# Hybrid maximum principle with regionally switching parameter

Térence Bayen, Anas Bouali, Loïc Bourdin

# Hybrid maximum principle with regionally switching parameter

Térence Bayen* Anas Bouali† Loïc Bourdin‡

April 12, 2022

**Abstract**

In this paper we consider a Mayer optimal control problem governed by a hybrid control system defined over a partition of the state space. We assume that the control system depends on a *regionally switching parameter* that remains constant in each region but that can change its value when the state position crosses interfaces. This new framework allows to deal, as a particular case, with control systems including non-control regions. In this paper our objective is to provide the corresponding necessary optimality conditions in a Pontryagin form. Our approach is based on a thorough sensitivity analysis of the hybrid control system under needle-like perturbations of the control and under convex perturbations of the parameter. To this aim we invoke implicit function arguments to deal with the interface crossings that are assumed to be transverse. The paper is concluded with a simple academic example showing that our framework allows to fill a gap in the literature.

**Keywords**: hybrid systems, optimal control, necessary optimality conditions, hybrid maximum principle, Pontryagin maximum principle, sensitivity analysis.

**AMS classification**: 34A38, 49K15, 93C57.

## 1 Introduction

Optimal control theory was developed at the end of the fifties with two major mathematical theorems, namely, the Hamilton-Jacobi-Bellman equation [12] (in short, HJB equation) and the Pontryagin Maximum Principle [52] (in short, PMP). The HJB equation focuses on sufficient optimality conditions, while the PMP represents, in some say, its counterpart about necessary optimality conditions. The objective of the present paper is to extend the latter in a new framework related to hybrid control systems.

To get into details of our new setting, we need to recall first the usual context of application of the PMP and some of its key issues. In many areas (such as engineering, biology, aeronautics, aerospace, etc.), cost functionals (like energies, time transfer, etc.) have to be minimized among solutions of a control system of the form

$$\dot{x}(t) = f(x(t), u(t), t), \qquad \text{a.e. } t \in [0, T],$$

where $x$ (resp. $u$) stands for the state position (resp. the control) and where $f : \mathbb{R}^n \times \mathbb{R}^m \times [0, T] \to \mathbb{R}^n$ is (usually) a $\mathrm{C}^1$ function called dynamics of the control system. The PMP has been developed especially to deal with such optimal control problems. It provides a so-called *Hamiltonian maximization condition* which guarantees, in particular, that the optimal control can be expressed as a feedback of the state position and a costate function.

**Initial motivation: control systems with non-control regions.** At this step, it is worth noticing that the literature usually presupposes that it is possible to change the control value $u(t)$ at any real time $t \in [0, T]$. In that case we speak of a *permanent control*. However, in several application models such as in automation, one can consider only a piecewise constant control (also known as *sampled-data control*) whose value can be modified only at certain instants (and remains constant otherwise). For instance, in the digital controller

*Avignon Université, Laboratoire de Mathéematiques d'Avignon (EA 2151) F-84018 terence.bayen@univ-avignon.fr
†Avignon Université, Laboratoire de Mathéematiques d'Avignon (EA 2151) F-84018 anas.bouali@univ-avignon.fr
‡Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France. loic.bourdin@unilim.fr

design for renewable energy systems [51], changes of the control value $u(t)$ are possible only at $t = k\tau$ for all $k \in \mathbb{N}$, for some fixed $\tau > 0$. To cope with such situations of sampled-data controls, the PMP has recently been adapted [19, 20, 21, 22] and the corresponding version provides a so-called *averaged Hamiltonian gradient condition* instead of the usual Hamiltonian maximization condition.

The initial motivation of the present work was to consider a related situation in which a control system possesses *non-control regions*. To illustrate this new concept, consider a finite partition of the state space

$$\mathbb{R}^n = \bigcup_{j=1}^{N} \overline{X_j},$$

in which the $X_j$ are disjoint nonempty connected open subsets of $\mathbb{R}^n$. Now consider that, according to the state position $x(t)$ in the above partition, it may be no longer possible to change the control value $u(t)$ in a permanent way. In other words, each region $X_j$ is either a *control region* (in which the control value $u(t)$ can be modified in a permanent way) or a *non-control region* (in which the control value $u(t)$ is frozen and remains constant as long as $x(t) \in X_j$). As an example of such a problematic, we can cite the problem of shadow effect in aerospace problems [35, 39, 42]. It is also of interest in population dynamics models when considering safety zones in which no change of the control value is applied for economic reasons (see the time crisis problem [11] in the context of viability [3]).

**Our approach: hybrid control systems with a regionally switching parameter.** In a control system including non-control regions, the control function can be seen as a permanent control in a control region, and as a constant parameter in a non-control region. Therefore, to address this setting, our approach was to write the control system as

$$\dot{x}(t) = \begin{cases} f(x(t), u(t), t) & \text{if } x(t) \text{ belongs to a control region,} \\ f(x(t), \lambda(t), t) & \text{if } x(t) \text{ belongs to a non-control region,} \end{cases} \qquad \text{a.e. } t \in [0, T],$$

where $\lambda$ is a *regionally switching parameter*, in the sense that $\lambda$ is a function that remains constant while the state position $x(t)$ stays inside a region, and can switch (that is, can change its value) only when the state position $x(t)$ goes from one region to another. Note that $\lambda$ is not necessarily constant over the whole interval $[0, T]$. It can have different values if different regions are visited (and can even have different values in the same region if the state position $x(t)$ quits and visits several times this region). Hence one can see $\lambda$ as a piecewise constant control but we insist on the fact that this framework strongly differs from the sampled-data control setting since the possibility (or not) of changing the value of $\lambda$ depends on the state position $x(t)$ (and not on the time variable $t$).

With the above approach, a crucial point is that the dynamics is no longer continuous at the interfaces between control regions and non-control regions. Therefore our framework falls into the domain of *hybrid control systems* (that is, control systems with discontinuous dynamics). Such control systems are naturally present in many areas such as aircraft planning [59, 61], motion planning [37] or population dynamics [47]. Note that situations of hybrid control systems are various. For example, the dynamics may change discontinuously (only) with respect to time (see, *e.g.*, [31]). A more general possibility is to consider a *switching law*, that may depend on several parameters, which governs the discontinuous changes of the dynamics. In that case one may speak of *switched systems* or *switching systems* (see, *e.g.*, [13, 29, 66]). To extend the PMP to such discontinuous frameworks, the so-called Hybrid Maximum Principle (in short, HMP) has been developed in various hybrid settings, in a series of papers such as [26, 27, 28, 38, 42, 50, 60] and references therein. In the present article, in the spirit of previous works such as [25, 26, 38, 42], we consider the hybrid setting where the dynamics changes discontinuously (only) according to the state position $x(t)$ in a given partition of the state space $\mathbb{R}^n$. In that context, in contrary to the classical PMP in which the costate function is absolutely continuous, the HMP provides a costate function that has discontinuity jumps when the state position $x(t)$ crosses interfaces.

As far as we know, the consideration of an additional regionally switching parameter $\lambda$ (as presented before) has never been considered in the literature. Furthermore, in contrary to a constant parameter (or to a sampled-data control), the framework of a regionally switching parameter cannot be easily covered by an augmentation technique or any other tricky procedure (see Remark 2.4 for details).

**Our main result.** Hence, in this work, we consider the general hybrid control system, involving both a permanent control and a regionally switching parameter, given by

$$\dot{x}(t) = h(x(t), \lambda(t), u(t), t), \qquad \text{a.e. } t \in [0, T],$$

where $\lambda$ is a regionally switching parameter (as presented above) and where $h : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0, T] \to \mathbb{R}^n$ is a hybrid dynamics, in the sense that it is defined regionally by

$$h(x, \lambda, u, t) := h_j(x, \lambda, u, t), \quad \text{when } x \in X_j,$$

where the $h_j : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0, T] \to \mathbb{R}^n$ are $\mathrm{C}^1$ functions. We insist here on the fact that control systems with non-control regions (which constitute our initial motivation) are a particular case of the above setting. The main objective of this paper is to provide first-order necessary optimality conditions in a PMP form for a Mayer optimal control problem

$$\text{minimize } \phi(x(T)),$$

among solutions to the above hybrid control system. Therefore our main result (Theorem 2.1) is called *hybrid maximum principle with regionally switching parameter*. As one can expect, given an optimal triplet $(x, \lambda, u)$, Theorem 2.1 asserts that $u$ satisfies the classical Hamiltonian maximization condition, while $\lambda$ satisfies the averaged Hamiltonian gradient condition over each region visited. Furthermore, as usual in the hybrid setting (as explained above), a jump of the costate function is observed at each crossing time.

Now let us discuss briefly the proof of Theorem 2.1 and the main difficulties encountered. In abstract optimization, to derive necessary optimality conditions, one has to perform a sensitivity analysis of the constraints. In our setting, this translates into a sensitivity analysis of the hybrid control system. To this aim, we consider a perturbation of the control (needle-like perturbation as in [7, 34, 53]) and of the regionally switching parameter (convex perturbation as in [19, 20]). Under such local perturbations, we obtain a perturbed trajectory, but also a perturbed crossing time. We stress that the major difficulty of this work lies in handling this perturbed crossing time. To prove its existence, we rely on implicit function arguments which require two regularity assumptions: left and right continuity of the nominal control at the crossing time, and a transverse crossing condition on the nominal trajectory. Such hypotheses are commonly used in the hybrid setting (see, *e.g.*, [9, 42, 50]). In addition, since the perturbed crossing time does perturb the next one, and so on, and so on, an inductive reasoning is required to prove the existence of the remaining perturbed crossing times. Once the sensitivity analysis is complete, the proof of our main result follows similar steps to the PMP's proof which is based on the construction of an adequate adjoint vector to maintain the constancy of the inner product with all variation vectors. Let us note that, as usual in the hybrid setting, since the variation vectors admit discontinuities at each crossing time (due to the perturbed ones, as explained above), the adjoint vector also admits jumps at each crossing time.

Finally, the main novelties of this work are the variation vectors obtained with convex perturbations of the regionally switching parameter (which lead to the averaged Hamiltonian gradient condition) and the applicability of our main result to control systems with non-control regions (which is developed in a companion proceeding [8]). Furthermore we emphasize that our goal was to provide a very complete and rigorous proof of the HMP. Therefore the proof is quite long and technical and it is postponed to the end of the paper. However, for pedagogical reasons and for the reader's convenience, we provide in Section 2.4 a short overview of the proof of Theorem 2.1.

**Some remarks.** Hereafter we provide a short list of comments that we thought important to highlight before starting the paper:

(i) In contrary to what is claimed above (for simplicity), we actually consider in the present work the possibility of a state partition that can be infinite, and also that can be not static (in other words, that can be time dependent). The first extension is trivial. However the second extension requires two basic continuity conditions on the time evolution of the partition (see Remark 2.1 for details).

(ii) In this paper we give a simple example showing, as noticed in [38], that a standard needle-like perturbation of the control (as used in the literature for non-hybrid control systems) can produce a non-admissible trajectory in the hybrid setting (see Item 2. in Section 2.4). This important issue leads us to consider the construction of *auxiliary controls* on which we perform needle-like perturbations to obtain admissible trajectories (see Section 2.4 for details).

(iii) The present paper does not cover *terminal state constraints* (that is, constraints on $x(0)$ and $x(T)$). In the classical non-hybrid setting, several methods have been developed in the literature to take into account such constraints. One can invoke the Ekeland variational principle [33] or some implicit function arguments (see, *e.g.*, [2, 58]). To the best of our knowledge, the Ekeland approach does not apply in the present hybrid setting for several reasons, while the method based on an implicit function argument should be adaptable but at the price of a heavy formalism (see Remark 2.11 for details). Since our main objective in this work was to focus on the new concept of regionally switching parameter and on the corresponding averaged Hamiltonian gradient condition, we decided to avoid the technicalities related to the presence of terminal state constraints which are already well known in the literature.

(iv) Section 2.5 is dedicated to a list of comments on our main result and its consequences, and also on possible extensions. For instance we discuss the behavior of the Hamiltonian function, and the possible extension to a general Bolza cost (instead of a Mayer cost) or to a free final time $T > 0$.

(v) By means of a simple academic example, we show in Section 3 how to use Theorem 2.1 and that the new framework considered in this paper fills a gap in the literature. Indeed, in that example, and as one can expect, the optimal solution associated with a regionally switching parameter has a better cost than the one associated with a constant parameter, but has a worse cost than the one considering a permanent control instead of the regionally switching parameter. We highlight that this example remains simple and academic. More concrete and complex examples will be the topics of further research works. Finally let us refer to [8] for the specification of our main result to (non-hybrid) optimal control problems with non-control regions.

**Organization of the paper.** The paper is structured as follows. Section 2 starts with notation and functional framework. Then we introduce a general Mayer optimal control problem governed by a hybrid control system including a regionally switching parameter. Our main result (Theorem 2.1) about the corresponding first-order necessary optimality conditions in a PMP form is stated right after. Next we give an overview of the proof of Theorem 2.1, as well as a list of comments and perspectives. Section 3 is dedicated to a simple academic example. The last three sections are dedicated to the quite long and technical proof of Theorem 2.1. Precisely, in the preliminary Sections 4 and 5, we provide a thorough sensitivity analysis of a non-hybrid and of a hybrid control system respectively. Based on these technical results, Section 6 is devoted to the complete proof of Theorem 2.1.

# 2 Main result

This section is dedicated to state our main result. To this aim, Section 2.1 is devoted to the required notations and functional framework. In Section 2.2, the hybrid optimal control problem with regionally switching parameter considered in this paper is presented, with terminology and assumptions. In Section 2.3, the corresponding hybrid maximum principle, which constitutes our main result, is provided (see Theorem 2.1). The proof of Theorem 2.1 is quite long and technical. Therefore it is postponed to the end of the paper. Nonetheless, for the reader's convenience, an overview of the proof of Theorem 2.1 is proposed in Section 2.4. Finally a list of general comments on Theorem 2.1 and its consequences, and also on possible relaxations and extensions is provided in Section 2.5.

## 2.1 Notations and functional framework

In this paper, for any positive integer $d \in \mathbb{N}^*$, we denote by $\langle \cdot, \cdot \rangle_{\mathbb{R}^d}$ (resp. $\| \cdot \|_{\mathbb{R}^d}$) the standard inner product (resp. Euclidean norm) of $\mathbb{R}^d$. For any subset $S \subset \mathbb{R}^d$, we denote by $\partial S$ the boundary of $S$ defined by $\partial S := \overline{S} \backslash \mathrm{Int}(S)$, where $\overline{S}$ and $\mathrm{Int}(S)$ stand for the closure and the interior of $S$ respectively. Furthermore, for any extended-real number $r \in [1, +\infty]$ and any nonempty real interval $I \subset \mathbb{R}$, we denote by:

- $\mathrm{L}^r(I, \mathbb{R}^d)$ the standard Lebesgue space of $r$-integrable functions defined on $I$ with values in $\mathbb{R}^d$, endowed with its usual norm $\| \cdot \|_{\mathrm{L}^r}$.

- $\mathrm{C}(I,\mathbb{R}^d)$ the standard space of continuous functions defined on $I$ with values in $\mathbb{R}^d$, endowed with its standard uniform norm $\|\cdot\|_{\mathrm{C}}$.
- $\mathrm{AC}(I,\mathbb{R}^d)$ the subspace of $\mathrm{C}(I,\mathbb{R}^d)$ of absolutely continuous functions.

If a function $\gamma : I \to \mathbb{R}^d$ admits left and right limits at some $\tau \in \mathrm{Int}(I)$, we denote by

$$\gamma^-(\tau) := \lim_{\substack{t\to\tau \\ t<\tau}} \gamma(t) \quad \text{and} \quad \gamma^+(\tau) := \lim_{\substack{t\to\tau \\ t>\tau}} \gamma(t).$$

Now take $I = [0,T]$ for some $T > 0$. Recall that a partition of the interval $[0,T]$ is a finite set $\mathbb{T} = \{t_k\}_{k=0,\ldots,N}$, for some positive integer $N \in \mathbb{N}^*$, such that $0 = t_0 < t_1 < \ldots < t_{N-1} < t_N = T$. In this paper:

- A function $\gamma \in \mathrm{L}^\infty([0,T],\mathbb{R}^d)$ is said to be *piecewise constant*, with respect to a partition $\mathbb{T} = \{t_k\}_{k=0,\ldots,N}$ of the interval $[0,T]$, if the restriction of $\gamma$ over each open interval $(t_{k-1}, t_k)$ is almost everywhere equal to a constant denoted by $\gamma_k \in \mathbb{R}^d$. If so, $\gamma$ is identified to the function $\gamma : [0,T] \to \mathbb{R}^d$ given by

$$\gamma(t) := \begin{cases} \gamma_k & \text{if } t \in [t_{k-1}, t_k) \text{ for all } k \in \{1,\ldots,N-1\}, \\ \gamma_N & \text{if } t \in [t_{N-1}, t_N] \text{ for } k = N, \end{cases}$$

  for all $t \in [0,T]$.

