Smooth regularization of bang-bang optimal control problems

Cristiana J. Silva, Emmanuel Trélat

Smooth regularization of bang-bang optimal control problems

C. J. Silva E. Trélat *†

Abstract

Consider the minimal time control problem for a single-input control-affine system $\dot{x} = X(x) + u_1 Y_1(x)$ in $\mathbb{R}^n$, where the scalar control $u_1(\cdot)$ satisfies the constraint $|u_1(\cdot)| \leqslant 1$. When applying a shooting method for solving this kind of optimal control problem, one may encounter numerical problems due to the fact that the shooting function is not smooth whenever the control is bang-bang. In this article we propose the following smoothing procedure. For $\varepsilon > 0$ small, we consider the minimal time problem for the control system $\dot{x} = X(x) + u_1^\varepsilon Y_1(x) + \varepsilon \sum_{i=2}^{m} u_i^\varepsilon Y_i(x)$, where the scalar controls $u_i^\varepsilon(\cdot)$, $i = 1, \ldots, m$, with $m \geqslant 2$, satisfy the constraint $\sum_{i=1}^{m} (u_i^\varepsilon(t))^2 \leqslant 1$. We prove, under appropriate assumptions, a strong convergence result of the solution of the regularized problem to the solution of the initial problem.

Keywords: Optimal control, bang-bang control, single shooting method, Pontryagin Maximum Principle.

1 Introduction

1.1 The optimal control problem

Consider the single-input control-affine system in $\mathbb{R}^n$

$$\dot{x} = X(x) + u_1 Y_1(x), \tag{1}$$

where X and Y_1 are smooth vector fields, and the control u_1 is a measurable scalar function satisfying the constraint

$$|u_1(\cdot)| \leqslant 1. \tag{2}$$

Let M_0 and M_1 be two compact subsets of $\mathbb{R}^n$. Assume that M_1 is reachable from M_0, that is, there exist a time $T > 0$ and a control function $u_1(\cdot) \in L^\infty(0, T)$ satisfying the constraint (2), such that the trajectory $x(\cdot)$, solution of (1) with $x(0) \in M_0$, satisfies $x(T) \in M_1$.

We consider the optimal control problem **(OCP)** of determining, among all solutions of (1)–(2) steering M_0 to M_1 in minimal time.

*Université d'Orléans, UFR Sciences, Math., Labo. MAPMO, UMR 6628, Route de Chartres, BP 6759, 45067 Orléans cedex 2, France. E-mail: `cristiana.silva@univ-orleans.fr`, `emmanuel.trelat@univ-orleans.fr`

†C. J. Silva is supported by FCT through the PhD fellowship SFRH/DB/27272/2006.

Assume that the subset M_1 is reachable from M_0; it follows that the latter optimal control problem admits a solution $x(\cdot)$, associated to a control $u_1(\cdot)$, on $[0, t_f]$, where $t_f > 0$ is the minimal time (see e.g. [5] for optimal control existence theorems). According to the Pontryagin maximum principle (see [22]), there exists a non trivial absolutely continuous mapping $p(\cdot) : [0, t_f] \to \mathbb{R}^n$, called *adjoint vector*, such that

$$\begin{aligned}\dot{p}(t) &= -\frac{\partial H}{\partial x}(x(t), p(t), p^0, u(t)) \\ &= -\left\langle p(t), \frac{\partial X}{\partial x}(x(t))\right\rangle - u_1(t)\left\langle p(t), \frac{\partial Y_1}{\partial x}(x(t))\right\rangle\end{aligned} \tag{3}$$

where the function $H(x, p, p^0, u) = \langle p, X + uY_1(x)\rangle + p^0$ is called the Hamiltonian, and the maximization condition

$$H(x(t), p(t), p^0, u(t)) = \max_{|w|\leqslant 1} H(x(t), p(t), p^0, w) \tag{4}$$

holds almost everywhere on $[0, t_f]$. Moreover, $\max_{|w|\leqslant 1} H(x(t), p(t), p^0, w) = 0$ for every $t \in [0, t_f]$. The quadruple $(x(\cdot), p(\cdot), p^0, u_1(\cdot))$ is called an *extremal*. The extremal is said *normal* whenever $p^0 \neq 0$, and in that case it is usual to normalize the adjoint vector so that $p^0 = -1$; otherwise it is said *abnormal*. It follows from (4) that

$$u_1(t) = \operatorname{sign}\langle p(t), Y_1(x(t))\rangle \tag{5}$$

for almost every t, provided the (continuous) switching function $\varphi(t) = \langle p(t), Y_1(x(t))\rangle$ does not vanish on any subinterval of $[0, t_f]$. In that case, $u_1(t)$ only depends on $x(t)$ and on the adjoint vector, and it follows from (3) that the extremal $(x(\cdot), p(\cdot), p^0, u_1(\cdot))$ is completely determined by the initial adjoint vector $p(0)$. The case where the switching function may vanish on a subinterval I is related to singular trajectories. In that case, derivating the relation $\langle p(t), Y_1(x(t))\rangle = 0$ on I leads to $\langle p(t), [X, Y_1](x(t))\rangle = 0$ on I, and a second derivation leads to $\langle p(t), [X, [X, Y_1]](x(t))\rangle + u_1(t)\langle p(t), [Y_1, [X, Y_1]](x(t))\rangle = 0$ on I, which permits, under generic assumptions on the vector fields X and Y_1 (see [7, 8, 9] for genericity results related to singular trajectories), to compute the singular control $u_1(\cdot)$ on I. Under such generic assumptions, the extremal $(x(\cdot), p(\cdot), p^0, u_1(\cdot))$ is still completely determined by the initial adjoint vector.

Note that, since $x(\cdot)$ is optimal on $[0, t_f]$, and since the control system under study is autonomous, it follows that $x(\cdot)$ is solution of the optimal control problem of steering the system (1)–(2) from $x_0 = x(0)$ to $x(t)$ in minimal time.

Remark 1.1 (Remark on shooting methods). Among the numerous numerical methods that exist to solve optimal control problems, the *shooting methods* consist in solving, via Newton-like methods, the two-point or multi-point boundary value problem arising from the application of the Pontryagin maximum principle. More precisely, a Newton method is applied in order to compute a zero of the *shooting function* associated to the problem (see e.g. [27]). For the minimal time problem **(OCP)**, optimal controls may be discontinuous, and it follows that the shooting function is not smooth on $\mathbb{R}^n$ in general. Actually it may be non differentiable on switching surfaces. This implies two difficulties when using a shooting method. First, if one does not know a priori the structure of the optimal control, then it may be very difficult to initialize properly the shooting method, and in general the iterates of the underlying Newton method will be unable to cross barriers generated by switching surfaces (see e.g. [16]). Second, the numerical computation of the shooting function and of its differential may be intricate since the shooting function is not continuously differentiable.

This observation is one of the possible motivations of the regularization procedure considered in this article. Indeed, the shooting functions related to the smooth optimal control problems described next are smooth, and in our main result we derive nice convergence properties.

1.2 The regularization procedure

Let ε be a positive real parameter and let $Y_2, \ldots, Y_m$ be $m-1$ arbitrary smooth vector fields on $\mathbb{R}^n$, where $m \geqslant 2$ is an integer. Consider the control-affine system

$$\dot{x}^\varepsilon(t) = X\left(x^\varepsilon(t)\right) + u_1^\varepsilon(t) Y_1\left(x^\varepsilon(t)\right) + \varepsilon \sum_{i=2}^{m} u_i^\varepsilon(t) Y_i\left(x^\varepsilon(t)\right), \tag{6}$$

where the control $u^\varepsilon(t) = (u_1^\varepsilon(t), \ldots, u_m^\varepsilon(t))$ satisfies the constraint

$$\sum_{i=1}^{m} \left(u_i^\varepsilon(t)\right)^2 \leqslant 1. \tag{7}$$

Consider the optimal control problem $\textbf{(OCP)}_\varepsilon$ of determining a trajectory $x^\varepsilon(\cdot)$, solution of (6)–(7) on $[0, t_f^\varepsilon]$, such that $x^\varepsilon(0) \in M_0$ and $x^\varepsilon(t_f^\varepsilon) \in M_1$, and minimizing the time of transfer t_f^ε. The parameter ε is viewed as a penalization parameter, and it is expected that any solution $x^\varepsilon(\cdot)$ of $\textbf{(OCP)}_\varepsilon$ tends to a solution $x(\cdot)$ of **(OCP)** as ε tends to zero. It is our aim to derive such a result.

According to the Pontryagin maximum principle, any optimal solution $x^\varepsilon(\cdot)$ of $\textbf{(OCP)}_\varepsilon$, associated with controls $(u_1^\varepsilon, \ldots, u_m^\varepsilon)$ satisfying the constraint (7), is the projection of an extremal $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), p^{0\varepsilon}, u^\varepsilon(\cdot))$ such that

$$\begin{aligned} \dot{p}^\varepsilon(t) &= -\frac{\partial H^\varepsilon}{\partial x}(x^\varepsilon(t), p^\varepsilon(t), p^{0\varepsilon}, u^\varepsilon(t)) \\ &= -\left\langle p^\varepsilon(t), \frac{\partial X}{\partial x}(x^\varepsilon(t)) \right\rangle - u_1^\varepsilon(t) \left\langle p^\varepsilon(t), \frac{\partial Y_1}{\partial x}(x^\varepsilon(t)) \right\rangle \\ &\quad - \varepsilon \sum_{i=2}^{m} u_i^\varepsilon(t) \left\langle p^\varepsilon(t), \frac{\partial Y_i}{\partial x}(x^\varepsilon(t)) \right\rangle \end{aligned} \tag{8}$$

where $H^\varepsilon(x, p, p^0, u) = \langle p, X(x) + u_1 Y_1(x) + \varepsilon \sum_{i=2}^{m} u_i Y_i(x) \rangle + p^0$ is the Hamiltonian, and

$$H(x^\varepsilon(t), p^\varepsilon(t), p^{0\varepsilon}, u^\varepsilon(t)) = \max_{\sum_{i=1}^{m} w_i^2 \leqslant 1} H(x^\varepsilon(t), p^\varepsilon(t), p^{0\varepsilon}, w) \tag{9}$$

almost everywhere on $[0, t_f^\varepsilon]$. Moreover, the maximized Hamiltonian is equal to 0 on $[0, t_f^\varepsilon]$. The maximization condition (9) turns into

$$\begin{aligned} & u_1^\varepsilon(t) \langle p^\varepsilon(t), Y_1(x^\varepsilon(t)) \rangle + \varepsilon \sum_{i=2}^{m} u_i^\varepsilon(t) \langle p^\varepsilon(t), Y_i(x^\varepsilon(t)) \rangle \\ &= \max_{\sum_{i=1}^{m} w_i^2 \leqslant 1} \left(w_1 \langle p^\varepsilon(t), Y_1(x^\varepsilon(t)) \rangle + \varepsilon \sum_{i=2}^{m} w_i \langle p^\varepsilon(t), Y_i(x^\varepsilon(t)) \rangle \right), \end{aligned} \tag{10}$$

