Regularization of chattering phenomena via bounded variation controls

Marco Caponigro[*] Roberta Ghezzi[†] Benedetto Piccoli[‡] Emmanuel Trélat[§]

Abstract

In control theory, the term *chattering* is used to refer to fast oscillations of controls, such as an infinite number of switchings over a finite time interval. In this paper we focus on three typical instances of chattering: the Fuller phenomenon, referring to situations where an optimal control features an accumulation of switchings in finite time; the Robbins phenomenon, concerning optimal control problems with state constraints, where the optimal trajectory touches the boundary of the constraint set an infinite number of times over a finite time interval; the Zeno phenomenon, for hybrid systems, referring to a trajectory that depicts an infinite number of location switchings in finite time. From the practical point of view, when trying to compute an optimal trajectory, for instance by means of a shooting method, chattering may be a serious obstacle to convergence.

In this paper we propose a general regularization procedure, by adding an appropriate penalization of the total variation. This produces a family of quasi-optimal controls whose associated cost converge to the optimal cost of the initial problem as the penalization tends to zero. Under additional assumptions, we also quantify quasi-optimality by determining a speed of convergence of the costs.

1 Introduction

Chattering phenomena in optimal control have been known since the first example presented in [1]. Roughly speaking, chattering refers to fast oscillations of the optimal control switching infinitely many times over a finite time interval. To explain this behavior, let us recall the famous example in [1], also known as *Fuller's problem.* Given $T > 0$ arbitrary, consider the control system in $\mathbb{R}^2$

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u, \tag{1}$$

with controls $u : [0, T] \to [-1, 1]$, and the optimal control problem which amounts to minimizing the cost functional

$$\int_0^T x_1^2(t)\, dt, \tag{2}$$

over all trajectories of (1) steering an (arbitrary) initial point (x_1^0, x_2^0) to the origin, i.e., such that

$$x_1(0) = x_1^0, \quad x_2(0) = x_2^0, \qquad x_1(T) = 0, \quad x_2(T) = 0.$$

[*]Conservatoire National des Arts et Métiers, Equipe M2N, 75003 Paris, France (marco.caponigro@cnam.fr).

[†]Institut de Mathématiques de Bourgogne, Université de Bourgogne-Franche Comté, Dijon, France (roberta.ghezzi@u-bourgogne.fr).

[‡]Department of Mathematical Sciences and Center for Computational and Integrative Biology, Rutgers University, Camden, NJ 08102, USA (piccoli@camden.rutgers.edu).

[§]Sorbonne Universités, UPMC Univ Paris 06, CNRS UMR 7598, Laboratoire Jacques-Louis Lions, Institut Universitaire de France, F-75005, Paris, France (emmanuel.trelat@upmc.fr).

It is well known that there exists a unique optimal control $u : [0, T] \to [-1, 1]$, satisfying

$$u(t) = \begin{cases} 1, & t \in (t_{2k}, t_{2k+1}), \quad k \in \mathbb{N}, \\ -1, & t \in (t_{2k+1}, t_{2k+2}), \quad k \in \mathbb{N}, \end{cases}$$

where $(t_k)_{k\in\mathbb{N}}$ is an increasing sequence depending on the initial condition (x_1^0, x_2^0) and converging to T. Although, at first sight, one could think that this strong oscillation property is a kind of aberration due to specific symmetries of the system, it turns out that such behavior is rather typical. Indeed, it was later shown in [2] that the set of single-input optimal control problems which have a control-affine Hamiltonian and whose solution is chattering is an open semi-algebraic set (see also [3]), showing therefore that chattering is a common phenomenon in optimal control.

Control problems presenting chattering properties have been found for a variety of problems: besides the ones mentioned above, a similar phenomenon also concerns state-constrained problems and hybrid systems. In [4], Robbins studied an optimal control problem with an inequality state constraint of third order and he showed that the optimal trajectory touches the constraint's boundary at an infinite sequence of isolated points converging to a point at the boundary, even if the optimal control has finite total variation. In the framework of hybrid systems, chattering is often called Zeno phenomenon and is due to trajectories whose discrete part jumps infinitely many times over a finite time interval (see, for instance, the examples in [5]).

Although chattering cannot be considered as a degeneracy phenomenon (see [6]), it may however cause some difficulties in theoretical and numerical aspects of optimal control.

From the theoretical point of view, due to the lack of a positive length interval where the control function is continuous when chattering occurs, finding necessary and sufficient optimality conditions becomes much more intricate (see [7] for state-constrained problems). Some results in this sense were proved in [3], yet the problem is not completely understood in other contexts, such as state-constrained problems or hybrid systems. Another delicate issue comes from the study of regularity properties of optimal syntheses [8, 9].

From the numerical point of view, chattering phenomena may be an obstacle to the convergence of numerical methods applied to optimal control problems, in particular when using indirect methods. Indeed, chattering implies ill-posedness of shooting methods (non-invertible Jacobian) [10, 11]. When chattering occurs it is therefore required to develop an adequate numerical method in order to compute a good approximation of the optimal control. This problem has been raised in [12, 13] for the optimal control of the attitude of a launcher, in which chattering may occur, depending on the terminal conditions under consideration. After having observed that chattering was indeed causing the failure of the shooting method, the authors have proposed two remedies: one is based on a specific homotopy combined with the shooting method, and the other consists of using a direct method with a finite number of arcs. However, on one hand these remedies remain specific to the problem studied thereof, and on the other hand there is no convergence result that would show and quantify the quasi-optimality property.

In this paper, we propose a general regularization procedure, consisting of penalizing the cost functional with a total variation term. The main idea comes from the fact that to avoid oscillations phenomena one needs to master the derivative of the control: indeed, for controls in the Sobolev space $W^{1,1}$ our penalization term coincides with the L^1-norm of the derivative of the control. Our method is actually more general, since $W^{1,1} \subsetneq BV$. Moreover, note that a more classical penalization in the L^2-norm of the control is not well-suited for our aim since it does not prevent chattering.

Our approach is valid for general classes of nonlinear optimal control problems. For a bang-bang scalar control, the total variation of the control is proportional to the number of switchings. In the case where the Fuller phenomenon occurs, the total variation is infinite. Hence, with such a penalization term, the optimal control does not chatter, and its numerical computation is then a priori feasible. Under appropriate assumptions of small-time local controllability, we prove in Theorem 1 that, if the weight on the total variation term in the cost functional tends to zero, then the regularized optimal control problem Γ-converges to the initial optimal control problem, meaning that the optimal cost and any optimal solution of the regularized problem converge respectively to the optimal cost and an optimal solution of the initial problem, as the parameter ε tends to zero. This shows that, when this total variation regularization is used, the optimal control that one may then compute numerically is quasi-optimal, with a good rate of optimality.

In order to quantify quasi-optimality, it remains to determine at what speed the cost of the regularized problem converges to the cost of the initial problem, as the weight of the total variation term tends to zero. This can be done by estimating explicitly the rate of convergence of the cost along suboptimal regimes obtained by suitable truncations of the chattering one in terms of switching times. In the existing literature, such results, related to truncation, were obtained in [14, 15] for small perturbations of the Fuller's problem. In those papers, the authors exhibited a sequence of suboptimal regimes for the specific optimal control problem (1)-(2), and they proved that the cost converges with the same rate as the sequence of switching times (of the chattering control). Our Theorem 2 establishes a polynomial rate of convergence for the cost, for general nonlinear optimal control problems, under appropriate controllability assumptions, and under the additional assumption of Hölder continuity of the time-optimal map, i.e. the map associating with every y the minimal time needed to steer the system from y to 0. Note that, for the specific case considered in [15], the rate of convergence is exponential as a function of the number of switchings. Likewise, for the class of systems considered in [2], the switching times converge exponentially to the final time. Whether a slower rate of convergence is "typical" remains an open question.

Finally, we treat by total variation regularization two other general cases where chattering occurs:

- For optimal control problems involving state-constraints, under adequate controllability assumptions, Theorem 3 provides a regularization result for optimal trajectories having an infinite sequence of contact points with the constraint's boundary (Robbins phenomenon). Here, the penalization term essentially counts the contact points with the constraint's boundary.

- For hybrid optimal control problems, Theorem 4 provides a convergence result to regularize the Zeno phenomenon, obtaining estimates of the cost convergence as the number of location switchings grows.

The paper is organized as follows. In Section 2 we present the main results on the regularization by total variation penalization of chattering phenomena (Fuller, Robbins and Zeno). Section 3 is devoted to prove the main results. In Appendix A we provide some additional results concerning the controllability condition required in Theorem 2. Finally, we provide in Appendix B an existence result for optimal control problems having a total variation term in the cost functional, without any convexity assumptions.

2 Main results

2.1 Regularization of the Fuller phenomenon

Let N and m be positive integers. Consider the control system

$$\dot{x} = f(x, u), \quad u \in \mathcal{U}, \tag{Σ}$$

where $f \in \mathcal{C}^\infty(\mathbb{R}^N \times \mathbb{R}^m, \mathbb{R}^N)$, $f(0,0) = 0$, and

$$\mathcal{U} = \{u(\cdot) \text{ measurable} \mid u(t) \in \mathbf{U} \text{ for a.e. } t\}, \tag{3}$$

with $\mathbf{U} \subset \mathbb{R}^m$ a measurable subset containing 0. Denote by

$$\mathcal{F} = \{f(\cdot, u) : \mathbb{R}^N \to \mathbb{R}^N \mid u \in \mathbf{U}\},$$

the family of vector fields associated with the dynamics of (Σ).

A control $u \in \mathcal{U}$ is called *admissible* if it steers the system (Σ) from a given (arbitrary) initial point to the origin in finite time denoted $t(u)$.

Given an initial state $x_0 \in \mathbb{R}^N$, a function $L \in \mathcal{C}^0(\mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^m)$ (called Lagrangian), we consider the optimal control problem

$$\left\{ \begin{array}{l} \displaystyle\min_{u \in \mathcal{U}} \int_0^{t(u)} L(s, x(s), u(s))\, ds, \\ \dot{x} = f(x, u), \quad u \in \mathcal{U}, \\ x(0) = x_0, \quad x(t(u)) = 0. \end{array} \right. \tag{OCP}$$

The final time in (OCP) may be fixed or free. If it is fixed to some $T > 0$, then of course one has to replace $t(u)$ with T everywhere.

Throughout the section, we make the following assumptions:

- for every $(t, x) \in \mathbb{R} \times \mathbb{R}^N$, the set

  $$V(t, x) = \{(f(x, u), L(t, x, u) + \gamma) \mid u \in \mathbf{U}, \gamma \geqslant 0\} \tag{4}$$

  is convex;

- $\mathbf{U}$ is compact, and there exists $b > 0$ such that, for every admissible control $u \in \mathcal{U}$, we have

  $$t(u) + \|x_u(\cdot)\|_\infty \leqslant b. \tag{5}$$

The first assumption means that the epigraph of extended velocities is convex. It is satisfied, for example, for control-affine systems with control-affine or quadratic cost.

These are classical assumptions used to derive existence results (see, for instance, [16, 17, 18]). Under these assumptions, the optimal control problem (OCP) has at least one optimal solution $x^*(\cdot)$, associated with a control $u^* : [0, t(u^*)] \to \mathbf{U}$.

We introduce a regularization of (OCP) by adding to the cost functional a total variation term, penalizing oscillations, with a small weight ε.

Given any $\varepsilon \geqslant 0$, we consider the optimal control problem

$$\left\{ \begin{array}{l} \displaystyle\min_{u \in \mathcal{U}} \left(\int_0^{t(u)} L(s, x(s), u(s))\, ds + \varepsilon\, \mathrm{TV}(u) \right), \\ \dot{x} = f(x, u), \quad u \in \mathcal{U}, \\ x(0) = x_0, \quad x(t(u)) = 0. \end{array} \right. \tag{OCP$_\varepsilon$}$$

Here, $\mathrm{TV}(u)$ denotes the total variation of the function $u \in L^1([0, t(u)], \mathbb{R}^m)$ and it is defined by

$$\mathrm{TV}(u) = \sup_\phi \int_0^{t(u)} u(s) \cdot \phi'(s) ds,$$

the supremum being taken over all functions $\phi \in C^1_c([0,t(u)],\mathbb{R}^m)$ with compact support such that $\|\phi\|_{L^\infty} \leqslant 1$. If u is continuous the total variation is equivalent to

$$\sup \sum_{i=1}^{p} \|u(t_i) - u(t_{i-1})\|,$$

the supremum being taken over all possible partitions $0 = t_0 < t_1 < \cdots < t_p = t(u)$ of the interval $[0,t(u)]$. Notice that if $m = 1$ and if u is a piecewise constant function taking values in $\{0,1\}$, then $\mathrm{TV}(u)$ is simply equal to the number of switchings. A function with bounded total variation is said to have bounded variation, or BV.

The rationale for introducing the term $\varepsilon\,\mathrm{TV}(u)$ in the cost of $(\mathrm{OCP})_\varepsilon$ is to penalize highly oscillating controls in order to avoid *chattering* in the sense of Definition 1 below.

Definition 1. *By* chattering *control we mean a measurable function* $u : [0,t(u)] \to \mathbf{U}$ *such that there exists an increasing sequence* $\{t_n\}_{n\in\mathbb{N}}$ *converging to* $t(u)$ *with the property that* $\mathrm{TV}(u|_{[0,t_n]}) < +\infty$ *for every* $n \in \mathbb{N}$, *and*

$$\lim_{n\to+\infty} \mathrm{TV}(u|_{[0,t_n]}) = +\infty.$$

The optimal control problem $(\mathrm{OCP})_\varepsilon$ is seen as a regularization of (OCP). We are next going to prove that any optimal solution $(\mathrm{OCP})_\varepsilon$ converges uniformly to an optimal solution of (OCP), thus providing a quasi-optimal solution that does not chatter.