- A function $\gamma : [0,T] \to \mathbb{R}^d$ is said to be *piecewise absolutely continuous*, with respect to a partition $\mathbb{T} = \{t_k\}_{k=0,\ldots,N}$ of the interval $[0,T]$, if $\gamma$ is continuous at $0$ and $T$ and the restriction of $\gamma$ over each open interval $(t_{k-1}, t_k)$ admits an extension over $[t_{k-1}, t_k]$ that is absolutely continuous. If so, $\gamma$ admits left and right limits at each $t_k \in (0,T)$, denoted by $\gamma^-(t_k)$ and $\gamma^+(t_k)$ respectively.

In what follows we denote by $\mathrm{PC}^{\mathbb{T}}([0,T],\mathbb{R}^d)$ (resp. $\mathrm{PAC}^{\mathbb{T}}([0,T],\mathbb{R}^d)$) the space of all piecewise constant functions (resp. piecewise absolutely continuous functions) respecting a given partition $\mathbb{T}$ of $[0,T]$. Finally we denote by $\mathrm{PC}([0,T],\mathbb{R}^d)$ (resp. $\mathrm{PAC}([0,T],\mathbb{R}^d)$) the set of all piecewise constant functions (resp. piecewise absolutely continuous functions), independently of the partition considered.

Finally, as usual in the literature, when $(\mathcal{Z}, \mathrm{d}_{\mathcal{Z}})$ is a metric set, we denote by $\mathrm{B}_{\mathcal{Z}}(z,\varepsilon)$ (resp. $\overline{\mathrm{B}}_{\mathcal{Z}}(z,\varepsilon)$) the standard open (resp. closed) ball of $\mathcal{Z}$ centered at $z \in \mathcal{Z}$ and of radius $\varepsilon > 0$.

## 2.2 A hybrid optimal control problem with regionally switching parameter

Let $n$, $d$, $m \in \mathbb{N}^*$ be three fixed positive integers and $T > 0$ be a fixed positive real number. In this paper, in the spirit of [42], we consider a *time dependent partition* of $\mathbb{R}^n$ given by

$$\forall t \in [0,T], \qquad \mathbb{R}^n = \bigcup_{j\in\mathcal{J}} \overline{X_j(t)},$$

where $\mathcal{J}$ is a (possibly infinite) family of indexes and where, for all $t \in [0,T]$, the nonempty connected open subsets $X_j(t)$, called *regions*, are disjoint. This time dependent partition is furthermore assumed to satisfy two basic *continuity conditions* given by:

($\mathcal{C}1$) for all $j \in \mathcal{J}$ and all $x \in \mathrm{C}([a,b],\mathbb{R}^n)$ satisfying $x(t) \in X_j(t)$ over $[a,b]$, for some $0 \leq a \leq b \leq T$, there exists a uniform $\sigma > 0$ such that $\overline{\mathrm{B}}_{\mathbb{R}^n}(x(t),\sigma) \subset X_j(t)$ for all $t \in [a,b]$.

($\mathcal{C}2$) for all $t^c \in (0,T)$ and all $x \in \mathrm{C}([t^c-\delta, t^c+\delta],\mathbb{R}^n)$ satisfying $x(t) \in X_j(t)$ over $[t^c-\delta, t^c)$ and $x(t) \in X_{j'}(t)$ over $(t^c, t^c+\delta]$, for some $j$, $j' \in \mathcal{J}$ with $j \neq j'$ and some small $\delta > 0$, it holds that $x(t^c) \in \partial X_j(t^c) \cap \partial X_{j'}(t^c)$.

**Remark 2.1.** Note that the continuity conditions ($\mathcal{C}1$) and ($\mathcal{C}2$) are automatically satisfied whenever the partition is *static* (that is, independent of the time variable $t$). In contrast, when the partition is not static, the continuity conditions ($\mathcal{C}1$) and ($\mathcal{C}2$) guarantee, as one might expect, a kind of smooth and reasonable time evolution of the regions composing the partition. We refer to Figure 1 for illustrations. Note that it is not our aim here to use the technical tools from multi-valued analysis (see, *e.g.*, [5, 57]) to describe the

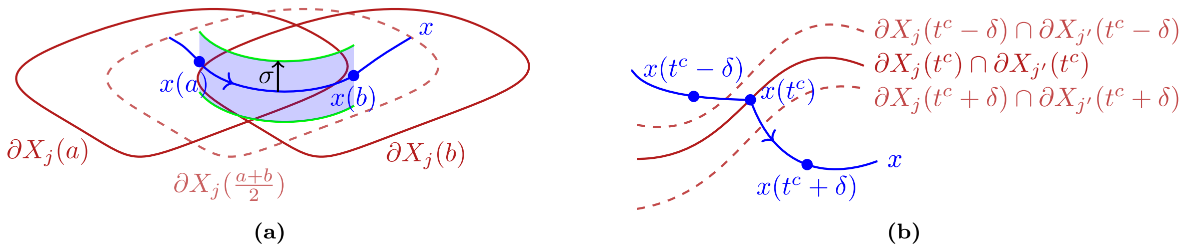


**Figure 1:** Illustrations of Condition ($\mathcal{C}1$) on the left, and of Condition ($\mathcal{C}2$) on the right.

continuity properties of the multi-valued functions $X_j : [0,T] \rightrightarrows \mathbb{R}^n$. Indeed the continuity conditions ($\mathcal{C}1$) and ($\mathcal{C}2$) are sufficient for our investigation. Precisely the present work focuses on optimal control problems involving hybrid control systems associated with the partition. Our aim is to derive necessary optimality conditions which are obtained, as usual in the literature, thanks to a thorough sensitivity analysis of the hybrid control system under various perturbations. In our approach, the continuity conditions ($\mathcal{C}1$) and ($\mathcal{C}2$) are used to construct appropriate perturbed trajectories which visit exactly (and in the same order) the same regions than the nominal one. We refer to Section 2.4 for details and illustrations.

Additionally to the above time dependent partition of $\mathbb{R}^n$, we consider a *hybrid dynamics* $h : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0,T] \to \mathbb{R}^n$ defined *regionally* by

$$\forall (x,\lambda,u,t) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0,T], \quad h(x,\lambda,u,t) := h_j(x,\lambda,u,t) \text{ when } x \in X_j(t),$$

where the maps $h_j : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0,T] \to \mathbb{R}^n$ are of class $\mathrm{C}^1$. Note that $h(x,\lambda,u,t)$ is not defined when $x \notin \cup_{j\in\mathcal{J}} X_j(t)$ but this fact will have no impact on the rest of this work. In this paper we focus on the *hybrid control system with regionally switching parameter* given by

$$\begin{cases} (x,\lambda,u) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{PC}([0,T],\mathbb{R}^d) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m), \\ \dot{x}(t) = h(x(t),\lambda(t),u(t),t), \quad \text{a.e. } t \in [0,T], \\ x(0) = x_{\text{init}}, \\ \lambda \text{ is a regionally switching parameter associated with } x, \end{cases} \tag{CS}$$

where the fixed initial condition $x_{\text{init}}$ belongs to $X_{j_1}(0)$ for some $j_1 \in \mathcal{J}$. In the control system (CS), as usual in the literature, $x \in \mathrm{AC}([0,T],\mathbb{R}^n)$ is called the *state* (or the *trajectory*) and $u \in \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ is called the *control*. In the literature, there are many references that discuss the additional presence of a constant parameter $\lambda \in \mathbb{R}^d$ (see, *e.g.*, [17]). The novelty of the present work lies in the consideration of a *regionally switching parameter* $\lambda \in \mathrm{PC}([0,T],\mathbb{R}^d)$ meaning, roughly speaking, that the parameter $\lambda$ remains constant while the trajectory $x$ stays inside a region, but is authorized to *switch* (that is, to change its value) when the trajectory $x$ crosses a boundary, going from one region to another. The precise definition of a solution to (CS) is given below.

**Definition 2.1** (Solution to (CS))**.** A triple $(x,\lambda,u) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{PC}([0,T],\mathbb{R}^d) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ is said to be a *solution* to (CS) if the following conditions are satisfied:

(i) There exists a partition $\mathbb{T} = \{t^c_k\}_{k=0,\dots,N}$ of the interval $[0,T]$ such that

$$\forall k \in \{1,\dots,N\}, \quad \exists j(k) \in \mathcal{J}, \quad \forall t \in (t^c_{k-1},t^c_k), \quad x(t) \in X_{j(k)}(t),$$

with $j(k) \neq j(k-1)$ for all $k \in \{2,\dots,N\}$, with $x(0) \in X_{j(1)}(0)$ and $x(T) \in X_{j(N)}(T)$.

(ii) $\lambda$ is a *regionally switching parameter* associated with $x$, that is, $\lambda \in \mathrm{PC}^{\mathbb{T}}([0,T],\mathbb{R}^d)$.

(iii) The state equation $\dot{x}(t) = h_{j(k)}(x(t),\lambda_k,u(t),t)$ is satisfied for almost every $t \in (t^c_{k-1},t^c_k)$ and all $k \in \{1,\dots,N\}$.

(iv) The initial condition $x(0) = x_{\text{init}}$ is satisfied (and thus $j(1) = j_1$).

In that case, for the ease of notations, we simply denote by $f_k := h_{j(k)}$ and $E_k := X_{j(k)}$ for all $k \in \{1, \dots, N\}$. With this system of notations we get that

$$\forall k \in \{1, \dots, N\}, \qquad \begin{cases} x(t) \in E_k(t), & \forall t \in (t^c_{k-1}, t^c_k), \\ \dot{x}(t) = f_k(x(t), \lambda_k, u(t), t), & \text{a.e. } t \in (t^c_{k-1}, t^c_k), \end{cases}$$

and $x(0) \in E_1(0)$, $x(T) \in E_N(T)$. Furthermore the times $t^c_k$, for $k \in \{1, \dots, N-1\}$, are called the *crossing times*, corresponding to the times at which the trajectory $x$ goes from the region $E_k$ to the region $E_{k+1}$, and thus $x(t^c_k) \in \partial E_k(t^c_k) \cap \partial E_{k+1}(t^c_k)$ from the continuity condition $(\mathcal{C}2)$. We refer to Figure 2 for an illustration.

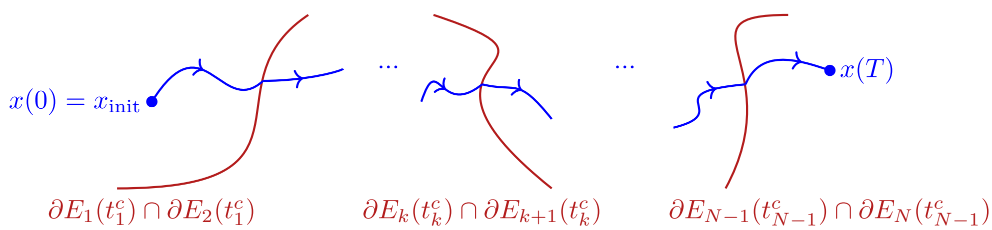


**Figure 2:** Illustration of Definition 2.1.

**Remark 2.2.** In control theory, the terminology *hybrid control systems* refers to control systems governed by dynamics that can change discontinuously. They arise in many application models such as in automatic control [6, 15, 49, 56], aircraft planning [59, 61], or motion planning [37], etc. They are also considered in population dynamics [47] and in viability theory [11] when dealing with a discontinuous Lagrange cost function to be minimized such as in time crisis problems [10]. Other examples can be found in the monographs [40, 63].

Note that situations of hybrid control systems are various. For example, the dynamics may change discontinuously (only) with respect to time (see, *e.g.*, [31]). A more general possibility is to consider a *switching law*, that may depend on several parameters, which governs the discontinuous changes of the dynamics. In that case one may speak of *switched systems* or *switching systems* (see, *e.g.*, [13, 29, 66]). In this paper, in the spirit of previous works such as [25, 26, 38, 42], we consider the case where the dynamics changes discontinuously (only) according to the position of the state in a given time dependent partition of the state space $\mathbb{R}^n$.

The main novelty of the present work is to take into account an additional *regionally switching parameter* that remains constant in each visited region but may switch at each boundary crossing. To the best of our knowledge, this framework is new and is motivated by several applications. Typically, in the satellite orbit transfer problem or in the aerospace domain, a controlled spacecraft may enter into *shadow zones* in which the control value cannot be modified anymore and thus remains constant (see, *e.g.*, [35, 39] and references therein). As explained in Introduction, such optimal control problems with *non-control regions* are particular cases of our setting. Note that, in population models, epidemiology, or in viability theory, it may be useful for a practitioner to stop changing the control value permanently in the complement of *crisis sets* [11], which also falls into the framework of non-control regions.

**Remark 2.3.** In this paper, note that the fixed initial condition $x_{\text{init}}$ belongs to a region (and not to a boundary) and, according to Definition 2.1, we deal (only) with trajectories $x$ whose final condition $x(T)$ also belongs to a region (and not to a boundary). These restrictions allow us, similarly to the continuity conditions $(\mathcal{C}1)$ and $(\mathcal{C}2)$ (see Remark 2.1), to avoid situations in which the sensitivity analysis of the hybrid control system (CS) would involve perturbed trajectories that would visit more regions than the nominal one. However we are confident that, at the price of a slightly more cumbersome analysis, the methodology developed in this paper (in particular the assumptions and techniques used to deal with the boundary crossings over the open time interval $(0, T)$) could be easily adapted to deal with terminal conditions that might belong to boundaries.

Our objective in the present work is to derive first-order necessary optimality conditions, in the form of a Pontryagin maximum principle, for the *hybrid optimal control problem with regionally switching parameter* given by

$$\begin{aligned} &\text{minimize} && \phi(x(T)), \\ &\text{subject to} && (x,\lambda,u) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{PC}([0,T],\mathbb{R}^d) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m) \text{ solution to (CS)}, \\ &&& (\lambda(t),u(t)) \in \Lambda \times \mathrm{U}, \quad \text{a.e. } t \in [0,T], \end{aligned} \tag{OCP}$$

where the Mayer cost function $\phi : \mathbb{R}^n \to \mathbb{R}$ is of class $\mathrm{C}^1$, the parameter constraint set $\Lambda$ is a nonempty convex subset of $\mathbb{R}^d$ and the control constraint set $\mathrm{U}$ is a nonempty closed subset of $\mathbb{R}^m$. We refer to Section 2.4 (Item 11) for comments on how the hypotheses made on $\Lambda$ and $\mathrm{U}$ are used in our approach, and to Remark 2.10 for possible relaxations.

**Remark 2.4.** Note that several versions of necessary optimality conditions for hybrid optimal control problems, in the form of a Pontryagin maximum principle, are already available in the literature (see, *e.g.*, [31, 38, 42] and references therein). However, to the best of our knowledge, none of them allows to deal with our framework. Indeed:

- First, note that a hybrid optimal control problem with an additional constant parameter $\lambda \in \mathbb{R}^d$ can be easily treated, thanks to the classical method of augmenting the control system with the equation $\dot{\lambda}(t) = 0_{\mathbb{R}^d}$ (see, *e.g.*, [17]). One obtains a necessary optimality condition written as an averaged Hamiltonian gradient condition over the whole interval $[0,T]$.

- Second, in the case where $\lambda \in \mathrm{PC}^{\mathbb{T}}([0,T],\mathbb{R}^d)$ is a piecewise constant control (also known as *sampled-data control*), with a given and fixed partition $\mathbb{T} = \{t_k\}_{k=0,\dots,N}$ (independent of the state), one can easily deduce necessary optimality conditions from the previous item. Indeed, using an adequate change of time variable (transforming all intervals $[t_{k-1},t_k]$ into a common interval $[0,1]$), all the values $\lambda_k \in \mathbb{R}^d$ become constant parameters and one can deduce an averaged Hamiltonian gradient condition over each interval $[t_{k-1},t_k]$ (see, *e.g.*, [20, 21]). Note that, in case of a free partition $\mathbb{T}$ (but with a fixed positive integer $N \in \mathbb{N}^*$), one can consider each $t_k$ as a parameter and derive a necessary optimality condition given by the continuity of the corresponding Hamiltonian function (see, *e.g.*, [19]).

The techniques and results presented above are well known in the non-hybrid setting, and can certainly be generalized to the hybrid setting with no major difficulty. However, even if considering a regionally switching parameter might seem as easy as dealing with sampled-data controls, it is not. The main technical issue here lies in the fact that the possibility of changing the parameter value depends on the state position (and not on the time variable). To the best of our knowledge, this point has never been discussed in the literature, and cannot be addressed easily with a technical trick. Hence the major contribution of this paper is to fill this gap in the literature.

