SIAM J. CONTROL OPTIM.
Vol. 49, No. 4, pp. 1498–1522

CONVERGENCE RESULTS FOR SMOOTH REGULARIZATIONS OF HYBRID NONLINEAR OPTIMAL CONTROL PROBLEMS*

T. HABERKORN† AND E. TRÉLAT†

Abstract. We consider a class of hybrid nonlinear optimal control problems having a discontinuous dynamics ruled by a partition of the state space. For this class of problems, some hybrid versions of the usual Pontryagin Maximum Principle are known. We introduce general regularization procedures, parameterized by a small parameter, smoothing the previous hybrid problems to standard smooth optimal control problems, for which we can apply the usual Pontryagin Maximum Principle. We investigate the question of the convergence of the resulting extremals as the regularization parameter tends to zero. Under some general assumptions, we prove that smoothing regularization procedures converge, in the sense that the solution of the regularized problem (as well as its extremal lift) converges to the solution of the initial hybrid problem. To illustrate our convergence result, we apply our approach to the minimal time low-thrust coplanar orbit transfer with eclipse constraint.

Key words. optimal control, hybrid control, regularization, shooting method, orbit transfer with eclipse

AMS subject classifications. 49J15, 34A38, 93C65

DOI. 10.1137/100809209

1. Introduction and main results.

1.1. Hybrid optimal control.

Let m and n be positive integers. In what follows, $t \in \mathbb{R}$ denotes the time variable, and we consider a time dependent partition of $\mathbb{R}^n$, such that $\mathbb{R}^n = \bigcup_{\alpha\in\mathcal{A}} \overline{X_\alpha(t)}$, where $\mathcal{A}$ is a countable set, and the subsets $X_\alpha(t)$ are disjoint and open with a piecewise C^1 boundary. Let Ω be a measurable subset of $\mathbb{R}^m$. For every $\alpha \in \mathcal{A}$, let $f_\alpha : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^n$ and $f^0_\alpha : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}$ be continuous mappings that are C^1 with respect to their second variable. For every $t \in \mathbb{R}$, every $x \in \mathbb{R}^n$, and every $u \in \Omega$, define $f(t, x, u) = f_\alpha(t, x, u)$ and $f^0(t, x, u) = f^0_\alpha(t, x, u)$ whenever $x \in X_\alpha(t)$. Consider the hybrid control system

$$\dot{x}(t) = f(t, x(t), u(t)), \tag{1.1}$$

where the control $u(\cdot)$ belongs to the class of measurable functions with values in Ω.

Let M_0 and M_1 be two compact subsets of $\mathbb{R}^n$. Assume that M_1 is reachable from M_0 for the control system (1.1), in the sense that there exist a time $t_f > 0$ and a control $u(\cdot) \in L^\infty([0, t_f], \Omega)$, such that the trajectory $x(\cdot)$, which is a solution of (1.1) with $x(0) \in M_0$, satisfies $x(t_f) \in M_1$. Consider the optimal control problem (HOCP) of steering the control system (1.1) from M_0 to M_1, and minimizing the cost function

$$C(t_f, u(\cdot)) = \int_0^{t_f} f^0(t, x(t), u(t))dt. \tag{1.2}$$

The final time t_f may or may not be fixed.

---

*Received by the editors September 20, 2010; accepted for publication (in revised form) March 28, 2011; published electronically July 5, 2011.
http://www.siam.org/journals/sicon/49-4/80920.html

†Université d'Orléans, Labo. MAPMO, CNRS, UMR 6628, Fédération Denis Poisson, FR 2964, Bat. Math., BP 6759, 45067 Orléans cedex 2, France (thomas.haberkorn@univ-orleans.fr, emmanuel.trelat@univ-orleans.fr).

In [8] this kind of hybrid optimal control problem is called an optimal control problem on stratified domains, and existence of optimal control and Cauchy uniqueness results are derived using a suitable modification of the usual Fillippov's arguments so as to handle the discontinuities of the dynamics and of the cost function. Note that, in that reference, the definition of stratified problem requires us to define the dynamics restricted to the boundary of some domain; however, we do not focus on that point here since we assume a transversal crossing throughout our article (see Definition 1.1). Another slight difference between the framework of [8] and ours is that our decomposition of $\mathbb{R}^n$ is assumed to be time dependent and our time horizon is finite; however, the existence and uniqueness results of [8] are easily extended to our context.

From now on, assume that (HOCP) has a solution $(x(\cdot), u(\cdot))$ defined on $[0, t_f]$, with $u(\cdot) \in L^\infty([0, t_f], \Omega)$.

For usual (smooth) optimal control problems, a well-known numerical method for computing the optimal trajectory, called the shooting method, consists of combining the necessary conditions derived from the Pontryagin Maximum Principle (PMP; see [22]) with a Newton method (see, e.g., [30]). Let us recall how this approach has been generalized to the hybrid framework. First, the PMP has been extended to a very general hybrid context in many references (see [7, 14, 16, 23, 27, 31] and references therein) with different proof approaches and presentations. However, it is not our aim to consider hybrid control systems in their full generality, and our (HOCP) is a specific hybrid optimal control problem in the sense that the state and control spaces do not vary and the change of dynamics (1.1) is ruled by the state position and is not directly controlled. In our case, all the versions of the Hybrid Maximum Principle (HMP) derived in the aforementioned references are equivalent, and we recall hereafter a statement of the HMP applied to the optimal solution $(x(\cdot), u(\cdot))$ of (HOCP).

According to the HMP, there exist $p^0 \leqslant 0$ and a piecewise absolutely continuous mapping $p(\cdot) : [0, t_f] \to \mathbb{R}^n$ called an *adjoint vector*, with $(p(\cdot), p^0) \neq (0, 0)$, such that the so-called *extremal* $(x(\cdot), p(\cdot), p^0, u(\cdot))$ is a solution of

$$\dot{x}(t) = \frac{\partial H}{\partial p}(t, x(t), p(t), p^0, u(t)), \quad \dot{p}(t) = -\frac{\partial H}{\partial x}(t, x(t), p(t), p^0, u(t)) \tag{1.3}$$

for almost every $t \in [0, t_f]$, where the Hamiltonian is defined by $H(t, x, p, p^0, u) = \langle p, f(t, x, u)\rangle + p^0 f^0(t, x, u)$, and the maximization condition

$$H(t, x(t), p(t), p^0, u(t)) = \max_{v \in \Omega} H(t, x(t), p(t), p^0, v) \tag{1.4}$$

holds almost everywhere on $[0, t_f]$. Moreover, if the final time t_f is free, then there holds $\max_{v\in\Omega} H(t_f, x(t_f), p(t_f), p^0, v) = 0$. In addition, for every time t_c such that the trajectory $x(\cdot)$ passes from the domain $X_\alpha(t_c)$ to some domain $X_\beta(t_c)$, a certain jump condition holds for the adjoint vector at the time t_c under some appropriate transversality condition. To make it explicit, we introduce hereafter the notion of a *regular crossing time*.

DEFINITION 1.1. *A regular crossing time of the trajectory $x(\cdot)$, which is a solution of* (1.1) *and associated with the control $u(\cdot)$, is a time t_c satisfying the following assumptions:*

1. *There exist exactly two elements α and β of $\mathcal{A}$ such that the point $x(t_c)$ belongs to the closure of the domains $X_\alpha(t_c)$ and $X_\beta(t_c)$;*
2. *there exists $\eta > 0$ such that $x(t) \in X_\alpha(t)$ for $t \in (t_c - \eta, t_c)$ and $x(t) \in X_\beta(t)$ for $t \in (t_c, t_c + \eta)$;*

3. *the boundary between the domains* $X_\alpha(t)$ *and* $X_\beta(t)$ *can be written as* $\{x \in \mathbb{R}^n \mid F(t,x) = 0\}$ *in a neighborhood of* $x(t_c)$ *and for* t *close to* t_c*, with a function* $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ *of class* C^1*;*
4. *the control* $u(\cdot)$ *is left- and right-continuous at* t_c*;*
5. $\langle \partial_x F(t_c, x(t_c)), f_\alpha(t_c, x(t_c), u(t_c^-))\rangle + \partial_t F(t_c, x(t_c)) \neq 0$*;*
6. $\langle \partial_x F(t_c, x(t_c)), f_\beta(t_c, x(t_c), u(t_c^+))\rangle + \partial_t F(t_c, x(t_c)) \neq 0$.

The two last items of this definition represent a transversality crossing condition. The following *jump condition* holds at every regular crossing time:

$$p(t_c^+) = p(t_c^-) + \frac{\langle p(t_c^-), f_\alpha(t_c^-) - f_\beta(t_c^+)\rangle + p^0(f_\alpha^0(t_c^-) - f_\beta^0(t_c^+))}{\langle \partial_x F(t_c, x(t_c)), f_\beta(t_c^+)\rangle + \partial_t F(t_c, x(t_c))} \partial_x F(t_c, x(t_c)). \tag{1.5}$$

Here, the short notation $f_\alpha(t_c^-)$ stands for $f_\alpha(t_c, x(t_c), u(t_c^-))$, and $f_\beta(t_c^+)$ stands for $f_\beta(t_c, x(t_c), u(t_c^+))$. The superscript + (resp., −) denotes the right (resp., left) limit.

Remark 1.2. Definition 1.1 also implies some regularity assumptions of the partition with respect to time t.

Remark 1.3. In the particular case where the partition of $\mathbb{R}^n$ into the domains X_α ($\alpha \in \mathcal{A}$) does not depend on time, the Hamiltonian remains continuous at each boundary crossing, that is, $H(t_c^-) = H(t_c^+)$.

Moreover, in that case, the last two items of Definition 1.1 mean that the left limit $\dot{x}(t_c^-)$ and right limit $\dot{x}(t_c^+)$ are transverse to $\partial X_\alpha(t_c)$, as illustrated in Figure 1.1. This transversality crossing assumption is crucial in our work. It discards the situations where the trajectory $x(\cdot)$ crosses a boundary between three domains or more, or hits tangentially a boundary of a domain.

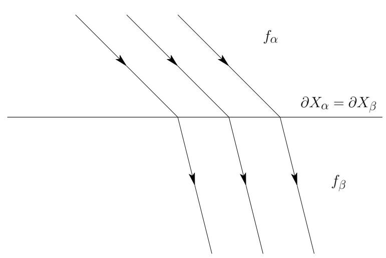

FIG. 1.1. *Transversality crossing assumption.*

The extremal $(x(\cdot), p(\cdot), p^0, u(\cdot))$ is said to be *normal* whenever $p^0 \neq 0$, and it is then usual to normalize the adjoint vector so that $p^0 = -1$; otherwise it is said to be *abnormal*.

As in the smooth case, it is then possible to derive from the HMP a (multiple) shooting method, and we refer the reader to, e.g., [13, 23, 27, 34] for some examples of applications. In those references, however, the examples are academic and either the computations can be made by hand or the control structure (and hence the sequence of dynamics) can be established beforehand. Recall that the shooting methods are usually referred to as indirect methods, since they are based on the preliminary use of theoretical necessary conditions (PMP or HMP). On the opposite, direct numerical methods consist of discretizing directly the optimal control problem so as to reduce it, after discretization, to some finite dimensional nonlinear optimization problem with constraints, with the dimension being as large as the discretization is fine. Direct methods can be developed as well for solving (HOCP) by approximating it by a nonlinear programming problem (see [26] for a survey); however, it is well known that, in general, such an approach cannot yield the degree of accuracy provided by an indirect (shooting) method. The main flaw of the indirect numerical approach is that shooting methods are known to be possibly hard to initialize because they rely on a Newton-like algorithm. In the hybrid context, it will be even harder to initialize because a multiple shooting has to be used to take into account the jump conditions (1.5) of the adjoint vector due to the crossings. Moreover, (1.5) is an implicit equation, because the computation of $f_\beta(t_c^+)$ in the right-hand side requires the computation of $u(t_c^+)$, which in turn requires the computation of $p(t_c^+)$ obtained from (1.4).

The method that we propose in this article consists of regularizing (HOCP) into a smooth optimal control problem, parameterized by $\varepsilon > 0$ and denoted $(\mathrm{OCP})_\varepsilon$, with the idea that, on the one hand, $(\mathrm{OCP})_\varepsilon$ can be expected to be easier to solve by a numerical shooting method than (HOCP), and, on the other hand, nice convergence properties can be expected as ε tends to zero.

The article is structured as follows. General regularization procedures are defined in section 1.2. They are used to smooth the hybrid problem under consideration into a family of usual smooth optimal control problems parameterized by a kind of penalization parameter ε, and for which the PMP can be applied. Our main result (Theorem 1.6), stated in section 1.3, asserts the convergence of the solution of $(\mathrm{OCP})_\varepsilon$ to the solution of (HOCP) (as well as their respective extremal lifts) as ε tends to zero, under appropriate assumptions whose relevance is discussed in a series of remarks. Section 2 is then devoted to the proof of the main result. Finally, in section 3 we illustrate our convergence result with a nonacademic application: the minimal time low-thrust coplanar orbit transfer around the Earth with eclipse constraint.

1.2. Regularization procedure. First, we define a concept of C^1 regularization of the extended hybrid dynamics (f, f^0).

DEFINITION 1.4. *The family $(f^\varepsilon, f^{0\varepsilon})_{\varepsilon>0}$ is called a C^1 regularization of the extended hybrid dynamics (f, f^0) if the mappings $f^\varepsilon : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^n$ and $f^{0\varepsilon} : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}$ are C^1, for every $\varepsilon > 0$, and if the following pointwise convergence properties hold: for all $t \in \mathbb{R}$, $\alpha \in \mathcal{A}$, $x \in X_\alpha(t)$, and $u \in \Omega$,*

$$f^\varepsilon(t,x,u) \underset{\varepsilon\to 0}{\longrightarrow} f_\alpha(t,x,u), \quad \frac{\partial f^\varepsilon}{\partial x}(t,x,u) \underset{\varepsilon\to 0}{\longrightarrow} \frac{\partial f_\alpha}{\partial x}(t,x,u),$$
$$f^{0\varepsilon}(t,x,u) \underset{\varepsilon\to 0}{\longrightarrow} f^0_\alpha(t,x,u), \quad \frac{\partial f^{0\varepsilon}}{\partial x}(t,x,u) \underset{\varepsilon\to 0}{\longrightarrow} \frac{\partial f^0_\alpha}{\partial x}(t,x,u).$$

Example 1.5. Let us provide an example of C^1 regularization. Examples of a regularization inside a domain $X_\alpha(\cdot)$ are classical (for instance, with a convolution). Let $\bar t \in \mathbb{R}$, and let $\bar x$ be a point belonging to the boundary of exactly two domains $X_\alpha(\bar t)$ and $X_\beta(\bar t)$ for some $\alpha, \beta \in \mathcal{A}$. Let $V \subset \mathbb{R}^n$ be a neighborhood of $\bar x$, and let $b^\varepsilon : \mathbb{R}\times V \to [0,1]$ be a C^1 (kind of Heaviside) function such that

$$b^\varepsilon(t,x) = \begin{cases} 1 & \text{if} \quad x \in X_\alpha(t) \text{ and } |t-t_c| + d(x,\Sigma) \geqslant \varepsilon, \\ 0 & \text{if} \quad x \in X_\beta(t) \text{ and } |t-t_c| + d(x,\Sigma) \geqslant \varepsilon, \end{cases}$$

where $d(\cdot,\cdot)$ is a distance in $\mathbb{R}^n$. For all $t \in \mathbb{R}$, $x \in V$, and $u \in \Omega$, define

$$\begin{aligned} f^\varepsilon(t,x,u) &= b^\varepsilon(t,x) f_\alpha(t,x,u) + (1 - b^\varepsilon(t,x)) f_\beta(t,x,u), \\ f^{0\varepsilon}(t,x,u) &= b^\varepsilon(t,x) f^0_\alpha(t,x,u) + (1 - b^\varepsilon(t,x)) f^0_\beta(t,x,u). \end{aligned} \tag{1.6}$$

This yields a local regularization on V. It is then easy to define a similar regularization for points belonging to the boundary of three or more domains. Then, to make it global it suffices to use, for instance, a partition of unity.

