := \mathcal{A}^{\varepsilon_0,\dagger} \circ a_\varepsilon$, we obtain a right pseudo inverse with constant $\hat{\alpha} := \alpha_{\varepsilon_0}\hat{a}$.

c) We claim that $C_\infty^{R,S}(\bar{u},\bar{y})$ is a dense subset, in the L^2 norm, of $C_\infty^S(\bar{u},\bar{y})$. Indeed, let $(\hat{v},\hat{z}) \in C_\infty^S(\bar{u},\bar{y})$, and set

$$b_\varepsilon := \Big((\hat{\mathcal{A}}_1^\varepsilon v, \ldots, \hat{\mathcal{A}}_r^\varepsilon v); \quad 0 \times g_i'(\bar{y}) z(v)(\mathcal{T}_{to}^i), \;\; i \in \{1:r\} \Big). \tag{99}$$

Consider the projection problem

$$\min_{v \in \mathcal{V}_2} \tfrac{1}{2} \int_0^T |v_t|^2 \mathrm{d}t; \quad \mathcal{A}^\varepsilon v = b_\varepsilon. \tag{100}$$

In view of step b), its unique solution denoted $\hat{v}^\varepsilon$ has an L^2 norm of order $\|b_\varepsilon\|$, (norm of $\hat{W}_2^\varepsilon \times \mathbb{R}^{|N_{to}|}$). Since $\|b_\varepsilon\| \to 0$ when $\varepsilon \downarrow 0$, we have that $\hat{v}^\varepsilon \to 0$ in $\mathcal{V}_2$. Similarly to step a), $\hat{v}^\varepsilon$ satisfies $\hat{v}_t^\varepsilon = -p_t^\varepsilon f_u(\bar{u}_t,\bar{y}_t)$, where the costate p^ε is solution of the equation with r.h.s. accounting for the state constraints corresponding to the constraints of (100), and is bounded as a function of time. The constraints of (100) are such that $\hat{v} - \hat{v}^\varepsilon \in C_\infty^{R,S}(\bar{u},\bar{y})$. Our claim follows.

d) We conclude by combining steps a) and c). ■

We recall that a continuous quadratic form Q defined over a Hilbert space X is a *Legendre form* (see, e.g., Bonnans and Shapiro, 2000, Ioffe, 1979), if it is weakly lower semi-continuous, and satisfies the following property: If $v_k \rightharpoonup v$ (weak convergence) in X, and $Q(v_k) \to Q(v)$, then $v_n \to v$ strongly.

Proof (Proof of Theorem 3). (i) *Second-order necessary condition.* Denote by $y(u)$ the state associated with control u. We remind that the function $\mu^{i,\tau}$ was defined in (73). The *reduced problem* (see Section 3.2.3 in Bonnans and Shapiro, 2000) is

$$\operatorname*{Min}_{u \in \mathcal{U}} J(u); \; g_i(y(u)) \le 0 \text{ on } \mathcal{I}_b^{i,\varepsilon}; \; \mu^{i,\tau}(y(u)) \le 0, \text{ for all } \tau \in \mathcal{T}_{to}^i, \, i \in \{1:r\}. \tag{101}$$

For u close to $\bar{u}$ in $\mathcal{U}$, we have that u is feasible for (1) iff it is feasible for (101). It follows that $\bar{u}$ is a local solution of (101), whose associated Lagrange multipliers of the reduced problem are the restriction of a Lagrange multiplier for the original formulation. The critical cones of various types coincide for the two problems. The Lagrangian function, associated with the reduced formulation (101), using notation (73), is

$$L(u,\eta) := J(u) + \sum_{i=1}^{r} \int_{\mathcal{I}_b^{i,\varepsilon}} g(y_t(u))\mathrm{d}\eta_{i,t} + \sum_{i=1}^{r} \sum_{\tau \in \mathcal{T}_{to}^i} [\eta_{i,\tau}]\mu^{i,\tau}(y(u)). \quad (102)$$

In view of (76), $\mathcal{J}(v, z(v))$ is the second order term in the Taylor expansion w.r.t. u of $L(\bar{u}, \bar{\eta})$. Since by Lemma 4 the derivative of constraints is onto, problem (101) is qualified. The standard second-order necessary conditions (see Section 3.2.2 in Bonnans and Shapiro, 2000) and the fact that the "σ-term" appearing in it vanishes for radial critical directions (see Remark 3.47 in the same reference), imply that

$$D^2L(\bar{u},\bar{\eta})(v,v) = \mathcal{J}(v,z) \geq 0, \quad \text{for any } (v,z) \in C_\infty^R(\bar{u},\bar{y}). \quad (103)$$

Since $z = z(v)$ when $(v,z) \in C_2(\bar{u},\bar{y})$ and $(v,z(v)) \to \mathcal{J}(v,z(v))$ is continuous in the L^2 norm, we conclude with Lemma 5.

(ii) *Second-order sufficient condition.* Let (86) hold, but not (88). Then, there exists a feasible sequence (u^k, y^k) such that $u^k \neq \bar{u}$, $u^k \to \bar{u}$ in $\mathcal{U}$ and

$$J(u^k) \leq J(\bar{u}) + o(\|u^k - \bar{u}\|_2^2). \quad (104)$$

Let $\sigma_k := \|u^k - \bar{u}\|_2$, $(v^k, z^k) := \sigma_k^{-1}(u^k - \bar{u}, y^k - \bar{y})$. Then, $\|v^k\|_2^2 = 1$, and extracting, if necessary, a subsequence, we may assume that $v^k \rightharpoonup \bar{v}$ in $L^2(0,T,I\!\!R^m)$, where by "$\rightharpoonup$" we denote weak convergence. Let z^k (respectively $\bar{z}$) denote the solution of the linearized state equation (74) with v^k (respectively $v = \bar{v}$). In view of the classical estimate

$$\|y^k - \bar{y} - \sigma_k z^k\|_\infty = O(\sigma_k^2), \quad (105)$$

we have that $z^k \rightharpoonup \bar{z}$ in $H^1(0,T,I\!\!R^n)$. Using (104) we deduce that

$$0 \geq \limsup_k \frac{J(u^k) - J(\bar{u})}{\sigma_k} = DJ(\bar{u})\bar{v}. \quad (106)$$

We easily obtain $g'(\bar{y}_t)\bar{z}_t \leq 0$ if $g(\bar{y}_t) = 0$ and $\int_0^T g'(\bar{y}_t)\bar{z}_t\mathrm{d}\bar{\eta}_t \leq 0$. Since

$$0 = \int_0^T H_u(\bar{u}_t,\bar{y}_t,\bar{p}_t)\mathrm{d}t = DJ(\bar{u})\bar{v} + \int_0^T g'(\bar{y}_t)\bar{z}_t\mathrm{d}\bar{\eta}_t, \quad (107)$$

we deduce that $DJ(\bar{u})\bar{v} = 0 = \int_0^T g'(\bar{y}_t)\bar{z}_t\mathrm{d}\bar{\eta}_t$. It follows that $(\bar{v},\bar{z}) \in C_2(\bar{u},\bar{y})$, and since $D_uL(\bar{u},\bar{\eta}) = 0$, using (105) we obtain

$$\limsup_k \mathcal{J}(v^k, z^k) = \limsup_k \frac{L(u^k, \bar{\eta}) - L(\bar{u}, \bar{\eta})}{\frac{1}{2}\sigma_k^2} \leq \limsup_k \frac{J(u^k) - J(\bar{u})}{\frac{1}{2}\sigma_k^2}. \quad (108)$$