Recall that the control system (Σ) is *small-time locally controllable* (STLC) at $x_0 \in \mathbb{R}^N$ if, for every $\delta > 0$, there exists a neighborhood $\mathcal{N}_\delta$ of x_0 such that every $x_1 \in \mathcal{N}_\delta$ can be reached by x_0 within time δ with a control $u \in \mathcal{U}$.

In the sequel, $\mathrm{Lie}_x\,\mathcal{F}$ denotes the Lie algebra of vector fields generated by $\mathcal{F}$ evaluated at x, that is $\mathrm{Lie}_x\,\mathcal{F} = \{V(x) \ \mid\ V \in \mathrm{Lie}\,\mathcal{F}\}$, where $\mathrm{Lie}\,\mathcal{F} = \mathrm{span}\{[f_1,[\ldots[f_{k+1},f_k]\ldots]] \mid f_i \in \mathcal{F}, k \in \mathbb{N}\}$.

Theorem 1. *Assume that* $\mathrm{Lie}_0\,\mathcal{F} = \mathbb{R}^N$, *and the control system* (Σ) *is small-time locally controllable at* 0. *Then, for every* $\varepsilon > 0$, *the optimal control problem* $(\mathrm{OCP})_\varepsilon$ *has at least one solution. Moreover, for every optimal solution* $x_\varepsilon(\cdot)$ *of* $(\mathrm{OCP})_\varepsilon$, *associated with a control* $u_\varepsilon : [0,t(u_\varepsilon)] \to \mathbf{U}$, *we have*

$$\lim_{\varepsilon\to 0} \int_0^{t(u_\varepsilon)} L(t,x_\varepsilon(t),u_\varepsilon(t))\,dt = \int_0^{t(u^*)} L(t,x^*(t),u^*(t))\,dt. \tag{6}$$

Theorem 1 establishes the existence of a non-chattering control u_ε which is quasi-optimal for (OCP) in the sense that the cost of u_ε converges to the optimal value of (OCP). In the case where the optimal control of (OCP) chatters and therefore cannot be computed by means of a shooting method, the total variation term in $(\mathrm{OCP})_\varepsilon$ plays the role of a regularization, and the control u_ε does not chatter and can be computed numerically. Theorem 1 establishes that u_ε is quasi-optimal, and hence it is reasonable to replace (OCP) by $(\mathrm{OCP})_\varepsilon$ when chattering occurs, in order to ensure the convergence of a shooting method.

We refer to [19] for a survey on methods for the numerical implementation of the TV term in the $(\mathrm{OCP})_\varepsilon$ problem and to [20] for numerical algorithms for the minimization of the total variation.

Remark 1. In general one cannot infer convergence of an optimal solution $x_\varepsilon(\cdot)$ of $(\mathrm{OCP})_\varepsilon$ to an optimal solution of the original problem (OCP). However, we have that for every sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ converging to 0 there exists a subsequence $(\varepsilon_{n_k})_{k\in\mathbb{N}}$ and an optimal trajectory x of (OCP) such that $x_{\varepsilon_{n_k}}(\cdot)$ converges uniformly to $x(\cdot)$ as $k \to \infty$ (see the proof of Lemma 6 below). Hence, in particular, if the optimal solution x^* of (OCP) is unique then one has that any optimal solution $x_\varepsilon(\cdot)$ of $(\mathrm{OCP})_\varepsilon$ converges uniformly to $x^*(\cdot)$.

Remark 2. The Lie algebra and small-time controllability assumptions, although generic, may be slightly weakened without altering the conclusion of the theorem: they can be replaced by assuming local controllability in a neighborhood of the origin in arbitrarily small time and with piecewise constant controls. The fact that the latter assumption is weaker follows from a well known result due to Krener (see for instance [21, Corollary 8.3]).

Remark 3. Note that smoothness of f is required in order to give a sense to Lie brackets. In contrast, we only need the Lagrangian function L to be continuous. Besides, L may depend on t but it is important that the dynamics f is autonomous (indeed in Lemma 5 below the reverse time dynamics is considered).

Theorem 1 establishes the convergence (6) of the costs. It is then interesting to derive a speed of convergence. This is possible under additional assumptions.

We need the following "strong" notion of controllability which requires a uniform bound on the total variation of the control and a steering time comparable with the minimum time. To this purpose, we define the time-optimal map $x_0 \mapsto \Upsilon(x_0)$ associated with the control system (Σ), by

$$\Upsilon(x_0) = \inf\{t > 0 \text{ such that } \dot{x} = f(x,u),\ x(0) = x_0,\ x(t) = 0,\ u \in \mathcal{U}\}. \tag{7}$$

Definition 2. *We say that the control system* (Σ) *satisfies* (Ω) *at* 0 *if*

(Ω_1) *the control system* (Σ) *is STLC at* 0*;*

(Ω_2) *there exist a neighborhood* $\mathcal{N}$ *of* 0 *and* $M \geqslant 1$ *such that, for every* $y \in \mathcal{N}$*, there exists* $u : [0,\tau_y] \to \mathbf{U}$ *such that* u *steers* y *to* 0 *in time* τ_y, $\tau_y \leqslant M\Upsilon(y)$, $\mathrm{TV}(u) \leqslant M$.

We provide in Appendix A some comments on Definition 2 and some results on the relationships between the properties (Ω), STLC, and the regularity of Υ.

In the sequel, $\mathcal{C}^{0,\alpha}$ designates the class of Hölder continuous functions with exponent α.

Theorem 2. *Assume that:*

(i) *the control system* (Σ) *satisfies* (Ω) *at* 0*;*

(ii) *the optimal control problem* (OCP) *admits an optimal solution* u^* *which is chattering and its sequence of switching times* $(t_n)_{n\in\mathbb{N}}$ *satisfies* $(t(u^*) - t_n) = \mathrm{O}(n^{-\beta})$ *for some* $\beta > 0$*;*

(iii) *the time-optimal map is* $\mathcal{C}^{0,\alpha}$ *for some* $\alpha \in (0,1]$ *in a neighborhood of* 0.

Then, for every $\varepsilon > 0$*, the optimal control problem* $(\mathrm{OCP})_\varepsilon$ *has at least one solution. Moreover, for every optimal solution* $x_\varepsilon(\cdot)$ *of* $(\mathrm{OCP})_\varepsilon$*, associated with a control* $u_\varepsilon : [0, t(u_\varepsilon)] \to \mathbf{U}$*, we have*

$$\int_0^{t(u_\varepsilon)} L(t, x_\varepsilon(t), u_\varepsilon(t))\,dt - \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\,dt = \mathrm{O}\left(\varepsilon^{\frac{\alpha\beta}{1+\alpha\beta}}\right). \tag{8}$$

If (ii) *is replaced by:*

(ii)′ *the optimal control problem* (OCP) *admits an optimal solution* u^* *with bounded total variation,*

then the conclusion holds with (8) *replaced by*

$$\int_0^{t(u_\varepsilon)} L(t, x_\varepsilon(t), u_\varepsilon(t))\,dt - \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\,dt = \mathrm{O}\left(\varepsilon\right).$$

Remark 4. For linear control systems and for driftless control-affine systems, (Ω) is related to controllability. Sufficient conditions guaranteeing that (Ω) holds true can be found in [22, 23] for single-input control systems. For more general control-affine systems, (Ω) is related to the *Exact State Space Linearizability Problem* (see Appendix A).

Remark 5. Assumption (*ii*) is verified for a large class of systems having an exponential rate of accumulation of switchings (see [2]). In this case the convergence rate is $O(\varepsilon^\gamma)$ for every $\gamma < 1$.

Remark 6. Sufficient conditions for Assumption (*iii*) have been established in [24, Theorem 3.3, 3.10, 3.12], where the authors provide an estimate on the Hölder exponent.

Example 1. *We consider the regularized Fuller's problem, that is, for* $\varepsilon \geqslant 0$ *the optimal control problem of minimizing*

$$\int_0^T x^2(t)dt + \varepsilon TV(u) \tag{9}$$

with dynamics $\ddot{x} = u$, $u \in [-1, 1]$ *and constraints* $x(0) = 0, \dot{x}(0) = 1$, $x(T) = 0 = \dot{x}(T)$. *We denote by* $(u_\varepsilon, x_\varepsilon)$ *a solution associated with the parameter* ε *and by* x_0 *the optimal trajectory of the unperturbed Fuller's problem. We compare, in Table 1, for different values of the parameter* ε*, the total variation of the associated optimal controls and the difference with the optimal cost. In Figure 1 we compare the controls associated with the values of* ε *in Table 1 and in Figure 2 the corresponding trajectories.*

Table 1: Regularized Fuller's problem

ε	$\mathrm{TV}(u_\varepsilon)$	$\int x_\varepsilon^2 - x_0^2$
0	$+\infty$	0
10^{-7}	5.0970	$1.6 * 10^{-8}$
10^{-6}	4.7514	$3 * 10^{-7}$
10^{-5}	3.3599	$5.3 * 10^{-6}$
10^{-4}	3.2557	$1.4 * 10^{-5}$
10^{-3}	3.0159	$2.9 * 10^{-4}$
10^{-2}	1.4897	$4.9 * 10^{-3}$

2.2 Regularization of the Robbins phenomenon for problems with state constraints

In this section, we consider the general optimal control problem (OCP) of the previous section with additional state constraints. Namely, let

$$\mathcal{C} = \{x \in \mathbb{R}^N \mid h_1(x) \geqslant 0, \dots, h_l(x) \geqslant 0\}, \tag{10}$$

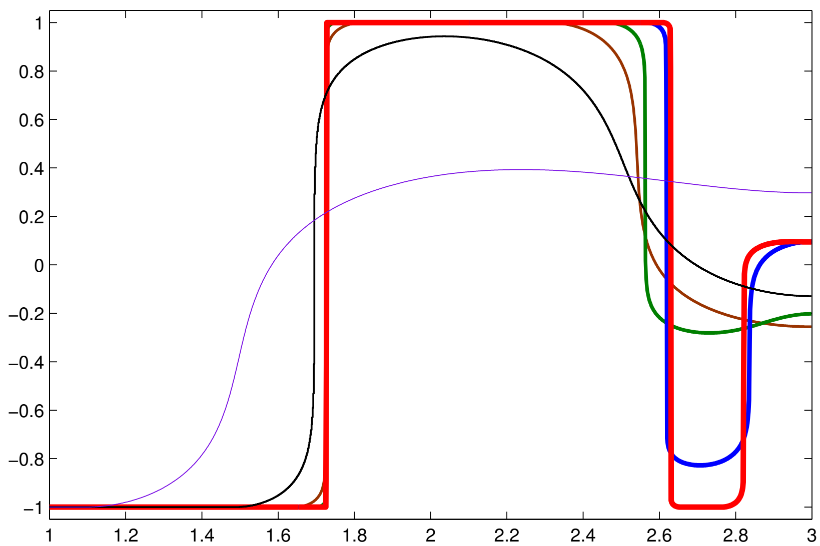

Figure 1: Optimal controls of the perturbed Fuller's problem associated with different values of ε ranging from 10^{-7} to 10^{-2} as in Table 1. Thicker lines corresponds to smaller values of ε.

where $h_1, \dots, h_l$ are continuous functions and consider the optimal control problem

$$\left\{\begin{array}{l} \min\limits_{u \in \mathcal{U}} \displaystyle\int_0^{t(u)} L(s, x(s), u(s))\, ds, \\ \dot{x} = f(x, u), \quad u \in \mathcal{U}, \\ x(t) \in \mathcal{C}, \quad t \in [0, t(u)], \\ x(0) = x_0, \quad x(t(u)) = 0. \end{array}\right. \tag{OCPS}$$

The notations and the assumptions on the dynamics are the same as in Section 2.1. In particular, we assume that the epigraph of extended velocities (4) is convex, that $\mathbf{U}$ is compact, that (5) holds true, and that there exists at least one admissible trajectory satisfying the constraints. Under these assumptions, the optimal control problem (OCPS) has at least one optimal solution $x^*(\cdot)$, associated with a control $u^* : [0, t(u^*)] \to \mathbf{U}$ (see [16, 17, 18]).

In [4], an instance of (OCPS) is provided where the final point 0 lies on the boundary $\partial\mathcal{C}$, the solution u^* is $\mathcal{C}^1$-smooth and the trajectory $x^*(\cdot)$ corresponding to u^* touches $\partial\mathcal{C}$ at a sequence of isolated points converging to the final point 0. In other words, the optimal trajectory is a concatenation of an infinite number of arcs contained in the interior of $\mathcal{C}$ and accumulating at the final point. We call this phenomenon the Robbins phenomenon.

To regularize this chattering effect, one needs to find suboptimal controls whose trajectories touch $\partial\mathcal{C}$ on a finite set. However, introducing the total variation of the control as a penalization term, as in Section 2, does not suffice to prevent the solution of the regularized problem from possibly intersecting $\partial\mathcal{C}$ infinitely many times. We next design a penalization term that rather counts the number of contact points with $\partial\mathcal{C}$.