## 2.3 Hybrid maximum principle with regionally switching parameter

Our main result (Theorem 2.1) is based on some regularity assumptions made on the behavior of the optimal triple $(x,\lambda,u)$ at the crossing times $t_k^c$. These hypotheses are precised in the next definition.

**Definition 2.2** (Regular solution to (CS))**.** Following the notations introduced in Definition 2.1, a solution $(x,\lambda,u) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{PC}([0,T],\mathbb{R}^d) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ to (CS) is said to be *regular* if there exist $0 < \delta \leq \frac{1}{3} \min_{k=1,\dots,N} |t_k - t_{k-1}|$ and $\nu > 0$ such that:

($\mathcal{A}1$) At each crossing time $t_k^c$, the control $u$ is continuous over $[t_k^c - \delta, t_k^c)$ and over $(t_k^c, t_k^c + \delta]$, and admits left and right limits at $t_k^c$, denoted by $u^-(t_k^c)$ and $u^+(t_k^c)$ respectively.

($\mathcal{A}2$) At each crossing time $t_k^c$, there exists a $\mathrm{C}^1$ function $F_k : \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_k^c),\nu) \times [t_k^c - \delta, t_k^c + \delta] \to \mathbb{R}$ such that

$$\forall (y,t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_k^c),\nu) \times [t_k^c - \delta, t_k^c + \delta], \qquad \left\{ \begin{aligned} y \in E_k(t) \quad &\Leftrightarrow \quad F_k(y,t) < 0, \\ y \in \partial E_k(t) \cap \partial E_{k+1}(t) \quad &\Leftrightarrow \quad F_k(y,t) = 0, \\ y \in E_{k+1}(t) \quad &\Leftrightarrow \quad F_k(y,t) > 0. \end{aligned} \right.$$

In particular it holds that $F_k(x(t_k^c), t_k^c) = 0$.

($\mathcal{A}3$) At each crossing time $t_k^c$, the *transverse conditions* given by

$$\langle \nabla_x F_k(x(t_k^c), t_k^c), (f_k)^-(t_k^c) \rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t_k^c), t_k^c) > 0,$$

$$\langle \nabla_x F_k(x(t_k^c), t_k^c), (f_{k+1})^+(t_k^c) \rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t_k^c), t_k^c) > 0,$$

where $(f_k)^-(t_k^c) := f_k(x(t_k^c), \lambda_k, u^-(t_k^c), t_k^c)$ and $(f_{k+1})^+(t_k^c) := f_{k+1}(x(t_k^c), \lambda_{k+1}, u^+(t_k^c), t_k^c)$, are both satisfied. We refer to Figure 3 for a geometrical illustration.

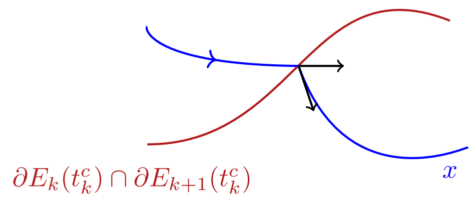


**Figure 3:** Geometrical illustration of a transversal boundary crossing (see Assumption ($\mathcal{A}3$)).

**Remark 2.5.** Let us comment the assumptions presented in Definition 2.2. In this paper, necessary optimality conditions for (OCP) will be obtained thanks to a sensitivity analysis of (CS) with respect to (local) perturbations of the regionally switching parameter and the control. However, considering such a local perturbation in a given region $E_k$ leads, as usual, to a perturbed trajectory, but also to a perturbed crossing time between the two consecutive regions $E_k$ and $E_{k+1}$. Our approach to guarantee the existence of this perturbed crossing time relies on the application of an implicit function theorem (see Lemma 5.1) which requires Assumption ($\mathcal{A}2$) to benefit a local description (in space and time) of the boundary between $E_k$ and $E_{k+1}$. This gives us an explicit function whose regularity is guaranteed by Assumption ($\mathcal{A}1$) and whose the invertibilty of the partial derivative (with respect to time) is guaranteed by the first transverse condition in Assumption ($\mathcal{A}3$). Finally, the second transverse condition in Assumption ($\mathcal{A}3$) allows us to guarantee that the perturbed trajectory enters in the next open region $E_{k+1}$. We then proceed by induction, region after region. We refer to Section 2.4 for details and illustrations.

Before stating the main result of this paper we just need to recall some basics, such as the standard definition of the (convex) *normal cone* to $\Lambda$ at a point $\lambda \in \Lambda$ given by

$$\mathrm{N}_\Lambda[\lambda] := \{\lambda'' \in \mathbb{R}^d \mid \forall \lambda' \in \Lambda,\ \langle \lambda'', \lambda' - \lambda \rangle_{\mathbb{R}^d} \leq 0\},$$

and the usual definition of the *Hamiltonian* $H : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^n \times [0,T] \to \mathbb{R}$ associated with the optimal control problem (OCP) given by

$$\forall (x, \lambda, u, p, t) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^n \times [0,T], \quad H(x, \lambda, u, p, t) := \langle p, h(x, \lambda, u, t) \rangle_{\mathbb{R}^n}.$$

We are now in a position to state the main result of this paper.

**Theorem 2.1** (Hybrid maximum principle with regionally switching parameter)**.** If $(x, \lambda, u) \in \mathrm{AC}([0,T], \mathbb{R}^n) \times \mathrm{PC}([0,T], \mathbb{R}^d) \times \mathrm{L}^\infty([0,T], \mathbb{R}^m)$ is a solution to (OCP), that is moreover a regular solution to (CS), then, following the notations introduced in Definitions 2.1 and 2.2, there exists an *adjoint vector* $p \in \mathrm{PAC}^{\mathbb{T}}([0,T], \mathbb{R}^n)$ (also called *costate*) such that:

(i) The *adjoint equation* $\dot{p}(t) = -\nabla_x f_k(x(t), \lambda_k, u(t), t)^\top p(t)$ is satisfied for almost every $t \in (t_{k-1}^c, t_k^c)$ and all $k \in \{1, \ldots, N\}$.

(ii) The *final condition* $p(T) = -\nabla \phi(x(T))$ is satisfied.

(iii) At each crossing time $t_k^c$, the *adjoint discontinuity condition*

$$p^+(t_k^c) - p^-(t_k^c) = -\frac{\langle p^+(t_k^c), (f_{k+1})^+(t_k^c) - (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_k(x(t_k^c), t_k^c), (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t_k^c), t_k^c)} \nabla_x F_k(x(t_k^c), t_k^c), \tag{AD}$$

is satisfied.

(iv) The *Hamiltonian maximization condition*

$$u(t) \in \underset{v \in \mathrm{U}}{\arg\max}\, H(x(t), \lambda_k, v, p(t), t), \tag{HM}$$

holds true for almost every $t \in (t_{k-1}^c, t_k^c)$ and all $k \in \{1, \dots, N\}$.

(v) The *averaged Hamiltonian gradient condition*

$$\int_{t_{k-1}^c}^{t_k^c} \nabla_\lambda H(x(s), \lambda_k, u(s), p(s), s)\, \mathrm{d}s \in \mathrm{N}_\Lambda[\lambda_k], \tag{AHG}$$

holds true for all $k \in \{1, ..., N\}$.

The proof of Theorem 2.1 is quite long and technical. Therefore it is postponed to Section 6, after the two preliminary Sections 4 and 5 that are dedicated to sensitivity analyses of non-hybrid and hybrid control systems respectively. Nonetheless, for the reader's convenience, an overview of the proof of Theorem 2.1 is proposed in the next Section 2.4. Finally a list of comments on Theorem 2.1 and its consequences, and also on possible relaxations and extensions is provided in Section 2.5.

## 2.4 Overview of the proof of Theorem 2.1

This section is dedicated to an overview of the proof of Theorem 2.1. For the reader's convenience, our presentation is divided into twelve major items in which we take care to highlight at which point of the proof the continuity conditions ($\mathcal{C}1$) and ($\mathcal{C}2$) and the regularity assumptions ($\mathcal{A}1$), ($\mathcal{A}2$) and ($\mathcal{A}3$) are used.

Before, we would like to emphasize a crucial point: Item 2 provides a simple example showing that a standard needle-like perturbation of the control may be not admissible in the hybrid setting, in the sense that the corresponding perturbed trajectory may not uniformly converge to the nominal one, or may not be a global solution to the control system. This counterexample reveals an erroneous assertion in [42, beginning of Section 2.1.1] and highlights interesting comments given in [38, pp. 1872]. As a conclusion, handling needle-like perturbations of the control in the hybrid setting requires a careful attention.

1. From the point of view of abstract optimization, sensitivity analysis of constraints (with respect to given parameters) plays a fundamental role in order to derive necessary optimality conditions. In optimal control theory, this translates into a sensitivity analysis of the control system with respect to perturbations of the control $u$. To derive the classical Pontryagin maximum principle, the standard method is to consider a *needle-like perturbation* defined by $u^\alpha(t) := v$ for all $t \in (\tau - \alpha, \tau)$ and $u^\alpha(t) := u(t)$ elsewhere, for all $\alpha > 0$ and where $v \in \mathbb{R}^m$ and $\tau \in (0, T)$ are fixed. Then one has to identify the corresponding variation vector, that is the uniform limit of $\frac{x^\alpha - x}{\alpha}$ when $\alpha \to 0$, where $x^\alpha$ stands for the perturbed trajectory associated with the perturbed control $u^\alpha$ (see Figure 4), as the solution to a linearized control system.

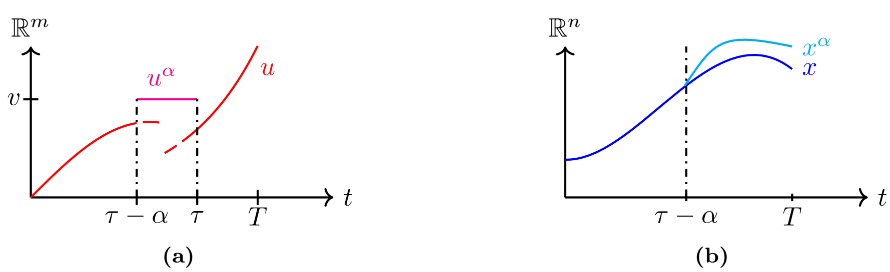


**Figure 4:** Illustrations of a needle-like perturbation (left) and the corresponding perturbed trajectory (right).

2. However a needle-like perturbation may be not admissible in the hybrid setting, in the sense that the corresponding perturbed trajectory $x^\alpha$ does not necessarily converge uniformly to $x$ over $[0,T]$ when $\alpha \to 0$, or even may be not defined globally over the whole interval $[0,T]$. Let us provide a simple counterexample which highlights this issue which is not encountered in the classical non-hybrid setting. Consider $T=2$, $n=m=1$ and the static partition $\mathbb{R} = \overline{X_1} \cup \overline{X_2}$, where $X_1 = \{y \in \mathbb{R} \mid y < 1\}$ and $X_2 = \{y \in \mathbb{R} \mid y > 1\}$. Now consider the hybrid control system given by

$$\dot{x}(t) = \begin{cases} +u(t) & \text{if } x(t) \in X_1, \\ -u(t) & \text{if } x(t) \in X_2, \end{cases}$$

with the initial condition $x_{\mathrm{init}} = 0$. By taking the control $u(t) = +1$ over $[0,1)$ and $u(t) = -1$ over $(1,2]$, we get the corresponding trajectory $x$ given by $x(t) = t$ over $[0,2]$, with $t_1^c = 1$ as unique crossing time. Note that all conditions considered in this paper are satisfied, including the regularity assumptions $(\mathcal{A}1)$, $(\mathcal{A}2)$ and $(\mathcal{A}3)$. Now we apply needle-like perturbations of the control $u$ at some $\tau \in (0,1)$ and we refer to Figure 5 for illustrations.

   (i) If $v = -1$ we get a perturbed trajectory $x^\alpha$ satisfying $x^\alpha(t) \in X_1$ over the whole interval $[0,2]$ and thus $x^\alpha$ does not uniformly converge to $x$ over $[0,2]$ when $\alpha \to 0$.

   (ii) If $v = 2$ we get a perturbed trajectory $x^\alpha$ defined over $[0, \tilde{t}(\alpha))$ for some $\tilde{t}(\alpha) < 1$. Note that $x^\alpha$ is not defined over $[\tilde{t}(\alpha), 2]$ since, by contradiction, one would obtain $\dot{x}^\alpha(\tilde{t}(\alpha)^-) = +1$ and $\dot{x}^\alpha(\tilde{t}(\alpha)^+) = -1$ implying that $x^\alpha$ does not enter into the open region $X_2$ over $(\tilde{t}(\alpha), 1)$. In that context, note that different approaches can be explored, such as differential inclusions (see, *e.g.*, [4]) and sliding modes (see, *e.g.*, [62]), to consider a generalized notion of solution to the hybrid control system. However these approaches would not solve the issue presented in Item (i) anyway, and thus we will not go any further in that direction.

   The reason of this feature in the hybrid setting lies in the fact that standard needle-like perturbations of the control $u$ do not take into account the perturbation of the next crossing time.

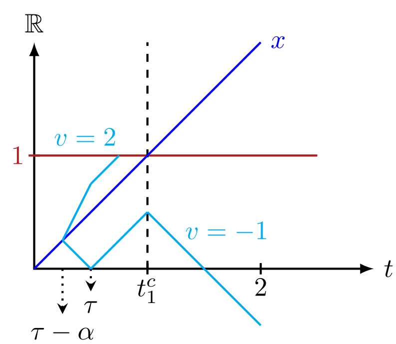


**Figure 5:** Illustration of the counterexample of Item 2.

3. We are now in a position to provide an overview of the proof of Theorem 2.1. Let $(x, \lambda, u)$ be a solution to (OCP), that is moreover a regular solution to (CS). To overcome the difficulty of handling needle-like perturbations in the hybrid setting (as discussed in Item 2), we shall introduce, for all $k \in \{1, \dots, N\}$, an *auxiliary control*, denoted by $\tilde{u}_k$, that coincides with the control $u$ over $(t_{k-1}^c, t_k^c)$ and that is continuously extended to a constant function outside $(t_{k-1}^c, t_k^c)$ thanks to Assumption $(\mathcal{A}1)$ (see Figure 6 and the exact definition of $\tilde{u}_k$ in Section 5.1). In the sequel we will apply needle-like perturbations only to auxiliary controls $\tilde{u}_k$ (and not to the nominal control $u$).

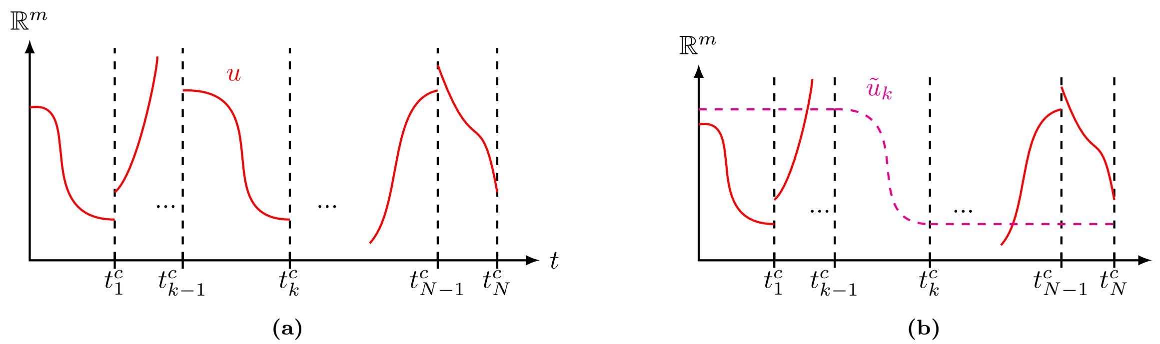


**Figure 6:** Illustration of an auxiliary control $\tilde{u}_k$. In this illustration, for simplicity, we have chosen a control $u$ that is continuous over each $(t_{k-1}^c, t_k^c)$ but it is not necessary. We only know that $u$ satisfies the continuity properties given in Assumption $(\mathcal{A}1)$.

4. Now let us fix $k \in \{1, ..., N\}$ (from now and until Item 11). The pair $(\lambda_k, \tilde{u}_k)$ allows us to define the *auxiliary non-hybrid trajectory*, denoted by $\tilde{z}_k$, as the unique solution to the non-hybrid state equation defined with the dynamics $f_k$ only (that is, with the dynamics $f_k$ all over $\mathbb{R}^n$, even outside $E_k$) and with the constant parameter $\lambda_k$ only (that is, with the constant parameter $\lambda_k$ all over $[0, T]$, even outside $(t_{k-1}^c, t_k^c)$), together with the initial condition $\tilde{z}_k(t_{k-1}^c) = x(t_{k-1}^c)$. Observe that $\tilde{z}_k$ represents an extension of the nominal trajectory $x$ as illustrated in Figure 7.