In view of (104), and since by (87) $\mathcal{J}(v^k, z^k)$ is weakly l.s.c., we deduce that $\mathcal{J}(\bar{v}, \bar{z}) \leq \liminf_k \mathcal{J}(v^k, z^k) \leq 0$. As $(\bar{v}, \bar{z}) \in C_2(\bar{u}, \bar{y})$, (86) implies $(\bar{v}, \bar{z}) = 0$, so that $\mathcal{J}(\bar{v}, \bar{z}) = \lim_k \mathcal{J}(v^k, z^k)$. By (87), $\mathcal{J}$ is a Legendre form, and hence, $v^k \to \bar{v}$ strongly in $L^2(0, T, \mathbb{R}^m)$, which is impossible since $\bar{v} = 0$ and v^k is of unit norm. ∎

Remark 2 (i) An improvement w.r.t. Bonnans and Hermant (2009a) is that we do not need the hypothesis below (which will, however, be essential for checking the well-posedness of the shooting algorithm):

$$g_i^{(q_i+1)}(\bar{u}_{\tau_{en}}^-, \bar{y}_\tau) \neq 0, \; g_i^{(q_i+1)}(\bar{u}_{\tau_{ex}}^+, \bar{y}_\tau) \neq 0, \quad \text{for all } \tau_{en} \in \mathcal{T}_{en}^i, \;\; \tau_{ex} \in \mathcal{T}_{ex}^i. \quad (109)$$

(ii) Another improvement in the necessary conditions is that we have relaxed the hypothesis of strict complementarity on boundary arcs used in Bonnans and Hermant (2009a). However, we had to work with a subset of the critical cone. Proving that $\mathcal{J}(v, z) \geq 0$ for all critical direction, without strict complementarity, is an interesting open problem.

(iii) For nonessential touch points $\tau \in \mathcal{T}_{to}^i$ we can avoid the reducibility condition; see Bonnans and Hermant (2007, 2009a).

4.2. The alternative tangent quadratic problems

For the study of the well-posedness of the shooting algorithm, i.e., the fact that the Jacobian at the solution is invertible (which ensures the convergence of Newton's method, and the stability of the solution under a small perturbation) we will need the *alternative tangent quadratic cost* function:

$$\begin{aligned} \mathcal{J}_q(v, z) \; := \; & \int_0^T H^q_{(u,y)^2}(\bar{u}_t, \bar{y}_t, \bar{p}_t^q, \bar{\eta}_t^q)(v_t, z_t)^2 \mathrm{d}t \;\; + \phi''(\bar{y}_T)(z_T, z_T) \\ + & \sum_{i=1}^r \sum_{\tau \in \mathcal{T}^i} \sum_{j=1}^{q_i} \nu_\tau^{i,j} g_{i,yy}^{(j-1)}(\bar{y}_\tau)(z_\tau)^2 - \sum_{i=1}^r \sum_{\tau \in \mathcal{T}_{to}^i} [\bar{\eta}_{i,\tau}] \frac{(g_{i,y}^{(1)}(\bar{y}_\tau) z_\tau)^2}{g_i^{(2)}(\bar{u}_\tau, \bar{y}_\tau)}, \end{aligned} \quad (110)$$

and the set of constraints:

$$\begin{aligned} \dot{z}_t = f'(\bar{u}_t, \bar{y}_t)(v_t, z_t) \qquad & \text{on } [0, T]; \quad z_0 = 0, && (111) \\ g_{i,y}^{(j)}(\bar{y}_\tau) z_\tau = 0 \qquad & j \in \{0 : (q_i - 1)\}, \; \tau \in \mathcal{T}_{en}^i, \quad i \in \{1 : r\}, && (112) \\ DG_i(\bar{u}_t, \bar{y}_t)(v_t, z_t) = 0 \qquad & t \in \mathcal{I}_b^i, \quad i \in \{1 : r\}, && (113) \\ g_i'(\bar{y}_\tau) z_\tau = 0 \qquad & \tau \in \mathcal{T}_{to}^i, \;\; i \in \{1 : r\}. && (114) \end{aligned}$$

Let the *alternative tangent linear quadratic problem* (PQ_q) be defined by:

$$(\mathbf{PQ}_q) \qquad \min_{(v,z)\in\mathcal{V}\times\mathcal{Z}} \tfrac{1}{2}\mathcal{J}_q(v,z) \qquad \text{subject to (111)-(114).} \tag{115}$$

We denote

$$\begin{cases} \nu_{\mathcal{T}_{en}} & = \left(\nu_\tau^{i,j},\ \tau\in\mathcal{T}_{en}^i,\ 1\le i\le r,\ 1\le j\le q_i\right). \\ \nu_{\mathcal{T}_{to}} & = \left(\nu_\tau^{i},\ \tau\in\mathcal{T}_{to}^i,\ 1\le i\le r\right). \end{cases} \tag{116}$$

Lemma 6 *Let $(\bar u,\bar y)$ be a Pontryagin extremal, with $\bar u$ continuous, satisfying* (24)-(25)*, with classical and alternative multipliers $(\bar p,\bar\eta)$ and $(\bar p^q,\bar\eta^q,\nu_{\mathcal{T}_{en}},\nu_{\mathcal{T}_{to}})$. Then the quadratic cost functions $\mathcal{J}$ and $\mathcal{J}_q$, defined respectively in* (77) *and* (110)*, are equal to each other over the space of linearized trajectories $(v,z)\in\mathcal{V}\times\mathcal{Z}$ satisfying the linearized state equation* (74)*.*

Proof. Let $(v,z)\in\mathcal{V}\times\mathcal{Z}$ satisfy (74). Denote the difference of quadratic costs as $\Delta := \mathcal{J}(v,z)-\mathcal{J}_q(v,z)$. We observe that the terms corresponding to the touch points and to the final time vanish. Writing

$$\mathrm{d}\bar\eta_t = \bar\eta_t^0\mathrm{d}t + \sum_{\tau\in\mathcal{T}}[\bar\eta_\tau]\delta_\tau \tag{117}$$

where $\bar\eta_t^0$ is the density of $\bar\eta$ over each arc (well defined in view of Proposition 1), and using (27), we may write Δ as a sum over the components of the state constraint: $\Delta=\sum_{i=1}^r\Delta_i$, where

$$\begin{aligned} \Delta_i \quad := \quad & \sum_{j=1}^{q_i}\int_0^T g_{i,y}^{(j-1)}(\bar y_t)f''(\bar u_t,\bar y_t)(v_t,z_t)^2\eta_t^{i,j}\mathrm{d}t + \int_0^T g_i''(\bar y_t)(z_t,z_t)\eta_t^0\mathrm{d}t \\ & - \int_0^T D^2g^{(q_i)}(\bar u_t,\bar y_t)(v_t,z_t)^2\eta_t^{i,q_i}\mathrm{d}t + \sum_{\tau\in\mathcal{T}}\nu_\tau^i g_i''(\bar y_\tau)(z_\tau,z_\tau) \\ & - \sum_{\tau\in\mathcal{T}_{en}^i}\sum_{j=1}^{q_i}\nu_\tau^{i,j}g_{i,yy}^{(j-1)}(\bar y_\tau)(z_\tau,z_\tau). \end{aligned} \tag{118}$$