and two cases may occur: either the maximum is attained in the interior of the domain, or it is attained on the boundary. In the first case, there must hold $\langle p^\varepsilon(t), Y_i(x^\varepsilon(t)) \rangle = 0$, for every $i \in \{1,\dots,m\}$; in particular, if the m functions $t \mapsto \langle p^\varepsilon(t), Y_i(x^\varepsilon(t))\rangle$, $i=1,\dots,m$, do not vanish simultaneously, then the maximum is attained on the boundary of the domain. Throughout the article, we assume that the integer m and the vector fields $Y_2,\dots,Y_m$ are chosen such that

$$\mathrm{Span}\{Y_i \mid i=1,\dots,m\} = \mathbb{R}^n. \tag{11}$$

Under this assumption, the maximization condition (10) yields

$$\begin{aligned} u_1^\varepsilon(t) &= \frac{\langle p^\varepsilon(t), Y_1(x^\varepsilon(t))\rangle}{\sqrt{\langle p^\varepsilon(t), Y_1(x^\varepsilon(t))\rangle^2 + \varepsilon^2 \sum_{i=2}^m \langle p^\varepsilon(t), Y_i(x^\varepsilon(t))\rangle^2}}, \\ u_i^\varepsilon(t) &= \frac{\varepsilon\langle p^\varepsilon(t), Y_i(x^\varepsilon(t))\rangle}{\sqrt{\langle p^\varepsilon(t), Y_1(x^\varepsilon(t))\rangle^2 + \varepsilon^2 \sum_{i=2}^m \langle p^\varepsilon(t), Y_i(x^\varepsilon(t))\rangle^2}}, \quad i=2,\dots,m, \end{aligned} \tag{12}$$

for almost every $t \in [0, t_f^\varepsilon]$, and moreover the control functions $u_i^\varepsilon(\cdot)$, $i=1,\dots,m$ are smooth functions of t (so that the above formula holds actually for every $t \in [0,t_f^\varepsilon]$). Indeed, to prove this fact, it suffices to prove that the functions $t \mapsto \langle p^\varepsilon(t), Y_i(x^\varepsilon(t)\rangle$, $i=1,\dots,m$ do not vanish simultaneously. The argument goes by contradiction: if these functions would vanish simultaneously, then, using the assumption (11), this would imply that $p^\varepsilon(t)=0$ for some t; combined with the fact that the maximized Hamiltonian is equal to zero along any extremal, it would follow that $p^{0\varepsilon}=0$, and this would raise a contradiction since the adjoint vector $(p^\varepsilon(\cdot), p^{0\varepsilon})$ of the maximum principle must be nontrivial.

Remark 1.2. The assumption (11) requires that $m \geqslant n$. One may however wish to choose $m=2$, i.e., to add only one new vector field Y_2, in the regularization procedure. In that case, the assumption (11) does not hold whenever $n>3$, and then two problems may occur: first, in the maximization condition (10) the maximum is not necessarily obtained at the boundary, i.e., the expressions (12) do not necessarily hold, and second, the controls $u_i^\varepsilon(\cdot)$, $i=1,\dots,m$ are not necessarily continuous (the continuity is used in a crucial way in the proof of our main result). These two problems are however not likely to occur, and we provide in Section 3 some comments on the generic validity of (12) and on the smoothness of the regularized controls, in the case $m=2$.

From (12), it is expected that $u_1^\varepsilon(\cdot)$ converges to $u_1(\cdot)$ and $u_i^\varepsilon(\cdot)$, $i=2,\dots,m$, tend to zero, in some topology to specify. This fact is derived rigorously in our main result.

Theorem 1. *Assume that the problem* ***(OCP)*** *has a unique solution $x(\cdot)$, defined on $[0,t_f]$, associated with a control $u_1(\cdot)$ on $[0,t_f]$. Moreover, assume that $x(\cdot)$ has a unique extremal lift (up to a multiplicative scalar), that is moreover normal, and denoted $(x(\cdot),p(\cdot),-1,u_1(\cdot))$.*

Then, under the assumption (11), there exists $\varepsilon_0>0$ such that, for every $\varepsilon\in(0,\varepsilon_0)$, the problem ***(OCP)***$_\varepsilon$ *has at least one solution $x^\varepsilon(\cdot)$, defined on $[0,t_f^\varepsilon]$ with $t_f^\varepsilon \leqslant t_f$, associated with a smooth control $u^\varepsilon=(u_1^\varepsilon,\dots,u_m^\varepsilon)$ satisfying the constraint (7), 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[1] to $x(\cdot)$, and $p^\varepsilon(\cdot)$ converges uniformly to $p(\cdot)$ on $[0,t_f]$;*

[1]We consider any continuous extension of $x^\varepsilon(\cdot)$ on $[0,t_f]$.

- $u_1^\varepsilon(\cdot)$ *converges weakly*[2] *to* $u_1(\cdot)$ *for the weak* $L^1(0,t_f)$ *topology.*

If the control u_1 *is moreover bang-bang, i.e., if the (continuous) switching function* $\varphi(t) = \langle p(t), Y_1(x(t))\rangle$ *does not vanish on any subinterval of* $[0,t_f]$*, then* $u_1^\varepsilon(\cdot)$ *converges to* $u_1(\cdot)$ *and* $u_i^\varepsilon(\cdot)$*,* $i=2,\dots,m$*, converge to* 0 *almost everywhere on* $[0,t_f]$*, and thus in particular for the strong* $L^1(0,t_f)$ *topology.*

Remark 1.3. We provide in Section 3 some further comments and two examples with numerical simulations in order to illustrate Theorem 1. The first example is the Rayleigh problem, on which the minimal time trajectory is bang-bang, and almost everywhere convergence of the regularized control can be observed, accordingly to our main result. Our second example involves a singular arc and we prove and observe that oscillations appear, so that the regularized control weakly converges, but fails to converge almost everywhere.

Remark 1.4. It is assumed that the problem **(OCP)** has a unique solution $x(\cdot)$, having a unique extremal lift that is normal. Such an assumption holds true whenever the minimum time function (the value function of the optimal control problem) enjoys differentiability properties (see e.g. [2, 10] for a precise relationship, see also [4, 23, 24, 26] for results on the size of the set where the value function is differentiable).

If one removes these uniqueness assumptions, then the following result still holds, provided that every extremal lift of every solution of **(OCP)** is normal. Consider the topological spaces $\mathcal{X} = C^0([0,t_f],\mathbb{R}^n)$, endowed with the uniform convergence topology, and $\mathcal{Y} = L^\infty(0,t_f;[-1,1])$, endowed with the weak star topology. In the following statement, the space $\mathcal{X}\times\mathcal{X}\times\mathcal{Y}$ is endowed with the resulting product topology. For every $\varepsilon\in(0,\varepsilon_0)$, let $x^\varepsilon(\cdot)$ be a solution of $\mathbf{(OCP)}_\varepsilon$, and let $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), -1, u^\varepsilon(\cdot))$ be a (normal) extremal lift of $x^\varepsilon(\cdot)$. Then, every closure point in $\mathcal{X}\times\mathcal{X}\times\mathcal{Y}$ of the family of triples $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), u_1^\varepsilon(\cdot))$ is a triple $(\bar{x}(\cdot), \bar{p}(\cdot), \bar{u}_1(\cdot))$, where $\bar{x}(\cdot)$ is an optimal solution of **(OCP)**, associated with the control $\bar{u}_1(\cdot)$, having as a normal extremal lift the 4-tuple $(\bar{x}(\cdot), \bar{p}(\cdot), -1, \bar{u}_1(\cdot))$. The rest of the statement of Theorem 1 still holds with an obvious adaptation in terms of closure points.

Remark 1.5. When applying a shooting method to the problem $\mathbf{(OCP)}_\varepsilon$, one is not ensured to determine an optimal solution, but only an extremal solution that is not necessarily optimal[3]. Notice however that the arguments of the proof of Theorem 1 permit to prove the following statement. Assume that there is no abnormal extremal among the set of extremals obtained by applying the Pontryagin maximum principle to the problem **(OCP)**; then, for $\varepsilon>0$ small enough, every extremal solution of $\mathbf{(OCP)}_\varepsilon$ is normal, and, using the notations of the previous remark, every closure point of such extremal solutions is a normal extremal solution of **(OCP)**.

Remark 1.6. There is a large literature dealing with optimal control problems depending on some parameters, involving state, control or mixed constraints, using a stability and sensitivity analysis in order to investigate the dependence of the optimal solution with respect to parameters (see e.g. [11, 12, 13, 14, 15, 17, 18, 19, 20, 21] and references therein). In the sensitivity approach, under second order sufficient conditions, results are derived that prove that the solutions of the parametrized problems, as well as the associated Lagrange multipliers, are Lipschitz continuous or directionally differentiable functions of the parameter. We stress however that Theorem 1 cannot be derived from these former works. Indeed, in these references, the results rely on second order sufficient conditions and certain regularity

[2] It means that $\int_0^{t_f} u_1^\varepsilon(t)g(t)dt \to \int_0^{t_f} u_1(t)g(t)dt$ as $\varepsilon\to 0$, for every $g\in L^1(0,t_f)$, and where the function $u_1^\varepsilon(\cdot)$ is extended continuously on $[0,t_f]$.

[3] This fact is well known, due to the fact that the Pontryagin maximum principle is a only first order necessary condition for optimality; sufficient conditions do exist but this is outside of the scope of that paper.

conditions on the initial problem. In our work we do not assume any second order sufficient condition; our approach is different from the usual sensitivity analysis and is rather, in some sense, a topological approach.

2 Proof of the main result

2.1 Preliminaries, Pontryagin maximum principle

In this subsection, we recall elements of a standard proof of the maximum principle using needle-like variations (see [22]), which are needed to derive our main result.

Consider a general control system

$$\dot{x}(t) = f(x(t), u(t)),\ x(0) = x_0, \tag{13}$$

where $x_0 \in \mathbb{R}^n$ is fixed, $f : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$ is smooth, the control u is a bounded measurable function taking its values in a measurable subset Ω of $\mathbb{R}^m$. A control function $u \in L^\infty([0, t_f], \mathbb{R}^m)$ is said *admissible* on $[0, t_f]$ if the trajectory $x(\cdot)$, solution of (13) associated to u and such that $x(0) = x_0$, is well defined on $[0, t_f]$, and the *end-point mapping* E is then defined by

$$E(x_0, t_f, u) = x(t_f).$$

The set of admissible controls on $[0, t_f]$ is denoted $\mathcal{U}_{t_f,\mathbb{R}^m}$, and the set of admissible controls on $[0, t_f]$ taking their values in Ω is denoted $\mathcal{U}_{t_f,\Omega}$. The set $\mathcal{U}_{t_f,\mathbb{R}^m}$, endowed with the standard topology of $L^\infty([0, t_f], \mathbb{R}^m)$, is open, and the end-point mapping is smooth on $\mathcal{U}_{t_f,\mathbb{R}^m}$.

Let $x_1 \in \mathbb{R}^n$. Consider the optimal control problem $(\mathcal{P})$ of determining a trajectory solution of (13) steering x_0 to x_1 in minimal time[4]. In other words, this is the problem of minimizing t_f among all admissible controls $u \in L^\infty([0, t_f], \Omega)$ satisfying the constraint $E(x_0, t_f, u) = x_1$.