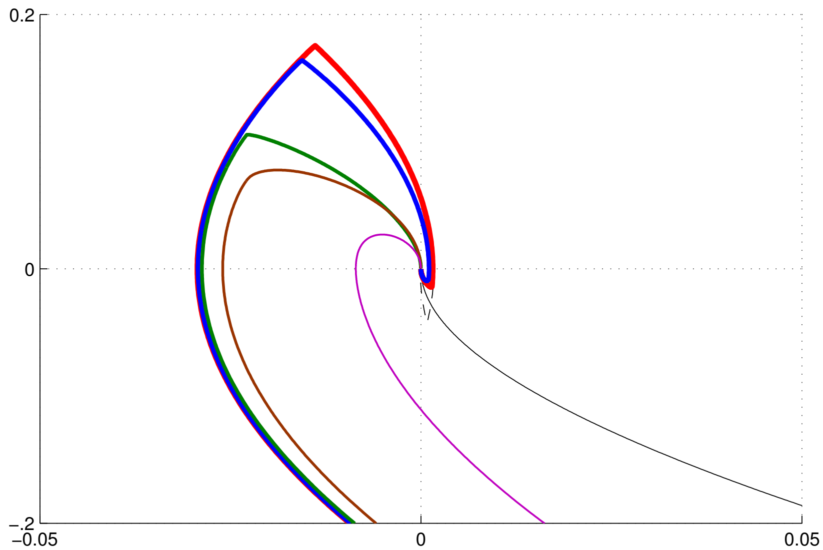

Figure 2: Optimal solutions of the perturbed Fuller's problem associated with different values of ε ranging from 10^{-7} to 10^{-2} as in Table 1. Thicker lines corresponds to smaller values of ε. The optimal Fuller's trajectory x_0 is represented by the dashed line.

Let $\mathbf{1}_{\partial\mathcal{C}} : \mathbb{R}^N \to \{0,1\}$ be the indicator function of $\partial\mathcal{C}$, defined by

$$\mathbf{1}_{\partial\mathcal{C}}(x) = \begin{cases} 1 & \text{if } x \in \partial\mathcal{C}, \\ 0 & \text{if } x \notin \partial\mathcal{C}. \end{cases}$$

Given an admissible control $u : [0, t(u)] \to \mathbf{U}$ with corresponding trajectory $x(\cdot)$, we define the function $X_u : [0, t(u)] \to \{0,1\}$ by

$$X_u(t) = \mathbf{1}_{\partial\mathcal{C}}(x(t)).$$

For every $\varepsilon \geqslant 0$, we consider the optimal control problem

$$\left\{\begin{array}{l} \min\limits_{u\in\mathcal{U}} \left(\displaystyle\int_0^{t(u)} L(s, x(s), u(s))\, ds + \varepsilon\, \mathrm{TV}(X_u) \right), \\ \dot{x} = f(x,u), \quad u \in \mathcal{U}, \\ x(t) \in \mathcal{C}, \quad t \in [0, t(u)], \\ x(0) = x_0, \quad x(t(u)) = 0. \end{array}\right. \tag{OCPS$_\varepsilon$}$$

In the sequel, we consider the reachable set from 0 with trajectories lying in the interior $\mathring{\mathcal{C}}$ of the constraint set $\mathcal{C}$ defined by (10): let $\mathcal{A}^{\mathcal{C}}(0, (0,\delta), f)$ be the set of points accessible from 0 in time $t \in (0,\delta)$ by trajectories $x(\cdot)$ of the control system (Σ) such that $x(s) \in \mathring{\mathcal{C}}$ for every $s \in (0,\delta)$. We denote by x^* an optimal trajectory of (OCPS) and by u^* the associated optimal control.

Theorem 3. *Assume that $0 \in \partial\mathcal{C}$ and that:*

(*i*) *for every* $\delta > 0$*, there exists a neighborhood* $\mathcal{N}$ *of* 0 *such that* $\mathcal{N} \cap \mathring{\mathcal{C}} \subset \mathcal{A}^{\mathcal{C}}(0, (0, \delta), -f)$*;*

(*ii*) *there exists a sequence of times* t_n *converging to* $t(u^*)$*, with*

$$x^*([0, t(u^*)]) \cap \partial\mathcal{C} \ \subset \ \{0\} \cup \{x^*(t_n) \ \mid \ n \in \mathbb{N}\};$$

Then, for every $\varepsilon > 0$*, the optimal control problem* $(\mathrm{OCPS})_\varepsilon$ *has at least one solution. Moreover, for every optimal solution* $x_\varepsilon(\cdot)$ *of* $(\mathrm{OCPS})_\varepsilon$*, associated with a control* $u_\varepsilon : [0, t(u_\varepsilon)] \to \mathbf{U}$*, we have*

$$\lim_{\varepsilon \to 0} \int_0^{t(u_\varepsilon)} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt = \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\, dt. \tag{11}$$

Remark 7. In analogy with Remark 1 one has that if, moreover, the solution x^* of (OCPS) is unique then any solution of $(\mathrm{OCPS})_\varepsilon$ converges to x^* as ε tends to 0.

Remark 8. Assumption (*i*) is an adaptation of the classical small-time local attainability (STLA) property (see [25, 26]), but we require here that the admissible trajectories stay in the interior of the constraint $\mathcal{C}$. Hence Assumption (*i*) may be seen as a generalization to nonlinear systems of the notion of small-time controllability with respect to a cone. Controllability with respect to a cone has been studied for linear control systems in [27] (see also [28]).

Remark 9. Assumption (*ii*) implies, in particular, that the interior of $\mathcal{C}$ is non-empty and prevents x^* to have boundary arcs.

2.3 Regularization of the Zeno phenomenon for hybrid problems

In this section, we present a regularization technique for hybrid systems, where the dynamics involve a continuous and a discrete part. In the spirit of Theorem 1 we design a perturbed optimal control problem with a penalization on the number of switchings of a trajectory, in order to rule out the so-called *Zeno* trajectories. The problem of finding necessary and sufficient conditions for the existence of Zeno trajectories of a hybrid system has been firstly addressed in [29] in which the authors dealt with the regularization of two specific hybrid systems: water tank and bouncing ball. Exploiting the specific geometry of the system, they introduced a family of regularized problems whose solution is “close to” the Zeno trajectory. Their idea was either to introduce an additional variable whose role is to delay of ε the time at which a switch takes place, or to introduce a spatial hysteresis. We refer to [30, 5] for a large number of examples of Zeno hybrid systems from the areas of modelling, simulation, verification, and control as well as for a list of references on the subject. We also refer to [31, 32, 33] where conditions for the existence of Zeno solutions have been established. The Zeno phenomenon for hybrid systems is related to so-called Zeno equilibria, which are invariant under the discrete (but not under the continuous) dynamics. See also [34] for asymptotic stability of Zeno equilibra.

Let us introduce some basic notions on hybrid systems, without control (see, e.g., [29]). A *hybrid system* is a collection $\mathcal{H} = (Q, X, f, E, G, R)$ where

- Q is a finite set;
- $X = \{X_q\}_{q \in Q}$ is a collection of subsets $X_q \subset \mathbb{R}^N$ called *locations*;
- $f = \{f_q\}_{q \in Q}$ is a collection of smooth vector fields f_q on X_q for every $q \in Q$;
- $E \subset Q \times Q$ is a subset of edges;
- G maps an edge $(q, q') \in E$ to a subset $G(q, q') \subset X_q$ called *guard set*;

- R maps a pair $((q,q'),x) \in E \times X_q$ to a subset $R((q,q'),x) \subset X_{q'}$.

A *trajectory* (or *execution*) of $\mathcal{H}$ is a triple $(\tau, q(\cdot), x(\cdot))$, where

- $\tau = \{\tau_i\}_{i=0}^M$ is an increasing sequence of positive numbers such that $\tau_0 = 0$ and $1 \leqslant M \leqslant \infty$. We set $I = [0, \tau_M]$ if $M < +\infty$, $I = [0, \tau_M)$ if $M = \infty$;
- $q : I \to Q$ is such that $q(t) = q_i$ constant on $[\tau_i, \tau_{i+1})$ for every $i = 0, \dots, M-1$;
- for every $i = 0, \dots, M-1$, $x_i(\cdot) = x|_{(\tau_i, \tau_{i+1})}$ is an absolutely continuous function in (τ_i, τ_{i+1}), which can be continuously extended to $[\tau_i, \tau_{i+1}]$, and such that $x_i(t) \in X_{q_i}$;
- for almost every $t \in (\tau_i, \tau_{i+1})$,
$$\dot{x}_i = f_{q_i}(x_i); \tag{12}$$
- for every $i = 0, \dots, M-1$, one has $(q_i, q_{i+1}) \in E$ and $x_i(\tau_{i+1}) \in G(q_i, q_{i+1})$ and, for every $i = 0, \dots, M-2$, one has $x_{i+1}(\tau_{i+1}) \in R((q_i, q_{i+1}), x_i(\tau_{i+1}))$.

In general a *Zeno trajectory* in a hybrid system is a trajectory presenting an infinite number of discrete events in a finite amount of time. Here we restrict our analysis to Zeno trajectories switching locations an infinite amount of times in a finite horizon. Namely, we say that a trajectory $(\tau, q(\cdot), x(\cdot))$ is *Zeno* if $M = +\infty$ and $\tau_\infty < +\infty$.

Given a hybrid system $\mathcal{H}$, a *Lagrangian* for $\mathcal{H}$ is a family $L = \{L_q\}_{q \in Q}$, with $L_q : \mathbb{R} \times X_q \to \mathbb{R}$, $L_q \geqslant 0$, such that, for every trajectory $(t, q(\cdot), x(\cdot))$ of $\mathcal{H}$ and every $i = 0, \dots, M-1$, the function $t \mapsto L_{q_i}(t, x_i(t))$ is continuous in (t_i, t_{i+1}). Given a Lagrangian for $\mathcal{H}$, we define the corresponding hybrid cost functional C by

$$C(\tau, q(\cdot), x(\cdot)) = \sum_{i=0}^{M-1} \int_{t_i}^{t_{i+1}} L_{q_i}(t, x_i(t))\, dt.$$

Let $(q_0, x_0) \in Q \times X_{q_0}$ be fixed. We consider the hybrid optimization problem

$$\begin{cases} \min C(\tau, q(\cdot), x(\cdot)), \\ (\tau, q(\cdot), x(\cdot)) \text{ trajectory of } \mathcal{H}, \\ q(0) = q_0, \quad x(0) = x_0. \end{cases} \tag{HP}$$

Let $Q = \{q_1, \dots, q_k\}$. We define $h : Q \to \{1, \dots, k\}$ by $h(q_i) = i$. For every $\varepsilon \geqslant 0$, we consider the optimization problem

$$\begin{cases} \min C(\tau, q(\cdot), x(\cdot)) + \varepsilon\, \mathrm{TV}(h \circ q(\cdot)), \\ (\tau, q(\cdot), x(\cdot)) \text{ trajectory of } \mathcal{H}, \\ q(0) = q_0, \quad x(0) = x_0. \end{cases} \tag{HP$_\varepsilon$}$$

Casting Theorem 1 in the language of hybrid systems, we obtain the following result.

Theorem 4. *Let $\mathcal{H}$ be a hybrid system such that X_q is a compact submanifold for every $q \in Q$, and such that the sets $G(q,q'), R((q,q'),x)$ are compact for every $((q,q'),x) \in E \times X_q$, with $q \in Q$. Let L be a Lagrangian for $\mathcal{H}$ with corresponding cost functional C. Assume that there exists a solution $(\tau^*, q^*(\cdot), x^*(\cdot))$ to* (HP) *which is Zeno. For every $\varepsilon > 0$, the problem* (HP)$_\varepsilon$ *has at least one solution. Moreover, for any solution $(\tau^\varepsilon, q^\varepsilon(\cdot), x^\varepsilon(\cdot))$ of* (HP)$_\varepsilon$*, we have*

$$\lim_{\varepsilon \to 0} C(\tau^\varepsilon, q^\varepsilon(\cdot), x^\varepsilon(\cdot)) = C(\tau^*, q^*(\cdot), x^*(\cdot)). \tag{13}$$

The compactness assumption on X_q can be slightly weakened, and replaced by compactness of trajectories in each location. The main idea of the proof is to interpret the role of the discrete part of the hybrid system in (12) as a control. Since there are no final conditions, the proof is simplified with respect to the ones of Theorem 1 and of Theorem 2.

The rate of convergence in (13) can be determined in the case where the rate of convergence of the switching times along the Zeno trajectory is known. We refer to Remark 11 in Section 3.4 (end of the proof of Theorem 4) for a precise statement.

Remark 10. In the definition of hybrid systems, one may add a control. We do not provide the details. For such hybrid optimal control problems, assuming moreover that, in each location, the epigraph of extended velocities (defined by (4)) is convex, that $\mathbf{U}$ is compact and that (5) holds true, the conclusion of Theorem 4 still holds true. In other words we have exactly the conclusion of Theorem 1 in the hybrid framework, including the convergence of trajectories.

3 Proofs

3.1 Proof of Theorem 1

Before going into technical details, let us outline the proof of Theorem 1. First, the local controllability assumption implies the existence of an optimal solution $(u_\varepsilon, x_\varepsilon)$ of $(\mathrm{OCP})_\varepsilon$ for any $\varepsilon \geqslant 0$. Second, thanks to the assumptions (4) on the extended velocity sets and on the equiboundedness of trajectories (5), there exists an admissible control w and a positive measurable function γ such that the family $t \mapsto (f(x_\varepsilon(t), u_\varepsilon(t)), L(t, x_\varepsilon(t), u_\varepsilon(t)))$ converges to $t \mapsto (f(x_w(t), w(t)), L(t, x_w(t), w(t)) + \gamma(t))$ for the weak star topology of L^∞. Third, we use the optimality of u_ε to prove that w is optimal for (OCP). Fourth, we establish that $\gamma = 0$, which implies that the Lagrangian cost along u_ε converges to the Lagrangian cost at u^*. This fact is proved thanks to Lemma 6 which exhibits a sequence of admissible controls v_n for which $\mathrm{TV}(v_n) < +\infty$ and whose Lagrangian costs converge to the cost of u^*. To construct v_n, we use a topological result (Lemma 5), providing admissible controls steering any point of a neighborhood of the origin to 0, with controls having bounded total variation.

We start by presenting the two auxiliary lemmas mentioned above, and then we proceed to the proof of the theorem. Note that the two lemmas do not require the convexity assumption of (4) nor the *a priori* estimate (5) on trajectories.