So, it suffices to check that $\Delta_i=0$. Since $g_i^{(j)}(\bar u_t,\bar y_t)=g_{i,y}^{(j-1)}(\bar y_t)f(\bar u_t,\bar y_t)$, for $j=1$ to q_i, we have that

$$\begin{aligned} D^2g_i^{(j)}(\bar u_t,\bar y_t)(v_t,z_t)^2 = {} & g_{i,yyy}^{(j-1)}(f(\bar u_t,\bar y_t),z_t,z_t) \\ & + 2g_{i,yy}^{(j-1)}(z,f'(\bar u_t,\bar y_t)(v_t,z_t)) + \ g_{i,y}^{(j-1)}f''(\bar u_t,\bar y_t)(v_t,z_t)^2. \end{aligned} \tag{119}$$

In addition, by the linearized state equation (74), we have, for all $j\in\{1:q_i\}$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left[g_{i,yy}^{(j-1)}(\bar y_t)(z_t,z_t)\right] = \quad & g_{i,yyy}^{(j-1)}(\bar y_t)(f(\bar u_t,\bar y_t),z_t,z_t) \\ & +2g_{i,yy}^{(j-1)}(\bar y_t)(z,f'(\bar u_t,\bar y_t)(v_t,z_t)), \end{aligned}$$

which gives by (119), for $j \in \{1 : q_i\}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[g_{i,yy}^{(j-1)}(\bar{y}_t)(z_t, z_t)\right] = D^2 g_i^{(j)}(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2 - g_{i,y}^{(j-1)}(\bar{y}_t) f''(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2. \tag{120}$$

Since $g_{i,u}^{(j-1)}(\bar{u}_t, \bar{y}_t) \equiv 0$ for $j \in \{1 : q_i\}$, we have

$$g_{i,yy}^{(j-1)}(\bar{y}_t)(z_t, z_t) = D^2 g_i^{(j-1)}(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2, \quad j \in \{1 : q_i\}. \tag{121}$$

Multiplying (120) by $\eta^{i,j}$, integrating over $[0, T]$, integrating by parts the left-hand side (recall that $\dot{\eta}^{i,j} = -\eta^{i,j-1}$), and using (121) we obtain, for $j \in \{1 : q_i\}$:

$$\begin{aligned}
&\int_0^T D^2 g_i^{(j-1)}(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2 \eta_t^{i,j-1} \mathrm{d}t - \sum_{\tau \in \mathcal{T}} [\eta_\tau^{i,j}] g_{i,yy}^{(j-1)}(\bar{y}_\tau)(z_\tau, z_\tau) \\
&= \int_0^T D^2 g_i^{(j)}(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2 \eta_t^{i,j} \mathrm{d}t - \int_0^T g_{i,y}^{(j-1)}(\bar{y}_\tau) f''(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2 \eta_t^{i,j} \mathrm{d}t.
\end{aligned}$$

Adding the above equalities for $j \in \{1 : q_i\}$, we get after simplification by the terms $\int_0^T D^2 g_i^{(j)}(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2 \eta^{i,j} \mathrm{d}t$ for $j \in \{1 : (q_i - 1)\}$ that:

$$\begin{aligned}
&\int_0^T g_i''(\bar{y}_t)(z_t, z_t)\eta_t^0 \mathrm{d}t - \sum_{\tau \in \mathcal{T}} \sum_{j=1}^{q_i} [\eta_\tau^{i,j}] g_{i,yy}^{(j-1)}(\bar{y}_\tau)(z_\tau, z_\tau) = \\
&\int_0^T D^2 g_i^{(q_i)}(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2 \eta_t^{i,q_i} \mathrm{d}t - \sum_{j=1}^{q_i} \int_0^T g_{i,y}^{(j-1)}(\bar{y}_\tau) f''(\bar{u}_t, \bar{y}_t)(v_t, z_t)^2 \eta_t^{i,j} \mathrm{d}t.
\end{aligned}$$

Substituting into (118) gives:

$$\Delta_i = \sum_{\tau \in \mathcal{T}} \left(\nu_\tau^i g_i''(\bar{y}_\tau)(z_\tau, z_\tau) + \sum_{j=1}^{q_i} \left(([\eta_\tau^{i,j}] - \nu_\tau^{i,j}) g_{i,yy}^{(j-1)} f''(\bar{u}_t, \bar{y}_t)(v_t, z_t) \right) \right),$$

implying $\Delta_i = 0$ in view of (45), as was to be proved. ∎

Note that the proof of Lemma 6 is similar to the one given for a scalar state constraint in Bonnans and Hermant (2007), combined with the junction conditions on the classical and alternative costate.

5. Well-posedness of the shooting algorithm

We recall that the shooting equations were defined as (48)-(61). We will now check that, under suitable assumptions, the shooting equations have an invertible Jacobian. Therefore, using Newton's method, we can (provided the starting point is close enough to the solution) compute its solution with a very high accuracy. In this sense the algorithm is well-posed.

We consider a Pontryagin extremal $(\bar{u},\bar{y},\bar{p},\bar{\eta})$ with finitely many junction points, and the associated alternative costate $\bar{p}^q$ and multiplier $\bar{\eta}^q$. We denote, e.g., by $g_i(\bar{u},\bar{y})(\mathcal{T}^i_{en})$, the vector in $I\!R^{N^i_b}$ of components $g_i(\bar{u}_\tau,\bar{y}_\tau)$, for $\tau\in\mathcal{T}^i_{en}$. By $g^{(0:q_i-1)}_{i,y}(y)z(\mathcal{T}^i_{en})$ we denote the vector in $I\!R^{qN^i_b}$ of components $g^{(j)}_{i,y}(\bar{y}_\tau)z_\tau$, $0\le j\le q_i-1$, $\tau\in\mathcal{T}^i_{en}$ (ordered in a convenient way).

As the vector of shooting parameters we choose

$$\theta := (p_0,\nu_{\mathcal{T}_{en}},\nu_{\mathcal{T}_{to}},\mathcal{T}_{en},\mathcal{T}_{ex},\mathcal{T}_{to}). \tag{122}$$

Here, p_0 is the initial value of the adjoint equation, $\nu_{\mathcal{T}_{en}}$ and $\nu_{\mathcal{T}_{to}}$ are the multipliers associated with the entry and touch conditions, respectively, defined in (116), and $\mathcal{T}_{en}$, $\mathcal{T}_{ex}$, and $\mathcal{T}_{to}$ are the junction times. These parameters define uniquely the variables (u,y,p^q,η^q) as the solution of (48)-(52) (without the bars on variables) as well as the junction conditions for the costate (56)-(58).