For every $t \geqslant 0$, define the *accessible set* $A_\Omega(x_0, t)$ as the image of the mapping $E(x_0, t, \cdot) : \mathcal{U}_{t,\Omega} \to \mathbb{R}^n$, with the agreement $A_\Omega(x_0, 0) = \{x_0\}$. Moreover, define

$$A_\Omega(x_0, \leqslant t) = \bigcup_{0 \leqslant s \leqslant t} A_\Omega(x_0, s).$$

The set $A_\Omega(x_0, \leqslant t)$ coincides with the image of the mapping $E(x_0, \cdot, \cdot) : [0, t] \times \mathcal{U}_{t,\Omega} \to \mathbb{R}^n$.

Let u be a minimal time control on $[0, t_f]$ for the problem $(\mathcal{P})$, and denote by $x(\cdot)$ the trajectory solution of (13) associated to the control u on $[0, t_f]$. Then the point $x_1 = x(t_f)$ belongs to the boundary of $A_\Omega(x_0, \leqslant t_f)$. This geometric property is at the basis of the proof of the Pontryagin maximum principle.

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

Needle-like variations. Let $t_1 \in [0, t_f)$ and $u_1 \in \Omega$. For $\eta_1 > 0$ such that $t_1 + \eta_1 \leqslant t_f$, the needle-like variation $\pi_1 = \{t_1, \eta_1, u_1\}$ of the control u is defined by

$$u_{\pi_1}(t) = \begin{cases} u_1 & \text{if } t \in [t_1, t_1 + \eta_1], \\ u(t) & \text{otherwise.} \end{cases}$$

[4]Note that we consider here a problem with fixed extremities, for simplicity of presentation. All what follows however easily extends to the case of initial and final subsets.

The control u_{π_1} takes its values in Ω. It is not difficult to prove that, if $\eta_1 > 0$ is small enough, then the control u_{π_1} is admissible, i.e., the trajectory $x_{\pi_1}(\cdot)$ associated with u_{π_1} and starting from $x_{\pi_1}(0) = x_0$ is well defined on $[0, t_f]$. Moreover, $x_{\pi_1}(\cdot)$ converges uniformly to $x(\cdot)$ on $[0, t_f]$ whenever η_1 tends to 0.

Recall that t_1 is a Lebesgue point of the function $t \mapsto f(x(t), u(t))$ on $[0, t_f)$ whenever

$$\lim_{h\to 0} \frac{1}{h} \int_{t_1}^{t_1+h} f(x(t), u(t))dt = f(x(t_1), u(t_1)),$$

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 $\eta_1 > 0$ small enough, and u_{π_1} be a needle-like variation of u, with $\pi_1 = \{t_1, \eta_1, u_1\}$. For every $t \in [t_1, t_f]$, define the variation vector $v_{\pi_1}(t)$ as the solution on $[t_1, t_f]$ of the Cauchy problem

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

Then, it is not difficult to prove that

$$x_{\pi_1}(t_f) = x(t_f) + \eta_1 v_{\pi_1}(t_f) + \mathrm{o}(\eta_1) \tag{14}$$

(see e.g. [22] for details).

Note that, for every $\alpha > 0$, the variation $\{t_1, \alpha\eta_1, u_1\}$ generates the variation vector αv_{π_1}, it follows that the set of variation vectors at time t is a cone.

For every $t \in (0, t_f]$, the *first Pontryagin cone* at time t, denoted $K(t)$, is the smallest closed convex cone containing all variation vectors $v_{\pi_1}(t)$ for all Lebesgue points t_1 such that $0 < t_1 < t$.

An immediate iteration leads to the following result. Let $t_1 < t_2 < \cdots < t_p$ be Lebesgue points of the function $t \mapsto f(x(t), u(t))$ on $(0, t_f)$, and $u_1, \ldots, u_p$ be points of Ω. 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 $v_{\pi_i}(\cdot)$ the associated variation vectors, defined as above. Define the variation

$$\pi = \{t_1, \ldots, t_p, \eta_1, \ldots, \eta_p, u_1, \ldots, u_p\}$$

of the control u on $[0, T]$ by

$$u_\pi(t) = \begin{cases} u_i & \text{if } t_i \leqslant t \leqslant t_i + \eta_i, \quad i = 1, \ldots, p, \\ u(t) & \text{otherwise.} \end{cases} \tag{15}$$

Let $x_\pi(\cdot)$ be the solution of (13) corresponding to the control u_π on $[0, t_f]$ and such that $x_\pi(0) = x_0$. Then,

$$x_\pi(t_f) = x(t_f) + \sum_{i=1}^{p} \eta_i v_{\pi_i}(t_f) + \mathrm{o}\Big(\sum_{i=1}^{p} \eta_i\Big). \tag{16}$$

The variation formula (16) shows that every combination with positive coefficients of variation vectors (taken at distinct Lebesgue points) provides the point $x(t) + v_\pi(t)$, where

$$v_\pi(t) = \sum_{i=1}^{p} \eta_i v_{\pi_i}(t), \tag{17}$$

which belongs, up to the remainder term, to the accessible set $A_\Omega(x_0,t)$ at time t for the system (13) starting from the point x_0. In this sense, the first Pontryagin cone serves as an estimate of the accessible set $A_\Omega(x_0,t)$.

Since we deal with a minimal time problem, we must rather consider the set $A_\Omega(x_0,\leqslant t)$, which leads to introduce also oriented time variations, as follows. Assume first that $x(\cdot)$ is differentiable[5] at time t_f. Let $\delta>0$ small enough; then, with the above notations,

$$x_\pi(t_f-\delta)=x(t_f)+\sum_{i=1}^{p}\eta_i v_{\pi_i}(t_f)-\delta f(x(t_f),u(t_f))+\mathrm{o}\Big(\delta+\sum_{i=1}^{p}\eta_i\Big). \tag{18}$$

Then, define the cone $K_1(t_f)$ as the smallest closed convex cone containing $K(t_f)$ and the vector $-f(x(t_f),u(t_f))$.
If $x(\cdot)$ is not differentiable at time t_f, then the above construction is slightly modified, by replacing $f(x(t_f),u(t_f))$ with any closure point of the corresponding difference quotient in an obvious way.

Conic implicit function theorem. We next provide a *conic implicit function theorem*, which is at the basis of the proof of the maximum principle (see e.g. [1] for a proof).

Lemma 2.1. *Let $C\subset I\!R^m$ be a convex subset of $I\!R^m$ with nonempty interior, of vertex 0, and $F:C\to I\!R^n$ be a Lipschitzian mapping such that $F(0)=0$ and F is differentiable in the sense of Gâteaux at 0. Assume that $dF(0).\mathrm{Cone}(C)=I\!R^n$, where $\mathrm{Cone}(C)$ stands for the (convex) cone generated by elements of C. Then 0 belongs to the interior of $F(V\cap C)$, for every neighborhood V of 0 in $I\!R^m$.*

Lagrange multipliers and Pontryagin maximum principle. We next restrict the end-point mapping to time and needle-like variations. Let p be a positive integer. Set

$$\mathbb{R}_+^{p+1}=\{(\delta,\eta_1,\dots,\eta_p)\in\mathbb{R}^{p+1}\mid \delta\geqslant 0,\eta_1\geqslant 0,\dots,\eta_p\geqslant 0\}.$$

Let $t_1<\cdots<t_p$ be Lebesgue points of the function $t\mapsto f(x(t),u(t))$ on $(0,t_f)$, and $u_1,\dots,u_p$ be points of Ω. Let V be a small neighborhood of 0 in $\mathbb{R}^p$. Define the mapping $F:V\cap\mathbb{R}_+^{p+1}\to\mathbb{R}^n$ by

$$F(\delta,\eta_1,\dots,\eta_p)=x_\pi(t_f-\delta),$$

where π is the variation $\pi=\{t_1,\dots,t_p,\eta_1,\dots,\eta_p,u_1,\dots,u_p\}$ and $\delta\geqslant 0$ is small enough so that $t_p<t_f-\delta$. If V is small enough, then F is well defined; moreover this mapping is clearly Lipschitzian, and $F(0)=x(t_f)$. From (18), F is Gâteaux differentiable on the conic neighborhood $V\cap\mathbb{R}_+^{p+1}$ of 0.

If the cone $K_1(t_f)$ would coincide with $\mathbb{R}^n$, then there would exist $\delta\geqslant 0$, an integer p and variations $\pi_i=\{t_i,\eta_i,u_i\}$, $i=1,\dots,p$, such that $F_0'\mathbb{R}_+^{p+1}=\mathbb{R}^n$, and then Lemma 2.1 would imply that the point $x(t_f)$ would belong to the interior of the accessible set $A_\Omega(x_0,\leqslant t_f)$, which would raise a contradiction.

Therefore the convex cone $K_1(t_f)$ is not equal to $\mathbb{R}^n$. As a consequence, there exists $\psi\in\mathbb{R}^n\setminus\{0\}$ called *Lagrange multiplier* such that $\langle\psi,v(t_f)\rangle\leqslant 0$ for every variation vector $v(t_f)\in K(t_f)$ and $\langle\psi,f(x(t_f),u(t_f))\rangle\geqslant 0$ (at least whenever $x(\cdot)$ is differentiable at time t_f; otherwise replace $f(x(t_f),u(t_f))$ with any closure point of the corresponding difference quotient).

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

These inequalities then permit to prove the maximum principle (see [22]), according to which the trajectory $x(\cdot)$, associated to the optimal control u, is the projection of an *extremal* $(x(\cdot), p(\cdot), p^0, u(\cdot))$ (called extremal lift), where $p^0 \leqslant 0$ and $p(\cdot) : [0, t_f] \to \mathbb{R}^n$ is a nontrivial absolutely continuous mapping called *adjoint vector*, such that

$$\dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t), u(t)), \ \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t), u(t)),$$

almost everywhere on $[0, t_f]$, where $H(x, p, u) = \langle p, f(x, u)\rangle + p^0$ is the Hamiltonian, and $H(x(t), p(t), p^0, u(t)) = M(x(t), p(t), p^0)$ almost everywhere on $[0, t_f]$, where $M(x(t), p(t), p^0) = \max_{w\in\Omega} H(x(t), p(t), p^0, w)$. Moreover, the function $t \mapsto M(x(t), p(t), p^0)$ is identically equal to zero on $t \in [0, t_f]$.

The relation between the above Lagrange multiplier and $(p(\cdot), p^0)$ is that the adjoint vector can be constructed so that

$$\psi = p(t_f) \quad \text{and} \quad p^0 = -\max_{w\in\Omega}\langle \psi, f(x(t_f), w)\rangle. \tag{19}$$

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

If $p^0 < 0$ the extremal is said *normal*, and in this case, since the Lagrange multiplier is defined up to a multiplicative scalar, it is usual to normalize it so that $p^0 = -1$. If $p^0 = 0$ the extremal is said *abnormal*.

Remark 2.1. 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 $\max_{w\in\Omega}\langle \psi, f(x(t_f), w)\rangle = 0$. In that case, one has $p(t_f) = \psi$, up to a multiplicative scalar.

Define the *first extended Pontryagin cone* $\tilde{K}(t)$ along $x(\cdot)$ as the smallest closed convex cone containing $K_1(t)$ and $f(x(t), u(t))$ (at least whenever $x(\cdot)$ is differentiable at time t; otherwise replace $f(x(t), u(t))$ with any closure point of the corresponding difference quotient).