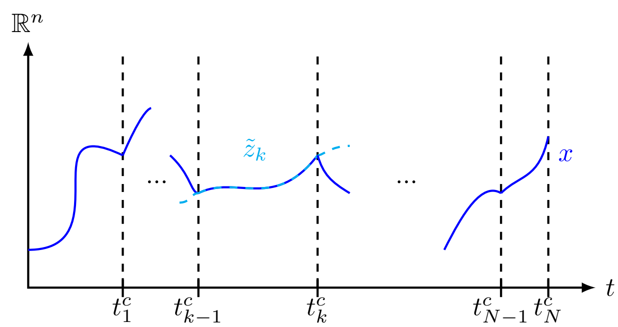


**Figure 7:** Illustration of the auxiliary non-hybrid trajectory $\tilde{z}_k$.

5. Now we will consider either a basic convex perturbation of $\lambda_k$ given by $\lambda_k + \alpha(\overline{\lambda}_k - \lambda_k)$ for some $\overline{\lambda}_k \in \Lambda$, either a classical needle-like perturbation of the auxiliary control $\tilde{u}_k$ for some $\tau \in (t_{k-1}^c, t_k^c)$ and some $v \in \mathrm{U}$ (see Figure 8). In both cases, this gives us a *perturbed auxiliary non-hybrid trajectory* denoted by $\tilde{z}_k^\alpha$. Since we deal here with a classical non-hybrid setting (with the dynamics $f_k$ only), we can use standard results from the literature such as the uniform convergence of $\tilde{z}_k^\alpha$ to $\tilde{z}_k$ when $\alpha \to 0$, and the existence of the corresponding variation vector, denoted by $w_k$, solution to a linearized control system with an initial condition at $t_{k-1}^c$ reduced to $0_{\mathbb{R}^n}$.

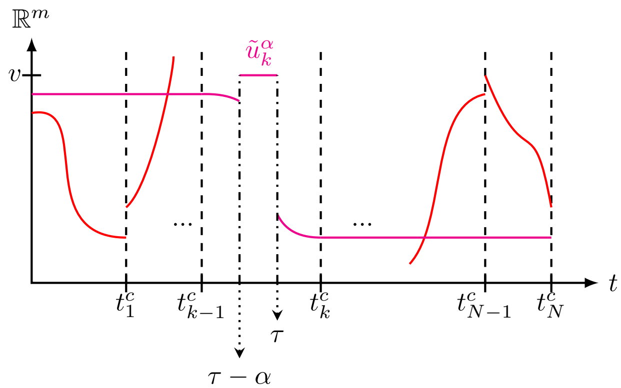


**Figure 8:** Illustration of a needle-like perturbation of $\tilde{u}_k$ (recall Figure 6).

6. The next step is to prove that the trajectory $\tilde{z}_k^\alpha$ crosses the boundary $\partial E_k \cap \partial E_{k+1}$ at a perturbed crossing time $\tilde{t}_k(\alpha)$ (see Figure 9). To this aim we invoke an implicit function theorem (see Lemma 5.1) to the map $G_k : (\alpha, t) \mapsto F_k(\tilde{z}_k^\alpha(t), t)$ that can be applied thanks to the regular assumptions $(\mathcal{A}1)$, $(\mathcal{A}2)$ and $(\mathcal{A}3)$ and the construction of $\tilde{u}_k$. In particular note that $\nabla_t G_k$ is invertible at $(0, t_k^c)$ thanks to the first transverse condition in Assumption $(\mathcal{A}3)$.

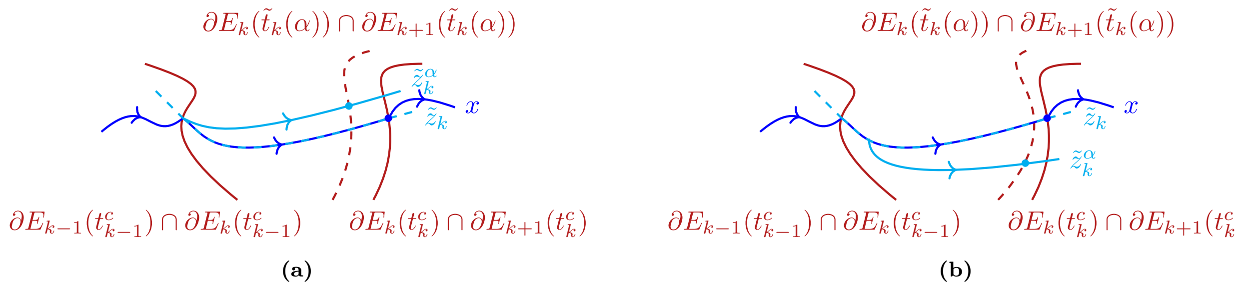


**Figure 9:** Plot of $\tilde{z}_k^\alpha$ under a convex perturbation of $\lambda_k$ (left). Plot of $\tilde{z}_k^\alpha$ under a needle-like perturbation of $\tilde{u}_k$ (right). In both cases $\tilde{z}_k^\alpha$ crosses the boundary $\partial E_k \cap E_{k+1}$ at some time $\tilde{t}_k(\alpha)$.

7. From the construction of the trajectory $\tilde{z}_k^\alpha$ , it can be proved that $\tilde{z}_k^\alpha$ stays inside $E_k$ over $(t_{k-1}^c, \tilde{t}_k(\alpha))$. Indeed, thanks to Assumption $(\mathcal{A}3)$, one can prove by contradiction that there exist $t_{k-1}^c < s'_k < s_k < \min\{t_k^c, \tilde{t}_k(\alpha)\}$, uniformly with respect to $\alpha$, such that $\tilde{z}_k^\alpha$ has values in $E_k$ over $(t_{k-1}^c, s'_k)$ and over $(s_k, \tilde{t}_k(\alpha))$ (see Lemmas 5.3 and 5.6 for technical details). Then, from Condition $(\mathcal{C}1)$ and the uniform convergence of $\tilde{z}_k^\alpha$ to $\tilde{z}_k = x$ over $[s'_k, s_k]$, we obtain that $\tilde{z}_k^\alpha$ has values in $E_k$ over $[s'_k, s_k]$ also.

8. After having considered perturbations in the region $E_k$ (see Items 4 and 5) and the consequences in the region $E_k$ only (see Items 6 and 7), our aim now is to analyze the resulting perturbations in the next regions $E_{k+1}, \ldots, E_N$. For the reader's convenience, we will detail here only the passage to the region $E_{k+1}$ (the other regions are treated with a basic induction, see Item 10). Similarly to Item 4, the pair $(\lambda_{k+1}, \tilde{u}_{k+1})$ allows us to define the auxiliary non-hybrid trajectory, denoted by $\tilde{z}_{k+1}$, as the unique solution to the non-hybrid state equation considered with the dynamics $f_{k+1}$ only and with the constant parameter $\lambda_{k+1}$ only, together with the initial condition $\tilde{z}_{k+1}(t_k^c) = x(t_k^c)$. Now, in contrary to Item 5 (in which we have proceeded either to a perturbation of the parameter, either to a perturbation of the control), we will consider here the perturbation of the initial time $t_k^c$ by $\tilde{t}_k(\alpha)$ (constructed in Item 6) and the perturbation of the initial condition $x(t_k^c)$ by $\tilde{z}_k^\alpha(\tilde{t}_k(\alpha))$. This gives us the perturbed auxiliary non-hybrid trajectory $\tilde{z}_{k+1}^\alpha$. This construction will allow us to proceed to a concatenation of the perturbed auxiliary non-hybrid trajectories $\tilde{z}_k^\alpha$ and $\tilde{z}_{k+1}^\alpha$ (see Figure 10).

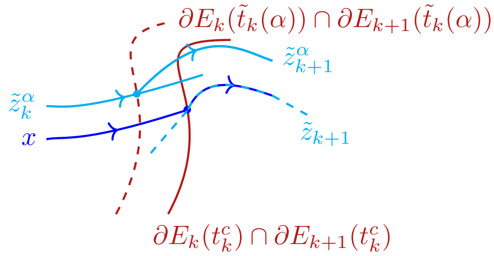


**Figure 10:** Perturbed auxiliary non-hybrid trajectory $\tilde{z}^\alpha_{k+1}$ under perturbations of the initial time and of the initial condition.

Since we deal here with a classical non-hybrid setting (with the dynamics $f_{k+1}$ only), we can use standard results from the literature such as the uniform convergence of $\tilde{z}^\alpha_{k+1}$ to $\tilde{z}_{k+1}$ when $\alpha \to 0$, and the existence of the corresponding variation vector, denoted by $w_{k+1}$, solution to a linearized control system with an initial condition at $t^c_k$ given by $w_k(t^c_{k-1})$ plus an additional term due to the perturbations of the initial time and of the initial condition. Finally, similarly to Item 6, we prove that $\tilde{z}^\alpha_{k+1}$ crosses the boundary $\partial E_{k+1} \cap \partial E_{k+2}$ at a perturbed crossing time $\tilde{t}_{k+1}(\alpha)$.

9. Using similar arguments to Item 7, it can be proved that the trajectory $\tilde{z}^\alpha_{k+1}$ stays inside $E_{k+1}$ over $(\tilde{t}_k(\alpha), \tilde{t}_{k+1}(\alpha))$.

10. Finally we proceed by induction, region after region, in order to construct the perturbed auxiliary non-hybrid trajectories $\tilde{z}^\alpha_q$ and the corresponding variation vectors $w_q$ for all $q \in \{k, ..., N\}$. Then we construct a "global" perturbed trajectory $x^\alpha$ of $x$ over the whole time interval $[0,T]$ (resp. a "global" variation vector $w$) by concatenation of the perturbed auxiliary non-hybrid trajectories $\tilde{z}^\alpha_q$ over $[\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)]$ (resp. of the variation vectors $w_q$ over $[t^c_{q-1}, t^c_q)$). This construction allows to guarantee several properties. First $x^\alpha$ visits exactly (and in the same order) the same regions that the nominal trajectory $x$. Second $x^\alpha$ converges uniformly to $x$ over $[0,T]$ when $\alpha \to 0$. Third the "global" variation vector $w$ corresponds to the variation vector associated with the "global" perturbed trajectory $x^\alpha$ of $x$. It is worth mentioning that, as reported in Item 8, the "global" variation vector $w$ has a discontinuity jump at each crossing time $t^c_q$.

11. From convexity of $\Lambda$, note that the convex perturbation of $\lambda_k$ belongs to $\Lambda$. Similarly, since $v \in \mathrm{U}$ and from the construction of $\tilde{u}_k$ and the closedness of $\mathrm{U}$, note that the needle-like perturbation of $\tilde{u}_k$ has values in $\mathrm{U}$. Therefore the constraints of Problem (OCP) are satisfied and thus, from optimality of the triple $(x, \lambda, u)$, it is clear that $\frac{\phi(x^\alpha(T)) - \phi(x(T))}{\alpha} \geq 0$ which leads to $\langle \nabla\phi(x(T)), w(T)\rangle_{\mathbb{R}^n} \geq 0$ when $\alpha \to 0$. One has to note that this last inequality is satisfied for any variation vector $w$ constructed as in the previous items, and thus is satisfied for any $\overline{\lambda}_k \in \Lambda$, any $v \in \mathrm{U}$, any $\tau \in (0,T)$ and for any $k \in \{1, \ldots, N\}$.

12. To conclude the proof, the method is now very similar to the standard non-hybrid setting found in the literature. The idea is to construct an adjoint vector $p$ which guarantees the constancy of the inner product between the adjoint vector $p$ and any variation vector $w$ constructed as in the previous items. To this aim we define $p$ as solution to the opposite of the transpose of the linearized control system satisfied by the variation vectors $w$ (which corresponds exactly to the adjoint equation in Theorem 2.1). On the other hand, to handle the discontinuity jumps of the variation vectors $w$ at each crossing time, we impose appropriate discontinuity jumps on $p$ (which correspond exactly to the adjoint discontinuity jumps in Theorem 2.1). Finally, imposing the final condition $p(T) = -\nabla\phi(x(T))$, we obtain that $\langle p(T), w(T)\rangle_{\mathbb{R}^n} \leq 0$ for any variation vector $w$. Using the classical Duhamel formula and thanks to the constancy of the inner product between the adjoint vector $p$ and any variation vector $w$, this last inequality can be rewritten as the averaged Hamiltonian gradient condition in Theorem 2.1 (if we have considered a variation vector $w$ associated with a convex perturbation of the parameter) or as the Hamiltonian maximization condition in Theorem 2.1 (if we have considered a variation vector $w$ associated with a needle-like perturbation of an auxiliary control). The proof is complete.

**Remark 2.6.** From the above overview of the proof of Theorem 2.1, we observe that a complete sensitivity analysis of non-hybrid state equations under perturbations of the parameter, of the control, but also of the initial time and of the initial condition, should be carried out. This is exactly the aim of the preliminary Section 4.

## 2.5 A list of comments on Theorem 2.1

This section is dedicated to a list of comments on Theorem 2.1 and its consequences, and also on possible relaxations and extensions.

**Remark 2.7.** Consider the framework of Theorem 2.1 and let us discuss the structure of the adjoint vector $p \in \mathrm{PAC}^{\mathbb{T}}([0,T],\mathbb{R}^n)$ that is piecewise absolutely continuous, respecting the same partition $\mathbb{T} = \{t^c_k\}_{k=0,\ldots,N}$ associated with the solution $(x,\lambda,u)$ (see Definition 2.1). Hence the restriction of $p$ over each open interval $(t^c_{k-1},t^c_k)$ admits an extension over $[t^c_{k-1},t^c_k]$ that is absolutely continuous, satisfying the adjoint equation provided in Theorem 2.1. Furthermore, at each crossing time $t^c_k$, the adjoint vector $p$ admits a discontinuity jump satisfying the equality (AD). Note that (AD) is written in a backward way, in the sense that $p^-(t^c_k)$ is expressed explicitly in terms of $p^+(t^c_k)$. Nevertheless we emphasize that (AD) can also be written in a forward way as

$$p^+(t^c_k) - p^-(t^c_k) = -\frac{\langle p^-(t^c_k), (f_{k+1})^+(t^c_k) - (f_k)^-(t^c_k)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_k(x(t^c_k),t^c_k), (f_k)^+(t^c_k)\rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t^c_k),t^c_k)} \nabla_x F_k(x(t^c_k),t^c_k).$$

On the other hand, due to the fact that a fixed initial condition is considered in Problem (OCP) with no final state constraint, let us note that the necessary optimality conditions imposed to the adjoint vector $p$ in Theorem 2.1 imply its uniqueness. However this property would not be true with general *terminal state constraints* as evoked in Item (i) of Remark 2.11.

**Remark 2.8.** Consider the framework of Theorem 2.1. Note that the Hamiltonian system

$$\dot{x}(t) = \nabla_p H(x(t),\lambda_k,u(t),p(t),t), \qquad \dot{p}(t) = -\nabla_x H(x(t),\lambda_k,u(t),p(t),t),$$

is satisfied for almost every $t \in (t^c_{k-1},t^c_k)$ and all $k \in \{1,\ldots,N\}$. As usual in the literature let us introduce the *Hamiltonian function* $\mathcal{H} : [0,T] \to \mathbb{R}$ defined by

$$\mathcal{H}(t) := H(x(t),\lambda(t),u(t),p(t),t),$$

for almost every $t \in [0,T]$. Using similar arguments as in [34, Theorem 2.6.1 pp. 71], one can prove from the above Hamiltonian system, from (HM) and from the piecewise constancy of the parameter $\lambda$, that $\mathcal{H}$ is equal almost everywhere over each interval $(t^c_{k-1},t^c_k)$ to an absolutely continuous function which satisfies

$$\dot{\mathcal{H}}(t) = \nabla_t H(x(t),\lambda_k,u(t),p(t),t),$$

for almost every $t \in (t^c_{k-1},t^c_k)$ and all $k \in \{1,\ldots,N\}$. Therefore we write $\mathcal{H} \in \mathrm{PAC}^{\mathbb{T}}([0,T],\mathbb{R})$ and one can easily obtain from simple computations that the discontinuity jumps of $\mathcal{H}$ are given by

$$\mathcal{H}^+(t^c_k) - \mathcal{H}^-(t^c_k) = \frac{\langle p^-(t^c_k), (f_{k+1})^+(t^c_k) - (f_k)^-(t^c_k)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_k(x(t^c_k)), (f_{k+1})^+(t^c_k)\rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t^c_k),t^c_k)} \nabla_t F_k(x(t^c_k),t^c_k).$$

As in Remark 2.7, we emphasize that the above formula can be rewritten in terms of $p^+(t^c_k)$ (instead of $p^-(t^c_k)$). Finally, from the results presented in this remark, we deduce that:

(i) If the partition is static, then the discontinuity jumps of $\mathcal{H}$ are reduced to zero and thus $\mathcal{H} \in \mathrm{C}([0,T],\mathbb{R})$.