Lemma 5. *Assume that* $\mathrm{Lie}_0\,\mathcal{F} = \mathbb{R}^N$ *and that the control system* (Σ) *is small-time locally controllable at* 0*. Then, there exists a neighborhood* $\mathcal{N}$ *of* 0 *such that, for every* $y \in \mathcal{N}$*, there exists a piecewise constant control* $w_y : [0, \tau_y] \to \mathbf{U}$ *steering* (Σ) *from* y *to* 0 *in time* τ_y*, with* $\lim \tau_y = 0$ *as* $y \to 0$*.*

Proof. Since (Σ) is STLC at 0, by [35, Theorem 5.3 a-d] we have that the reversed control system

$$\dot{x} = -f(x, u), \quad u \in \mathcal{U}, \tag{$-\Sigma$}$$

associated with the dynamics $-f$ is also STLC at 0. As a consequence the time optimal map $\bar{\Upsilon}$ associated with system ($-\Sigma$), namely $x_0 \mapsto \bar{\Upsilon}(x_0) = \inf\{t > 0 \mid \dot{x} = -f(x, u), x(0) = 0, x(t) = x_0\}$ is continuous at 0 (see [24, Theorem 2.2]). We denote by $\mathcal{A}(x, [0, T), -f)$, respectively by $\mathcal{A}(x, [0, T], -f)$, the set of points accessible from x in time $t < T$, respectively $t \leqslant T$, by trajectories of the control system ($-\Sigma$). We set $\mathcal{A}(x, -f) = \cup_{T>0}\mathcal{A}(x, [0, T), -f)$. By definition of STLC there exists a neighborhood $\mathcal{N}$ of 0 such that $\mathcal{N} \subset \mathcal{A}(0, -f)$. Let $y \in \mathcal{N}$. By definition of the time optimal map, we have that $y \in \mathcal{A}(0, [0, \bar{\Upsilon}(y)], -f) \subset \mathcal{A}(0, [0, 2\bar{\Upsilon}(y)), -f)$. In particular this implies (see [35, Theorem 5.5]) that y is *normally reachable* (see [35, Definition 3.6]) from 0 in time less than $2\bar{\Upsilon}(y)$

for the control system ($-\Sigma$). Namely, there exist $q = q(y) \in \mathbb{N}$, $u_1, \dots u_q \in \mathbf{U}$ and positive numbers $t_1, \dots, t_q$ with $t_1 + \dots + t_q < 2\bar{\Upsilon}(y)$, such that $y = \exp(-t_q f(\cdot, u_q)) \circ \dots \circ \exp(-t_1 f(\cdot, u_1))(0)$. Here, $\exp(tV)$ designates the flow at time t of the vector field V. Since f is autonomous, we obtain $\exp(t_1 f(\cdot, u_1)) \circ \dots \circ \exp(t_q f(\cdot, u_q))(y) = 0$. Setting $\tau_y = t_1 + \dots + t_q$ and defining $w_y : [0, \tau_y] :\to \mathbf{U}$ by

$$w_y(t) = \begin{cases} u_1, & t \in [0, t_1], \\ u_2, & t \in [t_1, t_1 + t_2], \\ \vdots & \\ u_q, & t \in [t_1 + \dots + t_{q-1}, \tau_y], \end{cases}$$

the lemma follows. Notice that the continuity of $\bar{\Upsilon}$ at 0 ensures $\tau_y \to 0$ as $y \to 0$. □

Lemma 6. *Let $u : [0, t(u)] \to \mathbf{U}$ be a measurable control steering x^0 to 0. Then there exists a countable family of controls $u_n : [0, t(u_n)] \to \mathbf{U}$ such that $\mathrm{TV}(u_n) < +\infty$ for every $n \in \mathbb{N}$, u_n steers the control system (Σ) from x^0 to 0 in time $t(u_n)$ and*

$$\lim_{n \to +\infty} \|u_n - u\|_{L^1} = 0.$$

Here, the L^1 norm is on $[0, +\infty)$, by extending u (resp., u_n) by 0 for $t > t(u)$ (resp., $t > t(u_n)$). Recall that $f(0, 0) = 0$, and thus this extension does not have any impact on admissible trajectories.

Proof. Consider a sequence of functions $v_n : [0, t(u)] \to \mathbf{U}$ with $\mathrm{TV}(v_n) < +\infty$ for every $n \in \mathbb{N}$ converging to u in $L^1([0, t(u)], \mathbb{R}^m)$ for the strong topology and consider the associated solutions $y_n(\cdot)$ of the Cauchy problem $\dot{y}_n = f(y_n, v_n)$, $y_n(0) = x^0$. Then the sequence $y_n(\cdot)$ converges uniformly to the trajectory $x_u(\cdot)$ associated with the control u (see for instance [17, Theorem 3.4.1]). In particular, $y_n(t(u))$ converge to 0 as n tends to $+\infty$. By Lemma 5, for n sufficiently large, there exists a control $w_n : [0, \tau_n] \to \mathbf{U}$ which is piecewise constant, of bounded variation, steering $y_n(t(u))$ to 0 in time τ_n and such that $\tau_n \to 0$ as $n \to \infty$. Define

$$u_n(t) = \begin{cases} v_n(t), & t \in [0, t(u)), \\ w_n(t - t(u)), & t \in [t(u), t(u) + \tau_n), \\ 0, & t > t(u) + \tau_n. \end{cases}$$

By construction, u_n steers x^0 to 0 in time $t(u_n) = t(u) + \tau_n$ and, for every n, one has $\mathrm{TV}(u_n) < +\infty$. We extend u to $[0, +\infty)$ by setting $u(t) = 0$ for $t > t(u)$. Then

$$\begin{aligned} \int_0^{+\infty} |u_n(s) - u(s)| \, ds &= \int_0^{t(u)} |v_n(s) - u(s)| \, ds + \int_{t(u)}^{t(u)+\tau_n} |w_n(s - T)| \, ds \\ &\leqslant \int_0^{t(u)} |v_n(s) - u(s)| \, ds + \tau_n \max_{z \in \mathbf{U}} |z|, \end{aligned}$$

which converges to zero since $\tau_n \to 0$ as $n \to \infty$ and since v_n tends to u (strongly) in L^1. □

Let us now prove Theorem 1. The proof follows the lines of [18, Theorem 5.14 and 6.15].

Proof of Theorem 1. First of all, by Lemma 5, there exists a piecewise constant control $u : [0, t(u)] \to \mathbf{U}$ steering x_0 to 0. In particular u has bounded total variation. Therefore, the existence of an optimal solution of $(\mathrm{OCP})_\varepsilon$ follows from Theorem 16 in Appendix B.

Let $x_\varepsilon(\cdot)$ be any optimal solution of $(\text{OCP})_\varepsilon$, associated with a control $u_\varepsilon : [0, t(u_\varepsilon)] \to \mathbf{U}$. Set $\tilde{x}_\varepsilon(t) = (x_\varepsilon(t), \int_0^t L(s, x_\varepsilon(s), u_\varepsilon(s))ds)$. Then the triple $(\tilde{x}_\varepsilon, u_\varepsilon, \gamma_\varepsilon)$ with $\gamma_\varepsilon \equiv 0$ is a solution of

$$\min_{u\in\mathcal{U},\gamma\geqslant 0} \left(\int_0^{t(u)} (L(s,x,u) + \gamma(s))ds + \varepsilon \mathrm{TV}(u) \right)$$

subject to

$$\dot{x} = f(x,u), \quad \dot{x}_{N+1} = L(t,x,u) + \gamma, \tag{14}$$

with initial conditions $x(0) = x_0$, $x_{N+1}(0) = 0$, and final conditions $x(t(u)) = 0$, $x_{N+1}(t(u)) \geqslant 0$. Denote by $\tilde{f}(t,x,u,\gamma) = (f(x,u), L(t,x,u)+\gamma)$ the augmented dynamics of (14) which are convex for u in $\mathbf{U}$ and $\gamma \geqslant 0$ by Assumption (4).

Thanks to Assumption (5), the sequence $t(u_\varepsilon)$ is bounded and converges, up to some subsequence, to $t_1 > 0$ as ε tends to 0. Hence, given $\delta > 0$ there exists $\varepsilon_0 > 0$ such that $|t(u_\varepsilon) - t_1| < \delta$ for every $\varepsilon \in [0, \varepsilon_0]$ in the chosen subsequence. Since $f(0,0) = 0$, we extend x_ε and u_ε to $[t(u_\varepsilon), t_1 + \delta]$ by 0. By Assumption (5), the trajectories $x_\varepsilon(\cdot)$ are uniformly bounded, and hence the family of functions $s \mapsto \tilde{f}(s, x_\varepsilon(s), u_\varepsilon(s), 0)$ is bounded in $L^\infty([0, t_1+\delta], \mathbb{R}^{N+1})$. Thus by Banach–Alaoglu Theorem [36] it converges, up to some subsequence, to some function $g \in L^\infty([0, t_1+\delta], \mathbb{R}^{N+1})$ for the weak star topology. We define

$$\tilde{x}(t) = \tilde{x}_0 + \int_0^t g(s)\, ds, \quad \tilde{x}_0 = (x_0, 0).$$

By construction, $t \mapsto \tilde{x}(t)$ is absolutely continuous. Moreover, the family $\tilde{x}_\varepsilon(t)$ converges uniformly, up to subsequences to $\tilde{x}(\cdot)$ on $[0, t_1+\delta]$. By the convexity assumption (4) the absolutely continuous function $\tilde{x}(\cdot)$ is also a trajectory of (14) (see, for instance [17, Corollary 3.3.2]), in particular there exists an admissible control $w : [0, t(w)] \to \mathbf{U}$ and a nonnegative measurable function $\gamma : [0, t(w)] \to \mathbb{R}$ such that $t \mapsto \tilde{x}(t) := (x_w(t), \int_0^t L(s, x_w(s), w(s)) + \gamma(s)ds)$ is the associated solution of (14).

It remains to prove that $x_w(\cdot)$ is optimal for (OCP). For every admissible control $v \in \mathcal{U}$ satisfying $\mathrm{TV}(v) < +\infty$, we have (note that $\gamma(\cdot) \geqslant 0$)

$$\begin{aligned}\int_0^{t(w)+\delta} L(t, x_w(t), w(t))\, dt &\leqslant \int_0^{t(w)+\delta} (L(t, x_w(t), w(t)) + \gamma(t))\, dt \\ &\leqslant \limsup_{\varepsilon\to 0} \left(\int_0^{t(w)+\delta} L(t, x_\varepsilon(t), u_\varepsilon(t)) + \varepsilon\, \mathrm{TV}(u_\varepsilon) \right) \\ &\leqslant \int_0^{t(v)} L(t, x_v, v)\, dt + \int_{t(w)}^{t(w)+\delta} (L(t, x_w(t), w(t)) + \gamma(t))\, dt.\end{aligned} \tag{15}$$

Hence, for every $\delta > 0$ and every admissible v as above, we have

$$\int_0^{t(w)} L(t, x_w(t), w(t))\, dt \leqslant \int_0^{t(v)} L(t, x_v, v)\, dt + \int_{t(w)}^{t(w)+\delta} \gamma(t)\, dt.$$

For δ converging to 0, we conclude that

$$\int_0^{t(w)} L(t, x_w(t), w(t))\, dt \leqslant \int_0^{t(v)} L(t, x_v, v)\, dt,$$

for every admissible control $v \in \mathcal{U}$ satisfying $\mathrm{TV}(v) < +\infty$ Using Lemma 6 and the dominated convergence theorem, we infer that the inequality above also holds for any possible admissible control $v \in \mathcal{U}$ (not necessarily of bounded variation). Therefore, w is the optimal control solution of (OCP).

Finally, to prove (6), it suffices to show that $\gamma = 0$. By optimality of u_ε, we have

$$\begin{aligned}\int_0^{t(u_\varepsilon)} L(s, x_\varepsilon(s), u_\varepsilon(s))\, ds &\leqslant \int_0^{t(u_\varepsilon)} L(s, x_\varepsilon(s), u_\varepsilon(s))\, ds + \varepsilon\, \mathrm{TV}(u_\varepsilon) \\ &\leqslant \int_0^{t(v)} L(s, x_v(s), v(s)))\, ds + \varepsilon\, \mathrm{TV}(v),\end{aligned}$$

for any admissible control v such that $\mathrm{TV}(v) < +\infty$. Letting ε tend to 0, we deduce that

$$\int_0^{t(w)} (L(s, x_w(s), w(s)) + \gamma(s))\, ds \leqslant \int_0^{t(v)} L(s, x_v(s), v(s)))\, ds.$$

Finally, since w is optimal for (OCP), we conclude that $\gamma = 0$. □

3.2 Proof of Theorem 2

We start with the following lemma.