Let $(f^\varepsilon, f^{0\varepsilon})_{\varepsilon>0}$ be a C^1 regularization of the extended hybrid dynamics (f, f^0). Fix $R > 0$ sufficiently large. Consider the optimal control problem $(\mathrm{OCP})_\varepsilon$ of steering the control system

$$\dot{x}^\varepsilon(t) = f^\varepsilon(t, x^\varepsilon(t), u^\varepsilon(t)) \tag{1.7}$$

from M_0 to M_1, with controls $u^\varepsilon(\cdot) \in L^\infty([0, t_f^\varepsilon], \Omega)$, under the additional compact state constraint

$$\max_{t\in[0,t_f^\varepsilon]} \|x^\varepsilon(t)\|_{\mathbb{R}^n} \leqslant R, \tag{1.8}$$

and minimizing the cost function

$$C^\varepsilon(t_f^\varepsilon, u^\varepsilon(\cdot)) = \int_0^{t_f^\varepsilon} f^{0\varepsilon}(t, x^\varepsilon(t), u^\varepsilon(t))dt. \tag{1.9}$$

If the final time t_f of (HOCP) is fixed, then we set $t_f^\varepsilon = t_f$. If the final time t_f of (HOCP) is free, then the final time t_f^ε of $(\mathrm{OCP})_\varepsilon$ is free as well; however, for the problem $(\mathrm{OCP})_\varepsilon$ to be well defined we have to bound t_f^ε, and, for instance, we impose that $0 \leqslant t_f^\varepsilon \leqslant t_f + 10$.

The additional constraint (1.8) is necessary to derive an existence result for $(\mathrm{OCP})_\varepsilon$. Adding such a constraint is a classical fact in any penalization procedure. If R is chosen large enough, then, under the assumptions of our main result below, the constraint (1.8) is actually not active, and hence R plays no further role. In particular it does not affect the numerical process resulting from our main result.

For this regularized optimal control problem $(\mathrm{OCP})_\varepsilon$, anticipating the fact that the constraint (1.8) will not (a posteriori) be active, the usual PMP implies that every optimal solution $(x^\varepsilon(\cdot), u^\varepsilon(\cdot))$ defined on $[0, t_f^\varepsilon]$ is the projection of an extremal $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), p^{0\varepsilon}, u^\varepsilon(\cdot))$ solution of

(1.10)
$$\dot{x}^\varepsilon(t) = \frac{\partial H^\varepsilon}{\partial p}(t, x^\varepsilon(t), p^\varepsilon(t), p^{0\varepsilon}, u^\varepsilon(t)), \quad \dot{p}^\varepsilon(t) = -\frac{\partial H^\varepsilon}{\partial x}(t, x^\varepsilon(t), p^\varepsilon(t), p^{0\varepsilon}, u^\varepsilon(t))$$

for almost every $t \in [0, t_f^\varepsilon]$, where $H^\varepsilon(t, x, p, p^0, u) = \langle p, f^\varepsilon(t, x, u)\rangle + p^0 f^{0\varepsilon}(t, x, u)$ is the Hamiltonian, $p^{0\varepsilon}$ is a nonpositive real number, and

$$H^\varepsilon(t, x^\varepsilon(t), p^\varepsilon(t), p^{0\varepsilon}, u^\varepsilon(t)) = \max_{v\in\Omega} H^\varepsilon(t, x^\varepsilon(t), p^\varepsilon(t), p^{0\varepsilon}, v) \tag{1.11}$$

almost everywhere on $[0, t_f^\varepsilon]$. Moreover, if the final time is free, then the maximized Hamiltonian is equal to 0 at t_f^ε.

Convergence properties of the solutions of $(\mathrm{OCP})_\varepsilon$ towards solutions of (HOCP) can be expected under appropriate assumptions. In particular, the convergence of the associated adjoint vectors, although quite difficult to prove, can be expected. This is the object of the next section, containing the main results of our paper.

1.3. Main results. Let $(f^\varepsilon, f^{0\varepsilon})_{\varepsilon>0}$ be a C^1 regularization of the extended hybrid dynamics (f, f^0) of (HOCP). We make the following assumptions:

(H_1) Ω is a compact and convex subset of $\mathbb{R}^m$;

(H_2) the problem (HOCP) has a unique solution $(x(\cdot), u(\cdot))$ defined on $[0, t_f]$;

(H_3) the optimal trajectory $x(\cdot)$ has a unique extremal lift (up to a multiplicative scalar) on every subinterval of $[0, t_f]$, which is, moreover, normal. This solution of HMP is denoted $(x(\cdot), p(\cdot), -1, u(\cdot))$;

(H_4) every time t such that $x(t)$ belongs to the boundary of some domain of the partition of the state space is a regular crossing time;

(H_5) the sets

$$\{(f_\alpha(t,x,u), f_\alpha^0(t,x,u)) \mid u \in \Omega\}, \ \left\{\left(\frac{\partial f_\alpha}{\partial x}(t,x,u), \frac{\partial f_\alpha^0}{\partial x}(t,x,u)\right) \mid u \in \Omega\right\},$$
$$\{(f^\varepsilon(t,x,u), f^{0\varepsilon}(t,x,u)) \mid u \in \Omega\}, \ \left\{\left(\frac{\partial f^\varepsilon}{\partial x}(t,x,u), \frac{\partial f^{0\varepsilon}}{\partial x}(t,x,u)\right) \mid u \in \Omega\right\}$$

are convex, for every $t \in \mathbb{R}$, every $x \in \mathbb{R}^n$, every $\alpha \in \mathcal{A}$, and every $\varepsilon > 0$.

Finally, we make one of the two following assumptions. Assume that either

(H_6) the C^1 regularization is such that every optimal control $u^\varepsilon(\cdot)$ of $(\mathrm{OCP})_\varepsilon$ is continuous for every $\varepsilon > 0$,

or

(H_6') the system is control-affine,[1] and $u(t)$ is an extremal point of Ω for almost every $t \in [0, t_f]$.

THEOREM 1.6. *Under assumptions* (H_1)–(H_5) *and either* (H_6) *or* (H_6'), *there exists* $\varepsilon_0 > 0$ *such that, for every* $\varepsilon \in (0, \varepsilon_0)$, *the problem* $(\mathrm{OCP})_\varepsilon$ *has at least one solution* $(x^\varepsilon(\cdot), u^\varepsilon(\cdot))$ *defined on* $[0, t_f^\varepsilon]$, *every extremal lift of which is normal. Let* $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), -1, u^\varepsilon(\cdot))$ *be such a normal extremal lift. Then, as* ε *tends to* 0,

- t_f^ε *converges to* t_f;
- $x^\varepsilon(\cdot)$ *converges uniformly*[2] *to* $x(\cdot)$;
- $\dot{x}^\varepsilon(\cdot)$ *converges to* $\dot{x}(\cdot)$ *in* L^∞ *for the weak star topology;*[3]
- $p^\varepsilon(\cdot)$ *converges uniformly to* $p(\cdot)$ *on every closed subinterval of* $[0, t_f]$ *that does not contain any crossing time.*

Moreover, under assumption (H_6') *(and thus, in the specific case of a control-affine system),* $u^\varepsilon(\cdot)$ *converges to* $u(\cdot)$ *for the strong topology of* L^1.

Remark 1.7. Assumption (H_6) is strong and depends of course on the regularization process. In particular, bang-bang controls for the regularized optimal control problems are not covered by this assumption. Bang-bang regularized controls are allowed under assumption (H_6'), but then our result must be specified to control-affine systems. Note by the way that, under assumption (H_6), Theorem 1.6 can be applied as well to control-affine systems, and in this case $f_\alpha(t,x,\cdot)$ and $\frac{\partial f_\alpha}{\partial x}(t,x,\cdot)$ are convex in u (however, assumption (H_5) still makes sense since we consider a general nonlinear cost).

Remark 1.8. Note that, for a general control system, we are not able to derive any convergence property of the controls $u^\varepsilon(\cdot)$ without any further assumption. For example, if we assume in addition, first, that for every $\varepsilon > 0$, the maximization condition (1.11) can be made explicit so that $u^\varepsilon(t) = \varphi^\varepsilon(t, x^\varepsilon(t), p^\varepsilon(t))$ for almost every

[1] The control system (1.1) is said to be *control-affine* in the case where the dynamics f is affine with respect to the control variable; that is, for every $\alpha \in \mathcal{A}$ there exist $m+1$ vector fields $f_{\alpha,0}, \ldots, f_{\alpha,m}$ of class C^1 on $\mathbb{R}^n$ such that $f_\alpha(t,x,u) = f_{\alpha,0}(t,x) + \sum_{i=1}^m u_i f_{\alpha,i}(t,x)$, where $u = (u_1, \ldots, u_m)$. Note that no further assumption is made on the cost: under assumption (H_6') the system is control-affine but the cost under consideration is still a general nonlinear cost.

[2] If $t_f^\varepsilon < t_f$, then we consider any continuous extension of $x^\varepsilon(\cdot)$ on $[0, t_f]$. The same remark holds for $\dot{x}^\varepsilon(\cdot)$, $p^\varepsilon(\cdot)$ and $u^\varepsilon(\cdot)$ in the next items.

[3] It means that $\int_0^{t_f} \langle \dot{x}^\varepsilon(t), g(t)\rangle dt \to \int_0^{t_f} \langle \dot{x}(t), g(t)\rangle dt$ as $\varepsilon \to 0$, for every $g(\cdot) \in L^1([0, t_f], \mathbb{R}^n)$, and where $\dot{x}^\varepsilon(\cdot)$ is extended continuously on $[0, t_f]$ if needed.

$t \in [0, t_f^\varepsilon]$; second, that the maximization condition (1.4) can be made explicit so that $u(t) = \varphi(t, x(t), p(t))$ for almost every $t \in [0, t_f]$; and third, that the (measurable) functions $\varphi^\varepsilon(\cdot,\cdot,\cdot)$ and $\varphi(\cdot,\cdot,\cdot)$ are such that $\varphi^\varepsilon(t, x, p)$ converges (pointwise) to $\varphi(t, x, p)$ as ε tends to 0 for almost every $(t, x, p) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$, then it follows immediately from the theorem that $u^\varepsilon(\cdot)$ converges to $u(\cdot)$ as ε tends to 0 for the strong topology of L^1. This is the case, e.g., when considering the minimal time problem or a quadratic criterion for control-affine systems.

Remark 1.9. As mentioned previously, if the final time t_f of (HOCP) is fixed, then $t_f^\varepsilon = t_f$ for every $\varepsilon > 0$.

Remark 1.10. As explained above, we stress the fact that the real number R is necessary to derive an existence result for $(\mathrm{OCP})_\varepsilon$, but plays no further role in the analysis or in the numerical process. Since it is assumed that the solution of (HOCP) is unique, it is possible to choose, for instance, $R = 2\max_{t\in[0,t_f]} \|x(t)\|_{\mathbb{R}^n}$.

Remark 1.11. The assumptions (H_2) and (H_3) on the uniqueness of the solution of (HOCP) and on the uniqueness of its extremal lift are related to the differentiability properties of the value function (see, for instance, [4, 12], and see [10, 24, 25, 29] for results on the size of the set where the value function is differentiable). These assumptions can be weakened as follows. If we replace assumptions (H_2) and (H_3) with the assumption "every extremal lift of every solution of (HOCP) is normal," then the conclusion of Theorem 1.6 still holds, except that the convergence properties must be written in terms of closure points (for the appropriate topologies).

Remark 1.12. It is well known that, since the PMP is only a necessary condition for optimality, the application of a shooting method to $(\mathrm{OCP})_\varepsilon$ may lead only to an extremal solution that is not necessarily optimal. However, we have the following result (the proof of which follows along the lines of the proof of our main result), which is slightly more general than Theorem 1.6. Assume that there is no abnormal extremal solution of the HMP applied to (HOCP). Then, every extremal lift of every solution of $(\mathrm{OCP})_\varepsilon$ is normal for $\varepsilon > 0$ small enough. Moreover, as in Remark 1.11, every closure point of a family of such extremal solutions is a normal extremal solution of (HOCP) (for the evident topologies).

Remark 1.13. Note that, as in [28], we do not assume any second order sufficient condition, and Theorem 1.6 does not result from a sensitivity analysis.

Remark 1.14. Finally, note that Theorem 1.6 can be applied, in particular, for classical (nonhybrid) nonlinear control systems and provides convergence properties for a regularization process of any control system.

2. Proof of the main result. Our strategy of proof is the following. First, we recall the main steps of a proof of the HMP using needle-like variations, which are needed to derive our main result. Notice that there exists a simple proof of the HMP, due to [14], that consists of reducing the problem to a usual optimal control problem with mixed initial and final conditions, and then applying the usual PMP. This simple approach unfortunately is not adapted to our problem since the regularization of the jump conditions of the HMP implies a quite difficult asymptotic study of Dirac-type effects. The proof of the HMP based on needle-like variations is hence more adaptable to our problem but is more intricate than in the usual case. Indeed, when crossing the boundary of a domain, variation vectors have a jump. For the regularized problem, this turns into the difficulty of recovering, at the limit, this jump condition. In the proof of our main result, the derivation of weak convergence properties is easy and standard, using classical convexity arguments. The difficult point is to derive a strong convergence property for the adjoint vector as well as for the control functions.

To this aim, using the geometric interpretation of optimality, we derive convergence results for the Pontryagin cones, within the hybrid optimal control context. The general strategy of the proof is quite similar to the one developed in [28]; however, the proofs of the intermediate results are far more intricate due to the hybrid framework and the nonlinear features of the dynamics.

2.1. Preliminaries; hybrid maximum principle. In this subsection, we recall the main steps of a proof of the HMP using needle-like variations.

Consider (HOCP) and introduce the instantaneous cost function $x^0(\cdot)$, defined on $[0, t_f]$ and is the solution of $\dot{x}^0(t) = f^0(t, x(t), u(t)),\ x^0(0) = 0$, so that the cost $C(t_f, u(\cdot))$ of the initial trajectory $x(\cdot)$ is $C(t_f, u(\cdot)) = x^0(t_f)$. The extended state $\tilde{x} \in \mathbb{R}^{n+1}$ is defined by $\tilde{x} = (x, x^0)$, and the extended dynamics is defined by $\tilde{f}(t, \tilde{x}, u) = (f(t, x, u), f^0(t, x, u))$. Consider the extended hybrid control system in $\mathbb{R}^{n+1}$,

$$\dot{\tilde{x}}(t) = \tilde{f}(t, \tilde{x}(t), u(t)). \tag{2.1}$$

Let $x_0 \in \mathbb{R}^n$; a control function $u(\cdot) \in L^\infty([0, t_f], \mathbb{R}^m)$ is said to be *admissible* on $[0, t_f]$ if the trajectory $\tilde{x}(\cdot)$, which is the solution of (2.1) associated with u and such that $\tilde{x}(0) = \tilde{x}_0 = (x_0, 0)$, is well defined on $[0, t_f]$, and the extended *end-point mapping* $\tilde{E}$ is then defined by $\tilde{E}(\tilde{x}_0, t_f, u(\cdot)) = \tilde{x}(t_f)$. The set of admissible controls on $[0, t_f]$ is denoted $\mathcal{U}_{\tilde{x}_0, t_f, \mathbb{R}^m}$, and the set of admissible controls on $[0, t_f]$ taking their values in Ω is denoted $\mathcal{U}_{\tilde{x}_0, t_f, \Omega}$. The set $\mathcal{U}_{\tilde{x}_0, t_f, \mathbb{R}^m}$, endowed with the standard topology of $L^\infty([0, t_f], \mathbb{R}^m)$, is open.[4] For every $t \geqslant 0$, define the extended *accessible set* $\tilde{A}_\Omega(\tilde{x}_0, t)$ as the image of the mapping $\tilde{E}(\tilde{x}_0, t, \cdot) : \mathcal{U}_{\tilde{x}_0, t, \Omega} \to \mathbb{R}^{n+1}$. Let $(x(\cdot), u(\cdot))$ be a solution of (HOCP) defined on $[0, t_f]$. Then the point $\tilde{x}(t_f)$ belongs to the boundary of the set $\tilde{A}_\Omega(\tilde{x}_0, t_f)$. This geometric property is at the basis of the proof of the Maximum Principle.

We next recall the concepts of needle-like variations and of Pontryagin cone, adapted to the hybrid context, which will be of crucial importance in order to prove our main result, and which also permit us to derive a proof of the HMP.

In what follows, we assume that the optimal trajectory $x(\cdot)$ satisfies the transversality assumption (H_4).