With the above notations, the shooting mapping $\mathcal{F}$ is defined over a neighborhood in Θ of shooting parameters, associated with a regular Pontryagin extremal, into Θ, by:

$$\theta=\begin{pmatrix} p_0^\top\\ \nu_{\mathcal{T}_{en}}\\ \nu_{\mathcal{T}_{to}}\\ \mathcal{T}_{en}\\ \mathcal{T}_{ex}\\ \mathcal{T}_{to}\end{pmatrix}\mapsto\begin{pmatrix} p^q_T-\phi'(y_T) & \\ g_i^{\{0:(q_i-1)\}}(y(\mathcal{T}^i_{en})), & i\in\{1:r\}\\ g_i(y(\mathcal{T}^i_{to})), & i\in\{1:r\}\\ G_i(u(\mathcal{T}^{i-}_{en}),y(\mathcal{T}^i_{en})), & i\in\{1:r\}\\ G_i(u(\mathcal{T}^{i+}_{ex}),y(\mathcal{T}^i_{ex})), & i\in\{1:r\}\\ g_i^{(1)}(y(\mathcal{T}^i_{to})), & i\in\{1:r\}\end{pmatrix}. \tag{123}$$

By definition, a zero of the shooting mapping $\mathcal{F}$ provides a trajectory (u,y) that is a shooting extremal.

Let $(\bar{u},\bar{y},\bar{p},\bar{\eta})$ be a Pontryagin extremal such that $\bar{u}$ is continuous, and for which there are finitely many junction points, and such that touch points associated with state constraints of order $q_i>1$ are reducible. The associated shooting vector $\bar{\theta}$, element of the vector space Θ, satisfies $\mathcal{F}(\bar{\theta})=0$. It is easily checked that, in a neighborhood Θ_0 of $\bar{\theta}$, the shooting mapping $\mathcal{F}$ is well-defined and of class C^∞. Its directional derivative $\mathcal{M}:=D\mathcal{F}(\bar{\theta})\omega$ in a direction

$$\omega := (\pi_0,\gamma_{\mathcal{T}_{en}},\gamma_{\mathcal{T}_{to}},\sigma_{\mathcal{T}_{en}},\sigma_{\mathcal{T}_{ex}},\sigma_{\mathcal{T}_{to}})\in\Theta, \tag{124}$$

can be split into $\mathcal{M}=\begin{pmatrix}\mathcal{M}_\mathcal{Q}\\ \mathcal{M}_\mathcal{T}\end{pmatrix}$ given by :

$$\mathcal{M}_\mathcal{Q} := \begin{pmatrix} \pi_T-\phi''(\bar{y}_T)z_T & \\ g^{[0:(q_i-1)]}_{i,y}(\bar{y})z(\mathcal{T}^i_{en}), & i\in\{1:r\}\\ g_i'(\bar{y})z(\mathcal{T}^i_{to}), & i\in\{1:r\}\end{pmatrix} \tag{125}$$

$$\mathcal{M}_{\mathcal{T}} := \begin{pmatrix} G_i'(\bar u,\bar y)(v,z)(\mathcal{T}_{en}^{i-}) + \sigma_{\mathcal{T}_{en}^i} g^{(q_i+1)}(\bar u;\bar y)|_{t=\mathcal{T}_{en}^{i-}} \\ G_i'(\bar u,\bar y)(v,z)(\mathcal{T}_{ex}^{i+}) + \sigma_{\mathcal{T}_{ex}^i} g^{(q_i+1)}(\bar u,\bar y)|_{t=\mathcal{T}_{ex}^{i+}} \\ g_{i,y}^{(1)}(\bar y)z(\mathcal{T}_{to}^i) + \sigma_{\mathcal{T}_{to}^i} g^{(2)}(\bar u,\bar y)|_{t=\mathcal{T}_{to}^i} \end{pmatrix}. \tag{126}$$

In this expression, (v,z,π,ζ) represent the linearized control, state, costate and state constraint multiplier, which are the solutions of the following equations, where the arguments $(\bar u,\bar y,\bar p^q,\bar\eta^q)$ and t are omitted:

$$\dot z = f_y z + f_u v \quad \text{on } [0,T]\ ; \qquad z_0 = 0, \tag{127}$$
$$-\dot\pi = H_{yy}^q z + H_{yu}^q v + \pi f_y + \zeta G_y \quad \text{on } [0,T]\setminus\mathcal{T} \tag{128}$$
$$0 = H_{uy}^q z + H_{uu}^q v + \pi f_u + \zeta G_u \quad \text{a.e. on } [0,T] \tag{129}$$
$$0 = G_i'(\bar u_t,\bar y_t)(v,z) \quad \text{a.e. on } \mathcal{I}_b^i,\ i\in\{1:r\} \tag{130}$$
$$0 = \zeta_i \quad \text{on } [0,T]\setminus\mathcal{I}_b^i,\ i\in\{1:r\}. \tag{131}$$

The linearization of jump conditions on the costate being not obvious, we derive them in the following lemma:

Lemma 7 *The jump conditions on the linearized costate π are given by, for all $i\in\{1:r\}$:*

$$\begin{aligned} -[\pi_\tau] &= \sum_{j=1}^{q_i}\left(\nu_\tau^{i,j} z_\tau^\top g_{i,yy}^{(j-1)}(\bar y_\tau) - \gamma_\tau^{i,j} g_{i,y}^{(j-1)}(\bar y_\tau)\right) \\ &\qquad + \sigma_\tau \sum_{j=1}^{q_i-1} \nu_\tau^{i,j} g_{i,y}^{(j)}(\bar y_\tau),\ \tau\in\mathcal{T}_{en}^i, \end{aligned} \tag{132}$$

$$-[\pi_\tau] = 0,\ \tau\in\mathcal{T}_{ex}^i, \tag{133}$$

$$-[\pi_\tau] = \nu_\tau^i z_\tau^\top g_i''(\bar y_\tau) + \gamma_\tau^i g_i'(\bar y_\tau) + \sigma_\tau \nu_\tau^i g_{i,y}^{(1)}(\bar y_\tau),\ \tau\in\mathcal{T}_{to}^i. \tag{134}$$

Proof. a) We first recall the formula for the sensitivity of a jump of an autonomous piecewise smooth differential system w.r.t. the jump time. Consider the system

$$\dot x_t = F_i(x_t),\ \ t\in[0,T];\quad [x_\tau] = \Phi(x_\tau^-), \tag{135}$$

where $\tau\in(0,T)$ is the switching time, $i=1$ for the first arc $(t<\tau)$ and $i=2$ for the second arc $(t>\tau)$. Let $[F(x_\tau)] := F_2(x_\tau^+) - F_1(x_\tau^-)$. If τ is changed into $\tau+\varepsilon$ with say $\varepsilon>0$, we denote by y the new solution and by χ the derivative w.r.t. τ, we obtain

$$\begin{aligned} y_{\tau+\varepsilon}^- &= x_\tau^- + \varepsilon F_1(x_\tau^-) + o(\varepsilon) \\ \Phi(y_{\tau+\varepsilon}^-) &= \Phi(x_\tau^-) + \varepsilon\Phi'(x_\tau^-)F_1(x_\tau^-) + o(\varepsilon) \\ y_{\tau+\varepsilon}^+ &= x_\tau^- + \varepsilon F_1(x_\tau^-) + \Phi(x_\tau^-) + \varepsilon\Phi'(x_\tau^-)F_1(x_\tau^-) + o(\varepsilon) \\ &= x_\tau^+ + \varepsilon\left(F_1(x_\tau^-) + \Phi'(x_\tau^-)F_1(x_\tau^-)\right) + o(\varepsilon) \\ [y_{\tau+\varepsilon}] &= [x_\tau] + \varepsilon\Phi'(x_\tau^-)F_1(x_\tau^-) + o(\varepsilon). \end{aligned} \tag{136}$$