Note that $x(\cdot)$ does not admit any abnormal extremal lift on $[0, t_f]$ if and only if $\tilde{K}(t_f) = \mathbb{R}^n$.

The following remark easily follows from the above considerations.

Remark 2.2. For the optimal trajectory $x(\cdot)$, the following statements are equivalent:

- The trajectory $x(\cdot)$ has a unique extremal lift (up to a multiplicative scalar); moreover, the extremal lift is normal.
- $K_1(t_f)$ is a half-space and $\tilde{K}(t_f) = \mathbb{R}^n$.
- $K(t_f)$ is a half-space and $\max_{w\in\Omega}\langle \psi, f(x(t_f), w)\rangle > 0$.

This remark permits to translate the assumptions of our main result into geometric considerations.

2.2 Convergence results for the problem $(\mathrm{OCP})_\varepsilon$

This subsection contains the proof of Theorem 1, that follows from Lemmas 2.2–2.10.

From now on, assume that all assumptions of Theorem 1 hold. We denote the end-point mapping for the system (6) by

$$E(\varepsilon, x_0, t_f, u^\varepsilon) = x^\varepsilon(t_f),$$

where $x^\varepsilon(\cdot)$ is the solution of (6) associated with the control $u^\varepsilon(\cdot) = (u_1^\varepsilon(\cdot), \ldots, u_m^\varepsilon(\cdot))$ and such that $x^\varepsilon(0) = x_0$. By extension, the end-point mapping for the system (1) corresponds to $\varepsilon = 0$,

$$E(0, x_0, t_f, (u_1, 0, \ldots, 0)) = x(t_f),$$

where $x(\cdot)$ is the solution of (1) associated with the control $u_1(\cdot)$ and such that $x(0) = x_0$. It will be also denoted $E(x_0, t_f, u_1) = E(0, x_0, t_f, (u_1, 0, \ldots, 0)) = x(t_f)$.

In the sequel, we denote by $u_1(\cdot)$ the minimal time control steering the system (1) from M_0 to M_1 in time t_f.

We first derive the following existence result.

Lemma 2.2. *For every $\varepsilon > 0$[6], the problem $\boldsymbol{(OCP)}_\varepsilon$ admits at least one solution $x^\varepsilon(\cdot)$, associated with a control $u^\varepsilon(\cdot) = (u_1^\varepsilon(\cdot), \ldots, u_m^\varepsilon(\cdot))$ satisfying the constraint* (7) *on $[0, t_f^\varepsilon]$. Moreover, $0 \leqslant t_f^\varepsilon \leqslant t_f$.*

Proof. Knowing that the constrained minimization problem

$$\left\{\begin{array}{l} \min t_f \\ |u_1| \leqslant 1,\ E(0, x_0, t_f, (u_1, 0, \ldots, 0)) = x_1 \\ x_0 \in M_0,\ x_1 \in M_1 \end{array}\right.$$

has a solution, it is our aim to prove that the problem

$$\left\{\begin{array}{l} \min t_f^\varepsilon \\ u^\varepsilon = (u_1^\varepsilon, \ldots, u_m^\varepsilon),\ \displaystyle\sum_{i=1}^m (u_i^\varepsilon)^2 \leqslant 1,\ E(\varepsilon, x_0, t_f^\varepsilon, u^\varepsilon) = x_1 \\ x_0 \in M_0,\ x_1 \in M_1 \end{array}\right.$$

has a solution, for every $\varepsilon > 0$. First of all, we claim that, for every $\varepsilon > 0$, the subset M_1 is reachable from the subset M_0, i.e., it is possible to solve the equation

$$E(\varepsilon, x_0, t_f^\varepsilon, u^\varepsilon) = x_1$$

with a control $u^\varepsilon = (u_1^\varepsilon, \ldots, u_m^\varepsilon)$ satisfying the constraint $\sum_{i=1}^m (u_i^\varepsilon)^2 \leqslant 1$, and with some $x_0 \in M_0$ and $x_1 \in M_1$. Indeed, if $u_i^\varepsilon = 0$, $i = 2, \ldots, m$, then the system (6) coincides with the system (1), and it suffices to choose $u_1^\varepsilon = u_1$ and the corresponding initial and final points. The existence of a minimal time control steering the system (6) from M_0 to M_1 is then a standard fact to derive for such a control-affine system (see e.g. [5], and note that M_0 and M_1 are compact). Moreover, the minimal time t_f^ε for the problem $\mathbf{(OCP)}_\varepsilon$ is less or equal than the minimal time t_f for the initial problem. □

As explained in Section 1.2, for $\epsilon > 0$ fixed, and assuming that (11) is satisfied, it follows from the Pontryagin maximum principle applied to $\mathbf{(OCP)}_\varepsilon$ that $x^\varepsilon(\cdot)$ is the projection of an extremal $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), p^{0\varepsilon}, u^\varepsilon(\cdot))$ such that

$$\begin{aligned} \dot{p}^\varepsilon(t) = &- \left\langle p^\varepsilon(t), \frac{\partial X}{\partial x}(x^\varepsilon(t)) \right\rangle - u_1^\varepsilon(t) \left\langle p^\varepsilon(t), \frac{\partial Y_1}{\partial x}(x^\varepsilon(t)) \right\rangle \\ &- \varepsilon \sum_{i=2}^m u_i^\varepsilon(t) \left\langle p^\varepsilon(t), \frac{\partial Y_i}{\partial x}(x^\varepsilon(t)) \right\rangle \end{aligned}$$

[6]Note that ε is not needed to be small.

and

$$u_1^\varepsilon(t) = \frac{\langle p^\varepsilon(t), Y_1(x^\varepsilon(t))\rangle}{\sqrt{\langle p^\varepsilon(t), Y_1(x^\varepsilon(t))\rangle^2 + \varepsilon^2 \sum_{i=2}^m \langle p^\varepsilon(t), Y_i(x^\varepsilon(t))\rangle^2}},$$

$$u_i^\varepsilon(t) = \frac{\varepsilon\langle p^\varepsilon(t), Y_i(x^\varepsilon(t))\rangle}{\sqrt{\langle p^\varepsilon(t), Y_1(x^\varepsilon(t))\rangle^2 + \varepsilon^2 \sum_{i=2}^m \langle p^\varepsilon(t), Y_i(x^\varepsilon(t))\rangle^2}}, \quad i = 2, \dots, m.$$

We stress on the fact that the controls u_i^ε, $i = 1, \dots, m$, are continuous functions of t.

Lemma 2.3. *If $\varepsilon > 0$ tends to 0, then t_f^ε converges to t_f, $u_1^\varepsilon(\cdot)$ converges to $u_1(\cdot)$ in $L^\infty(0, t_f)$ for the weak star topology, and $x^\varepsilon(\cdot)$ converges to $x(\cdot)$ uniformly on $[0, t_f]$.*

Proof. Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be an arbitrary sequence of positive real numbers converging to 0 as n tends to $+\infty$. From Lemma 2.2, $0 \leqslant t_f^{\varepsilon_n} \leqslant t_f$, hence, up to a subsequence, $(t_f^{\varepsilon_n})_{n\in\mathbb{N}}$ converges to some $T \geqslant 0$ such that $T \leqslant t_f$. By definition, the sequence of controls $(u_1^{\varepsilon_n}(\cdot))_{n\in\mathbb{N}}$ is bounded in $L^\infty(0, t_f)$ (with the agreement that the function $u_1^{\varepsilon_n}(\cdot)$ is extended on $(t_f^{\varepsilon_n}, t_f]$ e.g. by 0). Therefore, up to subsequence, it converges weakly to some control $\bar{u}_1(\cdot) \in L^\infty(0, t_f)$ for the weak star topology. In particular, it converges weakly to $\bar{u}_1(\cdot) \in L^2(0, t_f)$ for the weak topology of $L^2(0, t_f)$. The limit control $\bar{u}_1(\cdot)$ satisfies $|\bar{u}_1(\cdot)| \leqslant 1$, almost everywhere on $[0, t_f]$. To prove this fact, consider the set

$$\mathcal{V} = \{g \in L^2(0, t_f) \ \mid \ |g(t)| \leqslant 1 \ \text{ almost everywhere on } [0, t_f]\}.$$

For every integer n, $u_1^{\varepsilon_n}(\cdot) \in \mathcal{V}$; moreover $\mathcal{V}$ is a convex closed (for the strong topology) subset of $L^2(0, t_f)$, and hence is a convex closed (for the weak topology) subset of $L^2(0, t_f)$. It follows that $\bar{u}_1 \in \mathcal{V}$.

Since M_0 and M_1 are compact, it follows that, up to a subsequence, $x^{\varepsilon_n}(0)$ converges to some $\bar{x}_0 \in M_0$, and $x^{\varepsilon_n}(t_f^{\varepsilon_n})$ converges to some $\bar{x}_1 \in M_1$.

Let $\bar{x}(\cdot)$ denote the solution of the system (1), associated with the control $\bar{u}_1(\cdot)$ on $[0, T]$, and such that $\bar{x}(0) = \bar{x}_0$. Since the control systems under consideration are control-affine, it is not difficult to prove that the weak convergence of controls implies the uniform convergence of corresponding trajectories (see [28] for details). In particular, it follows that $\bar{x}(T) = \bar{x}_1$.

Therefore, we have proved that the control $\bar{u}$ on $[0, T]$ steers the system (1) from M_0 to M_1 in time T. Since $T \leqslant t_f$ and the problem **(OCP)** has a unique solution, we infer that $T = t_f$, $\bar{u}_1 = u_1$ and $\bar{x}(\cdot) = x(\cdot)$.

To conclude, it suffices to remark that the above reasoning proves that $(t_f, u_1(\cdot), x(\cdot))$ is the unique closure point of $(t_f^{\varepsilon_n}, u_1^{\varepsilon_n}(\cdot), x^{\varepsilon_n}(\cdot))$, where $(\varepsilon_n)_{n\in\mathbb{N}}$ is any sequence of positive real numbers converging to 0. □

Remark 2.3. In one does not assume the uniqueness of the optimal solution of **(OCP)**, then the following statement still holds. If $\varepsilon > 0$ tends to 0, then t_f^ε still converges to the minimal time t_f, the family $(u_1^\varepsilon(\cdot))_\varepsilon$ has a closure point $\bar{u}_1(\cdot)$ in $L^\infty(0, t_f)$ for the weak star topology, and the family $(x^\varepsilon(\cdot))_\varepsilon$ has a closure point $\bar{x}(\cdot)$ in $C^0([0, t_f], \mathbb{R}^n)$ for the uniform convergence topology, where $\bar{x}(\cdot)$ is the solution of the system (1) corresponding to the control $\bar{u}_1(\cdot)$ on $[0, t_f]$, such that $\bar{x}(0) \in M_0$ and $\bar{x}(t_f) \in M_1$. This means that $\bar{x}(\cdot)$ is another possible solution of (**OCP**).

In other words, every closure point of a family of solutions of $\mathbf{(OCP)}_\varepsilon$ is a solution of (**OCP**).

The next lemma will serve as a technical tool to derive Lemma 2.5.