2.1.1. Needle-like variations. Let $t_1 \in [0, t_f)$ and $u_1 \in \Omega$. For $\eta > 0$ such that $t_1 + \eta \leqslant t_f$, the needle-like variation $\pi_1 = \{t_1, \eta, u_1\}$ of the control $u(\cdot)$ is defined by $u_{\pi_1}(t) = u_1$ if $t \in [t_1, t_1 + \eta]$, and $u(t)$ otherwise. It is not difficult to prove that, if $\eta > 0$ is small enough, then the control $u_{\pi_1}(\cdot)$ is admissible on $[0, t_f]$. Moreover, $\tilde{x}_{\pi_1}(\cdot)$ converges uniformly to $\tilde{x}(\cdot)$ on $[0, t_f]$ whenever η tends to 0.

Recall that t_1 is a Lebesgue point of the function $t \mapsto \tilde{f}_\alpha(t, \tilde{x}(t), u(t))$ on $[0, t_f]$ whenever $\lim_{h \to 0} \frac{1}{h} \int_{t_1}^{t_1+h} \tilde{f}_\alpha(t, \tilde{x}(t), u(t))dt = \tilde{f}_\alpha(t_1, \tilde{x}(t_1), u(t_1))$, and that almost every point of $[0, t_f]$ is a Lebesgue point.

Let t_1 be a Lebesgue point on $[0, t_f)$, let $\eta > 0$ small enough, and let $u_{\pi_1}(\cdot)$ be a needle-like variation of $u(\cdot)$, with $\pi_1 = \{t_1, \eta, u_1\}$. For every $t \geqslant t_1$, as long as the trajectory $x(\cdot)$ remains in $X_\alpha(\cdot)$, the variation vector $\tilde{v}_{\pi_1}(\cdot)$ (not depending on η) is

[4] The smoothness of the end-point mapping on the open set $\mathcal{U}_{\tilde{x}_0, t_f, \mathbb{R}^m}$ is a standard fact for usual control systems. It is easy to prove that this property extends to the hybrid framework under the transversality assumption (H_4). However, this fact is not used in this article.

defined as the solution of the Cauchy problem

$$\begin{aligned} \dot{\tilde{v}}_{\pi_1}(t) &= \frac{\partial \tilde{f}_\alpha}{\partial \tilde{x}}(t,\tilde{x}(t),u(t))\tilde{v}_{\pi_1}(t), \\ \tilde{v}_{\pi_1}(t_1) &= \tilde{f}_\alpha(t_1,\tilde{x}(t_1),u_1) - \tilde{f}_\alpha(t_1,\tilde{x}(t_1),u(t_1)). \end{aligned} \tag{2.2}$$

Then, it is not difficult to prove that

$$\tilde{x}_{\pi_1}(t) = \tilde{x}(t) + \eta \tilde{v}_{\pi_1}(t) + \mathrm{o}(\eta) \tag{2.3}$$

(see, e.g., [22] for details). In particular, this formula means the following. For t_1 and u_1 fixed, denote $\tilde{x}_\alpha(t,\eta) = \tilde{x}_{\pi_1}(t)$; then $\eta \mapsto \tilde{x}_\alpha(\cdot,\eta)$ is differentiable at 0, and

$$\frac{\partial \tilde{x}_\alpha}{\partial \eta}(t,0) = \tilde{v}_{\pi_1}(t). \tag{2.4}$$

Let us now explain how this definition of variation vector must be adapted in order to take into account the change of dynamics whenever $x(\cdot)$ leaves the domain $X_\alpha(\cdot)$. For t_1 and u_1 fixed, let $t(\eta)$ denote the first time at which $x_\alpha(t(\eta),\eta) \in \partial X_\alpha(t(\eta))$. Recall that, by assumption (H_4), $t_c = t(0)$ is a regular crossing time. Assume that $x(\cdot)$ passes from $X_\alpha(\cdot)$ into $X_\beta(\cdot)$. In a neighborhood of $x(t_c)$, the boundary $\partial X_\alpha(t_c)$ can be written as $\{x \mid F(t_c,x)=0\}$, where F is a C^1 function on $\mathbb{R}^{n+1}$. The function $(t,\eta) \mapsto x_\alpha(t,\eta)$ can be extended for $t \geqslant t(\eta)$, prolongating the trajectory $x_\alpha(\cdot)$ in the domain $X_\beta(\cdot)$ with the dynamics f_α (hence, $x_\alpha(\cdot)$ differs from the true trajectory for $t \geqslant t(\eta)$). With this notation, the crossing time $t(\eta)$ is characterized by the equality $F(t(\eta),x_\alpha(t(\eta),\eta)) = 0$, which holds for every $\eta \geqslant 0$ small enough. Using the transversality crossing assumption (H_4), we infer that the function $\eta \mapsto t(\eta)$ is differentiable at $\eta = 0$, and

$$\left\langle \partial_x F(t_c,x(t_c)), t'(0)\dot{x}(t_c^-) + \frac{\partial x_\alpha}{\partial \eta}(t_c,0) \right\rangle + t'(0)\partial_t F(t_c,x(t_c)) = 0,$$

where the superscript $-$ stands for the left limit, and hence

$$t'(0) = -\frac{\left\langle \begin{pmatrix} \partial_x F(t_c,x(t_c)) \\ 0 \end{pmatrix}, \tilde{v}_{\pi_1}(t_c^-) \right\rangle}{\langle \partial_x F(t_c,x(t_c)), f_\alpha(t_c^-,x(t_c),u(t_c^-))\rangle + \partial_t F(t_c,x(t_c))}. \tag{2.5}$$

Note again that assumption (H_4) implies that the denominator of the above expression does not vanish.

We seek an extension of the definition of the variation vector, based on (2.4) so as to keep the validity of the expansion (2.3). Let us express the jump that is generated by the crossing. To this aim, we use the usual transport property of differential equations for the end-point mapping. Let $\delta > 0$. One has $\tilde{x}_{\pi_1}(t(\eta)+\delta) = \tilde{E}(x_{\pi_1}(t(\eta)-\delta), 2\delta, u_{\pi_1}(\cdot)_{\mid [t(\eta)-\delta,t(\eta)+\delta]})$, and

$$\tilde{x}_{\pi_1}(t(\eta)+\delta) = \tilde{E}(\tilde{E}(\tilde{x}_{\pi_1}(t(\eta)-\delta),\delta,u_{\pi_1}(\cdot)_{\mid [t(\eta)-\delta,t(\eta)]}),\delta,u_{\pi_1}(\cdot)_{\mid [t(\eta),t(\eta)+\delta]})$$

for every $\eta > 0$ small enough. On every piece, the end-point mapping is differentiable; hence since $\eta \mapsto t(\eta)$ is differentiable we can derive the above equality with respect to η and take $\eta = 0$. Since we have it in mind to keep the formula (2.3), this yields

$$\begin{aligned} \tilde{v}_{\pi_1}(t_c+\delta) + t'(0)\dot{\tilde{x}}(t_c+\delta) &= \frac{\partial \tilde{E}}{\partial x}(\tilde{E}(\tilde{x}(t_c-\delta),\delta,u(\cdot)_{\mid [t_c-\delta,t_c]}),\delta,u(\cdot)_{\mid [t_c,t_c+\delta]}) \\ &\quad \circ \frac{\partial \tilde{E}}{\partial x}(\tilde{x}(t_c-\delta),\delta,u(\cdot)_{\mid [t_c-\delta,t_c]}).(t'(0)\dot{\tilde{x}}(t_c-\delta) + v_{\pi_1}(t_c-\delta)). \end{aligned}$$

Then, letting δ tend to 0 leads to $\tilde{v}_{\pi_1}(t_c^+) = \tilde{v}_{\pi_1}(t_c^-) + t'(0)(\dot{\tilde{x}}(t_c^-) - \dot{\tilde{x}}(t_c^+))$; that is,

$$\tilde{v}_{\pi_1}(t_c^+) - \tilde{v}_{\pi_1}(t_c^-) = -t'(0)(\tilde{f}_\beta(t_c^+, \tilde{x}(t_c), u(t_c^+)) - \tilde{f}_\alpha(t_c^-, \tilde{x}(t_c), u(t_c^-))),$$

which, by using (2.5), we denote as

$$\tilde{v}_{\pi_1}(t_c^+) = \tilde{v}_{\pi_1}(t_c^-) + \frac{\left\langle \begin{pmatrix} \partial_x F(t_c, x(t_c)) \\ 0 \end{pmatrix}, \tilde{v}_{\pi_1}(t_c^-) \right\rangle}{\langle \partial_x F(t_c, x(t_c)), f_\alpha(t_c^-) \rangle + \partial_t F(t_c, x(t_c))} (\tilde{f}_\beta(t_c^+) - \tilde{f}_\alpha(t_c^-)). \tag{2.6}$$

We conclude that variation vectors of the hybrid control system (2.1) are defined by (2.2), as long as the trajectory remains in the domain $X_\alpha(\cdot)$, and satisfy the jump condition (2.6) whenever the trajectory crosses a boundary (with a regular transverse crossing). With that definition, the variation formula (2.3) still holds.

Remark 2.1. At the crossing time t_c, the following remarkable identity holds:

$$\frac{\left\langle \begin{pmatrix} \partial_x F(t_c, x(t_c)) \\ 0 \end{pmatrix}, \tilde{v}_\pi(t_c^-) \right\rangle}{\langle \partial_x F(t_c, x(t_c)), f_\alpha(t_c^-) \rangle + \partial_t F(t_c, x(t_c))} = \frac{\left\langle \begin{pmatrix} \partial_x F(t_c, x(t_c)) \\ 0 \end{pmatrix}, \tilde{v}_\pi(t_c^+) \right\rangle}{\langle \partial_x F(t_c, x(t_c)), f_\beta(t_c^+) \rangle + \partial_t F(t_c, x(t_c))}. \tag{2.7}$$

Indeed, it follows immediately from the jump formula (2.6). It can also be proved by establishing the formula (2.5) with t_c^- replaced with t_c^+, and α replaced with β (it suffices to consider x_β instead of x_α). The formula (2.7) is a kind of Snell–Descartes formula at the crossing time t_c. It will be of technical use in our proof.

Note that, for every $\gamma > 0$, the variation $\{t_1, \gamma\eta, u_1\}$ generates the variation vector $\gamma\tilde{v}_{\pi_1}(\cdot)$. It follows that the set of variation vectors at time t is a cone.

DEFINITION 2.2. *For every $t \in (0, t_f]$, the* first Pontryagin cone *$\tilde{K}(t) \subset \mathbb{R}^{n+1}$ at $\tilde{x}(t)$ for the extended system is defined as the smallest closed convex cone containing all variation vectors $\tilde{v}_{\pi_1}(t)$ for all Lebesgue points t_1 such that $0 < t_1 < t$. The first Pontryagin cone $K(t) \subset \mathbb{R}^n$ at $x(t)$ for the initial system is defined similarly, considering the initial dynamics f instead of the extended dynamics $\tilde{f}$. Obviously, $K(t)$ is the projection on $\mathbb{R}^n$ of $\tilde{K}(t)$.*

An immediate iteration leads to the following result, as in the usual case. Let $t_1 < t_2 < \cdots < t_k$ be Lebesgue points of the function $t \mapsto \tilde{f}(t, \tilde{x}(t), u(t))$ on $(0, t_f)$, and let $u_1, \ldots, u_k$ be points of Ω. Assume that all points $x(t_i)$ do not belong to the boundary of any domain $X_\alpha(\cdot)$ (note that, due to assumption (H_4), the set of such times is of full Lebesgue measure). Let $\eta_1, \ldots, \eta_p$ be small enough positive real numbers. Consider the variations $\pi_i = \{t_i, \eta_i, u_i\}$, and denote by $\tilde{v}_{\pi_i}(\cdot)$ the associated variation vectors, defined as above. Define the variation $\pi = \{t_1, \ldots, t_k, \eta_1, \ldots, \eta_k, u_1, \ldots, u_k\}$ of the control $u(\cdot)$ on $[0, t_f]$ by

$$u_\pi(t) = \begin{cases} u_i & \text{if } t_i \leqslant t \leqslant t_i + \eta_i, \quad i = 1, \ldots, k, \\ u(t) & \text{otherwise.} \end{cases} \tag{2.8}$$

Let $\tilde{x}_\pi(\cdot)$ be the solution of (2.1) associated with the control $u_\pi(\cdot)$ on $[0, t_f]$ and such that $\tilde{x}_\pi(0) = \tilde{x}_0$. Then

$$\tilde{x}_\pi(t_f) = \tilde{x}(t_f) + \sum_{i=1}^{k} \eta_i \tilde{v}_{\pi_i}(t_f) + \mathrm{o}\left(\sum_{i=1}^{k} \eta_i\right). \tag{2.9}$$

The first Pontryagin cone serves as an estimate of the accessible set $\tilde{A}_\Omega(\tilde{x}_0, t_f)$ in a neighborhood of $\tilde{x}(t_f)$.

When dealing with a free final time problem, we have to introduce time variations, and we consider the accessible set $\tilde{A}_\Omega(\tilde{x}_0)$ defined as the union of all $\tilde{A}_\Omega(\tilde{x}_0, s)$ over all $s \geqslant 0$. Assume first that $\tilde{x}(\cdot)$ is differentiable[5] at time t_f. Let $\delta \in \mathbb{R}$ small enough; then, with the above notation,

$$\tilde{x}_\pi(t_f + \delta) = \tilde{x}(t_f) + \sum_{i=1}^{k} \eta_i \tilde{v}_{\pi_i}(t_f) + \delta \tilde{f}(t_f, \tilde{x}(t_f), u(t_f)) + \mathrm{o}\left(\delta + \sum_{i=1}^{k} \eta_i\right). \tag{2.10}$$

Define the cone $\tilde{K}_1(t_f)$ as the smallest closed convex cone containing $\tilde{K}(t_f)$ and the vectors $\pm \tilde{f}(t_f, \tilde{x}(t_f), u(t_f))$. Similarly, the cone $K_1(t_f)$ is defined as the smallest closed convex cone containing $K(t_f)$ and the vectors $\pm f(t_f, x(t_f), u(t_f))$.

If $\tilde{x}(\cdot)$ is not differentiable at time t_f, then the above construction is slightly modified by replacing $\tilde{f}(t_f, \tilde{x}(t_f), u(t_f))$ with any closure point of the corresponding difference quotient in an obvious way.

2.1.2. Conic implicit function theorem. We next recall a *conic implicit function theorem*, which is useful for deriving a proof of the maximum principle.

LEMMA 2.3 (see [1]). *Let $C \subset \mathbb{R}^m$ be a convex subset of $\mathbb{R}^m$ with nonempty interior, of vertex 0, and let $F : C \to \mathbb{R}^n$ be a Lipschitzian mapping such that $F(0) = 0$ and F is differentiable in the sense of Gâteaux at 0. Assume that $dF(0).\mathrm{Cone}(C) = \mathbb{R}^n$, where $\mathrm{Cone}(C)$ stands for the (convex) cone generated by elements of C. Then 0 belongs to the interior of $F(\mathcal{V} \cap C)$ for every neighborhood $\mathcal{V}$ of 0 in $\mathbb{R}^m$.*

We also recall the following similar result involving a continuous dependence on parameters, adapted to our parametric regularization framework.

LEMMA 2.4 (see [3]). *Consider a continuous mapping $F : (\varepsilon, x) \in \mathbb{R}_+ \times \mathbb{R}^m_+ \mapsto F(\varepsilon, x) \in \mathbb{R}^n$, with $F(0,0) = 0$, such that, for every $\varepsilon \geqslant 0$, F is strictly differentiable with respect to x at 0 and $\frac{\partial F}{\partial x}$ is continuous with respect to ε, and such that $\frac{\partial F}{\partial x}(0,0).\mathbb{R}^m_+ = \mathbb{R}^n$. Then, there exist $\varepsilon_0 > 0$, a neighborhood V of 0 in $\mathbb{R}^n$, and a continuous function $g : [0, \varepsilon_0] \times V \to \mathbb{R}^m_+$ such that $F(\varepsilon, g(\varepsilon, y)) = y$ for every $\varepsilon \in [0, \varepsilon_0]$ and every $y \in V$.*

2.1.3. Lagrange multipliers and hybrid maximum principle. We next restrict the end-point mapping to time and needle-like variations. Assume that the final time t_f is free.[6] Let k be a positive integer. Set $\mathbb{R} \times \mathbb{R}^k_+ = \{(\delta, \eta_1, \ldots, \eta_k) \in \mathbb{R}^{k+1} \mid \eta_1 \geqslant 0, \ldots, \eta_k \geqslant 0\}$. Let $t_1 < \cdots < t_k$ be Lebesgue points of the function $t \mapsto \tilde{f}(t, \tilde{x}(t), u(t))$ on $(0, t_f)$, and let $u_1, \ldots, u_k$ be points of Ω. Let $\mathcal{V}$ be a small neighborhood of 0 in $\mathbb{R}^k$. Define the mapping $G : \mathcal{V} \cap (\mathbb{R} \times \mathbb{R}^k_+) \to \mathbb{R}^{n+1}$ by $G(\delta, \eta_1, \ldots, \eta_k) = \tilde{x}_\pi(t_f + \delta) - \tilde{x}(t_f)$, where π is the variation $\pi = \{t_1, \ldots, t_k, \eta_1, \ldots, \eta_k, u_1, \ldots, u_k\}$ and $|\delta|$ is small enough so that $t_k < t_f + \delta$. If $\mathcal{V}$ is small enough, then G is well defined; moreover, this mapping is clearly Lipschitzian, and $G(0) = 0$. From (2.10), G is Gâteaux differentiable on the conic neighborhood $\mathcal{V} \cap (\mathbb{R} \times \mathbb{R}^k_+)$ of 0.