Lemma 7. *Assume that the control system* (Σ) *satisfies* (Ω) *at* 0*. Then, for every* $\eta > 0$ *sufficiently small, there exists an admissible control* $v_\eta : [0, t(v_\eta)] \to \mathbf{U}$ *satisfying* $\mathrm{TV}(v_\eta) < +\infty$*, whose corresponding trajectory is denoted by* $x_\eta(\cdot)$*, such that*

$$\lim_{\eta \to 0} \int_0^{t(v_\eta)} L(t, x_\eta(t), v_\eta(t))\, dt = \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\, dt, \tag{16}$$

and

$$\begin{aligned}\lim_{\eta\to 0} |t(v_\eta) - t(u^*)| &= \lim_{\eta \to 0} \|v_\eta - u^*\|_{L^1} \\ &= \lim_{\eta\to 0} \|x_\eta(\cdot) - x^*(\cdot)\|_\infty = 0.\end{aligned}$$

Moreover, under the additional assumption that the time-optimal map is $\mathcal{C}^{0,\alpha}$ *for some* $\alpha \in (0, 1]$ *in a neighborhood of* 0*, there exists* $C > 0$ *such that*

$$\int_0^{t(v_\eta)} L(t, x_\eta(t), v_\eta(t))\, dt - \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\, dt \leqslant C\eta^\alpha. \tag{17}$$

Proof. Let $\mathcal{N}$ be the neighborhood of 0 in $\mathbb{R}^N$ and M be the constant given by Definition 2. Without loss of generality we can assume that $\mathcal{N}$ is bounded. Fix η_0 such that $x^*(s) \in \mathcal{N}$, for every $s \geqslant t(u^*) - \eta_0$. By condition ($\Omega$), there exists a control w_η steering $x^*(t(u^*) - \eta)$ to 0 in time $\tau_\eta \leqslant M\Upsilon(x^*(t(u^*) - \eta))$ with $\mathrm{TV}(w_\eta) \leqslant M$. We define v_η by

$$v_\eta(t) = \begin{cases} u^*(t) & \text{for } t \in [0, t(u^*) - \eta), \\ w_\eta(t - t(u^*) + \eta) & \text{for } t \in [t(u^*) - \eta, t(u^*) - \eta + \tau_\eta), \\ 0, & \text{for } t > t(u^*) - \eta + \tau_\eta, \end{cases} \tag{18}$$

and let $x_\eta(\cdot)$ be the corresponding trajectory, starting from x_0 (see Figure 3). By construction, we have $\mathrm{TV}(v_\eta) \leqslant \mathrm{TV}(u^*|_{[0,t(u^*)-\eta]}) + M$. If $\mathrm{TV}(u^*) < +\infty$ or u^* is chattering in the sense of

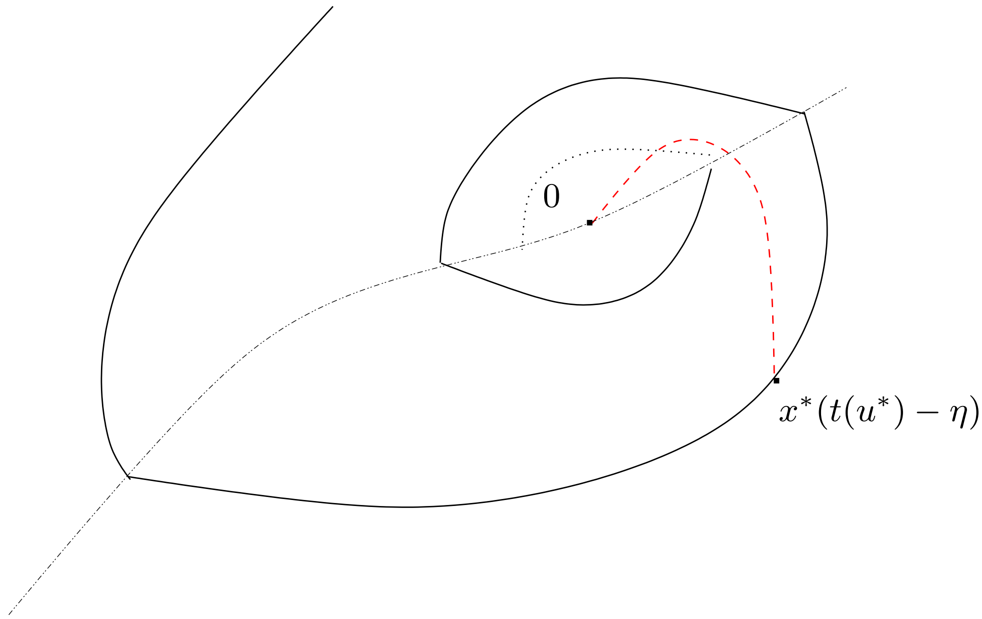

Figure 3: The trajectory $x_\eta(\cdot)$ (solid and dashed red line) associated with the control v_η built in Lemma 7. The solid line represents the optimal trajectory $x^*(\cdot)$.

Definition 1, then $\mathrm{TV}(v_\eta) < +\infty$. We have $\tau_\eta \to 0$ as $\eta \to 0$, since Υ is upper semi-continuous. Hence $v_\eta \to u^*$ almost everywhere, and for some subsequence, we have

$$\lim_{\eta\to 0} \|v_\eta - u^*\|_{L^1} = 0.$$

Now, set $\mathcal{X}_0 = \{x^*(t) \ \mid \ t \in [0, t(u^*) - \eta_0]\} \cup \overline{\mathcal{N}}$, $C_1 = \sup_{\mathcal{X}_0\times\mathbf{U}} |\partial_x f|$, and $C_2 = \sup_{\mathcal{X}_0\times\mathbf{U}} |\partial_u f|$. For every $t \geqslant 0$ and for every $\eta \in (0, \eta_0)$, we have

$$\begin{aligned}
&|x_\eta(t) - x^*(t)| \\
&\qquad = \left| \int_0^t f(x_\eta(s), v_\eta(s))\, ds - \int_0^t f(x^*(s), u^*(s))\, ds \right| \\
&\qquad \leqslant \int_0^t |f(x_\eta(s), v_\eta(s)) - f(x^*(s), v_\eta(s))|\, ds \\
&\qquad\quad + \int_0^t |f(x^*(s), v_\eta(s)) - f(x^*(s), u^*(s))|\, ds \\
&\qquad \leqslant C_1 \int_0^t |x_\eta(s) - x^*(s)|\, ds + C_2 \|u^* - v_\eta\|_{L^1},
\end{aligned}$$

and thus, by the Gronwall lemma, we get that $\|x_\eta(\cdot) - x^*(\cdot)\|_\infty \leqslant C_2 \|u^* - v_\eta\|_{L^1} e^{C_1 \bar{T}}$. In particular $\lim_{\eta\to 0} \|x_\eta(\cdot) - x^*(\cdot)\|_\infty = 0$.

Finally, let us prove (16). By continuity of L, there exist constants $c \in \mathbb{R}$ and $\bar{C} > 0$ such that $L(t, x^*(t), u^*(t)) \geqslant c$ for almost every $t \in [0, t(u^*)]$, and $|L(t, x, u)| \leqslant \bar{C}$ for almost every

$(t, x, u) \in [0, \bar{T}] \times \mathcal{X}_0 \times \mathbf{U}$. Then, we have

$$\begin{aligned} 0 &\leqslant \int_0^{t(u^*)-\eta+\tau_\eta} L(t, x_\eta(t), v_\eta(t))\, dt \\ &\quad - \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\, dt \\ &= \int_{t(u^*)-\eta}^{t(u^*)-\eta+\tau_\eta} L(t, x_\eta(t), v_\eta(t))\, dt \\ &\quad - \int_{t(u^*)-\eta}^{t(u^*)} L(t, x^*(t), u^*(t))\, dt \\ &\leqslant \bar{C}\tau_\eta - c\eta, \end{aligned}$$

which implies (16). To prove (17) it suffices to note that $\tau_\eta \leqslant M\Upsilon(x^*(t(u^*) - \eta)) \leqslant C\eta^\alpha$. □

We are now in a position to prove Theorem 2.

Proof of Theorem 2. Assumption (Ω) implies in particular the existence of bounded variation controls steering the control system (Σ) from any initial condition in the neighborhood $\mathcal{N}$ to the origin. Hence, from Theorem 16 in Appendix B, the problem $(\mathrm{OCP})_\varepsilon$ has at least one solution.

Let $x_\varepsilon(\cdot)$ be an arbitrary solution of $(\mathrm{OCP})_\varepsilon$, associated with a control $u_\varepsilon : [0, T_\varepsilon] \to \mathbf{U}$. Assume u^* chattering and let $n_0 \in \mathbb{N}$ be such that $x^*(t_n) \in \mathcal{N}$ for every $n \geqslant n_0$. We apply Lemma 7 with $\eta = t(u^*) - t_n$ and, for simplicity, we denote by u_n the control $v_{t(u^*)-t_n}$ and by τ_n the time $\tau_{t(u^*)-t_n}$. Note that $\mathrm{TV}(u_n) \leqslant c_1 n + M$, where c_1 is the diameter of $\mathbf{U}$. By optimality of u_ε for $(\mathrm{OCP})_\varepsilon$, we have

$$\begin{aligned} &\int_0^{T_\varepsilon} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt \\ &\qquad \leqslant \int_0^{T_\varepsilon} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt + \varepsilon\, \mathrm{TV}(u_\varepsilon) \\ &\qquad \leqslant \int_0^{t_n+\tau_n} L(t, x_n(t), u_n(t))\, dt + \varepsilon\, \mathrm{TV}(u_n) \\ &\qquad \leqslant \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\, dt \\ &\qquad\quad + C|t(u^*) - t_n|^\alpha + \varepsilon(c_1 n + M). \end{aligned} \tag{19}$$

Now, by Assumption *(ii)* made in the statement of the Theorem, we have $|t(u^*) - t_n|^\alpha = \mathrm{O}(n^{-\alpha\beta})$, and choosing $n = \mathrm{O}(\varepsilon^{-\frac{1}{1+\alpha\beta}})$, we infer that

$$\begin{aligned} &\int_0^{T_\varepsilon} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt - \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\, dt \\ &\qquad \leqslant C|t(u^*) - t_n|^\alpha + \varepsilon(n + M) = \mathrm{O}\left(\varepsilon^{\frac{\alpha\beta}{1+\alpha\beta}}\right). \end{aligned}$$

This concludes the proof in the chattering case. Let, now, u^* be of bounded variation. Let $\eta_0 > 0$ be such that $x^*(s) \in \mathcal{N}$ for every $s > T^* - \eta_0$. Apply Lemma 7 to obtain the control v_η. Then $TV(v_\eta) \leqslant TV(u^*) + M = \tilde{M}$. Using the same reasoning as in the chain of inequalitites (19) we deduce

$$\int_0^{T_\varepsilon} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt \leqslant \int_0^{t(u^*)} L(t, x^*(t), u^*(t))\, dt + C\eta + \tilde{M}\varepsilon.$$

Finally, the proof is concluded by choosing $\eta = O(\varepsilon)$. □

3.3 Proof of Theorem 3

We start with the following existence result.

Lemma 8. *Given any $\varepsilon > 0$, the problem* $(\mathrm{OCPS})_\varepsilon$ *has at least one solution.*

Proof. Let $\varepsilon > 0$ be fixed. First of all, remark that if there exists no admissible trajectory such that $\mathrm{TV}(X_u) < +\infty$, then the functional $u \mapsto \int_0^T L(s, x(s), u(s))\, ds + \varepsilon\, \mathrm{TV}(X_u)$ is infinite and there is nothing to prove. Otherwise, let $I < +\infty$ denote the infimum in $(\mathrm{OCPS})_\varepsilon$. We consider a minimizing sequence of admissible controls $u_n : [0, t(u_n)] \to \mathbf{U}$, with corresponding trajectories denoted by $x_n(\cdot)$, such that

$$\lim_{n\to\infty} \left(\int_0^{t(u_n)} L(s, x_n(s), u_n(s))\, ds + \varepsilon\, \mathrm{TV}(X_{u_n}) \right) = I.$$

Since the sequence $(t(u_n))_{n\in\mathbb{N}}$ is bounded by Assumption (5), we can assume that $t(u_n)$ converges (up to some subsequence) to some $t_\varepsilon > 0$. Using $f(0,0) = 0$ we extend u_n to $[t(u_n), t_\varepsilon + \delta]$ by 0 for $\delta > 0$. Reasoning as in the proof of Theorem 1, up to some subsequence, there exist a positive measurable function $\gamma : [0, t_\varepsilon + \delta] \to \mathbb{R}$ and a measurable control $w : [0, t_\varepsilon + \delta] \in \mathbf{U}$, with corresponding trajectory $x_w(\cdot)$, such that $x_n(\cdot)$ converges to $x_w(\cdot)$ uniformly on $[0, t_\varepsilon + \delta]$ and $L(\cdot, x_n(\cdot), u_n(\cdot))$ converges to $L(\cdot, x_w(\cdot), w(\cdot)) + \gamma(\cdot)$ in $L^\infty(0, t_\varepsilon + \delta)$ for the weak star topology. By uniform convergence of trajectories, $w : [0, t_\varepsilon] \to \mathbf{U}$ is admissible, that is, $x_w(t_\varepsilon) = 0$ and $x_w(t) \in \mathcal{C}$ for every t. Up to some subsequence, by dominated convergence, we can assume that $X_{u_n}(\cdot) = \mathbf{1}_{\partial\mathcal{C}}(x_n(\cdot))$ converges to $X_w(\cdot) = \mathbf{1}_{\partial\mathcal{C}}(x_w(\cdot))$ in $L^1(0, t_\varepsilon + \delta)$. Moreover, since $x_n(t) = 0$ on $[t(u_n), t_\varepsilon + \delta]$, we have $\mathrm{TV}(X_{u_n}|_{[0, t_\varepsilon+\delta]}) = \mathrm{TV}(X_{u_n}|_{[0, t(u_n)]})$. Therefore,

$$\begin{aligned}
&\int_0^{t_\varepsilon+\delta} L(t, x_n(t), u_n(t))\, dt + \varepsilon\, \mathrm{TV}(X_{u_n}|_{[0,t_\varepsilon+\delta]}) \\
&\qquad = \int_0^{t(u_n)} L(t, x_n(t), u_n(t))\, dt + \varepsilon\, \mathrm{TV}(X_{u_n}|_{[0,t(u_n)]}) \\
&\qquad\quad + \int_{t(u_n)}^{t_\varepsilon+\delta} L(t, x_n(t), u_n(t))\, dt.
\end{aligned}$$