Lemma 2.4. *Let $T > 0$, and let $(g_\varepsilon)_{\varepsilon>0}$ be a family of continuous functions on $[0,T]$ converging weakly to some $g \in L^2(0,T)$ as ε tends to 0, for the weak topology of $L^2(0,T)$. Then, for every $t \in (0,T)$, there exists a family $(t_\varepsilon)_{\varepsilon>0}$ of points of $[t,T)$ such that $t_\varepsilon \to t$ and $g_\varepsilon(t_\varepsilon) \to g(t)$ as $\varepsilon \to 0$.*

Proof. First of all, note that, since g_ε converges weakly to g on $[0,T]$, its restriction to any subinterval of $[0,T]$ converges weakly, as well, to the corresponding restriction of g. Let us prove that, for every $\beta > 0$, for every $\alpha > 0$ (small enough so that $t + \alpha \leqslant T$), there exists $\varepsilon_0 > 0$ such that, for every $\varepsilon \in (0, \varepsilon_0)$, there exists $t_\varepsilon \in [t, t+\alpha]$ such that $|g_\varepsilon(t_\varepsilon) - g(t)| \leqslant \beta$. The proof goes by contradiction. Assume that there exist $\beta > 0$ and $\alpha > 0$ such that, for every integer n, there exists $\varepsilon_n \in (0, 1/n)$ such that, for every $\tau \in [t, t+\alpha]$, there holds $|g_{\varepsilon_n}(\tau) - g(t)| \geqslant \beta$. Since g_{ε_n} is continuous, it follows that either $g_{\varepsilon_n}(\tau) \geqslant g(t) + \beta$ for every $\tau \in [t, t+\alpha]$, or $g_{\varepsilon_n}(\tau) \leqslant g(t) - \beta$ for every $\tau \in [t, t+\alpha]$. This inequality contradicts the weak convergence of the restriction of g_{ε_n} to $[t, t+\alpha]$ towards the restriction of g to $[t, t+\alpha]$. ☐

In what follows, we denote by $K(t)$, $K_1(t)$, $\tilde{K}(t)$, the Pontryagin cones along the trajectory $x(\cdot)$ solution of **(OCP)**, defined as in the previous subsection. Similarly, for every $\varepsilon > 0$, we denote by $K^\varepsilon(t)$, $K_1^\varepsilon(t)$, $\tilde{K}^\varepsilon(t)$ the Pontryagin cones along the trajectory $x^\varepsilon(\cdot)$, which is a solution of $\textbf{(OCP)}_\varepsilon$.

Lemma 2.5. *For every $v \in K(t_f)$, for every $\varepsilon > 0$, there exists $v^\varepsilon \in K^\varepsilon(t_f^\varepsilon)$ such that v^ε converges to v as ε tends to 0.*

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

$$\begin{aligned} \dot{v}_\pi(t) &= \left(\frac{\partial X}{\partial x}(x(t)) + u_1(t) \frac{\partial Y_1}{\partial x}(x(t)) \right) v_\pi(t) \\ v_\pi(t_1) &= (\bar{u}_1 - u_1(t_1)) Y_1(x(t_1)), \end{aligned} \tag{20}$$

where t_1 is a Lebesgue point of $[0, t_f)$, $\bar{u}_1 \in [-1,1]$, and the needle-like variation $\pi = \{t_1, \eta_1, \bar{u}_1\}$ of the control u_1 is defined by

$$u_{1,\pi}(t) = \begin{cases} \bar{u}_1 & \text{if } t \in [t_1, t_1 + \eta_1], \\ u_1(t) & \text{otherwise.} \end{cases}$$

For every $\varepsilon > 0$, consider the control $u^\varepsilon = (u_1^\varepsilon, \ldots, u_m^\varepsilon)$ of Lemma 2.2, solution of $\textbf{(OCP)}_\varepsilon$. It satisfies the constraint $\sum_{i=1}^m (u_i^\varepsilon)^2 \leqslant 1$. From Lemma 2.3, the continuous control function u_1^ε converges weakly to u_1 in $L^2(0, t_f)$. It then follows from Lemma 2.4 that, for every $\varepsilon > 0$, there exists $t_\varepsilon \geqslant t_1$ such that $t_\varepsilon \to t_1$ and $u_1^\varepsilon(t_\varepsilon) \to u_1(t_1)$ as $\varepsilon \to 0$.

For every $\varepsilon > 0$, consider the needle-like variation $\pi^\varepsilon = \{t_1^\varepsilon, \eta_1, (\bar{u}_1, 0, \ldots, 0)\}$ of the control $(u_1^\varepsilon, \ldots, u_m^\varepsilon)$ defined by[7], for $i = 2, \ldots, m$,

$$u_{1,\pi^\varepsilon}^\varepsilon(t) = \begin{cases} \bar{u}_1 & \text{if } t \in [t_1^\varepsilon, t_1^\varepsilon + \eta_1], \\ u_1^\varepsilon(t) & \text{otherwise,} \end{cases}$$

[7] Note that t_1^ε is a Lebesgue point of the function $t \mapsto X(x^\varepsilon(t)) + u_1^\varepsilon(t) Y_1(x^\varepsilon(t)) + \varepsilon \sum_{i=2}^m u_i^\varepsilon(t) Y_i(x^\varepsilon(t))$ since the controls u_i^ε are continuous functions of t.

and

$$u^{\varepsilon}_{i,\pi^{\varepsilon}}(t) = \begin{cases} 0 & \text{if } t \in [t^{\varepsilon}_1, t^{\varepsilon}_1 + \eta_1], \\ u^{\varepsilon}_i(t) & \text{otherwise,} \end{cases}$$

and define the variation vector $v_{\pi^{\varepsilon}}(\cdot)$ as the solution on $[t^{\varepsilon}_1, t^{\varepsilon}_f]$ of the Cauchy problem

$$\begin{aligned} \dot{v}_{\pi^{\varepsilon}}(t) &= \left(\frac{\partial X}{\partial x}(x^{\varepsilon}(t)) + u^{\varepsilon}_1(t) \frac{\partial Y_1}{\partial x}(x^{\varepsilon}(t)) + \varepsilon \sum_{i=2}^{m} u^{\varepsilon}_i(t) \frac{\partial Y_i}{\partial x}(x^{\varepsilon}(t)) \right) v_{\pi^{\varepsilon}}(t) \\ v_{\pi^{\varepsilon}}(t^{\varepsilon}_1) &= (\bar{u}_1 - u^{\varepsilon}_1(t^{\varepsilon}_1)) Y_1(x^{\varepsilon}(t^{\varepsilon}_1)) - \varepsilon \sum_{i=2}^{m} u^{\varepsilon}_i(t^{\varepsilon}_1) Y_i(x^{\varepsilon}(t^{\varepsilon}_1)). \end{aligned} \tag{21}$$

From Lemma 2.3, t^{ε}_f converges to t_f, $u^{\varepsilon}_1(\cdot)$ converges weakly to $u_1(\cdot)$, $x^{\varepsilon}(\cdot)$ converges uniformly to $x(\cdot)$; moreover, $\varepsilon u^{\varepsilon}_i(\cdot)$ converges weakly to 0, $\varepsilon u^{\varepsilon}_i(t^{\varepsilon}_1)$ converges to 0, for $i = 2, \ldots, m$, and $u^{\varepsilon}_1(t_1)$ converges to $u_1(t_1)$. As in the proof of Lemma 2.3, we infer the uniform convergence of $v^{\varepsilon}_{\pi}(\cdot)$ to $v_{\pi}(\cdot)$ (see [28] for details), and the conclusion follows. □

The next lemma will be useful in the proof of Lemma 2.7.

Lemma 2.6. *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 $M = \max_{u \in C} g(0, u)$. Then, $M(\varepsilon)$ tends to M as ε tends to 0.*

Proof. For every $\varepsilon > 0$, let $u_{\varepsilon} \in C$ such that $M(\varepsilon) = g(\varepsilon, u_{\varepsilon})$, and let $u \in C$ such that $M = g(0, u)$. Note that u_{ε} does not necessarily converge to u, however we will prove that $M(\varepsilon)$ tends to M, as ε tends to 0. Let $u_0 \in C$ be a closure point of the family $(u_{\varepsilon})_{\varepsilon > 0}$. Then, by definition of M, one has $g(0, u_0) \leqslant M$. On the other hand, since g is continuous, $g(\varepsilon, u)$ tends to $g(0, u) = M$ as ε tends to 0. By definition, $g(\varepsilon, u) \leqslant M(\varepsilon) = g(\varepsilon, u_{\varepsilon})$ for every $\varepsilon > 0$. Therefore, passing to the limit, one gets $M \leqslant g(0, u_0)$. It follows that $M = g(0, u_0)$. We have thus proved that the (bounded) family $(M(\varepsilon))_{\varepsilon > 0}$ of real numbers has a unique closure point, which is M. The conclusion follows. □

Lemma 2.7. *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 **(OCP)**$_{\varepsilon}$ is normal.*

Proof. We argue by contradiction. Assume that, for every integer n, there exist $\varepsilon_n \in (0, 1/n)$ and a solution $x^{\varepsilon_n}(\cdot)$ of **(OCP)**$_{\varepsilon_n}$ having an abnormal extremal lift $(x^{\varepsilon_n}(\cdot), p^{\varepsilon_n}(\cdot), 0, u^{\varepsilon_n}(\cdot))$. Set $\psi^{\varepsilon_n} = p^{\varepsilon_n}(t^{\varepsilon_n}_f)$, for every integer n. Then, from Remark 2.1, one has

$$\langle \psi^{\varepsilon_n}, v^{\varepsilon_n} \rangle \leqslant 0,$$

for every $v^{\varepsilon_n} \in K^{\varepsilon_n}(t^{\varepsilon_n}_f)$, and

$$\begin{aligned} M(\varepsilon_n) = \max_{\sum_{i=1}^m w_i^2 \leqslant 1} \Big(\left\langle \psi^{\varepsilon_n}, X(x^{\varepsilon_n}(t^{\varepsilon_n}_f)) \right\rangle &+ w_1 \left\langle \psi^{\varepsilon_n}, Y_1(x^{\varepsilon_n}(t^{\varepsilon_n}_f)) \right\rangle \\ &+ \varepsilon_n \sum_{i=2}^{m} w_i \left\langle \psi^{\varepsilon_n}, Y_i(x^{\varepsilon_n}(t^{\varepsilon_n}_f)) \right\rangle \Big) = 0, \end{aligned}$$

for every integer n. Since the final adjoint vector $(p^{\varepsilon_n}(t^{\varepsilon_n}_f), p^{0\,\varepsilon_n})$ is defined up to a multiplicative scalar, and $p^{0\,\varepsilon_n} = 0$, we assume that ψ^{ε_n} is a unit vector. Then, up to a subsequence, the sequence $(\psi^{\varepsilon_n})_{n \in \mathbb{N}}$ converges to some unit vector ψ. Using Lemmas 2.3, 2.5 and 2.6, we infer that

$$\langle \psi, v \rangle \leqslant 0,$$

for every $v \in K(t_f)$, and

$$M = \max_{|w_1| \leqslant 1} \left(\langle \psi, X(x(t_f)) \rangle + w_1 \langle \psi, Y_1(x(t_f)) \rangle \right) = 0.$$

It then follows from Remark 2.1 that the trajectory $x(\cdot)$ has an abnormal extremal lift. This is a contradiction since, by assumption, $x(\cdot)$ has a unique extremal lift, which is moreover normal. ☐

Remark 2.4. If we remove the assumption that the optimal trajectory $x(\cdot)$ has a unique extremal lift, which is moreover normal, then Lemma 2.7 still holds provided that every extremal lift of $x(\cdot)$ is normal.