Having in mind that the jump in χ remains at time τ, it follows that

$$[\chi_\tau] = \Phi'(x_\tau^-)F_1(x_\tau^-) - [F(x_\tau)]. \tag{137}$$

b) Derivation of (132). We linearize (57), taking into account the contribution of the different components of the shooting variables. In this way we obtain that the jump of π at $\tau \in \mathcal{T}_{en}^i$ is given by:

$$-[\pi_\tau] = \sum_{j=1}^{q_i} \nu_\tau^{i,j} z_\tau^\top g_{i,yy}^{(j-1)}(\bar{y}_\tau) + \sum_{j=1}^{q_i} \gamma_\tau^{i,j} g_{i,y}^{(j-1)}(\bar{y}_\tau) - \sigma_\tau \Delta_\tau, \tag{138}$$

where Δ_τ is the sensitivity coefficient on junction time. In view of (137) the latter is the difference of the derivative of the r.h.s. of (58) w.r.t. τ, and of the influence of the jump of the dynamics in the $[\bar{p}_\tau^q]$ (we skip the arguments of f):

$$\begin{aligned}
\Delta_\tau &= -\sum_{j=1}^{q_i} \nu_\tau^{i,j} g_{i,yy}^{(j-1)}(\bar{y}_\tau) f + [H_y^q(\bar{u}_\tau, \bar{y}_\tau, \bar{p}_\tau^q, \bar{\eta}_\tau^q)] \\
&= -\sum_{j=1}^{q_i} \nu_\tau^{i,j} \left(g_{i,yy}^{(j-1)}(\bar{y}_\tau) f + g_{i,y}^{(j-1)}(\bar{y}_\tau) f_y \right) + [\eta^{i,q_i}] G_{i,y}(\bar{u}_\tau, \bar{y}_\tau) \\
&= -\sum_{j=1}^{q_i} \nu_\tau^{i,j} g_{i,y}^{(j)}(\bar{u}_\tau, \bar{y}_\tau) + [\eta^{i,q_i}] G_{i,y}(\bar{u}_\tau, \bar{y}_\tau).
\end{aligned}$$

Since $\nu_\tau^{i,q_i} = [\eta^{i,q_i}]$, by (63)(ii), we obtain (132).

c) Derivation of (133): immediate consequence of step a) since the jump is zero.

d) Derivation of (134). We easily obtain

$$-[\pi_\tau] = \nu_\tau^i z_\tau^\top g_i''(\bar{y}_\tau) + \gamma_\tau^i g_i'(\bar{y}_\tau) - \sigma_\tau \Delta_\tau, \ \tau \in \mathcal{T}_{to}^i, \tag{139}$$

where again Δ_τ is the sensitivity coefficient on junction time, and by (137), we have

$$\Delta_\tau = -\nu_\tau^i g_i''(\bar{y}_\tau) f(\bar{u}_\tau, \bar{y}_\tau) - [H_y^q], \tag{140}$$

and since $[H_y^q] = -\nu_\tau^i g_i'(\bar{y}_\tau) f_y(\bar{u}_\tau, \bar{y}_\tau)$ and $g_{i,y}^{(1)} = g_i'' f + g_i' f_y$, the result follows. ■

COROLLARY 1 *Let $(\bar{u}, \bar{y})$ be a shooting extremal, with $\bar{u}$ continuous, satisfying* (24)-(25). *Assume that the second-order sufficient conditions* (87)-(86) *hold. Then there exists $\alpha > 0$, such that*

$$Q(v) := \mathcal{J}_q(v, z(v)) \geq \alpha \|v\|_2^2, \quad \forall v \in \operatorname{Ker} \mathcal{A}. \tag{141}$$

Proof. This is a consequence of Theorem 3 combined with Lemma 6. ■

The *nontangential conditions* (of appropriate order) at junction points are defined as

$$\text{(i)}\quad g_i^{(q_i+1)}(\bar{u},\bar{y})|_{t=\tau^-} \neq 0, \quad \text{for all } \tau \in \mathcal{T}_{en}^i \text{ and } i \in \{1:r\}, \tag{142}$$

$$\text{(ii)}\quad g_i^{(q_i+1)}(\bar{u},\bar{y})|_{t=\tau^+} \neq 0, \quad \text{for all } \tau \in \mathcal{T}_{ex}^i \text{ and } i \in \{1:r\}, \tag{143}$$

$$\text{(iii)}\quad g_i^{(2)}(\bar{u},\bar{y})|_{t=\tau} \neq 0, \quad \text{for all } \tau \in \mathcal{T}_{to}^i \text{ and } i \in \{1:r\}. \tag{144}$$

Consider the linear equation with unknown ω parameterized as in (124):

$$D\mathcal{F}(\bar{\theta})\omega = \delta, \quad \delta := (a_T, b_{\mathcal{T}_{en}}, b_{\mathcal{T}_{to}}, c_{\mathcal{T}_{en}}, c_{\mathcal{T}_{ex}}, c_{\mathcal{T}_{to}}). \tag{145}$$

We will see that ω is closely related to the solutions of the quadratic optimal control problem (in which $c_{\mathcal{T}_{to}}$ appears, but not $c_{\mathcal{T}_{en}}$ and $c_{\mathcal{T}_{ex}}$):

$$(\mathcal{P}^\delta) \quad \begin{cases} \displaystyle\operatorname*{Min}_{v\in\mathcal{V}} \quad \tfrac{1}{2}\mathcal{J}_q(v,z) \;+\; a_T\cdot z_T \;+\; \sum_{i=1}^{r}\sum_{\tau\in\mathcal{T}_{to}^i} c_\tau \nu_\tau^i \frac{g_{i,y}^{(1)}(\bar{y}_\tau)z_\tau}{g_i^{(2)}(\bar{u}_\tau,\bar{y}_\tau)}, \\ \text{subject to} \quad (74) \text{ and } \mathcal{A}v = (0_{\prod_{i=1}^r L^2(\mathcal{I}_b^i)}, b_{\mathcal{T}_{en}}, b_{\mathcal{T}_{to}})^\top. \end{cases}$$

Here is our main result:

THEOREM 4 *Let $(\bar{u},\bar{y})$ be a shooting extremal with $\bar{u}$ continuous, satisfying* (24)-(25)*. Denote by $\bar{\theta}\in\Theta$ the corresponding vector of shooting parameters. Assume that:* (i) *The second-order sufficient conditions* (87)-(86) *hold.* (ii) *The nontangential conditions* (142)-(144) *hold.*

Then the Jacobian $D\mathcal{F}(\bar{\theta})$ of the shooting mapping is invertible, and the (unique) solution ω of (145) *is as follows. With the notations of Lemma 4, denote by (v^δ, w^δ) with $w^\delta = (\zeta^\delta, \lambda_{\delta,\mathcal{T}_{en}}, \lambda_{\delta,\mathcal{T}_{to}})$ the unique solution in $\mathcal{V}\times W$ of the first-order optimality system of the problem $(\mathcal{P}^\delta)$.*