If the cone $\tilde{K}_1(t_f)$ were to coincide with $\mathbb{R}^{n+1}$, then there would exist a real number δ, an integer k, and variations $\pi_i = \{t_i, \eta_i, u_i\}$, $i = 1, \ldots, k$, such that $G'_0(\mathbb{R} \times \mathbb{R}^k_+) = \mathbb{R}^{n+1}$, and then Lemma 2.3 would imply that the point $\tilde{x}(t_f)$ belongs to the interior of the accessible set $\tilde{A}_\Omega(\tilde{x}_0)$, which would contradict the optimality of $x(\cdot)$.

[5]This holds true, e.g., whenever t_f is a Lebesgue point of the function $t \mapsto \tilde{f}(t, \tilde{x}(t), u(t))$.

[6]The case where t_f is fixed is simpler; in that case, it it not necessary to consider the real parameter δ, and one uses the cone $\tilde{K}$ instead of $\tilde{K}_1$.

Therefore the convex cone $\tilde{K}_1(t_f)$ is not equal to $\mathbb{R}^{n+1}$. As a consequence, there exists $\tilde{\psi} \in \mathbb{R}^{n+1} \setminus \{0\}$ called a *Lagrange multiplier* such that $\langle \tilde{\psi}, \tilde{v}(t_f)\rangle \leqslant 0$ for every variation vector $\tilde{v}(t_f) \in \tilde{K}(t_f)$ and $\langle \tilde{\psi}, \tilde{f}(t_f, \tilde{x}(t_f), u(t_f))\rangle = 0$ (at least whenever $\tilde{x}(\cdot)$ is differentiable at time t_f; otherwise replace $\tilde{f}(t_f, \tilde{x}(t_f), u(t_f))$ with any closure point of the corresponding difference quotient).

These inequalities then permit us to derive (as in the usual way; see [22]) the statement of the HMP presented in the first section. The relation with the above Lagrange multiplier $\tilde{\psi} = (\psi, \psi^0)$ is that the adjoint vector p can be constructed so that

$$\psi = p(t_f) \quad \text{and} \quad p^0 = \psi^0. \tag{2.11}$$

In particular, the Lagrange multiplier ψ is unique if and only if the trajectory $x(\cdot)$ admits a unique extremal lift (up to a multiplicative scalar).

Remark 2.5. The trajectory $x(\cdot)$ has an abnormal extremal lift $(x(\cdot), p(\cdot), 0, u(\cdot))$ on $[0, t_f]$ if and only if there exists a unit vector $\psi \in \mathbb{R}^n$ such that $\langle \psi, v\rangle \leqslant 0$ for every $v \in K(t_f)$ (and moreover, $\max_{w\in\Omega}\langle \psi, f(t_f, x(t_f), w)\rangle = 0$ whenever t_f is free). In that case, one has $p(t_f) = \psi$, up to a multiplicative scalar.

The following lemma easily follows from the above considerations.

LEMMA 2.6. *Assume that, in the optimal control problem, the final time t_f is free. For the optimal trajectory $x(\cdot)$, the following statements are equivalent:*

- *The trajectory $x(\cdot)$ has a unique extremal lift $(x(\cdot), p(\cdot), p^0, u(\cdot))$ (up to a multiplicative scalar), which is moreover normal, i.e., $p^0 < 0$;*
- *$\tilde{K}_1(t_f)$ is a half-space of $\mathbb{R}^{n+1}$ and $p^0 < 0$;*
- *$\tilde{K}_1(t_f)$ is a half-space of $\mathbb{R}^{n+1}$ and $K_1(t_f) = \mathbb{R}^n$.*

If the final time is fixed, then the above statement holds provided $\tilde{K}_1(t_f)$ is replaced with $\tilde{K}(t_f)$, and $K_1(t_f)$ is replaced with $K(t_f)$.

This important lemma permits us to translate the assumptions of our main result into geometric considerations.

2.2. Proof of Theorem 1.6. From now on, assume that assumptions (H_1)–(H_6) hold (with (H_6) possibly replaced with (H'_6) in the case of a control-affine system). We denote the end-point mapping for the extended regularized system by $\tilde{E}(\varepsilon, x_0, t, u^\varepsilon(\cdot)) = \tilde{x}^\varepsilon(t)$, where $\tilde{x}^\varepsilon(\cdot)$ is the solution of the extended regularized system

$$\dot{\tilde{x}}^\varepsilon(t) = \tilde{f}^\varepsilon(t, \tilde{x}^\varepsilon(t), u^\varepsilon(t)), \quad \tilde{x}^\varepsilon(0) = \tilde{x}_0 = (x_0, 0), \tag{2.12}$$

where $\tilde{f}^\varepsilon(t, \tilde{x}^\varepsilon, u^\varepsilon) = (f^\varepsilon(t, x^\varepsilon, u^\varepsilon), f^{0\varepsilon}(t, x^\varepsilon, u^\varepsilon))$. By extension, the end-point mapping for the hybrid system corresponds to $\varepsilon = 0$; that is, $\tilde{E}(0, x_0, t, u(\cdot)) = \tilde{x}(t)$, where $\tilde{x}(\cdot)$ is the solution of (2.1) associated with the control $u(\cdot)$ and such that $\tilde{x}(0) = \tilde{x}_0$. It will be also denoted $\tilde{E}(x_0, t, u(\cdot)) = \tilde{E}(0, x_0, t, u(\cdot)) = \tilde{x}(t)$.

In what follows, we denote by $(x(\cdot), u(\cdot))$ the (unique) solution of (HOCP). We assume that the final time t_f of (HOCP) is free (the case of a fixed final time is similar but simpler); as explained previously, we impose (for instance) $0 \leqslant t_f^\varepsilon \leqslant t_f + 10$.

The proof of Theorem 1.6 follows from the succession of results below. Proposition 2.7 provides an existence result for a solution to $(\mathrm{OCP})_\varepsilon$, as well as first convergence properties. Proposition 2.11 and Lemma 2.13 allow us to prove the existence of variation vectors of $(\mathrm{OCP})_\varepsilon$ that converge to variation vectors of (HOCP). Proposition 2.15 proves the normality of the extremal lifts of $(\mathrm{OCP})_\varepsilon$ and the boundedness of its adjoint vectors. With all those ingredients we are then able to prove the theorems.

Proposition 2.7. *There exists $\varepsilon_0 > 0$ such that, for every $\varepsilon \in (0, \varepsilon_0)$, the problem $(\mathrm{OCP})_\varepsilon$ admits at least one solution $(x^\varepsilon(\cdot), u^\varepsilon(\cdot))$ defined on $[0, t_f^\varepsilon]$. Moreover, t_f^ε converges to t_f, $x^\varepsilon(\cdot)$ converges to $x(\cdot)$ uniformly on $[0, t_f]$, $\tilde{f}^\varepsilon(\cdot, \tilde{x}^\varepsilon(\cdot), u^\varepsilon(\cdot))$ converges to $\tilde{f}(\cdot, \tilde{x}(\cdot), u(\cdot))$, and $\frac{\partial \tilde{f}^\varepsilon}{\partial \tilde{x}}(\cdot, \tilde{x}^\varepsilon(\cdot), u^\varepsilon(\cdot))$ converges to $\frac{\partial \tilde{f}}{\partial \tilde{x}}(\cdot, \tilde{x}(\cdot), u(\cdot))$ for the weak star topology of L^∞, as ε tends to 0.*

NO misprint there: it is not true that df^eps/dx converges to df/dx in weak star topology, because it is not even bounded!! The convergence is only true on every interval not containing a crossing time (and this is exactly what is required in the proof).
Thank you to Terence Bayen to have pointed this out.

In the specific case of a control-affine system, under assumption (H_6') instead of (H_6), the following additional convergence property holds: $u^\varepsilon(\cdot)$ converges to $u(\cdot)$ in $L^1([0, t_f], \mathbb{R}^m)$ as ε tends to 0 for the strong topology.

Remark 2.8. In particular, $\dot{x}^\varepsilon(\cdot)$ converges to $\dot{x}(\cdot)$ in L^∞ for the weak star topology, as ε tends to 0.

Proof. Knowing that the constrained minimization problem (HOCP) has a solution, let us first prove that the problem $(\mathrm{OCP})_\varepsilon$ has at least one solution for every $\varepsilon > 0$ small enough. We use a similar reasoning as in section 2.1.3. Let k be a positive integer, let $\varepsilon \geqslant 0$, and let $t_1 < \cdots < t_k$ be Lebesgue points of the function $t \mapsto f^\varepsilon(t, x^\varepsilon(t), u^\varepsilon(t))$. Let $u_1, \ldots, u_k$ be points of Ω, and let $\mathcal{V}$ be a neighborhood of 0 in $\mathbb{R}^{k+2}$. Consider the variation $\pi = \{t_1, \ldots, t_k, \eta_1, \ldots, \eta_k, u_1, \ldots, u_k\}$ of the control $u(\cdot)$, and define the associated mapping $\Gamma : \mathcal{V} \cap (\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+^k) \to \mathbb{R}^n$ by $\Gamma(\varepsilon, \delta, \eta_1, \ldots, \eta_k) = x_\pi^\varepsilon(t_f + \delta) - x(t_f)$.

From assumption (H_3), the unique extremal lift of $x(\cdot)$ is normal, and hence it follows from Lemma 2.6 that $K_1(t_f) = \mathbb{R}^n$. Therefore, there exist a real number δ, an integer k, and a variation $\pi = \{t_1, \ldots, t_k, \eta_1, \ldots, \eta_k, u_1, \ldots, u_k\}$ such that the associated mapping Γ satisfies $\frac{\partial \Gamma}{\partial(\delta, \eta_1, \ldots, \eta_k)}(0).(\mathbb{R} \times \mathbb{R}_+^k) = K_1(t_f) = \mathbb{R}^n$. Lemma 2.4 implies that there exists $\varepsilon_0 > 0$ such that, for every $\varepsilon \in [0, \varepsilon_0)$, there exists $\delta_\varepsilon \in \mathbb{R}$ and a variation $\pi_\varepsilon = \{t_1^\varepsilon, \ldots, t_k^\varepsilon, \eta_1^\varepsilon, \ldots, \eta_k^\varepsilon, u_1^\varepsilon, \ldots, u_k^\varepsilon\}$ such that $\Gamma(\varepsilon, \delta_\varepsilon, \eta_1^\varepsilon, \ldots, \eta_k^\varepsilon) = 0$, and moreover, $|\delta_\varepsilon|$ is small whenever ε_0 is small enough. In other words, for every $\varepsilon > 0$ small enough, the subset M_1 is reachable from the subset M_0 for the regularized control system (1.7), within a time $t_f^\varepsilon \in [0, t_f + 10]$, and with the control $u_{\pi_\varepsilon}(\cdot) \in L^\infty([0, t_f^\varepsilon], \Omega)$.

The existence of an optimal control steering the regularized system from M_0 to M_1 is then a standard fact to derive, using the convexity assumption (H_5) on the extended velocities, the compactness of M_0 and M_1 and the additional compactness assumption (1.8) (see, e.g., [11] for such existence results).

Let us now prove the convergence properties. Let $(\varepsilon_k)_{k \in \mathbb{N}}$ be an arbitrary sequence of positive real numbers converging to 0 as k tends to $+\infty$. Since $t_f^{\varepsilon_k} \in [0, t_f + 10]$, the sequence $(t_f^{\varepsilon_k})_{k \in \mathbb{N}}$ converges, up to a subsequence, to some $T \in [0, t_f + 10]$. Since M_0 and M_1 are compact, and $x^{\varepsilon_k}(0) \in M_0$ and $x^{\varepsilon_k}(t_f^{\varepsilon_k}) \in M_1$, the sequences $(x^{\varepsilon_k}(0))_{k \in \mathbb{N}}$ and $(x^{\varepsilon_k}(t_f^{\varepsilon_k}))_{k \in \mathbb{N}}$ converge, up to a subsequence, to some $\bar{x}_0 \in M_0$ and $\bar{x}_T \in M_1$, respectively.

For every integer k and almost every $t \in [0, t_f^{\varepsilon_k}]$, set $g_k(t) = f^{\varepsilon_k}(t, x^{\varepsilon_k}(t), u^{\varepsilon_k}(t))$. Assumption (H_1) and the state constraint (1.8) imply that the sequence $(g_k(\cdot))_{k \in \mathbb{N}}$ is bounded in L^∞; hence up to a subsequence it converges to some $g(\cdot) \in L^\infty([0, T], \mathbb{R}^n)$ for the weak star topology. For every $t \in [0, T]$, set $\bar{x}(t) = \bar{x}_0 + \int_0^t g(s) ds$. Since $x^{\varepsilon_k}(t) = \bar{x}_0 + \int_0^t g_k(s) ds$, it follows that $x^{\varepsilon_k}(\cdot)$ converges uniformly to $\bar{x}(\cdot)$, up to a subsequence. In particular, $\bar{x}(\cdot)$ satisfies the state constraint $\max_{t \in [0,T]} \|\bar{x}(t)\|_{\mathbb{R}^n} \leqslant R$.

Let us prove that there exists a control $\bar{u}(\cdot) \in L^\infty([0, T], \Omega)$ such that $\bar{x}(\cdot)$ is a solution of the control system (1.1), associated with the control $\bar{u}(\cdot)$. Note that, using assumptions (H_1) and (H_5), $g_k(t) \in V^{\varepsilon_k}(t, x^{\varepsilon_k}(t))$ for every integer k and almost every $t \in [0, t_f^{\varepsilon_k}]$, where $V^{\varepsilon_k}(t, x^{\varepsilon_k}(t)) = \{f^{\varepsilon_k}(t, x^{\varepsilon_k}(t), u) \mid u \in \Omega\}$ is a compact and convex subset of $\mathbb{R}^n$. To prove the statement, let us first prove that $g(t) \in V(t, \bar{x}(t))$

for almost every $t \in [0,T]$, where $V(t,\bar{x}(t)) = \{f(t,\bar{x}(t),u) \mid u \in \Omega\}$. Note that $V(t,\bar{x}(t))$ is a compact and convex subset of $\mathbb{R}^n$. It follows from the definition of a C^1 regularization (Definition 1.1) and from the convexity assumptions that the sequence of compact convex subsets $(V^{\varepsilon_k}(t,x^{\varepsilon_k}(t)))_{k\in\mathbb{N}}$ converges (in the usual sense of Hausdorff) to $V(t,\bar{x}(t))$ for almost every t. For every $\delta \geqslant 0$, set $\mathcal{V}_\delta = \{h(\cdot) \in L^2([0,T],\mathbb{R}^n) \mid h(t) \in V_\delta(t,\bar{x}(t))$ for almost every $t \in [0,t_f]\}$. Here $V_\delta(t,x(t))$ is the compact convex set consisting of the points of $\mathbb{R}^n$ that are at a distance of $V(t,x(t))$ less than or equal to δ. It is not difficult to see that, for every $\delta \geqslant 0$, $\mathcal{V}_\delta$ is a closed convex subset of $L^2([0,T],\mathbb{R}^n)$ for the strong topology, and thus for the weak topology as well. Let $\delta > 0$ be arbitrary. Note that $g_k(\cdot) \in \mathcal{V}_\delta$ whenever k is large enough. Since the sequence $(g_k(\cdot))_{k\in\mathbb{N}}$ converges up to a subsequence to $g(\cdot)$ for the weak star topology of L^∞, it converges up to a subsequence to $g(\cdot)$ for the weak topology of L^2 as well. Using the closedness of V_δ for this topology, we infer that $g(\cdot) \in V_\delta$. Since $\delta > 0$ is arbitrary, it follows that $g(\cdot) \in V_0$; that is, for almost every $t \in [0,T]$ there exists $\bar{u}(t) \in \Omega$ such that $g(t) = f(t,\bar{x}(t),\bar{u}(t))$. The fact that the function $\bar{u}(\cdot)$ can be chosen to be measurable on $[0,T]$ follows from a standard measurable selection lemma (see, e.g., [20, Lemma 3A, page 161]).