With the notations of Lemma 2.7, from now on we normalize the adjoint vector so that $p^{0\,\varepsilon} = -1$, for every $\varepsilon \in (0, \varepsilon_0)$.

Lemma 2.8. *In the setting of Lemma 2.7, the set of all possible $p^\varepsilon(t_f^\varepsilon)$, with $\varepsilon \in (0, \varepsilon_0)$, is bounded.*

Proof. The proof goes by contradiction. Assume that there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of positive real numbers converging to 0 such that $\|p^{\varepsilon_n}(t_f^{\varepsilon_n})\|$ tends to $+\infty$. Since the sequence $\left(\frac{p^{\varepsilon_n}(t_f^{\varepsilon_n})}{\|p^{\varepsilon_n}(t_f^{\varepsilon_n})\|}\right)_{n\in\mathbb{N}}$ is bounded in $\mathbb{R}^n$, up to a subsequence it converges to some unit vector ψ. Using the Lagrange multipliers property and (19), there holds

$$\langle p^{\varepsilon_n}(t_f^{\varepsilon_n}), v^{\varepsilon_n} \rangle \leqslant 0,$$

for every $v^{\varepsilon_n} \in K^{\varepsilon_n}(t_f^{\varepsilon_n})$, and

$$\begin{aligned} \max_{\sum_{i=1}^m w_i^2 \leqslant 1} \Big(\left\langle p^{\varepsilon_n}(t_f^{\varepsilon_n}), X(x^{\varepsilon_n}(t_f^{\varepsilon_n})) \right\rangle + w_1 \left\langle p^{\varepsilon_n}(t_f^{\varepsilon_n}), Y_1(x^{\varepsilon_n}(t_f^{\varepsilon_n})) \right\rangle \\ + \varepsilon_n \sum_{i=2}^m w_i \left\langle p^{\varepsilon_n}(t_f^{\varepsilon_n}), Y_i(x^{\varepsilon_n}(t_f^{\varepsilon_n})) \right\rangle \Big) = 1, \end{aligned}$$

for every integer n. Dividing by $\|p^{\varepsilon_n}(t_f^{\varepsilon_n})\|$, and passing to the limit, using Lemmas 2.3, 2.5 and 2.6, and Remark 2.1, the same reasoning as in the proof of the previous lemma yields that the trajectory $x(\cdot)$ has an abnormal extremal lift, which is a contradiction. ☐

Remark 2.5. Remark 2.4 applies as well to Lemma 2.8.

Lemma 2.9. *For every $\varepsilon \in (0, \varepsilon_0)$, let $x^\varepsilon(\cdot)$ be a solution of **(OCP)**$_\varepsilon$, and let $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), -1, u^\varepsilon(\cdot))$ be a (normal) extremal lift of $x^\varepsilon(\cdot)$. Then $p^\varepsilon(\cdot)$ converges uniformly*[8] *to $p(\cdot)$ on $[0, t_f]$ as ε tends to 0, where $(x(\cdot), p(\cdot), -1, u(\cdot))$ is the unique (normal) extremal lift of $x(\cdot)$.*

Proof. For every $\varepsilon > 0$, set $\psi^\varepsilon = p^\varepsilon(t_f^\varepsilon)$. The adjoint equation of the Pontryagin Maximum Principle is

$$\begin{aligned} \dot{p}^\varepsilon(t) = &- \left\langle p^\varepsilon(t), \frac{\partial X}{\partial x}(x^\varepsilon(t)) \right\rangle - u_1^\varepsilon(t) \left\langle p^\varepsilon(t), \frac{\partial Y_1}{\partial x}(x^\varepsilon(t)) \right\rangle \\ &- \varepsilon \sum_{i=2}^m u_i^\varepsilon(t) \left\langle p^\varepsilon(t), \frac{\partial Y_i}{\partial x}(x^\varepsilon(t)) \right\rangle, \end{aligned}$$

[8] We consider any continuous extension of $p^\varepsilon(\cdot)$ on $[0, t_f]$.

with $p^\varepsilon(t_f^\varepsilon) = \psi^\varepsilon$. Moreover, there holds

$$\langle \psi^\varepsilon, v^\varepsilon \rangle \leqslant 0,$$

for every $v^\varepsilon \in K^\varepsilon(t_f^\varepsilon)$, and

$$\max_{\sum_{i=1}^m w_i^2 \leqslant 1} \Big(\langle \psi^\varepsilon, X(x^\varepsilon(t_f^\varepsilon)) \rangle + w_1 \langle \psi^\varepsilon, Y_1(x^\varepsilon(t_f^\varepsilon)) \rangle + \varepsilon \sum_{i=2}^m w_i \langle \psi^\varepsilon, Y_i(x^\varepsilon(t_f^\varepsilon)) \rangle \Big) = 1.$$

From Lemma 2.8, the family of all ψ^ε, $0 < \varepsilon < \varepsilon_0$, is bounded. Let ψ be a closure point of that family, and $(\varepsilon_n)_{n\in\mathbb{N}}$ a sequence of positive real numbers converging to 0 such that ψ^{ε_n} tends to ψ. Using Lemma 2.3, and as in the proof of this lemma, we infer that the sequence $(p^{\varepsilon_n}(\cdot))_{n\in\mathbb{N}}$ converges uniformly to the solution $z(\cdot)$ of the Cauchy problem

$$\dot{z}(t) = -\left\langle z(t), \frac{\partial X}{\partial x}(x(t)) \right\rangle - u_1(t) \left\langle z(t), \frac{\partial Y_1}{\partial x}(x(t)) \right\rangle, \quad z(t_f) = \psi.$$

Moreover, passing to the limit as in the previous proof,

$$\langle \psi, v \rangle \leqslant 0,$$

for every $v \in K(t_f)$, and

$$\max_{|w_1| \leqslant 1} \left(\langle \psi, X(x(t_f)) \rangle + w_1 \langle \psi, Y_1(x(t_f)) \rangle \right) = 1.$$

It follows that $(x(\cdot), z(\cdot), -1, u_1(\cdot))$ is an extremal lift of $x(\cdot)$, and from the uniqueness assumption we infer that $z(\cdot) = p(\cdot)$. The conclusion follows. $\square$

Remark 2.6. If one removes the assumptions of uniqueness of the solution of **(OCP)** and uniqueness of the extremal lift, then the following result still holds, provided that every extremal lift of every solution of **(OCP)** is normal. Consider the topological spaces $\mathcal{X} = C^0([0, t_f], \mathbb{R}^n)$, endowed with the uniform convergence topology, and $\mathcal{Y} = L^\infty(0, t_f; [-1, 1])$, endowed with the weak star topology. In the following statement, the space $\mathcal{X} \times \mathcal{X} \times \mathcal{Y}$ is endowed with the resulting product topology. For every $\varepsilon \in (0, \varepsilon_0)$, let $x^\varepsilon(\cdot)$ be a solution of $\textbf{(OCP)}_\varepsilon$, and let $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), -1, u^\varepsilon(\cdot))$ be a (normal) extremal lift of $x^\varepsilon(\cdot)$. Then, every closure point of the family $(x^\varepsilon(\cdot), p^\varepsilon(\cdot), u_1^\varepsilon(\cdot))$ in $\mathcal{X} \times \mathcal{X} \times \mathcal{Y}$ is a triple $(\bar{x}(\cdot), \bar{p}(\cdot), \bar{u}_1(\cdot))$, where $\bar{x}(\cdot)$ is an optimal solution of **(OCP)**, associated with the control $\bar{u}_1(\cdot)$, having as a normal extremal lift the 4-tuple $(\bar{x}(\cdot), \bar{p}(\cdot), -1, \bar{u}_1(\cdot))$. This statement indeed follows from Remarks 2.3, 2.4 and 2.5.

Lemma 2.10. *If the control u_1 is moreover bang-bang, i.e., if the (continuous) switching function $\varphi(t) = \langle p(t), Y_1(x(t)) \rangle$ does not vanish on any subinterval of $[0, t_f]$, then $u_1^\varepsilon(\cdot)$ converges to $u_1(\cdot)$ and $u_i^\varepsilon(\cdot)$, $i = 2, \ldots, m$, converge to 0 almost everywhere on $[0, t_f]$, and thus in particular for the strong $L^1(0, t_f)$ topology.*

Proof. Using the expression (12) of the controls u_1^ε and u_i^ε, $i = 2, \ldots, m$, the expression (5) of the control u_1, and from Lemmas 2.3 and 2.9, it is clear that $u_1^\varepsilon(t)$ converges to $u_1(t)$ and $u_i^\varepsilon(t)$, $i = 2, \ldots, m$, converge to 0 as ε tends to 0, for almost every $t \in [0, t_f]$. Since the controls are bounded by 1, the strong L^1 convergence follows from the dominated convergence theorem. $\square$

3 Examples and further comments

3.1 Comments on the assumption (11)

Let us go on Remark 1.2, and consider the case $m = 2$, that is, consider only one arbitrary additional smooth vector field Y_2. For $\varepsilon > 0$ fixed, the maximization condition from the Pontryagin maximum principle applied to the problem $\mathbf{(OCP)}_\varepsilon$ is

$$\begin{aligned} & u_1^\varepsilon(t)\langle p^\varepsilon(t), Y_1(x^\varepsilon(t)\rangle + \varepsilon u_2^\varepsilon(t)\langle p^\varepsilon(t), Y_2(x^\varepsilon(t)\rangle \\ & = \max_{w_1^2+w_2^2\leqslant 1} \left(w_1\langle p^\varepsilon(t), Y_1(x^\varepsilon(t)\rangle + \varepsilon w_2\langle p^\varepsilon(t), Y_2(x^\varepsilon(t)\rangle\right) \end{aligned}$$

almost everywhere on $[0, t_f^\varepsilon]$. There are two cases: either the maximum is attained in the interior of the domain, or it is attained at the boundary. The proof of our main result requires this maximum to be attained at the boundary (see (12)), and the corresponding controls to be continuous. This fact depends on the choice of the vector field Y_2.

A simple example where this holds true is the case $Y_2 = X$. In that case it is indeed possible to ensure that both functions $t \mapsto \langle p^\varepsilon(t), Y_1(x^\varepsilon(t)\rangle$ and $t \mapsto \langle p^\varepsilon(t), Y_2(x^\varepsilon(t)\rangle$ do not vanish simultaneously for $\varepsilon > 0$ small enough (and this implies that the desired conclusion). To prove this assertion, we argue by contradiction and assume that, for every $n \in \mathbb{N}$, there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ converging to 0 and a sequence $(t^{\varepsilon_n})_{n\in\mathbb{N}}$ such that $\langle p^{\varepsilon_n}(t^{\varepsilon_n}), X(x^{\varepsilon_n}(t^{\varepsilon_n}))\rangle = \langle p^{\varepsilon_n}(t^{\varepsilon_n}), Y_2(x^{\varepsilon_n}(t^{\varepsilon_n}))\rangle = 0$. Combined with the fact that the Hamiltonian is constant along any extremal, and vanishes at the final time, these equalities imply that $p^{0\varepsilon_n} = 0$. This contradicts the conclusion of Lemma 2.7.