Then: $\pi_0 = \pi_0^\delta$, where π^δ is the solution on $[0,T]\backslash\mathcal{T}$ of (128) *with $(v^\delta,\zeta^\delta,z^\delta)$, final and jump conditions of π^δ being given by:*

$$\pi_T^\delta = (z_T^\delta)^\top \phi''(\bar{y}_T) + a_T^\top, \tag{146}$$

$$-\left[\pi_\tau^\delta\right] = \sum_{j=1}^{q_i} \nu_\tau^{i,j}(z_\tau^\delta)^\top g_{i,yy}^{(j-1)}(\bar{y}_\tau) + \sum_{j=1}^{q_i} \lambda_{\delta,\tau}^{i,j} g_{i,y}^{(j-1)}(\bar{y}_\tau), \;\; \tau\in\mathcal{T}_{en}^i, \tag{147}$$

$$-\left[\pi_\tau^\delta\right] = 0, \;\; \tau\in\mathcal{T}_{ex}^i, \tag{148}$$

$$\begin{aligned} -\left[\pi_\tau^\delta\right] = {} & \nu_\tau^i (z_\tau^\delta)^\top g_i''(\bar{y}_\tau) \;+\; \lambda_{\delta,\tau} g_i'(\bar{y}_\tau) \\ & + \nu_\tau^i g_{i,y}^{(1)}(\bar{y}_\tau) \frac{c_\tau - g_{i,y}^{(1)}(\bar{y}_\tau) z_\tau^\delta}{g_i^{(2)}(\bar{u}_\tau,\bar{y}_\tau)}, \;\; \tau\in\mathcal{T}_{to}^i. \end{aligned} \tag{149}$$

The variations of junction times are given by

$$\sigma_\tau = \frac{c_\tau - g_{i,y}^{(1)}(\bar{y}_\tau) z_\tau^\delta}{g_i^{(2)}(\bar{u}_\tau, \bar{y}_\tau)}, \quad \tau \in \mathcal{T}_{to}^i, \tag{150}$$

$$\sigma_\tau = \frac{c_\tau - G_i'(\bar{u}_\tau, \bar{y}_\tau)(v_\tau^{\delta,+}), z_\tau^\delta)}{\frac{\mathrm{d}}{\mathrm{d}t} G_i(\times\bar{u}, \bar{y})|_{t=\tau^+}}, \quad \tau \in \mathcal{T}_{ex}^i, \tag{151}$$

$$\sigma_\tau = \frac{c_\tau - G_i'(\bar{u}_\tau, \bar{y}_\tau)(v_\tau^{\delta,-}, z_\tau^\delta)}{\frac{\mathrm{d}}{\mathrm{d}t} G_i(\bar{u}, \bar{y})|_{t=\tau^-}}, \quad \tau \in \mathcal{T}_{en}^i, \tag{152}$$

and the following relations hold:

$$\gamma_\tau = \lambda_{\delta,\tau}, \quad \tau \in \mathcal{T}_{to}, \tag{153}$$

$$\gamma_\tau^1 = \lambda_{\delta,\tau}^1, \quad \gamma_\tau^j = \lambda_{\delta,\tau}^j - \nu_\tau^{i,j-1}\sigma_\tau, \quad j \in [2:q_i], \quad \tau \in \mathcal{T}_{en}^i. \tag{154}$$

Note that, since the functions of time $(v^\delta, \zeta^\delta, z_\delta, \pi^\delta)$ satisfy (127)-(131), by Proposition 1, they are of class C^∞ on $[0,T] \setminus \mathcal{T}$, and v^δ has limits when $t \to \tau^-$ and $t \to \tau^+$, for τ in, respectively, $\mathcal{T}_{en}$ and $\mathcal{T}_{ex}$. Therefore, (151)-(152) make sense.

Proof. Let $\delta \in \Theta$. By Theorem 3 and Lemma 4, the first-order optimality system of $(\mathcal{P}^\delta)$ has a unique solution and multipliers. One can easily check that (127)-(131) and (146)-(149) with

$$\begin{cases} g_{i,y}^{(j)}(\bar{y}_\tau) z_\tau^\delta = b_\tau^j, \ \tau \in \mathcal{T}_{en}^i, & i \in \{1:r\}, \ j \in \{1:(q_i-1)\}, \\ g_i'(\bar{y}_\tau) z_\tau^\delta = b_\tau, \ \tau \in \mathcal{T}_{to}^i, & i \in \{1:r\}, \end{cases} \tag{155}$$

constitute the first-order optimality system of $(\mathcal{P}^\delta)$, with $\lambda_{\delta,\mathcal{T}_{en}}$ and $\lambda_{\delta,\mathcal{T}_{to}}$ multipliers associated with (155). Denote by $(v^\delta, z^\delta, \pi^\delta, \zeta^\delta, \lambda_{\delta,\mathcal{T}_{en}}, \lambda_{\delta,\mathcal{T}_{to}})$ the solution.

Define $\sigma_\mathcal{T}$ by (150)-(152). Let $\gamma_{\mathcal{T}_{en}}$ and $\gamma_{\mathcal{T}_{to}}$ be related to $\lambda_{\delta,\mathcal{T}_{en}}$ and $\lambda_{\delta,\mathcal{T}_{to}}$ by (153)-(154). Using (150) and (154) in, respectively, (149) and (147), we find that the system of equations (127)-(131), (132)-(134), (146), (155) and (150)-(152) has a unique solution

$$(v^\delta, z^\delta, \pi^\delta, \zeta^\delta, \gamma_{\mathcal{T}_{en}}, \gamma_{\mathcal{T}_{to}}, \sigma_\mathcal{T}). \tag{156}$$

With Lemma 7, this implies that $D\mathcal{F}(\bar{\theta})\omega = \delta$ iff $\pi_0 = \pi_0^\delta$ and the remaining components of ω are determined by (150)-(154), as was to be proved. ∎

Remark 3 In view of the analysis of junction conditions in Bonnans and Hermant (2009b) and Maurer (1979), conditions (142)-(143) are typically not satisfied for boundary arcs of order greater than two. In that case the shooting algorithm is ill-posed, since the variations of corresponding times cannot be recovered as in (151)-(152). In fact, it is generally believed that boundary arcs of order greater than two are ill-posed. See on this subject Robbins (1980) and Milyutin (2000).