Repeating all previous arguments for the extended systems (replacing x with $\tilde{x}$ and x^ε with $\tilde{x}^\varepsilon$) permits us to show as well that $C^{\varepsilon_k}(t_f^{\varepsilon_k},u^{\varepsilon_k})$ converges to $C(T,\bar{u})$ as k tends to $+\infty$. Using the optimality of $C^{\varepsilon_k}(t_f^{\varepsilon_k},u^{\varepsilon_k})$ and the uniqueness assumption (H_2), we infer that $T = t_f$, $\bar{u}(\cdot) = u(\cdot)$ almost everywhere and $\bar{x}(\cdot) = x(\cdot)$ on $[0,t_f]$.

Similarly, using the convexity assumption (H_5), the previous argumentation can be developed to derive convergence properties for the sequence $\frac{\partial \tilde{f}^{\varepsilon_k}}{\partial \tilde{x}}(\cdot,\tilde{x}^{\varepsilon_k}(\cdot),u^{\varepsilon_k}(\cdot))$.

NO, see page 1510: this is done only on every interval not containing a crossing time.

Hence, at this step, we have proved that, up to a subsequence, $(t_f^{\varepsilon_k})_{k\in\mathbb{N}}$ converges to t_f, $(x^{\varepsilon_k}(\cdot))_{k\in\mathbb{N}}$ converges to $x(\cdot)$ uniformly on $[0,t_f]$, $(\tilde{f}^{\varepsilon_k}(\cdot,x^{\varepsilon_k}(\cdot),u^{\varepsilon_k}(\cdot)))_{k\in\mathbb{N}}$ converges to $\tilde{f}(\cdot,x(\cdot),u(\cdot))$, and $(\frac{\partial \tilde{f}^{\varepsilon_k}}{\partial \tilde{x}}(\cdot,x^{\varepsilon_k}(\cdot),u^{\varepsilon_k}(\cdot)))_{k\in\mathbb{N}}$ converges to $\frac{\partial \tilde{f}}{\partial \tilde{x}}(\cdot,x(\cdot),u(\cdot))$ in L^∞ for the weak star topology, as k tends to $+\infty$.

We now investigate the convergence of the sequence $(u^{\varepsilon_k}(\cdot))_{k\in\mathbb{N}}$ in the case of a control-affine system and under assumption (H_6') instead of (H_6). For a control-affine system, we infer from the above convergence properties that the sequence $(u^{\varepsilon_k}(\cdot))_{k\in\mathbb{N}}$ converges up to a subsequence to $u(\cdot)$ for the weak star topology of L^∞, and thus for the weak topology of L^2 as well. Besides, from assumption (H_6'), $u(t)$ is an extremal point of Ω for almost every $t \in [0,t_f]$. It then follows from [33, Corollary 1] that $u^{\varepsilon_k}(\cdot)$ converges strongly (up to a subsequence) to $u(\cdot)$ in L^1.

Since the sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ was arbitrary, the conclusion follows by uniqueness of the closure point. □

Remark 2.9. If one does not assume the uniqueness of the optimal solution of (HOCP), then we still have that every closure point (for the appropriate topologies) of a family of solutions of $(\mathrm{OCP})_\varepsilon$ is a solution of (HOCP).

Remark 2.10. The solution of $(\mathrm{OCP})_\varepsilon$ is not necessarily unique. However, all results that follow do not depend on the specific choice of a solution.

In what follows, let $(x^\varepsilon(\cdot),u^\varepsilon(\cdot))$ be a solution of $(\mathrm{OCP})_\varepsilon$ defined on $[0,t_f^\varepsilon]$ for every $\varepsilon \in (0,\varepsilon_0)$. Since $x^\varepsilon(\cdot)$ converges uniformly to $x(\cdot)$, if we choose $R > 0$ large enough, then the state constraint (1.8) is not active in $(\mathrm{OCP})_\varepsilon$, as announced in section 1.2. It then follows from the PMP applied to $(\mathrm{OCP})_\varepsilon$ that $x^\varepsilon(\cdot)$ is the projection of an extremal $(x^\varepsilon(\cdot),p^\varepsilon(\cdot),p^{0\varepsilon},u^\varepsilon(\cdot))$ satisfying (1.10) and (1.11).

In order to derive convergence properties for the adjoint vector, we come back to the geometric interpretation of the proof of the PMP and HMP in terms of Pontryagin cones, as explained formerly. For every $\varepsilon > 0$, we denote by $K^\varepsilon(t)$, $K_1^\varepsilon(t)$, $\tilde{K}^\varepsilon(t)$,

$\tilde{K}_1^\varepsilon(t)$ the Pontryagin cones along the trajectory $x^\varepsilon(\cdot)$. The following result states nice convergence properties for the Pontryagin cones.

PROPOSITION 2.11. *For every $\tilde{v} \in \tilde{K}(t_f)$ and for every $\varepsilon > 0$, there exists $\tilde{v}^\varepsilon \in \tilde{K}^\varepsilon(t_f^\varepsilon)$ such that $\tilde{v}^\varepsilon$ converges to $\tilde{v}$ as ε tends to* 0.

Proof. By construction of $\tilde{K}(t_f)$, it suffices to prove the lemma for a single needle-like variation. Assume that $\tilde{v} = \tilde{v}_\pi(t_f)$, where the variation vector $\tilde{v}_\pi(\cdot)$ is the solution on $[t_1, t_f]$ of the Cauchy problem

$$\dot{\tilde{v}}_\pi(t) = \frac{\partial \tilde{f}_\alpha}{\partial \tilde{x}}(t, \tilde{x}(t), u(t)).\tilde{v}_\pi(t), \quad \tilde{v}_\pi(t_1) = \tilde{f}(t_1, \tilde{x}(t_1), u_1) - \tilde{f}(t_1, \tilde{x}(t_1), u(t_1)), \tag{2.13}$$

as long as $x(t) \in X_\alpha(t)$, where t_1 is a Lebesgue point of $[0, t_f)$, $u_1 \in \Omega$, and the needle-like variation $\pi = \{t_1, \eta, u_1\}$ of the control $u(\cdot)$ is defined by $u_\pi(t) = u_1$ if $t \in [t_1, t_1 + \eta]$, and $u(t)$ otherwise. When $x(\cdot)$ crosses (transversally) the boundary of $X_\alpha(\cdot)$, the variation vector $\tilde{v}_\pi(\cdot)$ satisfies the jump condition (2.6).

In order to define a needle-like variation of the regularized control $u^\varepsilon(\cdot)$, we first need the following technical lemma.

LEMMA 2.12. *For almost every $t \in (0, t_f)$, there exists a family $(t^\varepsilon)_{\varepsilon>0}$ of points of $[t, t_f)$ such that $t^\varepsilon \to t$ and $\tilde{f}^\varepsilon(t^\varepsilon, \tilde{x}^\varepsilon(t^\varepsilon), u^\varepsilon(t^\varepsilon)) \to \tilde{f}(t, \tilde{x}(t), u(t))$ as $\varepsilon \to 0$, and such that t^ε is a Lebesgue point of the function $t \mapsto \tilde{f}^\varepsilon(t, \tilde{x}^\varepsilon(t), u^\varepsilon(t))$.*

Proof of Lemma 2.12. Set $h^\varepsilon(t) = \tilde{f}^\varepsilon(t, \tilde{x}^\varepsilon(t), u^\varepsilon(t))$ and $h(t) = \tilde{f}(t, \tilde{x}(t), u(t))$, and denote by $h^\varepsilon(t) = (h_1^\varepsilon(t), \dots, h_{n+1}^\varepsilon(t))$ and $h(t) = (h_1(t), \dots, h_{n+1}(t))$ their coordinates in $\mathbb{R}^{n+1}$. Let us prove that, for almost every $t \in (0, t_f)$, for every $\beta > 0$, and for every $\alpha > 0$ (small enough so that $t + \alpha < t_f$), there exists $\gamma > 0$ such that, for every $\varepsilon \in (0, \gamma)$, there exists $t^\varepsilon \in [t, t + \alpha]$ such that $\|h^\varepsilon(t^\varepsilon) - h(t)\| \leqslant \beta$ (here, $\|\cdot\|$ denotes a norm in $\mathbb{R}^{n+1}$). The proof goes by contradiction. Assume that there exists a measurable subset A of $(0, t_f)$ of positive measure such that, for every $t \in A$, there exist $\beta > 0$ and $\alpha > 0$ such that, for every integer k, there exist $\varepsilon_k \in (0, 1/k)$ and $i \in \{1, \dots, n+1\}$ such that, for every $s \in [t, t + \alpha]$, there holds $|h_i^{\varepsilon_k}(s) - h_i(t)| > \beta$.

We distinguish between two cases, depending on whether assumption (H_6) or (H_6') holds. First, under assumption (H_6), $h^\varepsilon(\cdot)$ is continuous for every $\varepsilon > 0$. It follows from the proof of the previous proposition that the family $(h^\varepsilon(\cdot))_{0<\varepsilon<\varepsilon_0}$ converges to $h(\cdot)$ in L^∞ for the weak star topology, and hence its restriction to any interval converges as well to the corresponding restriction of $h(\cdot)$. Since $h_i^{\varepsilon_k}(\cdot)$ is continuous, we infer that either $h_i^{\varepsilon_k}(s) \geqslant h_i(t) + \beta$ for every $s \in [t, t + \alpha]$, or $h_i^{\varepsilon_k}(s) \leqslant h_i(t) - \beta$ for every $s \in [t, t + \alpha]$. This inequality contradicts the weak convergence of the restriction to $[t, t + \alpha]$ of $h_i^{\varepsilon_k}(\cdot)$ towards the restriction to $[t, t + \alpha]$ of $h_i(\cdot)$. In the second case, under assumption (H_6') instead of (H_6) (and for a control-affine system), we have proved previously that the family $(u^\varepsilon(\cdot))_{0<\varepsilon<\varepsilon_0}$ converges to $u(\cdot)$ for the strong topology of L^1. Therefore, up to a subsequence the sequence $(u^{\varepsilon_k}(\cdot))_{k \in \mathbb{N}}$ converges almost everywhere to $u(\cdot)$. We infer that the sequence $(h_i^{\varepsilon_k}(\cdot))_{k \in \mathbb{N}}$ converges almost everywhere, up to a subsequence, to $h_i(\cdot)$. This raises a contradiction. ∎

From Lemma 2.12, for every $\varepsilon > 0$ small enough, there exists $t_1^\varepsilon \geqslant t_1$ such that $t_1^\varepsilon \to t_1$ and $\tilde{f}^\varepsilon(t_1^\varepsilon, \tilde{x}^\varepsilon(t_1^\varepsilon), u^\varepsilon(t_1^\varepsilon)) \to \tilde{f}(t_1, \tilde{x}(t_1), u(t_1))$ as $\varepsilon \to 0$. Consider then the needle-like variation $\pi^\varepsilon = \{t_1^\varepsilon, \eta, u_1\}$ of the control $u^\varepsilon(\cdot)$ defined by[7] $u_{\pi^\varepsilon}^\varepsilon(t) = u_1$ if $t \in [t_1^\varepsilon, t_1^\varepsilon + \eta]$, and $u_{\pi^\varepsilon}^\varepsilon(t) = u^\varepsilon(t)$ otherwise. Define the variation vector $\tilde{v}_{\pi^\varepsilon}(\cdot)$ as the

[7]Note that t_1^ε is a Lebesgue point of the function $t \mapsto \tilde{f}^\varepsilon(t, x^\varepsilon(t), u^\varepsilon(t))$.

solution on $[t_1^\varepsilon, t_f^\varepsilon]$ of the Cauchy problem

$$\tag{2.14} \begin{aligned} \dot{\tilde{v}}_{\pi^\varepsilon}(t) &= \frac{\partial \tilde{f}^\varepsilon}{\partial \tilde{x}}(t, \tilde{x}^\varepsilon(t), u^\varepsilon(t)).\tilde{v}_{\pi^\varepsilon}(t), \\ \tilde{v}_{\pi^\varepsilon}(t_1^\varepsilon) &= \tilde{f}^\varepsilon(t_1^\varepsilon, \tilde{x}^\varepsilon(t_1^\varepsilon), u_1) - \tilde{f}^\varepsilon(t_1^\varepsilon, \tilde{x}^\varepsilon(t_1^\varepsilon), u^\varepsilon(t_1^\varepsilon)). \end{aligned}$$

Since $\tilde{f}^\varepsilon(t_1^\varepsilon, \tilde{x}^\varepsilon(t_1^\varepsilon), u^\varepsilon(t_1^\varepsilon))$ converges to $\tilde{f}(t_1, \tilde{x}(t_1), u(t_1))$, it follows that $\tilde{v}_{\pi^\varepsilon}(t_1^\varepsilon)$ converges to $\tilde{v}_\pi(t_1)$. From Proposition 2.7, t_f^ε converges to t_f, $x^\varepsilon(\cdot)$ converges uniformly to $x(\cdot)$, and $\frac{\partial \tilde{f}^\varepsilon}{\partial \tilde{x}}(\cdot, \tilde{x}^\varepsilon(\cdot), u^\varepsilon(\cdot))$ converges to $\frac{\partial \tilde{f}}{\partial \tilde{x}}(\cdot, \tilde{x}(\cdot), u(\cdot))$ for the weak star topology of L^∞, as ε tends to 0. Therefore, $\tilde{v}^\varepsilon_{\pi^\varepsilon}(\cdot)$ converges uniformly to $\tilde{v}_\pi(\cdot)$ as long as $x(t) \in X_\alpha(t)$ (see, e.g., [32] for this kind of standard argument).

Note that Lemma 2.12 was used here to initialize the needle-like variation $u^\varepsilon_{\pi^\varepsilon}(\cdot)$ of the regularized control $u^\varepsilon(\cdot)$. The difficulty was that $u^\varepsilon_{\pi^\varepsilon}(\cdot)$ cannot be initialized at the time t_1 in general, since we do not know whether the simple convergence of $\tilde{f}^\varepsilon(t_1, \tilde{x}^\varepsilon(t_1), u^\varepsilon(t_1))$ to $\tilde{f}(t_1, \tilde{x}(t_1), u(t_1))$ holds or not, as ε tends to 0. This was the first main difficulty of the proof of Proposition 2.11 (already present in [28] but easier to overcome in that reference).

The second main difficulty occurs when $x(\cdot)$ crosses $\partial X_\alpha(\cdot)$, since the variation vector $\tilde{v}_\pi(\cdot)$ has a jump (2.6) at the crossing time t_c. Our aim is to prove that, at the limit, we recover this jump for $\tilde{v}_{\pi^\varepsilon}(\cdot)$. We first prove the following technical lemma.

LEMMA 2.13. *The function $q(\cdot)$ defined by*

$$q(s) = \frac{\left\langle \begin{pmatrix} \partial_x F(s, x(s)) \\ 0 \end{pmatrix}, \tilde{v}_\pi(s) \right\rangle}{\langle \partial_x F(x, x(s)), f(s, x(s), u(s)) \rangle + \partial_t F(x, x(s))}$$

is continuous in a neighborhood of the crossing time t_c. For every $\varepsilon \in (0, \varepsilon_0)$, define

$$q^\varepsilon(s) = \frac{\left\langle \begin{pmatrix} \partial_x F(s, x^\varepsilon(s)) \\ 0 \end{pmatrix}, \tilde{v}^\varepsilon_\pi(s) \right\rangle}{\langle \partial_x F(s, x^\varepsilon(s)), f^\varepsilon(s, x^\varepsilon(s), u^\varepsilon(s)) \rangle + \partial_t F(s, x^\varepsilon(s))}.$$

For every s close to t_c, $q^\varepsilon(s)$ converges to $q(s)$ as ε tends to 0.