More generally, and although such a statement may be nontrivial to derive, we conjecture that this fact holds true for generic vector fields Y_2 (see [7, 8, 9] for such genericity statements).

Note that, for generic triples of vector fields (X, Y_1, Y_2), this fact holds true. Indeed, to derive this statement it suffices to combine the fact that any totally singular minimizing trajectory must satisfy the Goh condition (see [1] and [3, Theorem 1.9] for details) and the fact that, for generic (in the strong sense of Whitney) triplets of vector fields (X, Y_1, Y_2), the associated control-affine system does not admit nontrivial Goh singular trajectories (see [9, Corollary 2.7]).

3.2 First example: the Rayleigh problem

To illustrate our results, we consider the minimal time control problem for the Rayleigh control system described in [18],

$$\begin{aligned} \dot{x}_1(t) &= x_2(t), \\ \dot{x}_2(t) &= -x_1(t) + x_2(t)(1.4 - 0.14x_2(t)^2) + u_1(t), \end{aligned} \tag{22}$$

with initial and final conditions

$$x_1(0) = x_2(0) = -5,\ x_1(t_f) = x_2(t_f) = 0, \tag{23}$$

and the control constraint

$$|u_1(\cdot)| \leqslant 4. \tag{24}$$

This optimal control problem has a unique solution, that has a unique extremal lift (up to a multiplicative scalar) which is moreover normal (see [18]).

We propose the regularized control system

$$\begin{aligned} \dot{x}_1^\varepsilon(t) &= x_2^\varepsilon(t) + \varepsilon u_2^\varepsilon(t)\,, \\ \dot{x}_2^\varepsilon(t) &= -x_1^\varepsilon(t) + x_2^\varepsilon(t)(1.4 - 0.14 x_2^\varepsilon(t)^2) + u_1^\varepsilon(t), \end{aligned} \tag{25}$$

with the same initial and final conditions, and where the control $u^\varepsilon(\cdot) = (u_1^\varepsilon(\cdot), u_2^\varepsilon(\cdot))$ satisfies the constraint

$$(u_1^\varepsilon(\cdot))^2 + (u_2^\varepsilon(\cdot))^2 \leqslant 16.$$

All assumptions of Theorem 1 are satisfied. A single shooting method is applied to both optimal control problems. The convergence results proved in Theorem 1 are illustrated on Figures 1, 2 and 3. In this example, the minimal time control solution of (22), (23), (24) is bang-bang, and we indeed observe, on the numerical simulations, the almost everywhere convergence of the regularized control.

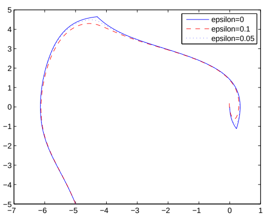

Figure 1: Trajectory

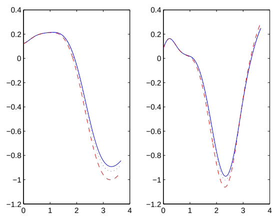

Figure 2: Adjoint vector

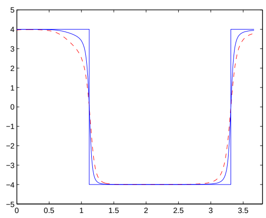

Figure 3: Control

3.3 Second example, involving a singular arc

In the example provided in this subsection, the minimal time control $u_1(\cdot)$ is singular. It is then not expected a priori that the regularized control $u_1^\varepsilon(\cdot)$ converges almost everywhere to $u_1(\cdot)$ along the singular arc. Our main result only asserts a weak convergence property along this arc. In the example presented below, the regularized control u_1^ε converges weakly to $u_1(\cdot)$ but not almost everywhere. We then provide some numerical simulations, on which we indeed observe that the almost everywhere convergence property fails along the singular arc, and we observe an oscillating property, which is a typical feature of weak convergence.

Consider the minimal time control problem for the system

$$\begin{aligned} \dot{x}_1(t) &= 1 - x_2(t)^2, \\ \dot{x}_2(t) &= u_1(t), \end{aligned} \tag{26}$$

with initial and final conditions

$$x_1(0) = x_2(0) = 0, \quad x_1(t_f) = 1,\ x_2(t_f) = 0, \tag{27}$$

and the control constraint

$$|u_1(\cdot)| \leqslant 1. \tag{28}$$

It is clear that the solution of this optimal control problem is unique, and is provided by the singular control $u_1(t) = 0$, for every $t \in [0, t_f]$, with $t_f = 1$. The corresponding trajectory is given by $x_1(t) = t$ and $x_2(t) = 0$.

We claim that this optimal trajectory has a unique extremal lift (up to a multiplicative scalar), which is moreover normal. Indeed, denoting by $p = (p_1, p_2)$ the adjoint vector, the Hamiltonian of the above optimal control problem is $H = p_1(1 - x_2^2) + p_2 u_1 + p^0$, and the differential equations of the adjoint vector are $\dot{p}_1 = 0$, $\dot{p}_2 = 2x_2 p_1$. Since $x_2(t) = 0$, it follows that the adjoint vector of any extremal lift of the optimal trajectory is constant. Moreover, the Hamiltonian vanishes at the final time, and hence there must hold $p_1(t) + p^0 = 0$, for every $t \in [0, t_f]$. Since the singular control $u_1(t) = 0$ is optimal and belongs to the interior of the domain of constraints (28), the maximization condition yields $\frac{\partial H}{\partial u_1} = 0$, and thus, $p_2(t) = 0$ for every $t \in [0, t_f]$. Then, since the adjoint vector is nontrivial, p^0 cannot be equal to 0, and up to a multiplicative scalar we assume that $p^0 = -1$. The assertion is thus proved, and the unique (normal) extremal lift is given by $(x_1(t), x_2(t), p_1(t), p_2(t), p^0, u_1(t)) = (t, 0, 1, 0, -1, 0)$.

We propose the following regularization of the problem (26)–(28). Let $g(\cdot)$ and $h(\cdot)$ be smooth functions, to be chosen; consider the minimal time control problem for the system

$$\begin{aligned} \dot{x}_1^\varepsilon(t) &= 1 - x_2^\varepsilon(t)^2 + \varepsilon u_2^\varepsilon(t) g(x_1^\varepsilon(t)), \\ \dot{x}_2^\varepsilon(t) &= u_1^\varepsilon(t) + \varepsilon u_2^\varepsilon(t) h(x_1^\varepsilon(t)), \end{aligned} \tag{29}$$

with initial and final conditions

$$x_1^\varepsilon(0) = x_2^\varepsilon(0) = 0, \quad x_1^\varepsilon(t_f^\varepsilon) = 1,\ x_2^\varepsilon(t_f^\varepsilon) = 0, \tag{30}$$

and the control constraint

$$(u_1^\varepsilon(\cdot))^2 + (u_2^\varepsilon(\cdot))^2 \leqslant 1. \tag{31}$$

Since the function g to be chosen below vanishes at some points, the assumption (11) does not hold everywhere. We claim however that, if the function g may only vanish on a subset of zero measure, and if $\varepsilon > 0$ is small enough, then the formula (12) holds, and the regularized controls are continuous, so that we are in the framework of Theorem 1.

Indeed, the Hamiltonian of this regularized optimal control problem is

$$H = p_1^\varepsilon(1 - (x_2^\varepsilon)^2) + p_2^\varepsilon u_1^\varepsilon + \varepsilon u_2^\varepsilon(p_1^\varepsilon g(x_1^\varepsilon) + p_2^\varepsilon h(x_1^\varepsilon)) + p^{0\varepsilon},$$

and the adjoint equations are

$$\begin{aligned} \dot{p}_1^\varepsilon(t) &= -\varepsilon u_2^\varepsilon(t)(p_1^\varepsilon(t) g'(x_1^\varepsilon(t)) + p_2^\varepsilon(t) h'(x_1^\varepsilon(t))), \\ \dot{p}_2^\varepsilon(t) &= 2x_2^\varepsilon(t) p_1^\varepsilon(t). \end{aligned}$$

It is not difficult to see that, for $\varepsilon > 0$ small enough, the optimal trajectory must be such that $\dot{x}_1^\varepsilon(t) > 0$; hence, $x_1^\varepsilon(\cdot)$ is an increasing function of t. Now, argue by contradiction, and assume that the optimal control takes its values in the interior of the domain (31), for $t \in I$, where I is a subset of $[0, t_f^\varepsilon]$ of positive measure. Then, the maximization condition yields $\frac{\partial H}{\partial u_1^\varepsilon} = \frac{\partial H}{\partial u_2^\varepsilon} = 0$, and hence $p_2^\varepsilon(t) = 0$ and $p_1^\varepsilon(t) g(x_1^\varepsilon(t)) + p_2^\varepsilon(t) h(x_1^\varepsilon(t)) = 0$, for $t \in I$. It follows that $p_1^\varepsilon(t) g(x_1^\varepsilon(t)) = 0$, for $t \in I$. Since the function g may only vanish on a subset of zero measure, and since $x_1^\varepsilon(\cdot)$ is increasing, it follows that there exists $t_1 \in I$ such that $g(x_1^\varepsilon(t_1)) \neq 0$, and therefore $p_1^\varepsilon(t_1) = p_2^\varepsilon(t_1) = 0$. Since the Hamiltonian vanishes almost everywhere, this yields moreover $p^{0\varepsilon} = 0$, which is a contradiction.

Therefore, under the above assumption on g, the formula (12) holds, and the optimal controls are given by

$$
\begin{aligned}
u_1^\varepsilon(t) &= \frac{p_2^\varepsilon(t)}{\sqrt{p_2^\varepsilon(t)^2 + \varepsilon^2\left(p_1^\varepsilon(t)g(x_1^\varepsilon(t)) + p_2^\varepsilon(t)h(x_1^\varepsilon(t))\right)^2}},\\
u_2^\varepsilon(t) &= \frac{\varepsilon\left(p_1^\varepsilon(t)g(x_1^\varepsilon(t)) + p_2^\varepsilon(t)h(x_1^\varepsilon(t))\right)}{\sqrt{p_2^\varepsilon(t)^2 + \varepsilon^2\left(p_1^\varepsilon(t)g(x_1^\varepsilon(t)) + p_2^\varepsilon(t)h(x_1^\varepsilon(t))\right)^2}},
\end{aligned}
\tag{32}
$$

for almost every $t \in [0, t_f^\varepsilon]$.