(ii) If the hybrid dynamics is autonomous, then $\mathcal{H}$ is constant over each interval $(t^c_{k-1},t_k)$ and thus $\mathcal{H} \in \mathrm{PC}^{\mathbb{T}}([0,T],\mathbb{R})$.

(iii) In the joint case where the partition is static and the hybrid dynamics is autonomous, then $\mathcal{H}$ is constant over $[0,T]$.

**Remark 2.9.** Note that the averaged Hamiltonian gradient condition (AHG) obtained in Theorem 2.1 was already derived in previous works [19, 20, 21, 22] in the context of *optimal sampled-data control problems* (that is, with piecewise constant controls but, compared to the present work, in the simpler context where the corresponding partition is independent of the state position). It is worth mentioning that (AHG) is implicit in general since $\lambda_k$ intervenes, not only in both sides of the equation, but moreover in the values of $x$ and $p$ along the interval $(t^c_{k-1}, t^c_k)$. Furthermore we do not know in advance the values of $t^c_{k-1}$ and $t^c_k$. However, as already seen in the previous works mentioned and as we will see in Section 3, the averaged Hamiltonian gradient condition (AHG) can be useful to determine the optimal values of regionally switching parameters. From a numerical point of view, when $\Lambda$ is closed, note that $\lambda_k$ can be expressed as the fixed point

$$\lambda_k = \mathrm{proj}_\Lambda \Big( \lambda_k + \int_{t^c_{k-1}}^{t^c_k} \nabla_\lambda H(x(s), \lambda_k, u(s), p(s), s) \, \mathrm{d}s \Big),$$

where $\mathrm{proj}_\Lambda : \mathbb{R}^d \to \mathbb{R}^d$ stands for the standard projection operator onto $\Lambda$. We also note that Item (i) in Remark 2.8 is in accordance with the main result obtained in [19] stating that, when the partition associated with sampled-data controls is free, then the corresponding necessary optimality condition coincides with the continuity of the Hamiltonian function $\mathcal{H}$. Finally, let us also mention that, under some additional appropriate convexity assumptions made on the dynamics (see, *e.g.*, [44]), it should be possible to derive an *averaged Hamiltonian maximization condition* of the form

$$\lambda_k \in \operatorname*{arg\,max}_{\mu \in \Lambda} \int_{t^c_{k-1}}^{t^c_k} H(x(s), \mu, u(s), p(s), s) \, \mathrm{d}s,$$

but this point is out of the scope of the present work.

**Remark 2.10.** In this paper we have considered a certain framework which is, of course, not the most general possible. In fact we have made certain choices to make the presentation and the notations as simple and pleasant to read as possible, while keeping the essence of our work. In this remark our aim is to gather a number of possible relaxations and extensions of Theorem 2.1. One can easily be convinced by the validity of these generalizations by reading the proof of Theorem 2.1 in Sections 4, 5 and 6 (or the overview of the proof provided in Section 2.4).

(i) The convexity hypothesis made on $\Lambda$ can be removed by using a generalized version of the normal cone. More precisely, instead of using basic convex perturbations of the form $\lambda_k + \alpha(\overline{\lambda}_k - \lambda_k)$ in the proof of Theorem 2.1, one can invoke a general perturbation $\tilde{\lambda}_k(\alpha)$ where $\tilde{\lambda}_k : [0,1] \to \Lambda$ is a continuous function satisfying $\tilde{\lambda}_k(0) = \lambda_k$ and that is differentiable at 0 with derivative denoted by $\tilde{\lambda}'_k(0)$. Therefore Theorem 2.1 remains valid by considering the generalized notion of normal cone to $\Lambda$ at some $\lambda \in \Lambda$ given by

$$\mathrm{N}^{gen}_\Lambda[\lambda] := \{ \lambda'' \in \mathbb{R}^d \mid \langle \lambda'', \tilde{\lambda}'(0) \rangle_{\mathbb{R}^d} \leq 0 \text{ for all continuous functions } \tilde{\lambda} : [0,1] \to \Lambda \text{ with } \tilde{\lambda}(0) = \lambda \text{ and differentiable at 0 with derivative denoted by } \tilde{\lambda}'(0) \}.$$

(ii) The closedness hypothesis made on U can be removed by assuming in Theorem 2.1 that all the limits $u^-(t^c_k)$ and $u^+(t^c_k)$ belong to U. Indeed, in our proof of Theorem 2.1, we only need that the auxiliary controls $\tilde{u}_k$ are with values in U.

(iii) The right continuity after each crossing time $t^c_k$, and the left continuity before the last crossing time $t^c_{N-1}$, of the control $u$ (see Definition 2.2) are useless in our proof of Theorem 2.1. We have adopted these hypotheses for the sake of simplicity of the presentation. However they can be removed.

(iv) Theorem 2.1 is stated for a solution to (OCP) in a global sense. However Theorem 2.1 remains valid for a solution to (OCP) in (only) a local sense to be precised.

(v) The $\mathrm{C}^1$-regularity of the map $\phi$ can be relaxed. Indeed only the differentiability of $\phi$ is required for our proof of Theorem 2.1. Similarly the $\mathrm{C}^1$-regularity of the maps $h_j$ can be relaxed. Indeed our proof of Theorem 2.1 (precisely the sensitivity analyses developed in the preliminary Sections 4 and 5) requires (only) that, for all $j \in \mathcal{J}$, the map $h_j$ is continuous, is differentiable with respect to its two first variables with $\nabla_x h_j$ and $\nabla_\lambda h_j$ continuous, and is Lipschitz continuous with respect to its three first variables on any compact subset of $\mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0,T]$. We refer to [14] where similar relaxed regularity assumptions have been considered.

(vi) A possible extension of our work is to consider, for each region $X_j$, a parameter constraint set $\Lambda_j \subset \mathbb{R}^d$ and a control constraint set $\mathrm{U}_j \subset \mathbb{R}^m$. Such a generalized context is interesting to impose the values of the regionally switching parameter and/or of the control in certain regions (for example by taking $\Lambda_j = \{0_{\mathbb{R}^d}\}$ and/or $\mathrm{U}_j = \{0_{\mathbb{R}^m}\}$ for some $j \in \mathcal{J}$). However, be careful, we emphasize that our proof is based on the construction of auxiliary controls (and, in some way, of auxiliary parameters) which would require that the parameter constraint sets $\Lambda_j$ are all subsets of the same space $\mathbb{R}^d$, and that the control constraint sets $\mathrm{U}_j$ are all subsets of the same space $\mathbb{R}^m$. Hence, to the best of our knowledge, one cannot extend our proof to the case of different $d_j$ and $m_j$.

(vii) A possible generalization of our work is to extend the space partition of $\mathbb{R}^n$ to a space-time partition of the form $\mathbb{R}^n \times [0,T] = \cup_{j\in\mathcal{J}} Y_j$. Such an extension would cover, in particular, the framework of *optimal sampled-data control problems* developed in [19, 20, 21, 22].

(viii) Using the classical technique of augmenting the state of the control system (see, *e.g.*, [17]), one can easily extend Theorem 2.1 to deal with Bolza costs, that is, when the cost of (OCP) is replaced by a cost of the form

$$\phi(x(T)) + \int_0^T L(x(s), \lambda(s), u(s), s)\, \mathrm{d}s,$$

where the hybrid Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0,T] \to \mathbb{R}$ is defined regionally by

$$\forall (x,\lambda,u,t) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0,T], \quad L(x,\lambda,u,t) := L_j(x,\lambda,u,t) \text{ when } x \in X_j(t),$$

where the maps $L_j : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times [0,T] \to \mathbb{R}$ are of class $\mathrm{C}^1$. In that context Theorem 2.1 remains valid by replacing the definition of the Hamiltonian $H$ by

$$H(x,\lambda,u,p,t) := \langle p, h(x,\lambda,u,t)\rangle_{\mathbb{R}^n} - L(x,\lambda,u,t),$$

for all $(x,\lambda,u,p,t) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^n \times [0,T]$, and by replacing (AD) by

$$\begin{aligned} p^+(t_k^c) - p^-(t_k^c) \\ &= -\frac{\langle p^+(t_k^c), (f_{k+1})^+(t_k^c) - (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n} - ((L_{k+1})^+(t_k^c) - (L_k)^-(t_k^c))}{\langle \nabla_x F_k(x(t_k^c), t_k^c), (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t_k^c), t_k^c)} \nabla_x F_k(x(t_k^c), t_k^c). \end{aligned}$$

(ix) One can also consider a hybrid control system of the form $\dot{x}(t) = h(x(t), \mu, \lambda(t), u(t), t)$ involving an additional constant parameter $\mu \in \mathbb{R}^{d'}$ for some $d' \in \mathbb{N}^*$. In that context we emphasize that $\mu$ is not a regionally switching parameter: it is constant over the whole interval $[0,T]$. Then one can consider the additional parameter constraint $\mu \in \mathrm{M}$ in Problem (OCP), where $\mathrm{M}$ is a nonempty convex subset of $\mathbb{R}^{d'}$. By adapting the proof of Theorem 2.1, one can easily see that Theorem 2.1 remains valid by replacing the definition of the Hamiltonian $H$ by

$$H(x,\mu,\lambda,u,p,t) := \langle p, h(x,\mu,\lambda,u,t)\rangle_{\mathbb{R}^n},$$

for all $(x,\mu,\lambda,u,p,t) \in \mathbb{R}^n \times \mathbb{R}^{d'} \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^n \times [0,T]$, and by adding the necessary optimality condition given by

$$\int_0^T \nabla_\mu H(x(s), \mu, \lambda(s), u(s), p(s), s)\, \mathrm{d}s \in \mathrm{N}_{\mathrm{M}}[\mu].$$

(x) One can also consider a version of Problem (OCP) with a *free final time* $T > 0$ and in which the Mayer cost is of the form $\phi(T, x(T))$. Such a framework is important to deal with *minimal time problems* (see, *e.g.*, [26, Section 3]). By using the classical technique of change of time variable $t = Ts$, one can transform the variable $T$ to optimize as a constant parameter. Thanks to the previous item and the results obtained in Remark 2.8, one can prove that Theorem 2.1 remains valid with the additional necessary optimality condition given by

$$\mathcal{H}(T) = \nabla_T \phi(T, x(T)),$$

where $\mathcal{H} : [0, T] \to \mathbb{R}$ is the Hamiltonian function introduced in Remark 2.8.

**Remark 2.11.** This last remark is dedicated to a nonexhaustive list of possible nontrivial perspectives:

(i) The present paper does not cover *terminal state constraints* (that is, constraints on $x(0)$ and $x(T)$) which are common in most applications of optimal control theory. This is clearly a criticism that can be made on the present work. In the classical non-hybrid setting, several methods have been developed in the literature to take into account such terminal state constraints. One can invoke the Ekeland variational principle [33] or some implicit function arguments (see, *e.g.*, [2, 58]). In one hand, it is worth mentioning that, to the best of our knowledge, the Ekeland approach does not apply in the present hybrid setting for two main reasons. First, this approach requires to define a continuous penalized functional on a $\mathrm{L}^1$-neighborhood of the optimal control $u$. However we have seen in Item 2 of Section 2.4 that such a construction is obstructed in the present hybrid setting. Second, the control sequence produced by the Ekeland variational principle (which converges in $\mathrm{L}^1$-norm to the optimal control $u$) would have no reason to satisfy the regularity assumption ($\mathcal{A}1$) and therefore the sensitivity analysis developed in the preliminary Section 5 may be not valid on the control sequence. On the other hand, we are confident that a method based on an implicit function argument could be adapted to the present hybrid setting. However this approach is also based on a separation argument within the so-called *Pontryagin convex cone* constructed thanks to the consideration of *multiple* needle-like perturbations of the control (see, *e.g.*, [1, 24, 30, 32, 46, 52] and references therein). In the present hybrid setting, this would have required the consideration of multiple needle-like perturbations of the control in each region simultaneously. This would have significantly increased the complexity of the analysis and the notations. Since our main objective in this work was to focus on the concept of *regionally switching parameter* and on the corresponding averaged Hamiltonian gradient condition (AHG), we decided to avoid the technicalities related to the presence of terminal state constraints which are already well known in the literature and to keep the reading of the technical proof of Theorem 2.1 as pleasant as possible. A fortiori, the same comment holds true for the consideration of running state constraints as developed in [17, 18, 20] and references therein.

(ii) The regularity assumptions introduced in Definition 2.2 are crucial to develop the sensitivity analysis of the hybrid control system in the preliminary Section 5, precisely to construct perturbed trajectories which visit exactly (and in the same order) the same regions than the nominal trajectory. To the best of our knowledge, an open question is how to obtain a hybrid maximum principle without these regularity assumptions. In that direction, note that a similar sensitivity analysis can be developed in a hybrid framework where Assumption ($\mathcal{A}3$) is removed, as it has been studied in the case of two static regions from a dynamic programming standpoint in [7].

(iii) As mentioned in Introduction, the original motivation of the present work was to deal with (non-hybrid) optimal control problems involving *non-control regions* (in which the control should remain constant). This framework is actually a particular case of the present work that has been developed in details in [8] and is motivated by applications related to aerospace, for example thrust problems with shadow zones causing inability to thrust while the spacecraft is passing through an eclipse, due to the low power generated by the solar panels (see [35, 41, 45, 64]). One could also consider a slightly different setting where, in non-control regions, the control is an affine feedback of the state (and thus is not necessarily constant). Again, this framework can be seen as a particular case of the present work and will be developed in details in a forthcoming research work.

(iv) In the field of mathematical epidemiology, hybrid frameworks provide an accurate description of some infectious diseases and their spread. We refer for instance to [47] where the authors take into account that the contact rate between members of the population changes throughout each season, or to [16] in which the authors provide a version of the SIR model that takes into consideration different control strategies (vaccination, isolation, culling, etc.). An interesting research perspective would be to consider a time crisis problem (such as in [9, 10]) related to a COVID-19 model, in order to provide better control strategies. To this aim, using the approach of optimal control problems with non-control regions presented in [8] (which is a particular case of the present work) is privileged. Moreover, since time crisis problems deal with Bolza costs with a discontinuous Lagrangian function, one can note that our main result (Theorem 2.1) tackles perfectly this discontinuity (see Remark 2.10).

(v) In this work we have investigated the necessary optimality conditions for hybrid optimal control problems with regionally switching parameter. However note that many other standard investigations from optimal control theory can be developed for that framework. First, one may develop existence results, by extending for example the classical Filippov theorem [36]. This would certainly require to introduce adequate differential inclusions (see, *e.g.*, [4]), in particular at the interfaces where the dynamics is not defined. Sufficient optimality conditions could also be investigated, at least in the case of LQ-problems (see related studies in [54, 55, 65] for switched systems). Also a complete extension of the Riccati theory in the present hybrid setting with regionally switching parameter constitutes an attractive perspective for future works (see [23] for a related study with sampled-data controls). Finally, from a numerical point of view, another perspective could be the formulation of a multiple shooting method as in [42] taking into account the averaged Hamiltonian gradient condition (AHG).

# 3 Application to an academic example

The objective of this section is to show that our work fills a gap in the literature. Precisely, based on the necessary optimality conditions derived in Theorem 2.1, we solve a hybrid optimal control problem involving a regionally switching parameter and we show (see Figure 12) that the corresponding optimal cost is strictly between, in one hand, the best cost that can be obtained when replacing the regionally switching parameter by a classical permanent control and, in the other hand, the best cost that can be obtained when replacing the regionally switching parameter by a classical constant parameter. The example studied in this section is a simple academic example whose only purpose is to fulfill the objective of this section. The application of Theorem 2.1 to concrete and sophisticated application models as evoked in Items (iii) and (iv) of Remark 2.11, in particular to optimal control problems with non-control regions (as specified in [8]), will be the focus of our forthcoming research works.