Proof of Lemma 2.13. The first part of the lemma follows from the Snell–Descartes-like formula (2.7) established in Remark 2.1. For the second part, we proceed similarly as we did in section 2.1 to derive the jump formula for the variation vector. Denote $\tilde{x}^\varepsilon(t, \eta) = \tilde{x}^\varepsilon_{\pi^\varepsilon}(t)$; one has $\tilde{x}^\varepsilon(t, \eta) = (x^\varepsilon(t, \eta), x^{0\varepsilon}(t, \eta))$. For every real number γ such that $|\gamma|$ is small enough, denote by $t^\varepsilon(\eta, \gamma)$ (also denoted $t^\varepsilon_{\eta,\gamma}$) the first time at which $F(t^\varepsilon_{\eta,\gamma}, x^\varepsilon(t^\varepsilon_{\eta,\gamma}, \eta)) = \gamma$, and denote by $t(\eta, \gamma)$ (also denoted $t_{\eta,\gamma}$) the first time at which $F(t_{\eta,\gamma}, x(t_{\eta,\gamma}, \eta)) = \gamma$. Using the implicit function theorem, the transversality crossing assumption (H_4) implies that, for $|\gamma|$ and ε small enough, the function $\eta \mapsto t^\varepsilon_{\eta,\gamma}$ is differentiable at 0, and

$$\begin{aligned} \Big\langle \partial_x F(t^\varepsilon_{0,\gamma}, x^\varepsilon(t^\varepsilon_{0,\gamma})), f^\varepsilon(t^\varepsilon_{0,\gamma}, x^\varepsilon(t^\varepsilon_{0,\gamma}), u^\varepsilon(t^\varepsilon_{0,\gamma})) \frac{\partial t^\varepsilon}{\partial \eta}(0, \gamma) + v^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_{0,\gamma}) \Big\rangle & \\ + \partial_t F(t^\varepsilon_{0,\gamma}, x^\varepsilon(t^\varepsilon_{0,\gamma})) \cdot \frac{\partial t^\varepsilon}{\partial \eta}(0, \gamma) = 0&, \end{aligned}$$

and moreover, the function $\gamma \mapsto t^\varepsilon(0, \gamma)$ converges to the function $\gamma \mapsto t(0, \gamma)$ in C^1 topology. Using again assumption (H_4), it is clear that the function $\gamma \mapsto t(0, \gamma)$, restricted to a neighborhood of 0, is a diffeomorphism. The result follows, using the change of variable $s = t(0, \gamma)$. ∎

To recover the jump formula as ε tends to 0, we next proceed again similarly as in section 2.1. Denote, as in the previous lemma, $\tilde{x}^\varepsilon(t,\eta) = \tilde{x}^\varepsilon_{\pi^\varepsilon}(t)$. Let $t^\varepsilon(\eta)$ (also denoted t^ε_η) denote the first time at which $x^\varepsilon(t^\varepsilon_\eta,\eta) \in \partial X_\alpha(t^\varepsilon_\eta)$, that is, $F(t^\varepsilon_\eta, x^\varepsilon(t^\varepsilon_\eta,\eta)) = 0$ (this corresponds to $\gamma = 0$ in the previous proof). Using the transversality crossing assumption (H_4) and the convergence of $x^\varepsilon(\cdot)$ to $x(\cdot)$, we get that the function $\eta \mapsto t^\varepsilon(\eta)$ is differentiable at $\eta = 0$ for $\varepsilon > 0$ small enough, and

$$\left\langle \partial_x F(t^\varepsilon_0, x^\varepsilon(t^\varepsilon_0)), t^{\varepsilon\prime}(0)\dot{x}^\varepsilon(t^\varepsilon_0) + \frac{\partial x^\varepsilon}{\partial \eta}(t^\varepsilon_0, 0)\right\rangle + t^{\varepsilon\prime}(0)\partial_t F(t^\varepsilon_0, x^\varepsilon(t^\varepsilon_0)) = 0,$$

where $t^\varepsilon_0 = t^\varepsilon(0)$, and hence, using Lemma 2.13, we get $t^{\varepsilon\prime}(0) = -q^\varepsilon(t)$. Let $\delta > 0$. One has $\tilde{x}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_\eta + \delta) = \tilde{E}(\varepsilon, x^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_\eta - \delta), 2\delta, u^\varepsilon_{\pi^\varepsilon}(\cdot)_{\,|\,[t^\varepsilon_\eta-\delta, t^\varepsilon_\eta+\delta]})$, and $\tilde{x}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_\eta + \delta) = \tilde{E}(\varepsilon, \tilde{E}(\varepsilon, \tilde{x}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_\eta - \delta), \delta, u^\varepsilon_{\pi^\varepsilon}(\cdot)_{\,|\,[t^\varepsilon_\eta-\delta, t^\varepsilon_\eta]}), \delta, u^\varepsilon_{\pi^\varepsilon}(\cdot)_{\,|\,[t^\varepsilon_\eta, t^\varepsilon_\eta+\delta]})$, for every $\eta > 0$ small enough. On every piece, the end-point mapping is differentiable, and hence, since $\eta \mapsto t^\varepsilon(\eta)$ is differentiable at 0, we can differentiate the above equality with respect to η and take $\eta = 0$. This yields

$$\begin{aligned}
&\tilde{v}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_0 + \delta) + t^{\varepsilon\prime}(0)\dot{\tilde{x}}^\varepsilon(t^\varepsilon_0 + \delta)\\
&\quad = \frac{\partial \tilde{E}}{\partial x}(\varepsilon, \tilde{E}(\varepsilon, \tilde{x}^\varepsilon(t^\varepsilon_0 - \delta), \delta, u^\varepsilon(\cdot)_{\,|\,[t^\varepsilon_0-\delta, t^\varepsilon_0]}), \delta, u^\varepsilon(\cdot)_{\,|\,[t^\varepsilon_0, t^\varepsilon_0+\delta]})\\
&\qquad \circ \frac{\partial \tilde{E}}{\partial x}(\varepsilon, \tilde{x}^\varepsilon(t^\varepsilon_0 - \delta), \delta, u^\varepsilon(\cdot)_{\,|\,[t^\varepsilon_0-\delta, t^\varepsilon_0]}) \cdot (t^{\varepsilon\prime}(0)\dot{\tilde{x}}^\varepsilon(t^\varepsilon_0 - \delta) + \tilde{v}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_0 - \delta)).
\end{aligned}$$

We first let ε tend to 0, and then let δ tend to 0. Using Lemma 2.13 and the fact that $t^{\varepsilon\prime}(0) = -q^\varepsilon(t)$, we infer that $\lim_{\delta\to 0}\lim_{\varepsilon\to 0}(\tilde{v}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_0 + \delta) - \tilde{v}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon_0 - \delta)) = q(t)(\tilde{f}_\beta(t^+) - \tilde{f}_\alpha(t^-))$, which corresponds exactly to (2.6). Hence, we have proved that

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0}(\tilde{v}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon(0) + \delta) - \tilde{v}^\varepsilon_{\pi^\varepsilon}(t^\varepsilon(0) - \delta)) = \tilde{v}_\pi(t_c^+) - \tilde{v}_\pi(t_c^-). \tag{2.15}$$

This formula shows that, when $x(\cdot)$ crosses $\partial X_\alpha(\cdot)$ and the variation vector $\tilde{v}_\pi(\cdot)$ has a jump (2.6) at the crossing time t_c, we recover this jump at the limit for $\tilde{v}_{\pi^\varepsilon}(\cdot)$. Notice that $\tilde{v}_{\pi^\varepsilon}(\cdot)$ is continuous, and the order of the limits in (2.15) cannot be switched.

We can end the proof of the proposition. Indeed, before the first crossing time, $\tilde{v}_{\pi^\varepsilon}(\cdot)$ converges uniformly to $\tilde{v}_\pi(\cdot)$ on every compact interval not containing this crossing time. At the crossing time, $\tilde{v}_{\pi^\varepsilon}(\cdot)$ has a jump, and the formula (2.15) shows that we recover this jump, at the limit, for $\tilde{v}_{\pi^\varepsilon}(\cdot)$. Then, beyond the first crossing time and before the second crossing time, one has uniform convergence as well, and the argument goes by iteration. ∎

Remark 2.14. The same convergence result holds for the Pontryagin cone K_1. We will also need the following statement, resulting obviously from the above proof: For every $t \in [0, t_f]$, for every $\tilde{v} \in \tilde{K}_1(t)$, and for every $\varepsilon > 0$ there exists $\tilde{v}^\varepsilon \in \tilde{K}_1^\varepsilon(t)$ such that $\tilde{v}^\varepsilon$ converges to $\tilde{v}$ as ε tends to 0.

PROPOSITION 2.15. *There exists $\varepsilon_0 > 0$ such that, for every $\varepsilon \in (0, \varepsilon_0)$, every extremal lift $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), p^{0\varepsilon}, u^\varepsilon(\cdot))$ of any solution $x^\varepsilon(\cdot)$ of $(\mathrm{OCP})_\varepsilon$ is normal.*

Furthermore, setting $p^{0\varepsilon} = -1$, the set $\{p^\varepsilon(t) \mid t \in [0, t^\varepsilon_f],\ 0 < \varepsilon < \varepsilon_0\}$ is bounded.

Proof. We argue by contradiction. Assume that, for every integer k, there exist $\varepsilon_k \in (0, 1/k)$ and a solution $x^{\varepsilon_k}(\cdot)$ of $(\mathrm{OCP})_{\varepsilon_k}$ having an abnormal extremal lift $(x^{\varepsilon_k}(\cdot), p^{\varepsilon_k}(\cdot), 0, u^{\varepsilon_k}(\cdot))$. Set $\psi^{\varepsilon_k} = p^{\varepsilon_k}(t_f^{\varepsilon_k})$ for every integer k. Then, from Remark 2.5, one has $\langle \psi^{\varepsilon_k}, v^{\varepsilon_k}\rangle \leqslant 0$ for every $v^{\varepsilon_k} \in K^{\varepsilon_k}(t_f^{\varepsilon_k})$, and, since the final time is free, $M(\varepsilon_k) = \max_{w\in\Omega}\langle \psi^{\varepsilon_k}, f^{\varepsilon}(t_f^{\varepsilon_k}, x^{\varepsilon_k}(t_f^{\varepsilon_k}), w)\rangle = 0$ for every integer k. Since the final adjoint vector $(p^{\varepsilon_k}(t_f^{\varepsilon_k}), p^{0\varepsilon_k})$ (here, $p^{0\varepsilon_k} = 0$) is defined up to a multiplicative scalar, we assume that ψ^{ε_k} is a unit vector for every integer k. Then, up to a subsequence, the sequence $(\psi^{\varepsilon_k})_{k\in\mathbb{N}}$ converges to some unit vector $\psi \in \mathbb{R}^n$. In order to pass to the limit, we use the following easy lemma (whose proof is omitted).

LEMMA 2.16. *Let m be a positive integer, g be a continuous function on $\mathbb{R}\times\mathbb{R}^m$, and C be a compact subset of $\mathbb{R}^m$. For every $\varepsilon > 0$, set $M(\varepsilon) = \max_{u\in C} g(\varepsilon, u)$, and let $M = \max_{u\in C} g(0, u)$. Then, $M(\varepsilon)$ tends to M as ε tends to 0.*

Using Proposition 2.7, Lemma 2.16, and Proposition 2.11, passing to the limit we infer that $\langle \psi, v\rangle \leqslant 0$ for every $v \in K(t_f)$ and that $M = \max_{w\in\Omega}\langle \psi, f(t_f, x(t_f), w)\rangle = 0$. It then follows from Remark 2.5 that the trajectory $x(\cdot)$ has an abnormal extremal lift. This contradicts assumption (H_3).

To derive the second part of the statement of the proposition, let us first prove the following lemma.

LEMMA 2.17. *Setting $p^{0\varepsilon} = -1$, the set of all possible $p^{\varepsilon}(t_f^{\varepsilon})$, with $\varepsilon \in (0, \varepsilon_0)$, is bounded.*

Proof of Lemma 2.17. We proceed again by contradiction. Assume that there exists a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ of positive real numbers converging to 0 such that $\|p^{\varepsilon_k}(t_f^{\varepsilon_k})\|$ tends to $+\infty$. Since the sequence $\left(\frac{p^{\varepsilon_k}(t_f^{\varepsilon_k})}{\|p^{\varepsilon_k}(t_f^{\varepsilon_k})\|}\right)_{k\in\mathbb{N}}$ is bounded in $\mathbb{R}^n$, up to a subsequence it converges to some unit vector ψ. Using the Lagrange multipliers property and (2.11), there holds $\langle p^{\varepsilon_k}(t_f^{\varepsilon_k}), v^{\varepsilon_k}\rangle \leqslant 0$ for every $v^{\varepsilon_k} \in K^{\varepsilon_k}(t_f^{\varepsilon_k})$, and

$$\max_{w\in\Omega}\left(\left\langle p^{\varepsilon_k}(t_f^{\varepsilon_k}), f^{\varepsilon_k}(t_f^{\varepsilon_k}, x^{\varepsilon_k}(t_f^{\varepsilon_k}), w)\right\rangle - f^{0\varepsilon_k}(t_f^{\varepsilon_k}, x^{\varepsilon_k}(t_f^{\varepsilon_k}), w)\right) = 0 \tag{2.16}$$

for every integer k. Dividing by $\|p^{\varepsilon_k}(t_f^{\varepsilon_k})\|$, passing to the limit, and using Proposition 2.7, Lemma 2.16, Proposition 2.11, and Remark 2.5, the same reasoning as above yields that the trajectory $x(\cdot)$ has an abnormal extremal lift, which is a contradiction. □

We now use the fact that, for every $\varepsilon \in (0, \varepsilon_0)$, the function $t \mapsto \langle \tilde{p}^{\varepsilon}(t), \tilde{v}^{\varepsilon}_{\pi^{\varepsilon}}(t)\rangle$ is constant[8] for every variation vector $\tilde{v}_{\pi^{\varepsilon}}(\cdot)$ along $x^{\varepsilon}(\cdot)$, where $\tilde{p}^{\varepsilon}(t) = (p^{\varepsilon}(t), -1)$. In particular, denoting $\tilde{v}^{\varepsilon}_{\pi^{\varepsilon}}(t) = (v^{\varepsilon}_{\pi^{\varepsilon}}(t), v^{0\varepsilon}_{\pi^{\varepsilon}}(t))$, this yields

$$\langle p^{\varepsilon}(t), v^{\varepsilon}_{\pi^{\varepsilon}}(t)\rangle - v^{0\varepsilon}_{\pi^{\varepsilon}}(t) = \langle p^{\varepsilon}(t_f^{\varepsilon}), v^{\varepsilon}_{\pi^{\varepsilon}}(t_f^{\varepsilon})\rangle - v^{0\varepsilon}_{\pi^{\varepsilon}}(t_f^{\varepsilon}) \tag{2.17}$$

for every $t \in [0, t_f^{\varepsilon}]$. Since $\tilde{K}_1^{\varepsilon}(t_f^{\varepsilon})$ is a cone, we assume that the variation vectors under consideration here are such that $\tilde{v}^{\varepsilon}_{\pi^{\varepsilon}}(t_f^{\varepsilon})$ is a unit vector of $\mathbb{R}^{n+1}$. In particular, using such variation vectors, Lemma 2.17 implies that the right-hand side of (2.17) remains uniformly bounded with respect to ε. Since the component $v^{0\varepsilon}_{\pi^{\varepsilon}}(\cdot)$ satisfies $\dot{v}^{0\varepsilon}_{\pi^{\varepsilon}}(t) = \frac{\partial f^{0\varepsilon}}{\partial x}(t, x^{\varepsilon}(t), u^{\varepsilon}(t))v^{\varepsilon}_{\pi^{\varepsilon}}(t)$ for almost every $t \in [0, t_f^{\varepsilon}]$, and since, from Proposition 2.7, $\frac{\partial f^{0\varepsilon}}{\partial x}(\cdot, x^{\varepsilon}(\cdot), u^{\varepsilon}(\cdot))$ converges to $\frac{\partial f^{0}}{\partial x}(\cdot, x(\cdot), u(\cdot))$ for the weak star topology of L^{∞} as ε tends to 0, it follows easily that the set of all $v^{0\varepsilon}_{\pi^{\varepsilon}}(t)$ under consideration, for $t \in [0, t_f^{\varepsilon}]$ and $0 < \varepsilon < \varepsilon_0$, is bounded. Therefore, we infer from (2.17) that the set of all $\langle p^{\varepsilon}(t), v^{\varepsilon}_{\pi^{\varepsilon}}(t)\rangle$, over all possible variation vectors such that $\tilde{v}^{\varepsilon}_{\pi^{\varepsilon}}(t_f^{\varepsilon})$ is a unit vector, is uniformly bounded with respect to ε. From assumption (H_3) and Lemma 2.6, there holds $K_1(t_f) = \mathbb{R}^n$, and therefore, from Proposition 2.11, there holds $K_1^{\varepsilon}(t_f^{\varepsilon}) = \mathbb{R}^n$

NO: see page 1510. This reasoning must be done iteratively. We propagate boundedness on every interval not containing a crossing time. When meeting a crossing time, we have a jump, which is bounded. Whence the result.