We infer that

$$\begin{aligned}
&\limsup_{n\to\infty} \left(\int_0^{t_\varepsilon+\delta} L(t, x_n(t), u_n(t))\, dt + \varepsilon\, \mathrm{TV}(X_{u_n}|_{[0,t_\varepsilon+\delta]}) \right) \\
&\qquad \leqslant I + \limsup_{n\to\infty} \int_{t(u_n)}^{t_\varepsilon+\delta} L(t, x_n(t), u_n(t))\, dt \\
&\qquad = I + \int_{t_\varepsilon}^{t_\varepsilon+\delta} \left(L(t, x_w(t), w(t)) + \gamma(t) \right) dt.
\end{aligned}$$

Besides, by lower semicontinuity of $\mathrm{TV}(\cdot)$, we have $\mathrm{TV}(X_w) < +\infty$ and

$$\begin{aligned}\liminf_{n\to\infty}&\left(\int_0^{t_\varepsilon+\delta} L(t,x_n(t),u_n(t))\,dt + \varepsilon\,\mathrm{TV}(X_{u_n}|_{[0,t_\varepsilon+\delta]})\right)\\ &\geqslant \liminf_{n\to\infty}\int_0^{t_\varepsilon+\delta} L(t,x_n(t),u_n(t))\,dt + \varepsilon\,\mathrm{TV}(X_w|_{[0,t_\varepsilon+\delta]})\\ &\geqslant \int_0^{t_\varepsilon+\delta}(L(t,x_w(t),w(t))+\gamma(t))\,dt + \varepsilon\,\mathrm{TV}(X_w|_{[0,t_\varepsilon]}).\end{aligned}$$

Finally, we obtain that $I \geqslant \int_0^{t_\varepsilon}(L(t,x_w(t),w(t))+\gamma(t))\,dt+\varepsilon\,\mathrm{TV}(X_w|_{[0,t_\varepsilon]})$. Since w is admissible, we have $x_w(t_\varepsilon)=0$ and there holds $I \leqslant \int_0^{t_\varepsilon} L(t,x_w(t),w(t))\,dt+\varepsilon\,\mathrm{TV}(X_w|_{[0,t_\varepsilon]})$. Therefore, since $\gamma\geqslant 0$, we infer that $\int_0^{t_\varepsilon}\gamma(t)\,dt=0$ and $w:[0,t_\varepsilon]\to\mathbf{U}$ is optimal for $(\mathrm{OCPS})_\varepsilon$. □

Lemma 9. *Assume condition (i) of Theorem 3. Let $\bar u:[0,\bar t]\to\mathbf{U}$ be an admissible control such that the corresponding trajectory $\bar x(\cdot)$ satisfies $\bar x(t)\in\mathcal{C}$ for every t and $\{t \mid \bar x(t)\in\partial\mathcal{C}\} = \{\bar t, t_1, t_2,\dots\}$ with $\lim_{n\to\infty} t_n = \bar t$. Then, there exists a sequence $(u_k)_{k\in\mathbb{N}}$ of $\mathcal{U}$ of admissible controls, such that u_k converges to $\bar u$ in L^1, the corresponding trajectories $x_k(\cdot)$ satisfy $x_k(t)\in\mathcal{C}$ for every t, and $\mathrm{TV}(X_{u_k})<+\infty$.*

Proof. Fix $k>0$. Recall that condition (i) states that there exists a neighborhood $\mathcal{N}$ of 0 such that $\mathcal{N}\cap\mathring{\mathcal{C}}\subset\mathcal{A}^{\mathcal{C}}(0,(0,1/k),-f)$. By assumption, for almost every $\eta>0$, the point $x^\eta=\bar x(\bar t-\eta)$ belongs to the interior of $\mathcal{C}$. Hence for almost every $\eta>0$ sufficiently small we have $x^\eta\in\mathcal{N}\cap\mathring{\mathcal{C}}\subset\mathcal{A}^{\mathcal{C}}(0,(0,\eta),-f)$. Then, there exists a control $w^\eta:[0,\tau_\eta]\to\mathbf{U}$, with $\tau_\eta\leqslant\eta$, such that the solution $y(\cdot)$ of the Cauchy problem $\dot y=-f(y,w)$, $y(0)=0$, satisfies $y(t)\in\mathring{\mathcal{C}}$ for every $t\in(0,\tau_\eta]$ and $y(\tau_\eta)=x^\eta$. Reversing time, since the dynamics is autonomous, we get that $z(\tau_\eta)=0$, where $z(\cdot)$ is the solution of the Cauchy problem $\dot z=f(z,w)$, $z(0)=x^\eta$, and $z(t)\in\mathring{\mathcal{C}}$ for every $t\in[0,\tau_\eta)$. Let $T>0$. We extend the control $\bar u$ to $[\bar t,\bar t+T]$ by setting $\bar u=0$. We define

$$u_\eta(t)=\begin{cases}\bar u(t), & t\in[0,\bar t-\eta]\\ w^\eta(t-\bar t+\eta), & t\in[\bar t-\eta,\bar t-\eta+\tau_\eta]\\ 0, & t>\bar t-\eta+\tau_\eta.\end{cases}$$

Since τ_η converges to 0 as $\eta\to 0$, u_η converges to $\bar u$ in $L^1(0,\bar t+T)$. Therefore, the sequence of corresponding trajectories $x_\eta(\cdot)$ converges uniformly to $\bar x(\cdot)$ on $[0,\bar t+T]$ (see Figure 4). Thus $L(\cdot,x_\eta(\cdot),u_\eta(\cdot))$ converges to $L(\cdot,\bar x(\cdot),\bar u(\cdot))$ strongly in $L^\infty(0,\bar t+T)$. Set $T_\eta=\bar t-\eta+\tau_\eta$. By construction, we have $x_\eta(T_\eta)=0$ and $\mathrm{TV}(\mathbf{1}_{\partial\mathcal{C}}(x_\eta)|_{[0,T_\eta]})=\mathrm{TV}(\mathbf{1}_{\partial\mathcal{C}}(\bar x)|_{[0,\bar t-\eta]})<+\infty$. Finally, the convergences above imply that $\int_0^{T_\eta} L(t,x_\eta(t),u_\eta(t))\,dt$ converges to $\int_0^{\bar u} L(t,\bar x(t),\bar u(t))\,dt$ as $\eta\to 0$. The statement follows by taking a sequence $\eta=1/k$ for $k\in\mathbb{N}$ sufficiently large. □

We are now in a position to prove Theorem 3.

Proof of Theorem 3. Let $x_\varepsilon(\cdot)$ be any optimal solution of $(\mathrm{OCPS})_\varepsilon$, associated with a control u_ε (existence is ensured by Lemma 8). We make the same reasoning as in the proof of Theorem 1.

Let $t(u_\varepsilon)$ converge (up to some subsequence) to some $t_1>0$. Let $\delta>0$ be arbitrary. We extend u_ε to $[0,t_1+\delta]$ by 0. As in the previous proofs, there exists an admissible control $w:[0,t_1+\delta]\to\mathbf{U}$, with corresponding trajectory $x_w(\cdot)$, and a positive measurable function $\gamma:[0,t_1+\delta]\to\mathbb{R}$ such that $x_\varepsilon(\cdot)$ converges to $x_w(\cdot)$ uniformly on $[0,t_1+\delta]$, and $L(\cdot,x_\varepsilon(\cdot),u_\varepsilon(\cdot))$ converges to $L(\cdot,x_w(\cdot),w(\cdot))+\gamma(\cdot)$ in $L^\infty(0,t_1+\delta)$ for the weak star topology. Replacing the total variation of controls $t\mapsto u(t)$

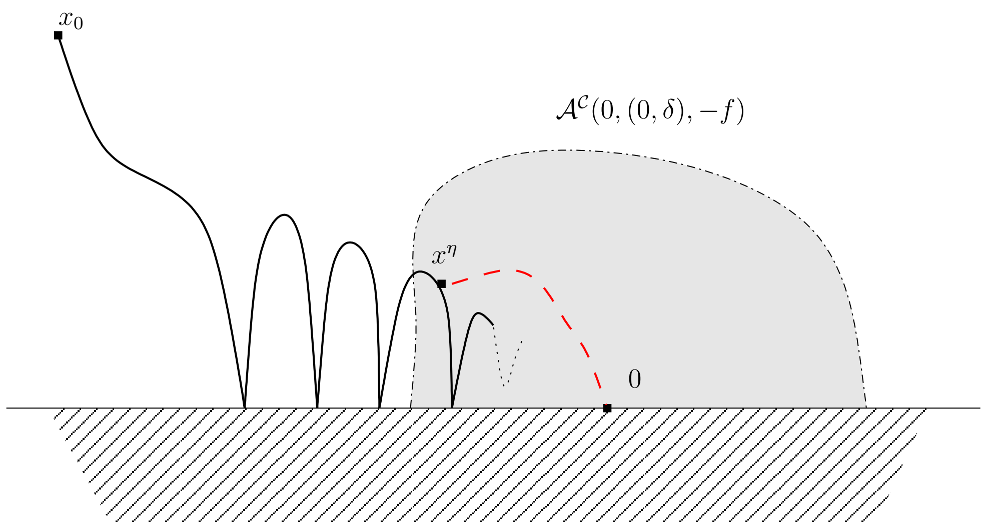

Figure 4: The trajectory $x_\eta(\cdot)$ associated with the control u_η.

with the total variation of $t \mapsto X_u(t) = \mathbf{1}_{\partial\mathcal{C}}(x_u(t))$ in (15), we get that, for every admissible control $v : [0, t(v)] \to \mathbf{U}$ such that $\mathrm{TV}(X_v) < +\infty$, there holds

$$\int_0^{t_1} L(t, x_w(t), w(t))\, dt \\ \leqslant \int_0^{t(v)} L(t, x_v(t), v(t))\, dt + \int_{t_1}^{t_1+\delta} \gamma(t)\, dt.$$

Since $\delta > 0$ is arbitrary, letting δ tend to zero we conclude that, for every v as above,

$$\int_0^{t_1} L(t, x_w(t), w(t))\, dt \leqslant \int_0^{t(v)} L(t, x_v(t), v(t))\, dt. \tag{20}$$

We apply Lemma 9 to $\bar{u} = u^*$ and we denote by u_k the corresponding sequence. Then, taking inequality (20) with $v = u_k$ and letting k tend to $+\infty$, we obtain that w is optimal for (OCPS). In order to establish (11), it remains to prove that $\gamma|_{[0,t_1]} \equiv 0$. To this aim, let $v : [0, t(v)] \to \mathbf{U}$ be an admissible control such that $\mathrm{TV}(X_v) < +\infty$. Then, by optimality of u_ε for $(\mathrm{OCPS})_\varepsilon$, we have

$$\begin{aligned}
\int_0^{t_1+T} &(L(t, x_w(t), w(t)) + \gamma(t))\, dt \\
&= \lim_{\varepsilon\to 0} \int_0^{t_1+T} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt \\
&\leqslant \limsup_{\varepsilon\to 0} \left(\int_0^{t_1+T} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt + \varepsilon\, \mathrm{TV}(X_{u_\varepsilon}) \right) \\
&\leqslant \int_0^{t(v)} L(t, x_v(t), v(t))\, dt \\
&\quad + \limsup_{\varepsilon\to 0} \int_{t(u_\varepsilon)}^{t_1+T} L(t, x_\varepsilon(t), u_\varepsilon(t))\, dt \\
&= \int_0^{t(v)} L(t, x_v(t), v(t))\, dt \\
&\quad + \int_{t_1}^{t_1+T} (L(t, x_w(t), w(t)) + \gamma(t))\, dt,
\end{aligned}$$

which gives $\int_0^{t_1} (L(t, x_w(t), w(t)) + \gamma(t))\, dt \leqslant \int_0^{t(v)} L(t, x_v(t), v(t))\, dt$. Again, let $(u_k)_{k\in\mathbb{N}}$ be the sequence provided by Lemma 9 with $\bar{u} = u^*$. Then, since w is optimal for (OCPS), the inequality above with $v = u_k$ implies that $\int_0^{t_1} \gamma(t)\, dt = 0$, which gives $\gamma_{[0,t_1]} = 0$. □

3.4 Proof of Theorem 4

We first prove an auxiliary lemma.