## 3.1 Presentation of the example

Take $T=8$, $n=d=m=1$, $x_{\mathrm{init}}=-1$ and the static partition $\mathbb{R}=\overline{X_1}\cup\overline{X_2}\cup\overline{X_3}$ where

$$X_1=\left\{y\in\mathbb{R}\mid y<-\frac{1}{2}\right\},\quad X_2=\left\{y\in\mathbb{R}\mid -\frac{1}{2}<y<\frac{1}{4}\right\},\quad X_3=\left\{y\in\mathbb{R}\mid y>\frac{1}{4}\right\}.$$

In this section we consider the hybrid optimal control problem with regionally switching parameter given by

$$\begin{aligned}
&\text{minimize} && -x(8),\\
&\text{subject to} && (x,\lambda,u)\in \mathrm{AC}([0,8],\mathbb{R})\times \mathrm{PC}([0,8],\mathbb{R})\times \mathrm{L}^\infty([0,8],\mathbb{R}),\\
& && \dot{x}(t)=h(x(t),\lambda(t),u(t)),\quad \text{a.e. } t\in[0,8],\\
& && x(0)=-1,\\
& && \lambda \text{ is a regionally switching parameter associated with } x,\\
& && \lambda(t)\in[-\tfrac{3}{2},\tfrac{3}{4}],\quad \text{a.e. } t\in[0,8],\\
& && u(t)\in[-1,1],\quad \text{a.e. } t\in[0,8],
\end{aligned}\tag{P$_{\mathrm{ex}}$}$$

where the (autonomous) hybrid dynamics $h : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is given by

$$h(x,\lambda,u) = \begin{cases} u(x-1)+\lambda & \text{if} \quad x \in X_1, \\ \lambda x + \frac{1}{2}u & \text{if} \quad x \in X_2, \\ u(x-1)+\lambda & \text{if} \quad x \in X_3, \end{cases}$$

for all $(x,\lambda,u) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. We refer to Figure 11 for an illustration of the setting of Problem ($\mathrm{P}_{\mathrm{ex}}$) in which the objective is to maximize the final value $x(8)$ starting from the initial condition $x_{\mathrm{init}} = -1$. Then let us recall that the Hamiltonian $H : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ associated with Problem ($\mathrm{P}_{\mathrm{ex}}$) is given by

$$H(x,\lambda,u,p) = \begin{cases} p(u(x-1)+\lambda) & \text{if} \quad x \in X_1, \\ p(\lambda x + \frac{1}{2}u) & \text{if} \quad x \in X_2, \\ p(u(x-1)+\lambda) & \text{if} \quad x \in X_3, \end{cases}$$

for all $(x,\lambda,u,p) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

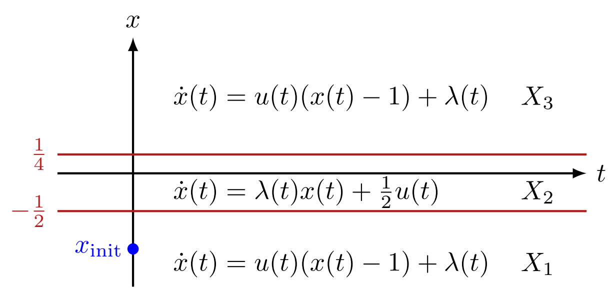


**Figure 11:** Illustration of the setting of Problem ($\mathrm{P}_{\mathrm{ex}}$). Here the objective is to maximize the final value $x(8)$ starting from the initial condition $x_{\mathrm{init}} = -1$.

Since existence results are out of the scope of the present work (see Item (v) of Remark 2.11), we assume here that ($\mathrm{P}_{\mathrm{ex}}$) has a solution $(x,\lambda,u)$ and we denote by $\mathbb{T}$ the corresponding partition. Our aim in the next section is to identify such a triple thanks to the necessary optimality conditions stated in Theorem 2.1. Therefore we assume furthermore that the regularity assumptions ($\mathcal{A}1$), ($\mathcal{A}2$) and ($\mathcal{A}3$) are fulfilled. Finally, to simplify the analysis and according to the nature of the objective of ($\mathrm{P}_{\mathrm{ex}}$), we assume that $x(8) > 1$, that $x$ is increasing over $[0,8]$, with exactly two crossing times $0 < t_1^c < t_2^c < 8$, and that

$$\begin{aligned} \forall t \in (0,t_1^c), &\quad (x(t),\lambda(t)) \in X_1 \times \{\lambda_1\}, \\ \forall t \in (t_1^c,t_2^c), &\quad (x(t),\lambda(t)) \in X_2 \times \{\lambda_2\}, \\ \forall t \in (t_2^c,8), &\quad (x(t),\lambda(t)) \in X_3 \times \{\lambda_3\}, \end{aligned}$$

for some $\lambda_1$, $\lambda_2$, $\lambda_3 \in [-\frac{3}{2}, \frac{3}{4}]$.

## 3.2 Application of Theorem 2.1

Let us denote by $p \in \mathrm{PAC}^{\mathbb{T}}([0,T],\mathbb{R})$ the adjoint vector provided in Theorem 2.1. Our aim in this section is to identify the triple $(x,\lambda,u)$ thanks to the necessary optimality conditions stated in Theorem 2.1. To this aim we reason backward in time.

**Analysis over the interval $(t_2^c, 8)$.** Since $x(t) \in X_3$ over $(t_2^c,8)$, the adjoint equation and the final condition give

$$\begin{cases} \dot{p}(t) = -u(t)p(t), & \text{a.e. } t \in (t_2^c,8), \\ p(8) = 1, \end{cases}$$

which implies that $p(t) > 0$ over $(t_2^c, 8]$. Since the averaged Hamiltonian gradient condition writes

$$\int_{t_2^c}^{8} p(s)\,\mathrm{d}s \in \mathrm{N}_{[-\frac{3}{2},\frac{3}{4}]}(\lambda_3),$$

we deduce that $\lambda_3 = \frac{3}{4}$. Since $p(t) > 0$ over $(t_2^c, 8]$, one can easily derive from the Hamiltonian maximization condition that

$$u(t) \in \underset{v\in[-1,1]}{\arg\max}\ v(x(t)-1),$$

for almost every $t \in (t_2^c, 8)$. Since $x(8) > 1$, $x(t_2^c) = \frac{1}{4}$ and $x$ is increasing over $[t_2^c, 8]$, there exists a unique $t^* \in (t_2^c, 8)$ such that $x(t^*) = 1$ and we obtain that

$$u(t) = \begin{cases} +1, & \text{a.e. } t \in (t^*, 8), \\ -1, & \text{a.e. } t \in (t_2^c, t^*). \end{cases}$$

We deduce that

$$p(t) = \begin{cases} e^{8-t}, & \text{for all } t \in [t^*, 8], \\ e^{8-t^*}e^{t-t^*}, & \text{for all } t \in (t_2^c, t^*], \end{cases} \qquad \text{and} \qquad x(t) = \begin{cases} \frac{3}{4}e^{t-t^*} + \frac{1}{4}, & \text{for all } t \in [t^*, 8], \\ -\frac{6}{4}e^{t_2^c - t} + \frac{7}{4}, & \text{for all } t \in [t_2^c, t^*]. \end{cases}$$

Finally, from the continuity of $x$ at $t^*$, we get that $t^* = t_2^c + \ln(2)$.

**Analysis over the interval $(t_1^c, t_2^c)$.** The adjoint discontinuity condition at $t_2^c$ writes $p^-(t_2^c) = p^+(t_2^c)\frac{(f_3)^+(t_2^c)}{(f_2)^-(t_2^c)}$ which implies, from Assumption $(\mathcal{A}3)$, that $p^-(t_2^c) > 0$ and, from $t^* = t_2^c + \ln(2)$, that $p^-(t_2^c) = \frac{3e^{8-t_2^c}}{2\lambda_2 + 4u^-(t_2^c)}$. Since $x(t) \in X_2$ over $(t_1^c, t_2^c)$, when adding the adjoint equation, we obtain that

$$\begin{cases} \dot{p}(t) = -\lambda_2 p(t), & \text{a.e. } t \in (t_1^c, t_2^c), \\ p^-(t_2^c) = \frac{3e^{8-t_2^c}}{2\lambda_2 + 4u^-(t_2^c)}, \end{cases}$$

which implies that $p(t) > 0$ over $(t_1^c, t_2^c)$. The Hamiltonian maximization condition leads to

$$u(t) \in \underset{v\in[-1,1]}{\arg\max}\ vp(t),$$

and thus to $u(t) = 1$ for almost every $t \in (t_1^c, t_2^c)$. We deduce that

$$p(t) = \frac{3e^{8-t_2^c}}{2\lambda_2 + 4}e^{\lambda_2(t_2^c - t)} \quad \text{and} \quad x(t) = \begin{cases} \frac{1}{2\lambda_2}\left(e^{\lambda_2(t-t_2^c)} - 1\right) + \frac{1}{4}e^{\lambda_2(t-t_2^c)}, & \text{if } \lambda_2 \neq 0, \\ \frac{t-t_2^c}{2} + \frac{1}{4}, & \text{if } \lambda_2 = 0, \end{cases}$$

for all $t \in (t_1^c, t_2^c)$. Since $x(t_1^c) = -\frac{1}{2}$ and $\lambda_2 \in [-\frac{3}{2}, \frac{3}{4}]$, we get that

$$t_2^c - t_1^c = \begin{cases} \frac{1}{\lambda_2}\ln\left(\frac{1+\frac{\lambda_2}{2}}{1-\lambda_2}\right), & \text{if } \lambda_2 \neq 0, \\ \frac{3}{2}, & \text{if } \lambda_2 = 0. \end{cases}$$

Since $\frac{3e^{8-t_2^c}}{2\lambda_2+4} > 0$, the averaged Hamiltonian gradient condition is equivalent to $\upsilon(\lambda_2) \in \mathrm{N}_{[-\frac{3}{2},\frac{3}{4}]}(\lambda_2)$ where

$$\upsilon(\lambda_2) := \begin{cases} \displaystyle\int_{t_1^c}^{t_2^c} \frac{1}{2\lambda_2} + \frac{1}{4} - \frac{1}{2\lambda_2}e^{\lambda_2(t_2^c - s)}\,\mathrm{d}s = \frac{1}{2\lambda_2^2}\left(1 + \frac{\lambda_2}{2}\right)\ln\left(\frac{1+\frac{\lambda_2}{2}}{1-\lambda_2}\right) - \frac{3}{4\lambda_2(1-\lambda_2)}, & \text{if } \lambda_2 \neq 0, \\ \displaystyle\int_{t_1^c}^{t_2^c} \frac{s - t_2^c}{2} + \frac{1}{4}\,\mathrm{d}s = -\frac{3}{16}, & \text{if } \lambda_2 = 0. \end{cases}$$

We find that:

- if $\lambda_2 = \frac{3}{4}$, then $\upsilon(\lambda_2) \simeq -1.916 < 0$, while $\mathrm{N}_{[-\frac{3}{2},\frac{3}{4}]}(\lambda_2) = \mathbb{R}_+$, which is a contradiction.
- if $\lambda_2 = -\frac{3}{2}$, then $\upsilon(\lambda_2) \simeq 0.072 > 0$, while $\mathrm{N}_{[-\frac{3}{2},\frac{3}{4}]}(\lambda_2) = \mathbb{R}_-$, which is a contradiction.

We deduce that $\lambda_2 \in (-\frac{3}{2}, \frac{3}{4})$ and thus $\mathrm{N}_{[-\frac{3}{2},\frac{3}{4}]}(\lambda_2) = \{0\}$. Solving the equation $\upsilon(\lambda_2) = 0$ over $(-\frac{3}{2}, \frac{3}{4})$, we find that $\lambda_2 \simeq -0.754$.

**Analysis over the interval** $(0, t_1^c)$**.** The adjoint discontinuity condition at $t_1^c$ writes $p^-(t_1^c) = p^+(t_1^c)\frac{(f_2)^+(t_1^c)}{(f_1)^-(t_1^c)}$ which implies, from Assumption $(\mathcal{A}3)$, that $p^-(t_1^c) > 0$ and, from $t_2^c - t_1^c = \frac{1}{\lambda_2}\ln(\frac{1+\frac{\lambda_2}{2}}{1-\lambda_2})$, one can obtain that $p^-(t_1^c) = \frac{3e^{8-t_2^c}}{8(\lambda_1 - \frac{3}{2}u^-(t_1^c))}$. Since $x(t) \in X_1$ over $(0, t_1^c)$, when adding the adjoint equation, we obtain that

$$\begin{cases} \dot{p}(t) = -u(t)p(t), & \text{a.e. } t \in (0, t_1^c), \\ p^-(t_1^c) = \frac{3e^{8-t_2^c}}{8(\lambda_1 - \frac{3}{2}u^-(t_1^c))}. \end{cases}$$

Following similar arguments as in the analysis over the interval $(t_2^c, 8)$, one can prove that $\lambda_1 = \frac{3}{4}$ and $u(t) = -1$ for almost every $t \in (0, t_1^c)$.

**Conclusion.** From the above analysis we obtain that

$$x(t) = \begin{cases} -\frac{11}{4}e^{-t} + \frac{7}{4}, & \text{for all } t \in [0, t_1^c], \\ -\frac{1}{2\lambda_2}((\lambda_2 - 1)e^{\lambda_2(t-t_1^c)} + 1), & \text{for all } t \in [t_1^c, t_2^c], \\ -\frac{6}{4}e^{-(t-t_2^c)} + \frac{7}{4}, & \text{for all } t \in [t_2^c, t^*], \\ +\frac{3}{4}e^{t-t^*} + \frac{1}{4}, & \text{for all } t \in [t^*, 8], \end{cases}$$

and

$$\lambda(t) = \begin{cases} \frac{3}{4}, & \text{for a.e. } t \in (0, t_1^c), \\ \lambda_2, & \text{for a.e. } t \in (t_1^c, t_2^c), \\ \frac{3}{4}, & \text{for a.e. } t \in (t_2^c, 8), \end{cases} \qquad u(t) = \begin{cases} -1, & \text{for a.e. } t \in (0, t_1^c), \\ +1, & \text{for a.e. } t \in (t_1^c, t_2^c), \\ -1, & \text{for a.e. } t \in (t_2^c, t^*), \\ +1, & \text{for a.e. } t \in (t^*, 8). \end{cases}$$

with $\lambda_2 \simeq -0.754$, $t_1^c = \ln(\frac{11}{9}) \simeq 0.2$, $t_2^c = t_1^c + \frac{1}{\lambda_2}\ln(\frac{1+\frac{\lambda_2}{2}}{1-\lambda_2}) \simeq 1.57$ and $t^* = t_2^c + \ln(2) \simeq 2.26$.

### 3.3 Comparisons with standard settings found in the literature

Our objective in this section is to emphasize that our work fills a gap in the literature. To this aim we will show on the present academic example that the optimal trajectory $x$ computed in the previous section (associated with a regionally switching parameter $\lambda$) is exactly between the optimal trajectory $x^\dagger$ when $\lambda$ is considered as a classical permanent control (that is, when $\lambda \in \mathrm{L}^\infty([0,8],\mathbb{R})$), and the optimal trajectory $\hat{x}$ when $\lambda$ is considered as a classical constant parameter (that is, when $\lambda \in \mathbb{R}$). Precisely:

- First, let us consider the case where $\lambda \in \mathrm{L}^\infty([0,8],\mathbb{R})$ is a classical permanent control in Problem $(\mathrm{P}_{\mathrm{ex}})$ (and not a regionally switching parameter that has to remain constant in each region) constrained to be with values in $[-\frac{3}{2}, \frac{3}{4}]$. By developing a similar analysis to the previous section, but by using a Hamiltonian maximization condition to determine the values $\lambda(t)$ over the whole interval $[0, T]$, we get that the optimal triple $(x^\dagger, \lambda^\dagger, u^\dagger)$ is given by

$$x^\dagger(t) = \begin{cases} -\frac{11}{4}e^{-t} + \frac{7}{4}, & \text{for all } t \in [0, (t_1^c)^\dagger], \\ -\frac{5}{6}e^{-\frac{3}{2}(t-(t_1^c)^\dagger)} + \frac{1}{3}, & \text{for all } t \in [(t_1^c)^\dagger, (t_1^*)^\dagger], \\ +\frac{2}{3}e^{\frac{3}{4}(t-(t_1^*)^\dagger)} - \frac{2}{3}, & \text{for all } t \in [(t_1^*)^\dagger, (t_2^c)^\dagger], \\ -\frac{3}{2}e^{-(t-(t_2^c)^\dagger)} + \frac{7}{4}, & \text{for all } t \in [(t_2^c)^\dagger, (t_2^*)^\dagger], \\ +\frac{3}{4}e^{t-(t_2^*)^\dagger} + \frac{1}{4}, & \text{for all } t \in [(t_2^*)^\dagger, 8], \end{cases}$$

and

$$\lambda^\dagger(t) = \begin{cases} +\frac{3}{4}, & \text{for a.e. } t \in (0, (t_1^c)^\dagger), \\ -\frac{3}{2}, & \text{for a.e. } t \in ((t_1^c)^\dagger, (t_1^*)^\dagger), \\ +\frac{3}{4}, & \text{for a.e. } t \in ((t_1^*)^\dagger, (t_2^c)^\dagger), \\ -\frac{3}{2}, & \text{for a.e. } t \in ((t_2^c)^\dagger, (t_2^*)^\dagger), \\ +\frac{3}{4}, & \text{for a.e. } t \in ((t_2^*)^\dagger, 8), \end{cases} \qquad u^\dagger(t) = \begin{cases} -1, & \text{for a.e. } t \in (0, (t_1^c)^\dagger), \\ +1, & \text{for a.e. } t \in ((t_1^c)^\dagger, (t_2^c)^\dagger), \\ -1, & \text{for a.e. } t \in ((t_2^c)^\dagger, (t_2^*)^\dagger), \\ +1, & \text{for a.e. } t \in ((t_2^*)^\dagger, 8). \end{cases}$$

with $(t_1^c)^\dagger = t_1^c \simeq 0.2$, $(t_1^*)^\dagger \simeq 0.81$, $(t_2^c)^\dagger \simeq 1.23$, $(t_2^*)^\dagger \simeq 2.07$. The detailed computations are left to the reader.