[8] Indeed, it is continuous and its derivative is everywhere equal to zero.

for $\varepsilon > 0$ small enough; hence the previous arguments imply that the set of all $p^\varepsilon(t)$ for $t \in [0, t_f^\varepsilon]$ is uniformly bounded with respect to ε. ☐

Remark 2.18. If we remove the assumption that the optimal trajectory $x(\cdot)$ has a unique extremal lift, which is, moreover, normal, then Proposition 2.15 still holds provided that every extremal lift of $x(\cdot)$ is normal.

We are now in position to end the proof of the theorems and to derive the convergence properties for the adjoint vector.

Recall that $(x(\cdot), p(\cdot), -1, u(\cdot))$ is the unique (normal) extremal lift of the solution $x(\cdot)$ of (HOCP), and that for every $\varepsilon \in (0, \varepsilon_0)$, $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), -1, u^\varepsilon(\cdot))$ is a (normal) extremal lift of a solution $x^\varepsilon(\cdot)$ of $(\mathrm{OCP})_\varepsilon$. Let us prove that $p^\varepsilon(\cdot)$ converges uniformly to $p(\cdot)$ on every closed subinterval of $[0, t_f]$ that does not contain any crossing time, as ε tends to 0. For every $\varepsilon \in (0, \varepsilon_0)$, set $\psi^\varepsilon = p^\varepsilon(t_f^\varepsilon)$. Recall that $p^\varepsilon(\cdot)$ is a solution of the Cauchy problem

$$\dot{p}^\varepsilon(t) = -\left\langle p^\varepsilon(t), \frac{\partial f^\varepsilon}{\partial x}(t, x^\varepsilon(t), u^\varepsilon(t)) \right\rangle - \frac{\partial f^{0\varepsilon}}{\partial x}(t, x^\varepsilon(t), u^\varepsilon(t)), \quad p^\varepsilon(t_f^\varepsilon) = \psi^\varepsilon,$$

for almost every $t \in [0, t_f^\varepsilon]$ and, moreover, $\langle \psi^\varepsilon, v^\varepsilon \rangle \leqslant 0$ for every $v^\varepsilon \in K^\varepsilon(t_f^\varepsilon)$ and $\max_{w\in\Omega}(\langle \psi^\varepsilon, f^\varepsilon(t_f^\varepsilon, x^\varepsilon(t_f^\varepsilon), w)\rangle - f^{0\varepsilon}(t_f^\varepsilon, x^\varepsilon(t_f^\varepsilon), w)) = 0$ for every $\varepsilon \in (0, \varepsilon_0)$. From Proposition 2.15, the family $(\psi^\varepsilon)_{0<\varepsilon<\varepsilon_0}$ is bounded; let ψ be a closure point of that family, and let $(\varepsilon_k)_{k\in\mathbb{N}}$ be a sequence of positive real numbers converging to 0 such that ψ^{ε_k} tends to ψ. Using Proposition 2.7, and as in the proof of this result, we infer that the sequence $(p^{\varepsilon_k}(\cdot))_{k\in\mathbb{N}}$ converges uniformly to the solution $z(\cdot)$ of the Cauchy problem

$$\dot{z}(t) = -\left\langle z(t), \frac{\partial f_\beta}{\partial x}(t, x(t), u(t)) \right\rangle - \frac{\partial f_\beta^0}{\partial x}(t, x(t), u(t)), \quad z(t_f) = \psi, \tag{2.18}$$

for almost every t belonging to some interval $[\tau, t_f]$, provided that the interval $[\tau, t_f]$ does not contain any crossing time (and assuming that $x(t) \in X_\beta(t)$ for every $t \in [\tau, t_f]$). Moreover, passing to the limit as in the previous proof, one gets $\langle \psi, v \rangle \leqslant 0$ for every $v \in K(t_f)$, and $\max_{w\in\Omega}(\langle \psi, f_\beta(t_f, x(t_f), w)\rangle - f_\beta^0(t_f, x(t_f), w)) = 0$. It follows that $(x(\cdot), z(\cdot), -1, u(\cdot))$ is an extremal lift of $x(\cdot)$ on the interval $[\tau, t_f]$, and from assumption (H_3) we infer that $z(\cdot)$ and $p(\cdot)$ coincide on the subinterval $[\tau, t_f]$.

We can proceed in reverse time in such a way as long as we do not encounter any crossing time. In order to take into account possible crossing times, we must use an additional argument. Let t be an element of $[0, t_f]$ that is not a crossing time. From Proposition 2.15, the family $(p^\varepsilon(t))_{0<\varepsilon<\varepsilon_0}$ is bounded and hence has a closure point $\bar{p}(t)$. To end the proof, it suffices to prove that $\bar{p}(t) = p(t)$ (indeed, once this fact is proved, the uniform convergence argument is similar to the one above). From Remark 2.14, for every variation vector $\tilde{v}_\pi(t) \in \tilde{K}_1(t)$, for every $\varepsilon > 0$ there exists $\tilde{v}_{\pi^\varepsilon}^\varepsilon(t) \in \tilde{K}_1(t)$ such that $\tilde{v}_{\pi^\varepsilon}^\varepsilon(t)$ converges to $\tilde{v}_\pi(t)$ as ε tends to 0. Setting $\tilde{p}^\varepsilon = (p^\varepsilon, -1)$, as in the proof of Lemma 2.17 we use the fact that the function $t \mapsto \langle \tilde{p}^\varepsilon(t), \tilde{v}_{\pi^\varepsilon}^\varepsilon(t)\rangle$ is constant, and this yields (2.17) at the time t. Passing to the limit, we get

$$\langle \bar{p}(t), v_\pi(t)\rangle - v_\pi^0(t) = \langle p(t_f), v_\pi(t_f)\rangle - v_\pi^0(t_f). \tag{2.19}$$

In order to prove that $\bar{p}(t) = p(t)$, we first need to derive the following lemma, which is an extension to the hybrid context of the well-known result that we used formerly.

LEMMA 2.19. *Denoting* $\tilde{p}(t) = (p(t), -1)$*, the function* $t \mapsto \langle \tilde{p}(t), \tilde{v}_\pi(t) \rangle$ *is constant on* $[0, t_f]$*, for every variation vector* $\tilde{v}_\pi(\cdot)$.

Proof of Lemma 2.19. It is easy to see that the derivative of the function $t \mapsto \langle \tilde{p}(t), \tilde{v}_\pi(t) \rangle$ is equal to 0 everywhere, using the differential equations satisfied by $\tilde{p}(\cdot)$ and $\tilde{v}_\pi(\cdot)$. Besides, this function is clearly continuous outside the crossing times. Hence, to prove the statement it suffices to see that the function remains continuous at crossing times. But this follows straightforwardly from a simple computation using the jump conditions (1.5) and (2.6). ∎

It follows from this lemma that $\langle p(t), v_\pi(t) \rangle - v_\pi^0(t) = \langle p(t_f), v_\pi(t_f) \rangle - v_\pi^0(t_f)$, and hence we infer from (2.19) that $\langle \bar{p}(t), v_\pi(t) \rangle = \langle p(t), v_\pi(t) \rangle$. Since this equality holds for every variation vector, and since $K_1(t) = \mathbb{R}^n$ (this follows from assumption (H_3) and Lemma 2.6), it follows that $\bar{p}(t) = p(t)$. This ends the proof of the theorem.

Remark 2.20. In the proof above, it is possible to replace assumptions (H_2) and (H_3) with the weaker assumption that every extremal lift of every solution of (HOCP) is normal. In that case, using the same arguments, we prove that every closure point of the family $(t_f^\varepsilon, x^\varepsilon(\cdot), \dot{x}^\varepsilon(\cdot), p^\varepsilon(\cdot))_{0<\varepsilon<\varepsilon_0}$ (for the evident topologies) can be written as $(T, \bar{x}(\cdot), \dot{\bar{x}}(\cdot), \bar{p}(\cdot))$, where $\bar{x}(\cdot)$ is a solution of (1.1) associated with a control $\bar{u}(\cdot) \in L^\infty([0, t_f], \Omega)$, such that $\bar{x}(0) \in M_0$ and $\bar{x}(t_f) \in M_1$. Moreover, $(\bar{x}(\cdot), \bar{u}(\cdot))$ is a solution of (HOCP) defined on $[0, T]$, having as a normal extremal lift the 4-tuple $(\bar{x}(\cdot), \bar{p}(\cdot), -1, \bar{u}(\cdot))$.

3. Application to the minimal time low-thrust coplanar orbit transfer with eclipse constraint.

3.1. Problem statement. We focus on the coplanar orbit transfer of a satellite around the Earth. The satellite is modeled as a mass point and is assumed to evolve in a central gravitational field. We neglect the gravitational perturbations such as the Earth's oblateness. The satellite follows the two-dimensional controlled Kepler equation $\ddot{q}(t) = -\frac{\mu}{r(t)^3} q(t) + \frac{T(t)}{m(t)}$, $\dot{m}(t) = -\beta \|T(t)\|$. Here, $q(t) \in \mathbb{R}^2$ denotes the Cartesian coordinates of the satellite in an inertial geocentric reference frame, $r(t) = \|q(t)\|$ is the distance to the Earth's center, $T(t) \in \bar{B}(0, T_{\max}) \subset \mathbb{R}^2$ is the thrust, $T_{\max}$ is the maximal allowed thrust, $m(t)$ is the mass, μ stands for the Earth's gravitational constant, and β is a coefficient depending on the thruster characteristics.

The objective is to realize a minimal time orbit transfer, for instance, from a low and/or highly eccentric initial orbit to a geostationary final one. The controllability aspects of that problem were studied in [5, 6, 9, 18].

We are interested in low-thrust engines, that is, with maximum thrust $T_{\max}$ small when compared with the mass of the satellite. Thus, the orbit transfer will require a lot of revolutions and the Cartesian coordinates are not well suited. Indeed, the evolution of those coordinates is large when compared to the small evolution of the orbit shape. We use the set $x = (P, e_x, e_y, L)$ of Gauss coordinates defined by $e_x = e\cos(\omega)$, $e_y = \sin(\omega)$, $L = \omega + \nu$, where P is the semilatus rectum of the osculating ellipsis, e is the eccentricity, ν is the true anomaly, and ω is the argument of perigee. In addition to this change of coordinates, we express the thrust T in the moving reference frame $(q/\|q\|, s)$, where s is the unit projection of $\dot{q}$ on the orthogonal of q. We also rewrite the control as $T(\cdot) = T_{\max} u(\cdot)$, where $u(\cdot)$ takes its values in the closed unit ball $\bar{B}(0, 1)$ of $\mathbb{R}^2$. The equations of motion can then be rewritten as

$$\dot{x}(t) = f_0(x(t)) + \frac{T_{\max}}{m(t)} \left(u_1(t) f_1(x(t)) + u_2(t) f_2(x(t)) \right), \quad \dot{m}(t) = -\beta T_{\max} \|u(t)\|,$$

where the vector fields f_0, f_1, and f_2 are defined by

$$f_0(x) = \sqrt{\frac{\mu}{P}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{W^2}{P} \end{pmatrix}, \quad f_1(x) = \sqrt{\frac{P}{\mu}} \begin{pmatrix} 0 \\ \sin L \\ \cos L \\ 0 \end{pmatrix}, \quad f_2(x) = \sqrt{\frac{P}{\mu}} \begin{pmatrix} 2P/W \\ \cos L + \frac{e_x + \cos L}{W} \\ \sin L + \frac{e_y + \sin L}{W} \\ 0 \end{pmatrix},$$

where $W = 1 + e_x \cos L + e_y \sin L$. We denote by $y = (x, m)$ the full state, and its dynamic will be written as $\dot{y}(t) = f(y(t), u(t))$.

With no additional constraints, this optimal control problem has already been widely studied (see, for instance, [5, 6, 9]). We propose here to add a constraint and use it to illustrate the convergence properties of our regularization approach. The problem has already been studied in [15] with a regularization approach but without the present theoretical background regarding convergence and hybrid necessary conditions.

The low-thrust orbit transfer is achieved thanks to electro-ionic thrusters that in practice cannot operate without any source of power. If this source of power is simply the sun, then the satellite is not actuated while standing in the shadow cone of the Earth. Considering the distance Sun-Earth and Earth-satellite, we can assume that the shadow cone of the Earth is a half cylinder.

Denote by $\Omega_c(t)$ the inclination of the shadow cone in the geocentric inertial reference frame. The frontier between light and shadow is given by the zeros of the function F_{r_E} defined by $F_{r_E}(t, x) = \sin^2(L - \Omega_c) - \left(\frac{r_E W}{P}\right)^2$, if $\cos(L - \Omega_c) > 0$, where r_E is the Earth radius. Note that we need only consider this function in the neighborhood of the frontier. If we further assume that the Earth orbit is circular, then $\Omega_c(t) = \Omega_{c0} + \frac{2\pi}{\text{one year}} t$, where Ω_{c0} is the inclination of the shadow cone at the initial time $t_0 = 0$.

3.2. Hybridization of the problem. The optimal control problem settled above is naturally an (HOCP). Instead of considering the control $u(\cdot)$ to be equal to zero in the shadow cone, we use a model with discontinuous vector fields f_1 and f_2, deciding that they are equal to zero in the shadow cone. With such a model, (HOCP) is written as follows. Let $\bar{f}_0$ be the vector field on $\mathbb{R}^5$ defined by $\bar{f}_0(y) = (f_0(x), 0)$. For every time t, denote by $X_0(t)$ the interior of the shadow cone and by $X_1(t)$ the interior of its complement. Then, we seek a trajectory $y(\cdot) = (x(\cdot), m(\cdot))$, the solution on $[0, t_f]$, of the hybrid control system

$$\dot{y}(t) = \begin{cases} f(y(t), u(t)) & \text{if} \quad x(t) \in \bar{X}_1(t) \subset \mathbb{R}^4, \\ \bar{f}_0(y(t)) & \text{if} \quad x(t) \in X_0(t) \subset \mathbb{R}^4, \end{cases} \tag{3.1}$$

associated with a control $u(\cdot)$ satisfying the constraint $\|u(\cdot)\| \leqslant 1$, starting from the initial conditions $P(0) = P_0$, $e_x(0) = e_{x0}$, $e_y(0) = e_{y0}$, $L(0) = L_0$, and joining in minimal time t_f the final conditions $P(t_f) = P_f$, $e_x(t_f) = e_{xf}$, $e_y(t_f) = e_{yf}$. Note that $L(t_f)$ and $m(t_f)$ are free.

Remark 3.1. Note that the partition of the state space depends only on the position variable q. This implies that the control vector fields f_1 and f_2 are orthogonal to the boundary at a crossing point.

From the HMP recalled in section 1, every minimizing trajectory $y(\cdot)$ is the projection of an extremal $(y(\cdot), p(\cdot), p^0, u(\cdot))$ satisfying (1.3) and (1.4), and the maximized Hamiltonian vanishes at the final time t_f. The adjoint vector has a jump

(1.5) every time the trajectory crosses (transversally) a boundary. Denoting $p = (p_P, p_{e_x}, p_{e_y}, p_L, p_m)$, notice that the component $p_m(\cdot)$ is continuous since the boundary depends only on x. The transversality conditions are $p_L(t_f) = 0$, $p_m(t_f) = 0$.

Remark 3.2. Due to a large number of revolutions, and hence of crossing times, it can be expected that the application of a shooting method to this problem, using the above necessary conditions, is difficult to carry out successfully. Indeed, we observe on our numerical simulations that the domain of convergence of the method is very small. This motivates the use of our regularization procedure, as explained next. Observe, however, that the usual orbit transfer problem in the absence of a shadow cone constraint has been solved by shooting methods (see, e.g., [5, 9, 17, 18, 21]).

3.3. Regularization of the hybrid optimal control problem. Our aim is to solve this problem using a shooting method, but, as explained formerly, directly trying to find a zero of the shooting function associated with (HOCP) fails in general. The idea is then to define a regularization $(\text{OCP})_\varepsilon$ of (HOCP) such that the solving of $(\text{OCP})_\varepsilon$ using a single shooting method is easier than for (HOCP). Then, using a finite decreasing sequence of positive ε_i, we use the solution of $(\text{OCP})_{\varepsilon_i}$ to initialize the solving of $(\text{OCP})_{\varepsilon_{i+1}}$. This iteration is performed until ε_i is small enough, and the last solution is used to initialize the solving of (HOCP). This process is justified by the convergence properties stated in our main results.