Let us prove that the controls $u_1^\varepsilon(\cdot)$ and $u_2^\varepsilon(\cdot)$ are smooth functions of t. For this purpose, we prove hereafter that the function $p_2^\varepsilon(\cdot)$ does not vanish on any subset of positive measure. Argue by contradiction and assume that there exists a subset I of $[0, t_f^\varepsilon]$ on which $p_2^\varepsilon(\cdot)$ vanishes. Then, on the one part, (32) implies that $u_1^\varepsilon(t) = 0$ and $u_2^\varepsilon(t) = \operatorname{sign}(p_1^\varepsilon(t)g(x_1^\varepsilon(t)) + p_2^\varepsilon(t)h(x_1^\varepsilon(t)))$, for almost every $t \in I$. On the other part, using the adjoint equations, we have $x_2^\varepsilon(t)p_1^\varepsilon(t) = 0$, for $t \in I$. The scalar $p_1^\varepsilon(t)$ cannot vanish, for any $t \in I$; indeed otherwise there would hold $p_1^\varepsilon(t) = p_2^\varepsilon(t) = 0$, and since the Hamiltonian vanishes, it would follow that $p^{0\varepsilon} = 0$, which is a contradiction with the normality of the extremal lift (see Lemma 2.7). Hence, $x_2^\varepsilon(t) = 0$ for $t \in I$, and thus, by differentiation, $u_1^\varepsilon(t) + \varepsilon u_2^\varepsilon(t) = 0$. This contradicts the equalities $u_1^\varepsilon(t) = 0$ and $u_2^\varepsilon(t) = \operatorname{sign}(p_1^\varepsilon(t)g(x_1^\varepsilon(t)) + p_2^\varepsilon(t)h(x_1^\varepsilon(t)))$.

From Theorem 1, we can assert that, as ε tends to 0,

- $x_1^\varepsilon(\cdot)$ (resp., $x_2^\varepsilon(\cdot)$) converges uniformly to $x_1(\cdot)$ (resp., $x_2(\cdot)$) on $[0,1]$,
- $p_1^\varepsilon(\cdot)$ (resp., $p_2^\varepsilon(\cdot)$) converges uniformly to $p_1(\cdot) = 1$ (resp., $p_2(\cdot) = 0$),
- $u_1^\varepsilon(\cdot)$ converges weakly to $u_1(\cdot) = 0$.

Let us next prove that, for certain choices of the functions $g(\cdot)$ and $h(\cdot)$, the regularized control $u_1^\varepsilon(\cdot)$ does not converge almost everywhere to $u_1(\cdot)$. We choose a smooth function $g(\cdot)$ defined on $\mathbb{R}$ that is strongly oscillating in the neighborhood of $1/2$, for instance,

$$g(x) = h(x)\sin\frac{1}{x - 1/2},$$

and a flat function h so that g is indeed smooth, for instance,

$$h(x) = \exp\left(\frac{-1}{(x-1/2)^2}\right).$$

If ε is small enough, then $x_1^\varepsilon(t)$ is close to t, $p_1^\varepsilon(t)$ is close to 1, $p_2^\varepsilon(t)$ is close to 0, and hence the sign of $u_2^\varepsilon(t)$, that is equal to the sign of

$$h(x_1^\varepsilon(t))\left(p_1^\varepsilon(t)\sin\frac{1}{x_1^\varepsilon(t) - 1/2} + p_2^\varepsilon(t)\right)$$

is close to the sign of $\sin\frac{1}{t-1/2}$. Therefore, the control $u_2^\varepsilon(\cdot)$ strongly oscillates between -1 and 1 for t close to $1/2$. Since $u_1^\varepsilon(\cdot)$ and $u_2^\varepsilon(\cdot)$ are continuous and satisfy $(u_1^\varepsilon(\cdot))^2 + (u_2^\varepsilon(\cdot))^2 = 1$, it follows that the control $u_1^\varepsilon(\cdot)$ strongly oscillates as well between -1 and 1 for t close to $1/2$.

This oscillation feature is similar to what happens with chattering controls, and illustrates the fact that $u_1^\varepsilon(\cdot)$ weakly converges to $u_1(\cdot) = 0$ as ε tends to 0, but does not converge almost everywhere.

Numerical simulations lead to Figures 4 and 5, on which we can observe the oscillating properties of the regularized controls. Note that these numerical simulations are difficult to obtain with the above function h, because of its flatness. First of all, in our numerical simulations we rather choose the function $h(x) = (x - 1/2)^3$, that is not so flat, but for which the system is however not smooth (but this does not change anything to the result). Second, it is difficult to make converge the shooting method for small values of ε, and we had to make use of a continuation method, starting with a large value of ε and making decrease step by step this value.

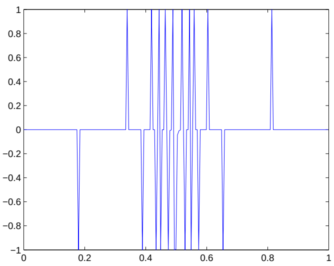

Figure 4: Control u_1^ϵ ($\varepsilon = 0.01$)

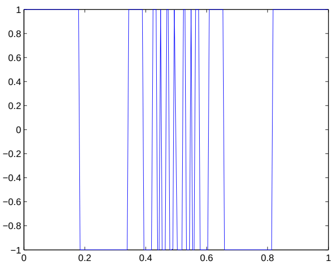

Figure 5: Control u_2^ϵ ($\varepsilon = 0.01$)

4 Conclusion

In this article, we described a smoothing procedure for the minimal time problem for single-input control-affine systems $\dot{x} = X(x) + u_1 Y_1(x)$ in $\mathbb{R}^n$ with the control constraint $|u_1(\cdot)| \leqslant 1$, which consists in adding new smooth vector fields $Y_2, \dots, Y_m$ and a small parameter $\varepsilon > 0$, so as to come up with the minimal time problem for the system $\dot{x} = X(x) + u_1^\varepsilon Y_1(x) + \sum_{i=2}^m \varepsilon u_i^\varepsilon Y_i(x)$, under the control constraint $\sum_{i=1}^m (u_i^\varepsilon(\cdot))^2 \leqslant 1$. Under appropriate assumptions, the optimal controls of the latter system, depending on ε, are smooth functions of t, and converge weakly to the optimal control of the initial system; moreover the associated trajectories converge uniformly. If the optimal control of the initial system is moreover bang-bang, then the convergence of the regularized control holds almost everywhere; this property may however fail whenever the bang-bang property does not hold. We provided examples and counterexamples to illustrate our result.

Finally, note that, in the present article, we focused on the minimal time problem for single-input control-affine systems. The extension of our procedure to a general optimization criterion seems reachable, however the extension to more general nonlinear control systems seems difficult. First, because it may be not obvious to generalize the nice expression (12) to more general situations. Second, because Lemma 2.2 does not hold a priori for general control systems, and it is not clear how to derive Lemma 2.3 and the next results. The regularization procedure is quite natural for control-affine systems but it is not clear how it should be adapted to more general control systems.

Acknowledgment. During the review process, and after the preprint of that article has appeared online, A. Rapaport has mentioned to us the paper [25] and the (unpublished) report of a practical stage supervised by him, in which the regularization procedure consid-

ered in this paper has been used, but without proving the convergence result derived in our paper. We thank warmly to A. Rapaport for his remarks and interest to this work.

References

[1] A. Agrachev and Y. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, 87, Control Theory and Optimization, II, Berlin: Springer-Verlag, 2004.

[2] J.-P. Aubin, H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.

[3] B. Bonnard, J.-B. Caillau, E. Trélat, *Second order optimality conditions in the smooth case and applications in optimal control*, ESAIM Control Optim. Calc. Var., 13, 2: 207–236, 2007.

[4] P. Cannarsa, C. Sinestrari, *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Progress in Nonlinear Differential Equations and their Applications, 58, Birkhäuser Boston, Inc., Boston, MA, 2004.

[5] L. Cesari, *Optimization – theory and applications. Problems with ordinary differential equations*, Applications of Mathematics, 17, New York: Springer-Verlag, 1983.

[6] N. Caroff, H. Frankowska, *Conjugate points and shocks in nonlinear optimal control*, Trans. Amer. Math. Soc., 348 (1996), no. 8, 3133–3153.

[7] Y. Chitour, F. Jean, E. Trélat, *Propriétés génériques des trajectoires singulières*, C. R. Math. Acad. Sci. Paris, 337, no. 1: 49–52, 2003.

[8] Y. Chitour, F. Jean, E. Trélat, *Genericity results for singular curves*, J. Differential Geom., 73, no. 1: 45–73, 2006.

[9] Y. Chitour, F. Jean, E. Trélat, *Singular trajectories of control-affine systems*, SIAM J. Control Optim., 47, no. 2: 1078–1095, 2008.

[10] F.H. Clarke, R. Vinter, *The relationship between the maximum principle and dynamic programming*, SIAM J. Control Optim., 25 (1987), no. 5, 1291–1311.

[11] A. L. Dontchev, *Perturbations, Approximations and Sensitivity Analysis of Optimal Control Systems*, Lecture Notes in Control and Information Sciences, 52, Springer-Verlag, Berlin/Heidelberg/New York, 1983.

[12] W. Hager, *Lipschitz continuity for constrained processes*, SIAM J. Control Optim., 17, 321–337, 1979.

[13] J. Kim, H. Maurer, *Sensitivity Analysis of Optimal Control Problems with Bang-Bang Controls*, Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Dec. 9-12, 2003, IEEE Control Society, 3281–3286, 2003.

[14] K. Malanowski, *Differential stability of solutions to convex, control constrained optimal control problems*, Appl. Math. Optim, 12, 1–14, 1984.

[15] K. Malanowski, *Second-order conditions in stability analysis for state constrained optimal control*, J. Global Optim. 40, no. 1-3: 161–168, 2008.

[16] P. Martinon, J. Gergaud, *Using switching detection and variational equations for the shooting method*, Optimal Cont. Appl. Methods, 28, no. 2: 95–116, 2007.

[17] H. Maurer, *Differential stability in optimal control problems*, Appl. Math. Optim. 5, 283–295, 1979.

[18] H. Maurer, D. Augustin, *Second order sufficient conditions and sensitivity analysis for the controlled Rayleigh problem*, Parametric Optimization and Related Topics, IV: 15–55, Frankfurt: J. Guddat et al., eds., Lang, 1997.

[19] H. Maurer, N.P. Osmolovskii, *Second order sufficient conditions for time-optimal bang-bang control problems*, SIAM Journal on Control and Optimization 42, 2239–2263, 2004.

[20] H. Maurer and H.J. Pesch, *Solution differentiability for parametric nonlinear control problems with control-state constraints*, Control and Cybernetics, 23, 201–227, 1994.

[21] H. Maurer, and G. Vossen, *Sufficient conditions and sensitivity analysis for optimal bangbang control problems with state constraints*, Proceedings of the 23rd IFIP Conference on System Modeling and Optimization, Cracow, Poland, 2007.

[22] L. Pontryagin, V. Boltyanskii, R. Gramkrelidze, E. Mischenko, *The mathematical theory of optimal processes*, Wiley Interscience, 1962.

[23] L. Rifford, E. Trélat, *Morse-Sard type results in sub-Riemannian geometry*, Math. Ann. 332 (2005), no. 1, 145–159.

[24] L. Rifford, E. Trélat, *On the stabilization problem for nonholonomic distributions*, J. Eur. Math. Soc., 11 (2009), no. 2, 223–255.

[25] P. Spinelli, G.S. Rakotonirainy, *Minimum time problem synthesis*, Systems and Control Letters, 10: 281–290, 1988.

[26] G. Stefani, *Regularity properties of the minimum-time map*, Nonlinear synthesis (Sopron, 1989), 270–282, Progr. Systems Control Theory, 9, Birkhuser Boston, Boston, MA, 1991.

[27] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York: Berlin-Heidelberg, 1983.

[28] E. Trélat, *Some properties of the value function and its level sets for affine control systems with quadratic cost*, J. Dyn. Cont. Syst., 6, no. 4: 511–541, 2000.