- Second, let us consider the case where $\lambda \in \mathbb{R}$ is a classical constant parameter in Problem $(\mathrm{P}_{\mathrm{ex}})$ (that cannot switch at boundary crossings) constrained to belong to $[-\frac{3}{2}, \frac{3}{4}]$. By developing a similar analysis to the previous section, but by using the averaged Hamiltonian gradient condition given in Item (ix) of Remark 2.10, we get that the optimal triple $(\hat{x}, \hat{\lambda}, \hat{u})$ is given by

$$\hat{x}(t) = \begin{cases} -\frac{11}{4}e^{-t} + \frac{7}{4}, & \text{for all } t \in [0, \hat{t_1^c}], \\ +\frac{5}{3}e^{\frac{3}{4}(t-\hat{t_1^c})} - \frac{2}{3}, & \text{for all } t \in [\hat{t_1^c}, \hat{t_2^c}], \\ -\frac{3}{2}e^{-(t-\hat{t_2^c})} + \frac{7}{4}, & \text{for all } t \in [\hat{t_2^c}, \hat{t_2^*}], \\ -\frac{3}{4}e^{-(t-\hat{t^*})} + 7/4, & \text{for all } t \in [\hat{t_2^*}, 8], \end{cases}$$

and

$$\hat{\lambda} = \frac{3}{4}, \qquad \hat{u}(t) = \begin{cases} -1, & \text{for a.e. } t \in (0, \hat{t_1^c}), \\ +1, & \text{for a.e. } t \in (\hat{t_1^c}, \hat{t_2^c}), \\ -1, & \text{for a.e. } t \in (\hat{t_2^c}, \hat{t^*}), \\ +1, & \text{for a.e. } t \in (\hat{t^*}, 8). \end{cases}$$

with $\hat{t_1^c} = t_1^c \simeq 0.2$, $\hat{t_2^c} \simeq 2.47$, and $\hat{t^*} = 3.16$. The detailed computations are left to the reader.

We refer to Figure 12 for the plots of the three trajectories $x^\dagger$, $x$ and $\hat{x}$. As expected, the trajectory $x^\dagger$ (associated with a classical permanent control) provides a better cost than the trajectory $x$ (associated with a regionally switching parameter) which provides a better cost than the trajectory $\hat{x}$ (associated with a classical constant parameter). This figure emphasizes that the present work allows to fill a gap in the literature.

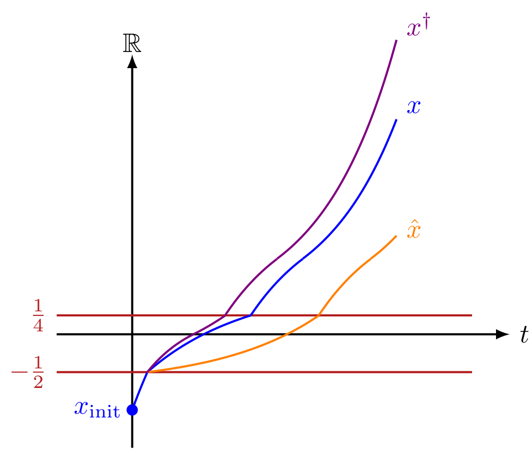


**Figure 12:** Trajectories $x^\dagger$, $x$ and $\hat{x}$ (zoom on the time interval $[0, \frac{7}{2}]$).

## 4 Preliminaries: sensitivity analysis in the non-hybrid context

As explained in the overview of the proof of Theorem 2.1 developed in Section 2.4, for the needs of the sensitivity analysis in the hybrid context performed in the next Section 5, we need to provide a complete sensitivity analysis of a general non-hybrid parameterized control system with respect to perturbations of the parameter, the control, the initial time and the initial condition. This is precisely the aim of the present section. We will work on the time interval $[0,T]$. Nevertheless our results can be trivially extended to any compact time interval $[a,b]$ with $a<b$. In fact note that we will use these results in the next Section 5 on compact subintervals of $[0,T]$.

Let $g:\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^m\times[0,T]\to\mathbb{R}^n$ be a general (non-hybrid) dynamics of class $\mathrm{C}^1$. For any quadruplet $\theta=(\lambda,u,r,y_r)\in\mathbb{R}^d\times\mathrm{L}^\infty([0,T],\mathbb{R}^m)\times[0,T]\times\mathbb{R}^n$, the Cauchy-Lipschitz theorem ensures the existence and the uniqueness of the maximal solution to the Cauchy problem

$$\begin{cases} \dot{y}(t)=g(y(t),\lambda,u(t),t), & \text{a.e. } t\in[0,T],\\ y(r)=y_r. \end{cases}$$

This maximal solution is denoted by $y(\cdot,g,\theta)$ and is defined over the maximal interval denoted by $I(g,\theta)\subset[0,T]$. Recall that the blow-up theorem ensures that, either $I(g,\theta)=[0,T]$ (in that case we speak of a global solution), either $y(\cdot,g,\theta)$ is unbounded over $I(g,\theta)$. In the sequel we denote by $\mathrm{Glob}(g)$ the set of all quadruplets $\theta$ such that $I(g,\theta)=[0,T]$.

For the technical needs of this section, for any quadruplet $\theta=(\lambda,u,r,y_r)\in\mathrm{Glob}(g)$ and any $R\geq\|u\|_{\mathrm{L}^\infty}$, we denote by $M(g,\theta,R)\geq 0$ a common bound of $\|g\|_{\mathbb{R}^n}$, $\|\nabla_x g\|_{\mathbb{R}^{n\times n}}$, $\|\nabla_\lambda g\|_{\mathbb{R}^{n\times d}}$ and $\|\nabla_u g\|_{\mathbb{R}^{n\times m}}$ over the compact set

$$K(g,\theta,R):=\Big\{(x,\mu,v,t)\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^m\times[0,T]\mid \|x-y(t,g,\theta)\|_{\mathbb{R}^n}\leq 1,\ \|\mu-\lambda\|_{\mathbb{R}^d}\leq 1,\ \|v\|_{\mathbb{R}^m}\leq R\Big\}.$$

Note that $(y(t,g,\theta),\lambda,u(t),t)\in K(g,\theta,R)$ for almost every $t\in[0,T]$. Since $K(g,\theta,R)$ is convex with respect to its first three components, one can easily get that

$$\|g(y_2,\mu_2,v_2,t)-g(y_1,\mu_1,v_1,t)\|_{\mathbb{R}^n}\leq M(g,\theta,R)(\|y_2-y_1\|_{\mathbb{R}^n}+\|\mu_2-\mu_1\|_{\mathbb{R}^d}+\|v_2-v_1\|_{\mathbb{R}^m}),$$

for all $(y_2,\mu_2,v_2,t),(y_1,\mu_1,v_1,t)\in K(g,\theta,R)$.

We are now in a position to state and prove the next continuous dependence result for the trajectory $y(\cdot,g,\theta)$ with respect to the quadruplet $\theta$.

**Lemma 4.1.** For any quadruplet $\theta=(\lambda,u,r,y_r)\in\mathrm{Glob}(g)$ and any $R\geq\|u\|_{\mathrm{L}^\infty}$, there exists $\varepsilon>0$ such that the neighborhood of $\theta$ given by

$$\mathcal{N}(g,\theta,R,\varepsilon):=\overline{\mathrm{B}}_{\mathbb{R}^d}(\lambda,\varepsilon)\times\Big(\overline{\mathrm{B}}_{\mathrm{L}^1}(u,\varepsilon)\cap\overline{\mathrm{B}}_{\mathrm{L}^\infty}(0_{\mathrm{L}^\infty},R)\Big)\times\Big([r-\varepsilon,r+\varepsilon]\cap[0,T]\Big)\times\overline{\mathrm{B}}_{\mathbb{R}^n}(y_r,\varepsilon),$$

is included in $\mathrm{Glob}(g)$. Furthermore, for all quadruplets $\theta'=(\lambda',u',r',y'_r)\in\mathcal{N}(g,\theta,R,\varepsilon)$, it holds that $(y(t,g,\theta'),\lambda',u'(t),t)\in K(g,\theta,R)$ for almost every $t\in[0,T]$. Finally the map

$$\begin{array}{rccc}\mathcal{F}: & \mathcal{N}(g,\theta,R,\varepsilon) & \to & \mathrm{C}([0,T],\mathbb{R}^n)\\ & \theta' & \mapsto & y(\cdot,g,\theta'),\end{array}$$

is Lipschitz continuous, in the sense that there exists $L(g,\theta,R)\geq 0$ such that

$$\|y(\cdot,g,\theta'')-y(\cdot,g,\theta')\|_{\mathrm{C}}\leq L(g,\theta,R)(\|\lambda''-\lambda'\|_{\mathbb{R}^d}+\|u''-u'\|_{\mathrm{L}^1}+|r''-r'|+\|y''_r-y'_r\|_{\mathbb{R}^n}),$$

for all $\theta'=(\lambda',u',r',y'_r)$, $\theta''=(\lambda'',u'',r'',y''_r)\in\mathcal{N}(g,\theta,R,\varepsilon)$.

*Proof.* Let $\theta=(\lambda,u,r,y_r)\in\mathrm{Glob}(g)$ and $R\geq\|u\|_{\mathrm{L}^\infty}$. In this proof, for the ease of notations, we denote by $M:=M(g,\theta,R)$. Let us fix $\varepsilon>0$ such that $\varepsilon(1+M(2+T))e^{MT}<1$ and let us prove that $\mathcal{N}(g,\theta,R,\varepsilon)\subset\mathrm{Glob}(g)$. To this aim let $\theta'=(\lambda',u',r',y'_r)\in\mathcal{N}(g,\theta,R,\varepsilon)$ and introduce the sets

$$\mathcal{I}_1:=\{t\in I(g,\theta')\cap[0,r']\mid \|y(t,g,\theta')-y(t,g,\theta)\|_{\mathbb{R}^n}>1\}$$
$$\text{and}\quad \mathcal{I}_2:=\{t\in I(g,\theta')\cap[r',T]\mid \|y(t,g,\theta')-y(t,g,\theta)\|_{\mathbb{R}^n}>1\}.$$

If $\mathcal{I}_1 \cup \mathcal{I}_2 = \emptyset$, then the solution $y(\cdot, g, \theta')$ is bounded over $I(g, \theta')$, and thus $\theta' \in \mathrm{Glob}(g)$ from the blow-up theorem. Therefore, by contradiction, let us assume that $\mathcal{I}_1 \cup \mathcal{I}_2 \neq \emptyset$. In the sequel we only deal with the case $\mathcal{I}_2 \neq \emptyset$ (the case where $\mathcal{I}_2 = \emptyset$, and thus $\mathcal{I}_1 \neq \emptyset$, is similar). From integral representations it holds that

$$y(t, g, \theta') - y(t, g, \theta) = (y'_r - y_r) + \int_{r'}^{t} g(y(s, g, \theta'), \lambda', u'(s), s) - g(y(s, g, \theta), \lambda, u(s), s) \, \mathrm{d}s$$
$$- \int_{r}^{r'} g(y(s, g, \theta), \lambda, u(s), s) \, \mathrm{d}s,$$

for all $t \in I(g, \theta')$. Now let $t_2 := \inf \mathcal{I}_2 \geq r'$. From continuity and definition of $t_2$ we know that $\|y(t_2, g, \theta') - y(t_2, g, \theta)\|_{\mathbb{R}^n} \geq 1$ and thus $r' < t_2$ since

$$\|y(r', g, \theta') - y(r', g, \theta)\|_{\mathbb{R}^n} \leq \|y'_r - y_r\|_{\mathbb{R}^n} + \left| \int_{r}^{r'} \|g(y(s, g, \theta), \lambda, u(s), s)\|_{\mathbb{R}^n} \, \mathrm{d}s \right|$$
$$\leq \|y'_r - y_r\|_{\mathbb{R}^n} + M|r' - r| \leq \varepsilon(1 + M) < 1.$$

From definition of $t_2$ we deduce that $\|y(t, g, \theta') - y(t, g, \theta)\|_{\mathbb{R}^n} \leq 1$ for all $t \in [r', t_2]$. Therefore, since moreover $\|\lambda' - \lambda\|_{\mathbb{R}^d} \leq \varepsilon < 1$ and $\|u'\|_{\mathrm{L}^\infty} \leq R$, we deduce that $(y(t, g, \theta'), \lambda', u'(t), t) \in K(g, \theta, R)$ for almost every $t \in [r', t_2]$. Hence, from integral representations, we get that

$$\|y(t, g, \theta') - y(t, g, \theta)\|_{\mathbb{R}^n}$$
$$\leq \|y'_r - y_r\|_{\mathbb{R}^n} + M|r' - r| + M \int_{r'}^{t} \|y(s, g, \theta') - y(s, g, \theta)\|_{\mathbb{R}^n} + \|\lambda' - \lambda\|_{\mathbb{R}^d} + \|u'(s) - u(s)\|_{\mathbb{R}^m} \, \mathrm{d}s$$
$$\leq \|y'_r - y_r\|_{\mathbb{R}^n} + M|r' - r| + M \int_{r'}^{t} \|y(s, g, \theta') - y(s, g, \theta)\|_{\mathbb{R}^n} \, \mathrm{d}s + MT\|\lambda' - \lambda\|_{\mathbb{R}^d} + M\|u' - u\|_{\mathrm{L}^1},$$

for all $t \in [r', t_2]$. From the Grönwall lemma we obtain that

$$\|y(t, g, \theta') - y(t, g, \theta)\|_{\mathbb{R}^n} \leq (\|y'_r - y_r\|_{\mathbb{R}^n} + M|r' - r| + MT\|\lambda' - \lambda\|_{\mathbb{R}^d} + M\|u' - u\|_{\mathrm{L}^1})e^{MT}$$
$$\leq \varepsilon(1 + M(2 + T))e^{MT} < 1,$$

for all $t \in [r', t_2]$, which raises a contradiction at $t = t_2$. Thus we have proved that $\mathcal{I}_1 \cup \mathcal{I}_2 = \emptyset$ which gives $\theta' \in \mathrm{Glob}(g)$ but also $(y(t, g, \theta'), \lambda', u'(t), t) \in K(g, \theta, R)$ for almost every $t \in [0, T]$. Hence the proofs of the first two parts of Lemma 4.1 are complete. Now let us prove the last part. To this aim let $\theta' = (\lambda', u', r', y'_r)$, $\theta'' = (\lambda'', u'', r'', y''_r) \in \mathcal{N}(g, \theta, R, \varepsilon)$. From integral representations it holds that

$$y(t, g, \theta'') - y(t, g, \theta') = (y''_r - y'_r) + \int_{r''}^{t} g(y(s, g, \theta''), \lambda'', u''(s), s) - g(y(s, g, \theta'), \lambda', u'(s), s) \, \mathrm{d}s$$
$$- \int_{r'}^{r''} g(y(s, g, \theta'), \lambda', u'(s), s) \, \mathrm{d}s,$$

for all $t \in [0, T]$. Using similar arguments than before (in particular using the Grönwall lemma), we get that

$$\|y(t, g, \theta'') - y(t, g, \theta')\|_{\mathbb{R}^n} \leq (\|y_{r''} - y_{r'}\|_{\mathbb{R}^n} + M|r'' - r'| + MT\|\lambda'' - \lambda'\|_{\mathbb{R}^d} + M\|u'' - u'\|_{\mathrm{L}^1})e^{MT},$$

for all $t \in [0, T]$, which concludes the proof of the last part of Lemma 4.1. □

In the next proposition we state a differentiability result for the trajectory $y(\cdot, g, \theta)$ with respect to perturbations of the quadruplet $\theta \in \mathrm{Glob}(g)$. As explained in the overview of the proof of Theorem 2.1 developed in Section 2.4, this proposition will be useful in the next Section 5 to construct perturbed trajectories of the hybrid control system (CS) which visit exactly (and in the same order) the same regions than a given nominal trajectory.