Lemma 10. *Let $\mathcal{H}$ be a hybrid system such that X_q is a compact submanifold for every $q \in Q$, and such that the sets $G(q, q'), R((q, q'), x)$ are compact for every $((q, q'), x) \in E \times X_q$, with $q \in Q$.*

Let L be a Lagrangian for $\mathcal{H}$ with corresponding cost functional $C(\cdot)$. Assume that $(\tau^*, q^*(\cdot), x^*(\cdot))$ is a Zeno solution of (HP). *Let $\tau^* = \{\tau_i^*\}_{i=0}^\infty$. Define the sequence of trajectories $(\tau^n, q^n(\cdot), x^n(\cdot))$ of $\mathcal{H}$ by*

- *$\tau^n = \{\tau_0^*, \tau_1^*, \ldots, \tau_n^*, \tau_\infty^*\}$;*
- *$q^n(t) = q^*(t)$ for every $t \in [0, \tau_n^*)$, $q^n(t) \equiv q^*(\tau_n^*)$ for $t \in [\tau_n^*, \tau_\infty^*]$;*
- *$x^n(t) = x^*(t)$ for every $t \in [0, \tau_n^*]$, and on $[\tau_n^*, \tau_\infty^*]$ the (continuous) trajectory $x^n(\cdot)$ is solution of $\dot{x}^n(t) = f_{q^*(\tau_n^*)}(x^n(t))$ almost everywhere.*

Then:

$$\sup_{\tau_0^* \leqslant t \leqslant \tau_n^*} ||x^n(t) - x^*(t)|| \leqslant \mathrm{O}(\tau_\infty^* - \tau_n^*), \tag{21}$$

$$C(\tau^n, q^n(\cdot), x^n(\cdot)) - C(\tau^*, q^*(\cdot), x^*(\cdot)) \leqslant \mathrm{O}(\tau_\infty^* - \tau_n^*). \tag{22}$$

Proof. Since $q^n(t)$ converges to $q^*(t)$ almost everywhere in $[0, \tau_\infty^*]$, by standard convergence results (see for instance [37, Theorem 1 p. 57]) we deduce (21). For (22), note that, since the Lagrangian is continuous, there exist positive constants $\tilde{c}$ and c, satisfying $\tilde{c} - c > 0$, such that

$\int_{\tau_n^*}^{\tau_\infty^*} L_{q^*(\tau_n^*)}(t, x^n(t))\, dt \leqslant \tilde{c}(\tau_\infty^* - \tau_n^*)$ for every n, and $L_{q^*(\tau_i^*)}(t, x^*(t)) \geqslant c$ almost everywhere in $[\tau_i^*, \tau_{i+1}^*]$ for every i. Therefore,

$$\begin{aligned}
0 &\leqslant C(\tau^n, q^n(\cdot), x^n(\cdot)) - C(\tau^*, q^*(\cdot), x^*(\cdot)) \\
&= \int_{\tau_n^*}^{\tau_\infty^*} L_{q^*(\tau_{n_i})}(t, x^n(t))\, dt - \sum_{i=n}^{\infty} \int_{\tau_i^*}^{\tau_{i+1}^*} L_{q^*(t)}(t, x^*(t))\, dt \\
&\leqslant \tilde{c}(\tau_\infty^* - \tau_n^*) - c \sum_{i=n}^{\infty} (\tau_{i+1}^* - \tau_i^*) = (\tilde{c} - c)(\tau_\infty^* - \tau_n^*).
\end{aligned}$$

This concludes the proof of (22). □

Consider the set $\mathcal{T}_M$ of trajectories of a hybrid system having at most M switchings. If $M < +\infty$ these trajectories are non-Zeno. We say that two non-Zeno trajectories have the *same history* if they visit the same locations in the same sequence. Having the same history is an equivalence relation in $\mathcal{T}_M$ and the number of equivalence classes in $\mathcal{T}_M$ is finite.

Let us now prove Theorem 4.

Proof of Theorem 4. By compactness of the location X_{q_0} there exists at least one trajectory starting at (q_0, x_0) and having no location switchings [1]. Hence, for every $\varepsilon > 0$ the functional to minimize in $(\mathrm{HP})_\varepsilon$ is finite. Since the hybrid cost functional C is bounded from below, there exists M_ε such that any optimal trajectory of $(\mathrm{HP})_\varepsilon$ has at most M_ε switchings. Hence any solution of

$$\begin{cases} \min C(\tau, q(\cdot), x(\cdot)) + \varepsilon\, \mathrm{TV}(h \circ q(\cdot)), \\ (\tau, q(\cdot), x(\cdot)) \in \mathcal{T}_{M_\varepsilon}, \\ q(0) = q_0, \quad x(0) = x_0, \end{cases} \qquad (\mathrm{HP})_\varepsilon'$$

where the minimization runs over all possible trajectories having only a finite number of switchings M_ε is also a solution of $(\mathrm{HP})_\varepsilon$. Now consider a minimizing sequence for $(\mathrm{HP})_\varepsilon'$. Then, up to some subsequence, we can assume that all trajectories have the same history. Hence the penalization term of total variation is constant along the chosen subsequence and the problem is then reduced to that of minimizing the Lagrangian cost $C(\cdot, \cdot, \cdot)$ among trajectories with a fixed history. Hence, by compactness, this problem has at least one solution, see [38, Theorem 1].

Let $(\tau^\varepsilon, q^\varepsilon(\cdot), x^\varepsilon(\cdot))$ be a solution of $(\mathrm{HP})_\varepsilon'$, then it is also a solution of $(\mathrm{HP})_\varepsilon$. We apply Lemma 10, and we consider the corresponding sequence $(\tau^n, q^n(\cdot), x^n(\cdot))$, which by construction has a finite number of location switchings. Then, by optimality and using (22),

$$\begin{aligned}
0 &\leqslant C(\tau^\varepsilon, q^\varepsilon(\cdot), x^\varepsilon(\cdot)) - C(\tau^*, q^*(\cdot), x^*(\cdot)) \\
&\leqslant C(\tau^\varepsilon, q^\varepsilon(\cdot), x^\varepsilon(\cdot)) - C(\tau^*, q^*(\cdot), x^*(\cdot)) + \varepsilon\, \mathrm{TV}(h \circ q^\varepsilon(\cdot)) \\
&\leqslant C(\tau^n, q^n(\cdot), x^n(\cdot)) - C(\tau^*, q^*(\cdot), x^*(\cdot)) + \varepsilon\, \mathrm{TV}(h \circ q^n(\cdot)) \\
&\leqslant \mathrm{O}(\tau_\infty^* - \tau_n^*) + \varepsilon n |Q|,
\end{aligned}$$

where $|Q|$ is the number of locations. Choose $n = \lfloor \varepsilon^{-1/2} \rfloor$. The convergence (13) follows by letting ε converge to 0. □

Remark 11. If the rate of convergence of τ_n^* to τ_∞^* is known, then it is possible to determine the rate of convergence in (13). For instance, if $\tau_n^* - \tau_\infty^* \leqslant \mathrm{O}(n^{-\beta})$ for some $\beta > 0$, then, for every $\alpha > 0$, we have

$$C(\tau^\varepsilon, q^\varepsilon(\cdot), x^\varepsilon(\cdot)) - C(\tau^*, q^*(\cdot), x^*(\cdot)) \leqslant \mathrm{O}\left(\varepsilon^{\min(1-\alpha, \alpha\beta)}\right).$$

[1] By definition, a vector field on a compact manifold with boundary is tangent to the boundary. The compactness of the location ensures that trajectories staying in one location are defined for all times.

A Further comments on condition (Ω)

The relation between condition (Ω) and small-time local controllability depends on the continuity of the time-optimal map Υ. Recall that $\Upsilon(y)$ is the minimal time needed to steer the control system (Σ) from y to 0.

Note that, in Definition 2, (Ω_2) does not imply (Ω_1) in general, as the following example shows.

Example 2. *Consider the control system*

$$\dot{x} = uf_1(x) + vf_2(x), \quad x = (x_1, x_2) \in \mathbb{R}^2, \quad (u, v) \in [-1, 1]^2,$$

where $f_1(x) = \partial_{x_1}$, $f_2(x) = h(x_1)\partial_{x_2}$, namely

$$\begin{cases} \dot{x}_1 &= u \\ \dot{x}_2 &= h(x_1)v. \end{cases}$$

with

$$h(x_1) = \begin{cases} 0, & \text{if } x_1 \in [-1, 1], \\ 1, & \text{if } x_1 \notin (-2, 2), \end{cases}$$

and h is a smooth function with $h(x_1) \in [0, 1]$ for every $x_1 \in \mathbb{R}$.

The control system is clearly not STLC at 0. However every point of $\mathbb{R}^2$ can be steered to the 0 with at most two switches. Moreover every point y in the open strip $\mathcal{N} = (-1, 1) \times \mathbb{R}$ can be steered to 0 with two switches in time $\tau_y \leqslant 4\Upsilon(y)$. Indeed consider for instance $y = (y_1, y_2) \in \mathcal{N}$ with $y_1 \geqslant 0, y_2 > 0$ (the other cases can be treated similarly). The control

$$(u(t), v(t)) = \begin{cases} (1, 0), & t \in [0, 2 - y_1) \\ (0, -1), & t \in [2 - y_1, 2 - y_1 + y_2) \\ (-1, 0), & t \in [2 - y_1 + y_2, 4 - y_1 + y_2] \end{cases}$$

steers y to 0 in time $\tau_y = 4 - y_1 + y_2 \leqslant 4 + y_2$ while $\Upsilon(y) \geqslant 1 + y_2$. Hence the control system satisfies condition (Ω_2).

In the example above, the time-optimal map is not continuous at 0. Indeed $\Upsilon(0) = 0$ while $\Upsilon((0, x_2)) \geqslant 2$ for every $x_2 \neq 0$. A relationship between (Ω_2) and (Ω_1) can be established depending on the continuity of the time-optimal map Υ.

Proposition 11. *The following conditions are equivalent.*

- (a) *Υ is continuous at 0.*
- (b) *Condition (Ω_2) implies condition (Ω_1).*

Proof. $(a) \Rightarrow (b)$. A stronger assertion actually holds, namely, (a) implies (Ω_1). Indeed, if Υ is continuous at 0, since $\Upsilon(0) = 0$, for every $\varepsilon > 0$, $\Upsilon^{-1}([0, \varepsilon))$ is a neighborhood of 0 and every point in $\Upsilon^{-1}([0, \varepsilon))$ can be steered to 0 in time less than ε for the control system (Σ).
$(b) \Rightarrow (a)$. This is a consequence of the classical fact that, if (Σ) is STLC at 0, then Υ is continuous at 0 (see [24, Theorem 2.2]). □

In the rest of this section, we present sufficient conditions ensuring (Ω). First, note that in the simple case of a driftless control-affine system, (Ω) is a consequence of the Lie Algebra Rank Condition. In this case the number of switchings needed to reach any point in a small neighborhood of 0 depends only on the step of the Lie algebra $\mathrm{Lie}(f_1, \ldots, f_m)$ at 0.

Proposition 12. *For a driftless control-affine system* $\dot{x} = \sum_{i=1}^{m} u_i f_i(x)$*, if* $\mathrm{Lie}_0(f_1, \ldots, f_m) = \mathbb{R}^N$*, then* (Ω) *is satisfied at* 0*.*

For control-affine systems with a drift, a sufficient condition comes from the classical result by Sussmann [23] in the single-input case. The main assumption in [23] (denoted by (Δ) in this reference) involves Lie brackets between the drift vector field and the controlled vector field (we also refer to [22] for more precise estimates on the number of switchings in a particular case). More precisely we have the following result.

Proposition 13. *Consider the single-input control-affine system* $\dot{x} = f(x) + ug(x)$*, where* f *and* g *are analytic vector fields in* $\mathbb{R}^N$*. If the condition* (Δ) *of [23] is satisfied, and if the control system is STLC at* 0*, then* (Ω) *holds true at* 0*.*

Proof. By [23], the system satisfies the bang-bang property with bounds on the number of switchings (BBBNS). More precisely, for every K compact and for every $T > 0$, there exists $n_0 \in \mathbb{N}^*$ such that, if $x(\cdot)$ is a time-optimal trajectory that is entirely contained in K and steers the control system from $x \in K$ to $y \in K$, then there exists a time-optimal trajectory steering as well the control system from x to y, which is moreover bang-bang with at most n_0 switchings, with n_0 depending on K and T. Since the control system is STLC at 0, the set $K = \{x \mid \Upsilon(x) \leqslant 1\}$ is a compact set containing 0 in its interior. Every $x \in K$ can be steered to 0 in time $\Upsilon(x)$ with at most n_0 switchings. □

Linear autonomous systems generically satisfy (Ω), as established next.

Proposition 14. *If the linear autonomous control system* $\dot{x} = Ax + Bu$ *satisfies the Kalman condition, then* (Ω) *holds true.*

Proof. It suffices to write the system in Brunowsky form (see, e.g., [37, Theorem 14, Section 5.2]). The time-optimal control of a cascade system has a number of switchings depending only on *Kronecker indices* (or *controllability indices*) of the system (see also [39]). □

As a consequence, we have the following sufficient condition for control-affine systems.

Proposition 15. *Consider the control affine system* $\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x)$*. We set*

$$G_i = \mathrm{span}\{\mathrm{ad}_f^k g_j \ \mid\ 0 \leqslant k \leqslant i,\ 1 \leqslant j \leqslant m\}.$$

Assume that:

(i) for every $1 \leqslant i \leqslant N-1$*, the distribution* G_i *has constant dimension near* 0*;*

(ii) the distribution G_{N-1} *has dimension* N*;*

(iii) for every $1 \leqslant i \leqslant N-2$*, the distribution* G_i *is involutive.*

Then (Ω) *holds true at* 0*.*

Proof. The result follows from Proposition 14 and from the fact that the *State Space Exact Linearization Problem* is solvable (see, e.g., [40, Theorem 5.2.3]). □

B An existence result

For every $\varepsilon \geqslant 0$, consider the optimal control problem

$$\left\{ \begin{array}{l} \min\limits_{u\in\mathcal{U}} \left(\displaystyle\int_0^{t(u)} L(s,x(s),u(s))\,ds + \varepsilon\,\mathrm{TV}(u) \right), \\ \dot{x} = f(t,x,u), \quad u \in \mathcal{U}, \\ x(t) \in \mathcal{C}, \quad t \in [0,t(u)], \\ x(0) \in M_0, \quad x(t(u)) \in M_1, \end{array} \right. \tag{OCPS$_\varepsilon$}$$

where

- $f : \mathbb{R} \times \mathbb{R}^N \times \mathbf{U} \to \mathbb{R}^N$ is measurable w.r.t. t, locally Lipschitz w.r.t. x,
- $L \in \mathcal{C}^0(\mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^m)$,
- $\mathcal{C} = \{x \in \mathbb{R}^N \mid h_1(x) \geqslant 0, \ldots, h_l(x) \geqslant 0\}$ for some $h_1, \ldots h_l \in \mathcal{C}^0(\mathbb{R}^N)$,
- $\mathbf{U} \subset \mathbb{R}^m$ is compact,
- M_0 and M_1 are compact subsets of $\mathcal{C}$.

Here, $\mathcal{U}$ is still defined by (3).