6. Numerical application: collision avoidance

We present here a numerical application of the shooting algorithm for two academic problems involving three state constraints. The latter model the problem of obstacle avoidance for two vehicles, assimilated to material points, with a constraint of keeping a minimum distance between them. The goal is to go from given initial positions to final ones by minimizing a compromise between the final time and the energy spent by the control. It is convenient in the examples to denote as e.g. $y(t)$ the dependence w.r.t. t. The problems under consideration are, for the first order dynamics

$$(\mathcal{P}_1) \quad \begin{cases} \min \displaystyle\int_0^{t_f} \left(1 + \mu \sum_{i=1}^{4} u_i^2\right) \mathrm{d}t \\ \dot{y}_i = u_i, \quad i = 1, \dots, 4, \quad y(0) = y^0, \quad y(t_f) = y^f, \quad g(y) \le 0, \end{cases}$$

and for the second order dynamics

$$(\mathcal{P}_2) \quad \begin{cases} \min \displaystyle\int_0^{t_f} \left(1 + \mu \sum_{i=1}^{4} u_i^2\right) \mathrm{d}t \\ \dot{y}_i = y_{i+4}, \quad \dot{y}_{i+4} = u_i, \quad i = 1, \dots, 4, \quad y(0) = y^0, \quad y(t_f) = y^f, \\ g(y) \le 0, \end{cases}$$

with $\mu > 0$. The Cartesian coordinates of the two vehicles in the plane are given respectively by (y_1, y_2) and (y_3, y_4). The state constraint g has three components: obstacles avoidance (the obstacles are modelled by two parabolas)

$$g_1(y) := -y_1 - b(y_2 - c)^2 - a \le 0, \qquad g_2(y) := y_3 - b(y_4 + c)^2 - a \le 0, \qquad (157)$$

where $a, b > 0$ and c are given parameters, and a minimum distance constraint between the two vehicles:

$$g_3(y) := \rho_{min}^2 - \left((y_1 - y_3)^2 + (y_2 - y_4)^2\right) \le 0, \qquad (158)$$

with $\rho_{min} > 0$. The final time t_f is free. Note that by the well-known change of time $s := t/t_f$, we recover the case of a fixed final time. Due to the constraint on the final state, the final condition in the shooting algorithm $p(t_f) - \phi'(y(t_f)) = 0$ is replaced by the condition $y(t_f) - y^f = 0$.

Remark 4 For the problems under consideration, the structure of the trajectory and initial values for the unknowns variables were guessed, making, if necessary, several tries. Methods that automatically determine the structure are presented in Bonnans and Hermant (2009b) and Hermant (2009b) (when there is, however, only one constraint and one control).

6.1. First order state constraints

We solve problem $(\mathcal{P}_1)$ using the shooting algorithm for the parameters

$$\mu = 0.5, \qquad a = 0.3, \qquad b = 0.7, \qquad c = -1, \qquad \rho_{min} = 1, \tag{159}$$

and initial and final conditions given by

$$y^0 = (-1,\ -4,\ 0.5,\ -4.5)^\top, \qquad y^f = (-0.5,\ 4,\ 1,\ 4)^\top.$$

Each component of the state constraint is active on a single boundary arc. The structure of the trajectory, composed by seven arcs, is given more precisely in Table 1, where the junction times are given at the beginning of arcs, and jump parameters are given at entry times.

Table 1. Structure of problem $(\mathcal{P}_1)$.

Arc	1	2	3	4	5	6	7
Active(s) constraint(s)	no	1	1, 3	3	2, 3	2	no
Junction time	0	2.82	2.95	3.07	5.49	5.62	5.71
Jump parameter	-	0.42	0.12	-	0.47	-	-

The solution is plotted in Fig. 1. The initial costate and final time are

$$\begin{aligned} p_0 &= \begin{pmatrix} -0.2412 \\ -1.0041 \\ 0.0040 \\ -0.9662 \end{pmatrix}, \\ t_f &= 8.3605, \end{aligned}$$

and the junction times and jump parameters are given in Table 1.

The three components of the alternative state constraint multiplier η_i^1, $i = 1$ to 3, are plotted along their respective boundary arcs. We check that the latter are decreasing, and hence the condition $\dot\eta = -\dot\eta^1 \geq 0$ of the minimum principle is satisfied. The trajectories of the two vehicles in the plane and the two obstacles given by (157), as well as the distance

$$d := \sqrt{(y_1 - y_3)^2 + (y_2 - y_4)^2} \tag{160}$$

between the two vehicles, are plotted in Fig. 2.

6.2. Second order state constraints

We solve problem $(\mathcal{P}_2)$ using the shooting algorithm for parameters given by (159) and initial and final conditions as follows:

$$\begin{aligned} y^0 &= (-1,\ -4,\ 0.5,\ -4.5,\ -0.6,\ 0.8,\ -0.5,\ 0.85)^\top, \\ y^f &= (-0.5,\ 4,\ 1,\ 4,\ 0,\ 1,\ 0,\ 1)^\top. \end{aligned}$$

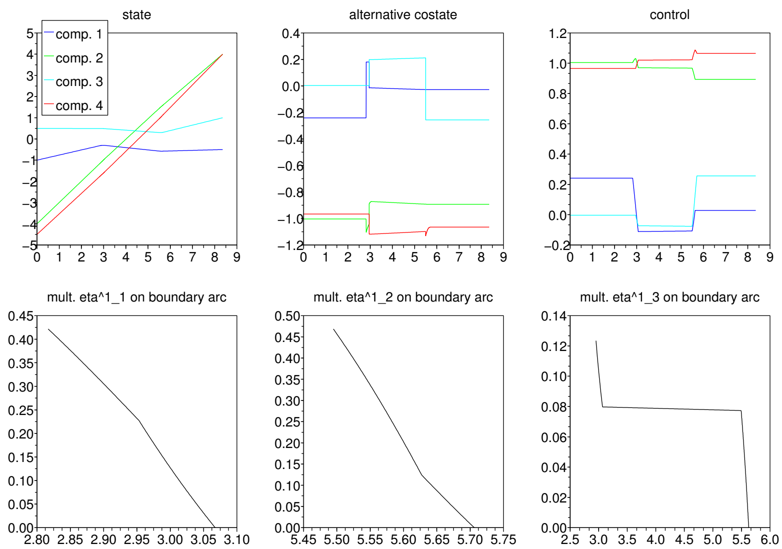

Figure 1. State, alternative costate, control, and alternative state constraint multiplier η^1 on boundary arcs in function of time for problem $(\mathcal{P}_1)$.

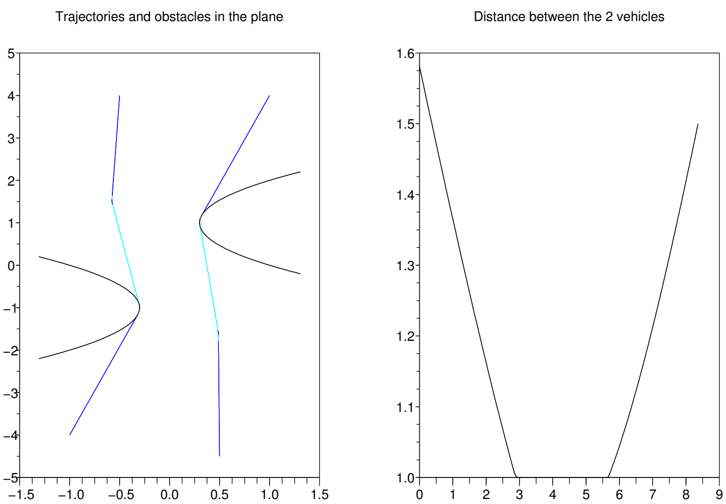

Figure 2. Trajectories and obstacles in the plane and distance between the two vehicles in function of time for problem $(\mathcal{P}_1)$.

.

Each of the two first state constraints are active at a single touch point, while the third is active at two touch points. The structure of the trajectory, composed by five arcs separated by four touch points (t.p.), is given more precisely in Table 2. We check that the jumps parameters of the costate at touch points are nonnegative.