We define a regularization of the original (HOCP) by thickening the light-shadow frontier as in [15]. The procedure is the following. Define an imaginary planet, with the same center than the Earth and with radius $r_\varepsilon = (1-\varepsilon)r_E,\ \varepsilon \in [0, 1)$. Figure 3.1 shows the two planets and their shadows. The plane consists of three domains $X_1^\varepsilon(t)$, $X_0^\varepsilon(t)$, and $X_2^\varepsilon(t)$, where $X_1^\varepsilon(t)$ denotes the domain completely exposed to the sun, $X_0^\varepsilon(t)$ is the imaginary shadow cone, and $X_2^\varepsilon(t)$ is the region in the shadow cone of the Earth but not in the imaginary shadow cone. The regularized system is defined as

$$\begin{aligned}
\dot{x}^\varepsilon(t) &= f_0^\varepsilon(x^\varepsilon(t)) + \frac{T_{\max}}{m(t)} b(\varepsilon, t, x^\varepsilon(t)) \left(u_1^\varepsilon(t) f_1(x^\varepsilon(t)) + u_2^\varepsilon(t) f_2(x^\varepsilon(t)) \right), \\
\dot{m}^\varepsilon(t) &= -\beta T_{\max} b(\varepsilon, t, x^\varepsilon(t)) \| u^\varepsilon(t) \|,
\end{aligned}$$

where the function $b(\cdot,\cdot,\cdot)$ is such that $b(\varepsilon, t, \cdot)$ is equal to 1 in $X_1^\varepsilon(t)$ and to 0 in $X_0^\varepsilon(t)$. We define $b(\varepsilon, t, \cdot)$ as a fifth-order polynomial based on the Euclidean distance from $x(t)$ to one of the frontiers of the shadow cones, equal to 1 on the boundary between

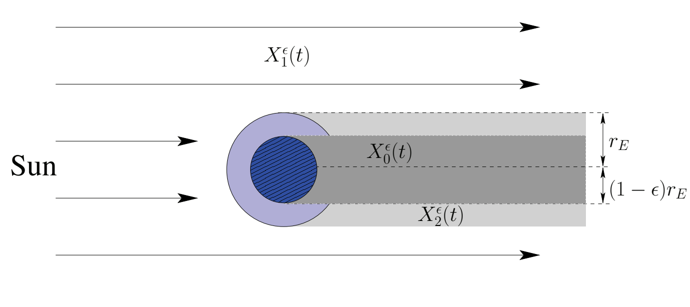

FIG. 3.1. *The two planets and their shadow cones.*

$X_1^\varepsilon(t)$ and $X_2^\varepsilon(t)$, equal to 0 on the boundary between $X_0^\varepsilon(t)$ and $X_1^\varepsilon(t)$, and such that its first and second derivatives with respect to x vanish on both boundaries.

According to the PMP, every minimizing trajectory $y^\varepsilon(\cdot)$ defined on $[0, t_f^\varepsilon]$, which is a solution of the resulting $(\text{OCP})_\varepsilon$, is the projection of an extremal $(y^\varepsilon(\cdot), p^\varepsilon(\cdot), p^{0\varepsilon}, u^\varepsilon(\cdot))$ solution of (1.10) and (1.11), with the same transversality conditions as (HOCP).

Note that the control domain $b(\varepsilon, \cdot, \cdot)\Omega$ of $(\text{OCP})_\varepsilon$ contains the control domain of (HOCP). It follows that any feasible control strategy of (HOCP) is also feasible for $(\text{OCP})_\varepsilon$, and therefore $t_f^\varepsilon \leqslant t_f$ for every $\varepsilon > 0$.

3.4. Numerical simulations. We consider the initial conditions $x(0) = (P_0, e_{x0}, e_{y0}, L_0) = (11.625\ Mm, 0.75, 0, \pi)$, $m(0) = 1500$ kg and the final conditions $x(t_f) = (P_f, e_{xf}, e_{yf}, L_f) = (42.165\ Mm, 0, 0, \text{free})$, $m(t_f)$ free. This corresponds to a low, highly eccentric initial orbit and to the geostationary final orbit. The various numerical parameters are $\mu = 5165.86248\ Mm^3/h^2$, $\beta \approx 0.028325\ h/Mm$, $r_E = 6.378\ Mm$, and $\Omega_{c0} = 0\ rad$. The value of the maximal thrust is chosen between 0.1 and 60 Newton in the next numerical simulations.

A single shooting method is used to solve the regularized and hybrid problems. The extremal flow is integrated using the adaptive step integrator *DOPRI* (see [19]). This integrator is useful to take into account the jump conditions since its dense output allows us to accurately locate the crossing times. The Newton-like method underlying our shooting method is the routine *HYBRD* of the *minpack* library.

We combine the shooting method with a continuation on the values of $T_{\max}$. The unconstrained problem is easily solved for a large value of $T_{\max}$, and the obtained solution is used to successfully initialize the solving of the regularized problem for a large value of ε, say $\varepsilon = 0.9$. This solution enables us to solve the regularized problem for any desired (low) value of $T_{\max}$ by a decreasing continuation on the parameter $T_{\max}$. Then, another decreasing continuation on the regularization parameter ε leads to the solution of the regularized problem with a small value of ε. If $\varepsilon > 0$ is small enough, then the latter solution falls into the domain of convergence of the shooting method associated with (HOCP) and is thus used to successfully initialize the shooting method on the hybrid problem. This illustrates the convergence properties derived in our main results.

We report in Table 3.1 the values of t_f^ε for several values of ε and of $T_{\max}$. As expected, the minimal time t_f^ε of $(\text{OCP})_\varepsilon$ converges to t_f as ε tends to 0. We report in Table 3.2 several numerical results illustrating the convergence of the state, adjoint vector, and control. It is important to notice in this table that it is not the magnitude of the differences but rather their decrease that illustrates the convergence. Graphic evidence of the convergence properties of $p^\varepsilon(\cdot)$ and $u^\varepsilon(\cdot)$ is illustrated in Figure 3.2 for $T_{\max} = 60\ N$. The reason for choosing such a large value of $T_{\max}$ is that the orbit transfer exhibits only one passage in the Earth's shadow cone, and the result is therefore more visible in the figure. Furthermore, we use the large value $\varepsilon = 0.9$ so as to observe clearly the continuity of $p^\varepsilon(\cdot)$. In this figure, we observe that $p^\varepsilon(\cdot)$ already mimics the behavior of $p(\cdot)$. Moreover, the adjoint vector of the hybrid problem has the expected jump.

The evolution of the zero of the shooting function with respect to ε is very smooth with respect to ε (it is nearly a straight line for the six unknowns), which is an additional hint of the nice convergence properties.

Table 3.1

$\varepsilon \setminus T_{\max}$	60 N	10 N	1 N	0.1 N
No cone	14.281	80.782	806.831	7985.138
0.9	14.337	81.307	810.273	8209.665
0.5	14.359	81.521	812.607	8298.558
0.1	14.383	81.750	815.143	8382.722
0	14.389	81.810	815.813	8406.773

Table 3.2

Componentwise $\|\cdot\|_\infty$-difference between $x^\varepsilon(\cdot)$ and $x(\cdot)$ (restricted to $[0, t_f^\varepsilon]$), the difference between initial adjoint vectors, and averaged L^1-difference of controls for $T_{\max} = 0.1\ N$.

$\cdot \setminus \varepsilon$	0.9	0.5	0.1	0.05
$\|P^\varepsilon(\cdot) - P(\cdot)_{\|[0,t_f^\varepsilon]}\|_\infty$	0.3799	0.1969	0.1529	0.0840
$\|e_x^\varepsilon(\cdot) - e_x(\cdot)_{\|[0,t_f^\varepsilon]}\|_\infty$	0.0270	0.0148	0.0034	0.0017
$\|e_y^\varepsilon(\cdot) - e_y(\cdot)_{\|[0,t_f^\varepsilon]}\|_\infty$	0.0080	0.0047	0.0011	0.0005
$\|L^\varepsilon(\cdot) - L(\cdot)_{\|[0,t_f^\varepsilon]}\|_\infty$	28.2347	28.5897	2.7437	1.2967
$\|m^\varepsilon(\cdot) - m(\cdot)_{\|[0,t_f^\varepsilon]}\|_\infty$	7.8880	4.4118	0.9477	0.4971
$\|p_P^\varepsilon(0) - p_P(0)\|$	95.3	78.6	8.19	4.97
$\|p_{e_x}^\varepsilon(0) - p_{e_x}(0)\|$	2886	2394	236.9	144.7
$\|p_{e_y}^\varepsilon(0) - p_{e_y}(0)\|$	59.3	40.4	6.84	3.55
$\|p_L^\varepsilon(0) - p_L(0)\|$	1	0.9	0.088	0.055
$\|p_m^\varepsilon(0) - p_m(0)\|$	0.6	0.5	0.05	0.029
$\|u^\varepsilon(\cdot) - u(\cdot)_{\|[0,t_f^\varepsilon]}\|_{L^1}/t_f^\varepsilon$	0.8857	0.9435	0.8762	0.4272

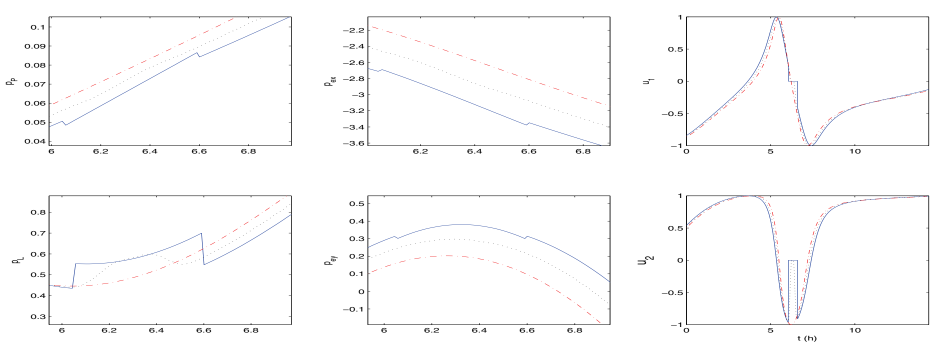

Fig. 3.2. *Control and zoom on the adjoint vector $p_x(\cdot)$ for the unconstrained (dashed), regularized ($\varepsilon = 0.9$, dotted), and hybrid (solid) problem, $T_{\max} = 60\ N$.*

REFERENCES

[1] A. Agrachev and Y. Sachkov, *Control Theory from the Geometric Viewpoint*, Encyclopaedia Math. Sci. 87, Control Theory and Optimization, II, Springer-Verlag, Berlin, 2004.

[2] M. Alamir, *A benchmark for optimal control problem solvers for hybrid nonlinear systems*, Automatica J. IFAC, 42 (2006), pp. 1593–1598.

[3] Ph. Antoine and H. Zouaki, *Etude locale de l'ensemble des points critiques d'un problème d'optimisation paramétré*, C. R. Acad. Sci. Paris Sér. I Math., 310 (1990), pp. 587–590.

[4] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser Boston, Inc., Boston, MA, 1990.

[5] B. Bonnard, J.-B. Caillau, and E. Trélat, *Geometric optimal control of elliptic Keplerian orbits*, Discrete Contin. Dyn. Syst. Ser. B, 5 (2005), pp. 929–956.

[6] B. Bonnard, L. Faubourg, and E. Trélat, *Mécanique céleste et contrôle des véhicules spatiaux*, Math. Appl. (Berlin) 51, Springer-Verlag, Berlin, 2006.
[7] M. S. Branicky, V. S. Borkar, and S. K. Mitter, *A unified framework for hybrid control: Model and optimal control theory*, IEEE Trans. Automat. Control, 43 (1998), pp. 31–45.
[8] A. Bressan and Y. Huong, *Optimal control problems on stratified domains*, Netw. Heterog. Media, 2 (2007), pp. 313–331.
[9] J.-B. Caillau and J. Noailles, *Coplanar control of a satellite around the earth*, ESAIM Cont. Optim. Calc. Var., 6 (2001), pp. 239–258.
[10] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Progr. Nonlinear Differential Equations Appl. 58, Birkhäuser Boston, Inc., Boston, MA, 2004.
[11] L. Cesari, *Optimization—Theory and Applications. Problems with Ordinary Differential Equations*, Appl. Math. 17, Springer-Verlag, New York, 1983.
[12] F. H. Clarke and R. B. Vinter, *The relationship between the maximum principle and dynamic programming*, SIAM J. Control Optim., 25 (1987), pp. 1291–1311.
[13] C. D'Apice, M. Garavello, R. Manzo, and B. Piccoli, *Hybrid optimal control: Case study of a car with gears*, Internat. J. Control, 76 (2003), pp. 1272–1284.
[14] A. V. Dmitruk and A. M. Kaganovich, *The hybrid maximum principle is a consequence of Pontryagin maximum principle*, Systems Control Lett., 57 (2008), pp. 964–970.
[15] Ch. Ferrier and R. Epenoy, *Optimal control for engines with electro-ionic propulsion under constraint of eclipse*, Acta Astronaut., 48 (2001), pp. 181–192.
[16] M. Garavello and B. Piccoli, *Hybrid necessary principle*, SIAM J. Control Optim., 43 (2005), pp. 1867–1887.
[17] J. Gergaud and T. Haberkorn, *Homotopy method for minimum consumption orbit transfer problem*, ESAIM Cont. Optim. Calc. Var., 12 (2006), pp. 294–310.
[18] J. Gergaud, T. Haberkorn, and P. Martinon, *Low thrust minimum fuel orbital transfer: An homotopic approach*, J. Guidance Cont. Dyn., 27 (2004), pp. 1046–1060.
[19] E. Hairer, S. P. Norsett, and G. Wanner, *Solving Ordinary Differential Equations* I. *Nonstiff Problems*, 2nd ed., Springer Ser. Comput. Math., Springer Verlag, 1993.
[20] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, John Wiley, New York, 1967.
[21] P. Martinon and J. Gergaud, *Using switching detection and variational equations for the shooting method*, Optimal Control Appl. Methods, 28 (2007), pp. 95–116.
[22] L. Pontryagin, V. Boltyanskii, R. Gramkrelidze, and E. Mischenko, *The mathematical theory of optimal processes*, Wiley Interscience, New York, 1962.
[23] P. Riedinger, C. Iung, and F. Kratz, *An optimal control approach for hybrid systems*, European J. Control, 9 (2003), pp. 449–458.
[24] L. Rifford and E. Trélat, *Morse-Sard type results in sub-Riemannian geometry*, Math. Ann., 332 (2005), pp. 145–159.
[25] L. Rifford and E. Trélat, *On the stabilization problem for nonholonomic distributions*, J. J. Eur. Math. Soc (JEMS), 11 (2009), pp. 223–255.
[26] S. Sager, *Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control*, J. Process Control, 19 (2009), pp. 1238–1247.
[27] M. S. Shaikh and P. E. Caines, *On the hybrid optimal control problem: Theory and algorithms*, IEEE Trans. Automat. Control, 52 (2007), pp. 1587–1603.
[28] C. J. Silva and E. Trélat, *Smooth regularization of bang-bang optimal control problems*, IEEE Trans. Automat. Control, 55 (2010), pp. 2488–2499.
[29] G. Stefani, *Regularity properties of the minimum-time map*, in Nonlinear Synthesis (Sopron, 1989), Progr. Systems Control Theory 9, Birkhäuser Boston, Inc., Boston, MA, 1991, pp. 270–282.
[30] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, Berlin, Heidelberg, New York, 1983.
[31] H. J. Sussmann, *A nonsmooth hybrid maximum principle*, in Stability and Stabilization of Nonlinear Systems (Ghent, 1999), Lecture Notes in Control and Inform. Sci. 246, Springer, London, 1999, pp. 325–354.
[32] E. Trélat, *Some properties of the value function and its level sets for affine control systems with quadratic cost*, J. Dyn. Cont. Syst., 6 (2000), pp. 511–541.
[33] A. Visintin, *Strong convergence results related to strict convexity*, Comm. Partial Differential Equations, 9 (1984), pp. 439–466.
[34] X. Xu and P. J. Antsaklis, *Optimal control of switched systems via nonlinear optimization based on direct differentiation of value functions*, Internat. J. Control, 75 (2002), pp. 1406–1426.