Theorem 16. *Assume that:*

(i) there exists $\bar{u} \in \mathcal{U}$ having bounded variation, steering the control system $\dot{x} = f(t,x,u)$ from M_0 to M_1, and whose corresponding trajectory satisfies the state constraint $x(t) \in \mathcal{C}$, for every $t \in [0,t(\bar{u})]$;

(ii) there exists $b > 0$ such that, for every $u \in \mathcal{U}$ steering the control system from M_0 to M_1, its corresponding trajectory x_u satisfies $t(u) + \|x_u(\cdot)\|_\infty \leqslant b$.

Then, for every $\varepsilon > 0$, the optimal control problem (OCPS)$_\varepsilon$ *has at least one solution.*

Note that existence is not ensured for $\varepsilon = 0$. The fact that $\varepsilon > 0$ is crucial here. The difference with usual existence theorems is that, in the proof below, we use in an instrumental way the total variation term. Note the remarkable fact that, in contrast to usual existence theorems (see [16]), we do not assume, here, that the set of extended velocities (4) is convex. This classical assumption can be removed thanks to the use of the total variation term.

Proof. The proof follows the lines of [18, Theorem 5.14 and 6.15], with an adaptation to the bounded variation context. Let

$$\delta = \inf \left(\int_0^{t(u)} L(s,x(s),u(s))\,ds + \varepsilon\,\mathrm{TV}(u) \right),$$

where the infimum is taken among all controls $u \in \mathcal{U}$ steering the control system from M_0 to M_1 and whose corresponding trajectory satisfies the state constraint $x(t) \in \mathcal{C}$, for every $t \in [0,t(u)]$. Let $x_n(\cdot)$ be a sequence of admissible trajectories, corresponding to a minimizing sequence of admissible controls $u_n : [0,t(u_n)] \to \mathbf{U}$, i.e.,

$$\lim_{n\to\infty} \left(\int_0^{t(u_n)} L(s,x_n(s),u_n(s))\,ds + \varepsilon\,\mathrm{TV}(u_n) \right) = \delta.$$

Using Assumptions (i) and (ii), for n sufficiently large we have

$$\varepsilon\,\mathrm{TV}(u_n) \leqslant \int_0^{t_(\bar{u})} L(s, x_{\bar{u}}(s), \bar{u}(s))\,ds + \varepsilon\,\mathrm{TV}(\bar{u}) + C,$$

for some constant $C \geqslant 0$, and since $t(u_n)$ is bounded by b, extending u_n by 0 for $t > t(u_n)$, we infer that the sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in the set $\mathrm{BV}([0,b],\mathbb{R}^m)$ of bounded variation functions from $[0,b]$ to $\mathbb{R}^m$. Since the embedding $\mathrm{BV}([0,b],\mathbb{R}^m) \hookrightarrow L^1([0,b],\mathbb{R}^m)$ is compact (see [41]), up to some subsequence, $(u_n)_{n\in\mathbb{N}}$, converges to some $u_\varepsilon \in L^1([0,b],\mathbb{R}^m)$ for the strong topology of L^1. Still up to some subsequence, $x_n(0)$ converge to some $x_\varepsilon^0 \in \mathbb{R}^N$, u_n converges to u_ε almost everywhere and $t(u_n)$ converges to $t(u_\varepsilon)$, and thus $u_\varepsilon : [0, t(u_\varepsilon)] \to \mathbf{U}$ takes values in $\mathbf{U}$.

Let us prove that $u_\varepsilon : [0, t(u_\varepsilon)] \to \mathbf{U}$ is a solution of $(\mathrm{OCPS})_\varepsilon$. By a standard Gronwall argument (see [37, Theorem 1 p. 57], or see [42, 18]), the convergence almost everywhere of u_n to u_ε implies that $x_n(\cdot)$ converges uniformly to $x_\varepsilon(\cdot)$, where $x_\varepsilon(\cdot)$ is the trajectory corresponding to the control u_ε and starting at x_ε^0. In particular, we get that $x_\varepsilon(t) \in \mathcal{C}$ for every $t \in [0, t(u_\varepsilon)]$ and, by compactness of M_0 and M_1, we obtain that $x_\varepsilon(t(u_\varepsilon)) \in M_1$. Hence u_ε is an admissible control. Moreover, $L(t, x_n(t), u_n(t))$ converges to $L(t, x_\varepsilon(t), u_\varepsilon(t))$ for almost every t. Hence, using Assumption (ii) and the dominated convergence theorem, we conclude that

$$\lim_{n\to\infty} \int_0^{t(u_n)} L(t, x_n(t), u_n(t))\,dt = \int_0^{t(u_\varepsilon)} L(t, x_\varepsilon(t), u_\varepsilon(t))\,dt. \tag{23}$$

On the other hand, by lower semicontinuity of the functional $\mathrm{TV}(\cdot)$, we have

$$\mathrm{TV}(u_\varepsilon) \leqslant \liminf_{n\to\infty} \mathrm{TV}(u_n). \tag{24}$$

Using (23), (24) and since u_n is a minimizing sequence, we infer that $\int_0^{t(u_\varepsilon)} L(t, x_\varepsilon(t), u_\varepsilon(t))\,dt + \varepsilon \mathrm{TV}(u_\varepsilon) \leqslant \delta$, which implies that u_ε is optimal. □

Acknowledgement

The authors are grateful to Jean-Michel Coron, Matthias Kawski, Andrey Sarychev, and Mikhail Il'ich Zelikin for useful discussions and valuable suggestions. The last author acknowledges the support of the ANR project Finite4SoS ANR-15-CE23-0007-01.

References

[1] A. T. Fuller, "Relay control systems optimized for various performance criteria," in *Proceedings of the 1st World Congress IFAC, Moscow, 1960, Butterworths, London (1961)*, pp. 510–519.

[2] I. Kupka, "The ubiquity of Fuller's phenomenon," in *Nonlinear controllability and optimal control*, ser. Monogr. Textbooks Pure Appl. Math. New York: Dekker, 1990, vol. 133, pp. 313–350.

[3] M. I. Zelikin and V. F. Borisov, *Theory of chattering control*, ser. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston Inc., 1994.

[4] H. Robbins, "Junction phenomena for optimal control with state-variable inequality constraints of third order," *J. Optim. Theory Appl.*, vol. 31, no. 1, pp. 85–99, 1980. [Online]. Available: http://dx.doi.org/10.1007/BF00934790

[5] J. Zhang, K. H. Johansson, J. Lygeros, and S. Sastry, "Zeno hybrid systems," *International Journal of Robust and Nonlinear Control*, vol. 11, no. 5, pp. 435–451, 2001. [Online]. Available: http://dx.doi.org/10.1002/rnc.592

[6] H. Schättler and U. Ledzewicz, *Geometric optimal control*, ser. Interdisciplinary Applied Mathematics. Springer, New York, 2012, vol. 38, theory, methods and examples. [Online]. Available: http://dx.doi.org/10.1007/978-1-4614-3834-2

[7] J. F. Bonnans and A. Hermant, "No-gap second-order optimality conditions for optimal control problems with a single state constraint and control," *Math. Program.*, vol. 117, no. 1-2, Ser. B, pp. 21–50, 2009. [Online]. Available: http://dx.doi.org/10.1007/s10107-007-0167-8

[8] U. Boscain and B. Piccoli, *Optimal syntheses for control systems on 2-D manifolds*, ser. Mathématiques & Applications (Berlin) [Mathematics & Applications]. Berlin: Springer-Verlag, 2004, vol. 43.

[9] B. Piccoli and H. J. Sussmann, "Regular synthesis and sufficiency conditions for optimality," *SIAM J. Control Optim.*, vol. 39, no. 2, pp. 359–410 (electronic), 2000. [Online]. Available: http://dx.doi.org/10.1137/S0363012999322031

[10] N. Bérend, J. F. Bonnans, J. Laurent-Varin, M. Haddou, and C. Talbot, "An Interior-Point Approach to Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 5, pp. 1228–1238, 2007.

[11] J. F. Bonnans and A. Hermant, "Well-posedness of the shooting algorithm for state constrained optimal control problems with a single constraint and control," *SIAM J. Control Optim.*, vol. 46, no. 4, pp. 1398–1430 (electronic), 2007. [Online]. Available: http://dx.doi.org/10.1137/06065756X

[12] J. Zhu, E. Trélat, and M. Cerf, "Minimum time control of the rocket attitude reorientation associated with orbit dynamics," *SIAM J. Cont. Optim.*, vol. 54, no. 1, pp. 391–422, 2016.

[13] ——, "Planar tilting maneuver of a spacecraft: singular arcs in the minimum time problem and chattering," *Discrete and Continuous Dynamical Systems - Series B*, vol. 21, no. 4, pp. 1347–1388, 2016.

[14] M. I. Zelikin and L. A. Manita, "Accumulation of control switchings in problems with distributed parameters," *J. Math. Sci. (N.Y.)*, vol. 151, no. 6, pp. 3506–3542, 2008.

[15] M. I. Zelikin and L. F. Zelikina, "Asymptotics of the deviation of a functional from its optimal value when the chattering is replaced by a suboptimal regime," *Russian Math. Surveys*, vol. 54, no. 3, pp. 662–664, 1999.

[16] L. Cesari, *Optimization—theory and applications*, ser. Applications of Mathematics (New York). New York: Springer-Verlag, 1983, vol. 17, problems with ordinary differential equations.

[17] A. Bressan and B. Piccoli, *Introduction to the mathematical theory of control*, ser. AIMS Series on Applied Mathematics, vol. 2.

[18] E. Trélat, *Contrôle optimal: théorie & applications.* Vuibert, 2008.

[19] G. Aubert and P. Kornprobst, *Mathematical problems in image processing*, 2nd ed., ser. Applied Mathematical Sciences. Springer, New York, 2006, vol. 147, partial differential equations and the calculus of variations, With a foreword by Olivier Faugeras.

[20] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1, pp. 89–97, Jan 2004. [Online]. Available: https://doi.org/10.1023/B:JMIV.0000011325.36760.1e

[21] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*, ser. Encyclopaedia of Mathematical Sciences. Berlin: Springer-Verlag, 2004, vol. 87.

[22] Y. Sharon and M. Margaliot, "Third-order nilpotent, nice reachability and asymptotic stability," *J. Differ. Equations*, vol. 233, pp. 136–150, 2007.

[23] H. J. Sussmann, "A bang-bang theorem with bounds on the number of switchings," *SIAM J. Control Optim.*, vol. 17, no. 5, pp. 629–651, 1979. [Online]. Available: http://dx.doi.org/10.1137/0317045

[24] R. M. Bianchini and G. Stefani, "Time-optimal problem and time-optimal map," *Rend. Sem. Mat. Univ. Politec. Torino*, vol. 48, no. 3, pp. 401–429 (1992), 1990.

[25] M. Krastanov and M. Quincampoix, "Local small time controllability and attainability of a set for nonlinear control system," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 6, pp. 499–516, 2001.

[26] A. Marigonda and T. T. Le, "Small-time local attainability for a class of control systems with state constraints," *ESAIM: Control, Optimization and Calculus of Variations*, 2016, to appear.

[27] M. I. Krastanov, "On the constrained small-time controllability of linear systems," *Automatica*, vol. 44, no. 9, pp. 2370–2374, 2008.

[28] B. Roszak and M. E. Broucke, "Necessary and sufficient conditions for reachability on a simplex," *Automatica*, vol. 42, no. 11, pp. 1913–1918, 2006.

[29] K. H. Johansson, M. Egerstedt, J. Lygeros, and S. Sastry, "On the regularization of Zeno hybrid automata," *Systems Control Lett.*, vol. 38, no. 3, pp. 141–150, 1999. [Online]. Available: http://dx.doi.org/10.1016/S0167-6911(99)00059-6

[30] J. Zhang, K. H. Johansson, J. Lygeros, and S. Sastry, *Dynamical systems revisited: Hybrid systems with Zeno executions*. Springer, 2000.

[31] A. D. Ames, A. Abate, and S. Sastry, "Sufficient conditions for the existence of zeno behavior," in *Decision and Control, 2005 European Control Conference*, 2005, pp. 696–701.

[32] M. Heymann, F. Lin, G. Meyer, and S. Resmerita, "Analysis of zeno behaviors in a class of hybrid systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 376–383, 2005.

[33] S. N. Simić, K. H. Johansson, S. Sastry, and J. Lygeros, "Towards a geometric theory of hybrid systems," in *Hybrid Systems: Computation and Control*. Springer, 2000, pp. 421–436.

[34] A. Lamperski and A. D. Ames, "Lyapunov theory for zeno stability," *IEEE Transactions on Automatic Control*, vol. 58, no. 1, pp. 100–112, 2013.

[35] K. A. Grasse, "On the relation between small-time local controllability and normal self-reachability," *Math. Control Signals Systems*, vol. 5, no. 1, pp. 41–66, 1992. [Online]. Available: http://dx.doi.org/10.1007/BF01211975

[36] W. Rudin, *Functional analysis*, 2nd ed., ser. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991.

[37] E. D. Sontag, *Mathematical control theory*, 2nd ed., ser. Texts in Applied Mathematics. Springer-Verlag, New York, 1998, vol. 6, deterministic finite-dimensional systems. [Online]. Available: http://dx.doi.org/10.1007/978-1-4612-0577-7

[38] B. Piccoli, "Hybrid systems and optimal control," in *Decision and Control, 1998. Proceedings of the 37th IEEE Conference on*, vol. 1. IEEE, 1998, pp. 13–18.

[39] E. B. Lee and L. Markus, *Foundations of optimal control theory.* John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[40] A. Isidori, *Nonlinear control systems.* Springer Science & Business Media, 1995.

[41] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, ser. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[42] E. Trélat, "Some properties of the value function and its level sets for affine control systems with quadratic cost," *J. Dynam. Control Systems*, vol. 6, no. 4, pp. 511–541, 2000. [Online]. Available: http://dx.doi.org/10.1023/A:1009552511132