Table 2. Structure of problem ($\mathcal{P}_2$).

Arc/t.p.	1	t.p.	2	t.p.	3	t.p.	4	t.p.	5
Active constraint	no	1	no	3	no	3	no	2	no
Junction time	-	2.82	-	2.93	-	3.93	-	4.06	-
Jump parameter	-	4.09	-	0.96	-	0.83	-	3.75	-

The solution is plotted in Fig. 3. The initial costate and final time are

$$\begin{aligned} p_0 &= \begin{pmatrix} -0.8916 \\ -0.0574 \\ -0.5752 \\ -0.1773 \\ -1.4389 \\ -0.2531 \\ -0.9448 \\ -0.4735 \end{pmatrix}, \\ t_f &= 6.4730, \end{aligned}$$

and the values of junction times and jump parameters are given in Table 2.

The trajectories of the two vehicles in the plane and the two obstacles given by (157), as well as the distance (160) between the two vehicles, are plotted in Fig. 4. A zoom is needed on the latter to see the two isolated contact points of the third constraint, given by (158).

References

Bonnans, J.F. and Hermant, A. (2007) Well-posedness of the shooting algorithm for state constrained optimal control problems with a single constraint and control. *SIAM J. Control Optimization* **46** (4), 1398–1430.

Bonnans, J.F. and Hermant, A. (2008) Stability and sensitivity analysis for optimal control problems with a first-order state constraint. *ESAIM: COCV* **14** (4), 825–863.

Bonnans, J.F. and Hermant, A. (2009a) No gap second order optimality conditions for optimal control problems with a single state constraint and control. *Mathematical Programming, Series B*, **117**, 21–50.

Bonnans, J.F. and Hermant, A. (2009b) Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints. *Annals of I.H.P. - Nonlinear Analysis* **26**, 561–598.

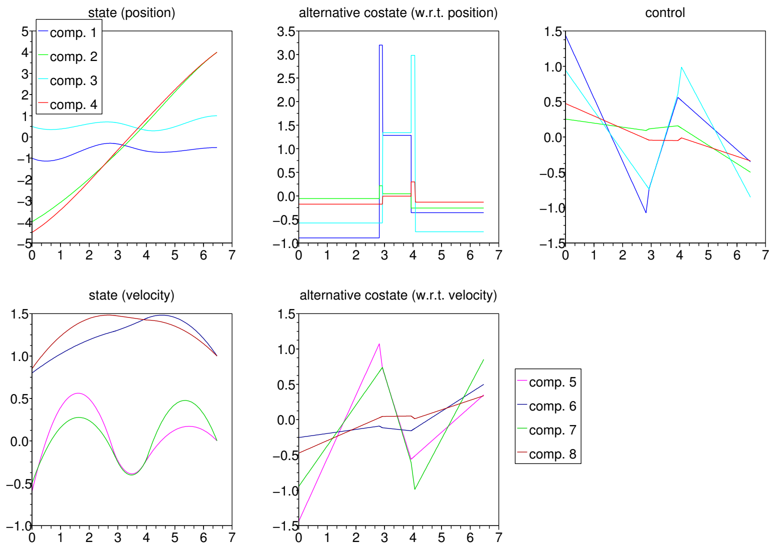

Figure 3. State, alternative costate, and control in function of time for problem $(\mathcal{P}_2)$.

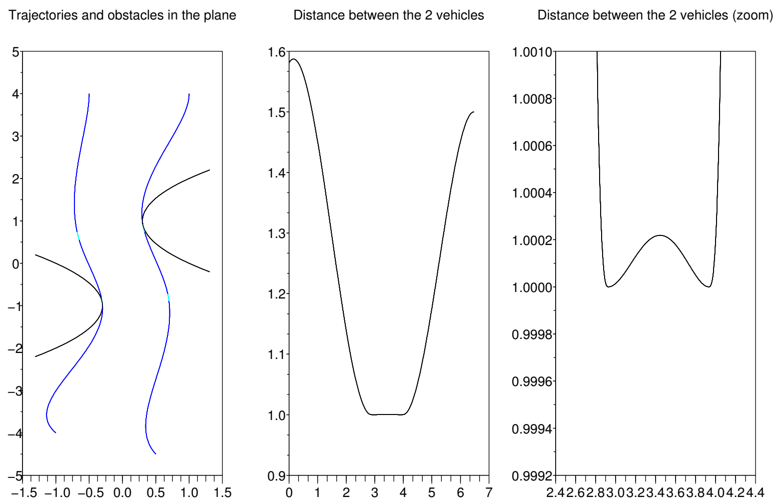

Figure 4. Trajectories and obstacles in the plane and distance between the two vehicles in function of time for problem $(\mathcal{P}_2)$.

Bonnans, J.F. and Shapiro, A. (2000) *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York.

Bryson, A.E., Denham, W.F. and Dreyfus, S.E. (1963) Optimal programming problems with inequality constraints I: necessary conditions for extremal solutions. *AIAA Journal* **1**, 2544–2550.

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

Hermant, A. (2008) Optimal control of the atmospheric reentry of a space shuttle by an homotopy method. Rapport de Recherche RR 6627, INRIA.

Hermant, A. (2009a) Stability analysis of optimal control problems with a second order state constraint. *SIAM J. Optimization* **20** (1), 104–129.

Hermant, A. (2009b) Homotopy algorithm for optimal control problems with a second-order state constraint. *Applied Mathematics and Optimization*, DOI:10.1007/s00245-009-9076-y.

Ioffe, A.D., and Tihomirov, V.M. (1979) *Theory of Extremal Problems*. North-Holland Publishing Company, Amsterdam. Russian Edition: Nauka, Moscow, 1974.

Jacobson, D.H., Lele, M.M. and Speyer, J.L. (1971) New necessary conditions of optimality for control problems with state-variable inequality contraints. *J. of Mathematical Analysis and Applications* **35**, 255–284.

Malanowski, K. (2007) Stability analysis for nonlinear optimal control problems subject to state constraints. *SIAM J. Optim.* **18** (3), 926–945 (electronic).

Malanowski, K. and Maurer, H. (1998) Sensitivity analysis for state constrained optimal control problems. *Discrete and Continuous Dynamical Systems* **4**, 241–272.

Malanowski, K. and Maurer, H. (2001) Sensitivity analysis for optimal control problems subject to higher order state constraints. *Ann. Oper. Res.* **101**, 43–73. Optimization with data perturbations, II.

Maurer, H. (1979) On the minimum principle for optimal control problems with state constraints. Schriftenreihe des Rechenzentrum 41, Universität Münster.

Milyutin, A.A. (2000) On a certain family of optimal control problems with phase constraint. *J. Math. Sci. (New York)*, **100** (5), 2564–2571. Pontryagin Conference, 1, Optimal Control (Moscow, 1998).

Opfer, G. and Oberle, H.J. (1988) The derivation of cubic splines with obstacles by methods of optimization and optimal control. *Numerische Mathematik* **52**, 17–31.

Robbins, H.M. (1980) Junction phenomena for optimal control with state-variable inequality constraints of third order. *J. of Optimization Theory and Applications* **31**, 85–99.