**Proposition 4.1.** Consider the perturbation of a quadruplet $\theta = (\lambda, u, r, y_r) \in \mathrm{Glob}(g)$ given by $\tilde{\theta}(\alpha) := (\tilde{\lambda}(\alpha), \tilde{u}(\alpha), \tilde{r}(\alpha), \tilde{y_r}(\alpha))$ for all $\alpha \in [0,1]$ where:

- $\tilde{\lambda} : [0,1] \to \mathbb{R}^d$ satisfies $\tilde{\lambda}(0) = \lambda$ and is differentiable at 0 with derivative denoted by $\tilde{\lambda}'(0)$.
- either $\tilde{u} : [0,1] \to \mathrm{L}^\infty([0,T], \mathbb{R}^m)$ is given by $\tilde{u}(\alpha) := u$ for all $\alpha \in [0,1]$ (no perturbation of the control), either $\tilde{u} : [0,1] \to \mathrm{L}^\infty([0,T], \mathbb{R}^m)$ is the needle-like perturbation of $u$ given by

$$\tilde{u}(\alpha)(t) := \begin{cases} v & \text{if } t \in [\tau - \alpha, \tau), \\ u(t) & \text{if } t \notin [\tau - \alpha, \tau), \end{cases} \tag{4.1}$$

  for almost every $t \in [0,T]$ and all $\alpha \in [0,1]$, where $v \in \mathbb{R}^m$ and $\tau \in (0,T]$ is a Lebesgue point of the map $g(y(\cdot, g, \theta), \lambda, u(\cdot), \cdot)$.
- either $\tilde{r} : [0,1] \to [0,T]$ is constantly equal to $r$ (no perturbation of the initial time), either $\tilde{r} : [0,1] \to [0,T]$ satisfies $\tilde{r}(0) = r$ and is differentiable at 0 with derivative denoted by $\tilde{r}'(0)$ (in that second context, assume that $r \in [0,T)$ is a Lebesgue point of the map $g(y(\cdot, g, \theta), \lambda, u(\cdot), \cdot)$ and, in case of needle-like perturbation of the control, assume furthermore that $r \neq \tau$).
- $\tilde{y_r} : [0,1] \to \mathbb{R}^n$ satisfies $\tilde{y_r}(0) = y_r$ and is differentiable at 0 with derivative denoted by $\tilde{y_r}'(0)$.

Then:

(i) There exists $0 < \overline{\alpha} \leq 1$ such that $\tilde{\theta}(\alpha) \in \mathrm{Glob}(g)$ for all $\alpha \in [0, \overline{\alpha}]$.

(ii) The perturbed trajectory $y(\cdot, g, \tilde{\theta}(\alpha))$ uniformly converges to $y(\cdot, g, \theta)$ over $[0,T]$ when $\alpha \to 0$.

(iii) The map

$$\begin{array}{rccl} \mathcal{P} : & [0, \overline{\alpha}] & \to & \mathrm{C}([\varsigma, T], \mathbb{R}^n) \\ & \alpha & \mapsto & y(\cdot, g, \tilde{\theta}(\alpha)), \end{array}$$

with $\varsigma := \tau$ in case of needle-like perturbation of the control and $\varsigma := 0$ otherwise, is differentiable at 0 and its derivative is equal to $w_{\tilde{\lambda}} + w_{\tilde{u}} + w_{(\tilde{r}, \tilde{y_r})}$, where $w_{\tilde{\lambda}}$, $w_{\tilde{u}}$ and $w_{(\tilde{r}, \tilde{y_r})}$ are the three variation vectors respectively defined as the unique maximal solutions (which are global) to the three linearized Cauchy problems given by

$$\begin{cases} \dot{w}(t) = \nabla_x g(y(t, g, \theta), \lambda, u(t), t) w(t) + \nabla_\lambda g(y(t, g, \theta), \lambda, u(t), t) \tilde{\lambda}'(0), & \text{a.e. } t \in [0,T], \\ w(r) = 0_{\mathbb{R}^n}, \end{cases}$$

$$\begin{cases} \dot{w}(t) = \nabla_x g(y(t, g, \theta), \lambda, u(t), t) w(t), & \text{a.e. } t \in [0,T], \\ w(\tau) = g(y(\tau, g, \theta), \lambda, v, \tau) - g(y(\tau, g, \theta), \lambda, u(\tau), \tau), \end{cases}$$

$$\begin{cases} \dot{w}(t) = \nabla_x g(y(t, g, \theta), \lambda, u(t), t) w(t), & \text{a.e. } t \in [0,T], \\ w(r) = \tilde{y_r}'(0) - \tilde{r}'(0) g(y(r, g, \theta), \lambda, u(r), r). \end{cases}$$

(iv) If furthermore the three functions $\tilde{\lambda}$, $\tilde{y_r}$ and $\tilde{r}$ are assumed to be continuous over $[0,1]$, then the map $(\alpha, t) \in [0, \overline{\alpha}] \times [0,T] \mapsto y(t, g, \tilde{\theta}(\alpha)) \in \mathbb{R}^n$ is continuous.

*Proof.* This proof is dedicated to the case of a needle-like perturbation of the control and of a perturbation of the initial time (the other cases are similar and simpler). Let $R \geq \|u\|_{\mathrm{L}^\infty} + \|v\|_{\mathbb{R}^m}$. As in the proof of Lemma 4.1, we denote by $M := M(g, \theta, R)$. Consider $\varepsilon > 0$ provided in Lemma 4.1. It is clear that $\tilde{\theta}(\alpha) \in \mathcal{N}(g, \theta, R, \varepsilon)$ for sufficiently small $\alpha > 0$. As a consequence, from Lemma 4.1, there exists $0 < \overline{\alpha} \leq 1$ such that $\tilde{\theta}(\alpha) \in \mathrm{Glob}(g)$ for all $\alpha \in [0, \overline{\alpha}]$ which concludes the proof of the first item. The second and fourth items are trivial consequences of the Lipschitz continuity provided in Lemma 4.1. Now our aim is to prove the third item. To this aim let us introduce

$$\chi^\alpha(t) := \frac{y(t, g, \tilde{\theta}(\alpha)) - y(t, g, \theta)}{\alpha} - w_{\tilde{\lambda}}(t) - w_{\tilde{u}}(t) - w_{(\tilde{r}, \tilde{y_r})}(t),$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$. Our aim is to prove that $\chi^\alpha$ uniformly converges to zero over $[\tau,T]$ when $\alpha \to 0$. To this aim we write $\chi^\alpha = \chi_1^\alpha + \chi_2^\alpha + \chi_3^\alpha$ where

$$\chi_1^\alpha(t) := \frac{\tilde{y}^\alpha(t) - \tilde{y}_1^\alpha(t)}{\alpha} - w_{\tilde{\lambda}}(t), \quad \chi_2^\alpha(t) := \frac{\tilde{y}_2^\alpha(t) - y(t)}{\alpha} - w_{\tilde{u}}(t), \quad \chi_3^\alpha(t) := \frac{\tilde{y}_1^\alpha(t) - \tilde{y}_2^\alpha(t)}{\alpha} - w_{(\tilde{r},\tilde{y}_r)}(t),$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$, where we use the notations

$$\tilde{y}^\alpha(t) := y(t,g,\tilde{\theta}(\alpha)), \quad \tilde{y}_1^\alpha(t) := y(t,g,\tilde{\theta}_1(\alpha)), \quad \tilde{y}_2^\alpha(t) := y(t,g,\tilde{\theta}_2(\alpha)), \quad y(t) := y(t,g,\theta),$$

and

$$\tilde{\theta}_1(\alpha) := (\lambda, \tilde{u}(\alpha), \tilde{r}(\alpha), \tilde{y}_r(\alpha)) \in \mathcal{N}(g,\theta,R,\varepsilon) \quad \text{and} \quad \tilde{\theta}_2(\alpha) := (\lambda, \tilde{u}(\alpha), r, y_r) \in \mathcal{N}(g,\theta,R,\varepsilon),$$

for all $t \in [0,T]$ and all $\alpha \in [0,\overline{\alpha}]$. From Lemma 4.1, for almost every $t \in [0,T]$ and all $\alpha \in [0,\overline{\alpha}]$, the five elements

$$(y(t),\lambda,u(t),t), \ (\tilde{y}^\alpha(t),\tilde{\lambda}(\alpha),u(\alpha)(t),t), \ (\tilde{y}^\alpha(t),\lambda,u(\alpha)(t),t), \ (\tilde{y}_1^\alpha(t),\lambda,\tilde{u}(\alpha)(t),t), \ (\tilde{y}_2^\alpha(t),\lambda,\tilde{u}(\alpha)(t),t),$$

belong to $K(g,\theta,R)$, but also their convex combinations. Also note that $\tilde{y}^\alpha$, $\tilde{y}_1^\alpha$ and $\tilde{y}_2^\alpha$ uniformly converge to $y$ over $[0,T]$ when $\alpha \to 0$ from the Lipschitz continuity provided in Lemma 4.1.

In what follows, as in the proof of Lemma 4.1, we will use integral representations and the Grönwall lemma to prove that $\chi_1^\alpha$, $\chi_2^\alpha$ and $\chi_3^\alpha$ uniformly converge to zero over $[\tau,T]$ when $\alpha \to 0$. To reduce the notation in integrands, we will use the notation $\rho(\cdot) := (y(\cdot,g,\theta),\lambda,u(\cdot),\cdot)$.

*Step 1:* Let us prove that $\chi_1^\alpha$ uniformly converges to zero over $[0,T]$ when $\alpha \to 0$. From integral representations it holds that

$$\chi_1^\alpha(t) = \chi_1^\alpha(r) + \int_r^t \frac{g(\tilde{y}^\alpha(s),\tilde{\lambda}(\alpha),\tilde{u}(\alpha)(s),s) - g(\tilde{y}^\alpha(s),\lambda,\tilde{u}(\alpha)(s),s)}{\alpha} - \nabla_\lambda g(\rho(s))\tilde{\lambda}'(0)\,\mathrm{d}s$$
$$+ \int_r^t \frac{g(\tilde{y}^\alpha(s),\lambda,\tilde{u}(\alpha)(s),s) - g(\tilde{y}_1^\alpha(s),\lambda,\tilde{u}(\alpha)(s),s)}{\alpha} - \nabla_x g(\rho(s)) w_{\tilde{\lambda}}(s)\,\mathrm{d}s,$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$. Using Taylor expansions with integral rest, we obtain that

$$\|\chi_1^\alpha(t)\|_{\mathbb{R}^n} \leq \|\chi_1^\alpha(r)\|_{\mathbb{R}^n} + \underbrace{\int_0^T \int_0^1 \left\|\nabla_\lambda g(\tilde{y}^\alpha(s),\lambda+\eta(\tilde{\lambda}(\alpha)-\lambda),\tilde{u}(\alpha)(s),s)\right\|_{\mathbb{R}^{n\times d}} \left|\frac{\tilde{\lambda}(\alpha)-\lambda}{\alpha} - \tilde{\lambda}'(0)\right| \mathrm{d}\eta\,\mathrm{d}s}_{\Gamma_1(\alpha)}$$
$$+ \underbrace{\int_0^T \int_0^1 \left\|\nabla_\lambda g(\tilde{y}^\alpha(s),\lambda+\eta(\tilde{\lambda}(\alpha)-\lambda),\tilde{u}(\alpha)(s),s) - \nabla_\lambda g(\rho(s))\right\|_{\mathbb{R}^{n\times d}} |\tilde{\lambda}'(0)|\,\mathrm{d}\eta\,\mathrm{d}s}_{\Gamma_2(\alpha)}$$
$$+ \left|\int_r^t \int_0^1 \|\nabla_x g(\tilde{y}_1^\alpha(s)+\eta(\tilde{y}^\alpha(s)-\tilde{y}_1^\alpha(s)),\lambda,\tilde{u}(\alpha)(s),s)\|_{\mathbb{R}^{n\times n}} \left\| \underbrace{\frac{\tilde{y}^\alpha(s)-\tilde{y}_1^\alpha(s)}{\alpha} - w_{\tilde{\lambda}}(s)}_{\chi_1^\alpha(s)} \right\|_{\mathbb{R}^n} \mathrm{d}\eta\,\mathrm{d}s\right|$$
$$+ \underbrace{\int_0^T \int_0^1 \|\nabla_x g(\tilde{y}_1^\alpha(s)+\eta(\tilde{y}^\alpha(s)-\tilde{y}_1^\alpha(s)),\lambda,\tilde{u}(\alpha)(s),s) - \nabla_x g(\rho(s))\|_{\mathbb{R}^{n\times n}} \|w_{\tilde{\lambda}}(s)\|_{\mathbb{R}^n}\,\mathrm{d}\eta\,\mathrm{d}s}_{\Gamma_3(\alpha)},$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$. The Grönwall lemma leads to

$$\|\chi_1^\alpha(t)\|_{\mathbb{R}^n} \leq \Big(\|\chi_1^\alpha(r)\|_{\mathbb{R}^n} + \Gamma_1(\alpha) + \Gamma_2(\alpha) + \Gamma_3(\alpha)\Big)e^{MT},$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$. From the differentiability of $\tilde{\lambda}(\cdot)$ at 0, the boundedness of $\|\nabla_x g\|_{\mathbb{R}^{n\times n}}$ and $\|\nabla_\lambda g\|_{\mathbb{R}^{n\times d}}$ over $K(g,\theta,R)$, the uniform convergences of $\tilde{y}^\alpha$ and $\tilde{y}_1^\alpha$ to $y$ over $[0,T]$ when $\alpha \to 0$ and from the dominated convergence theorem, we prove that $\Gamma_1(\alpha), \Gamma_2(\alpha)$ and $\Gamma_3(\alpha)$ converge to zero when $\alpha \to 0$. It remains to prove that $\|\chi_1^\alpha(r)\|_{\mathbb{R}^n}$ converges to zero when $\alpha \to 0$. From integral representations it holds that

$$\chi_1^\alpha(r) = \frac{1}{\alpha} \int_{\tilde{r}(\alpha)}^{r} g(\tilde{y}^\alpha(s), \tilde{\lambda}(\alpha), \tilde{u}(\alpha)(s), s) - g(\tilde{y}_1^\alpha(s), \lambda, \tilde{u}(\alpha)(s), s) \, \mathrm{d}s,$$

for all $\alpha \in (0, \overline{\alpha}]$, and, using similar arguments than before, we obtain that

$$\begin{aligned} \|\chi_1^\alpha(r)\|_{\mathbb{R}^n} \leq \frac{M}{\alpha} \left| \int_{\tilde{r}(\alpha)}^{r} \|\tilde{y}^\alpha(s) - \tilde{y}_1^\alpha(s)\|_{\mathbb{R}^n} + \|\tilde{\lambda}(\alpha) - \lambda\|_{\mathbb{R}^d} \, \mathrm{d}s \right| \\ \leq M \left| \frac{\tilde{r}(\alpha) - r}{\alpha} \right| (\|\tilde{y}^\alpha - \tilde{y}_1^\alpha\|_{\mathrm{C}} + \|\tilde{\lambda}(\alpha) - \lambda\|_{\mathbb{R}^d}), \end{aligned}$$

for all $\alpha \in (0, \overline{\alpha}]$, which concludes the proof of Step 1 from the differentiability of $\tilde{r}(\cdot)$ and the continuity of $\tilde{\lambda}(\cdot)$ at 0 and from the uniform convergences of $\tilde{y}^\alpha$ and $\tilde{y}_1^\alpha$ to $y$ over $[0, T]$ when $\alpha \to 0$. The proof of Step 1 is complete.

*Step 2:* Let us prove that $\chi_2^\alpha$ uniformly converges to zero over $[\tau, T]$ when $\alpha \to 0$. From integral representations it holds that

$$\chi_2^\alpha(t) = \chi_2^\alpha(\tau) + \int_\tau^t \frac{g(\tilde{y}_2^\alpha(s), \lambda, u(s), s) - g(y(s), \lambda, u(s), s)}{\alpha} - \nabla_x g(\rho(s)) w_{\tilde{u}}(s) \, \mathrm{d}s,$$

for all $t \in [\tau, T]$ and all $\alpha \in (0, \overline{\alpha}]$. Using a Taylor expansion with integral rest, we obtain that

$$\begin{aligned} \|\chi_2^\alpha(t)\|_{\mathbb{R}^n} &\leq \|\chi_2^\alpha(\tau)\|_{\mathbb{R}^n} \\ &+ \int_\tau^t \int_0^1 \|\nabla_x g(y(s) + \eta(\tilde{y}_2^\alpha(s) - y(s)), \lambda, u(s), s)\|_{\mathbb{R}^{n\times n}} \left\| \underbrace{\frac{\tilde{y}_2^\alpha(s) - y(s)}{\alpha} - w_{\tilde{u}}(s)}_{\chi_2^\alpha(s)} \right\|_{\mathbb{R}^n} \mathrm{d}\eta \, \mathrm{d}s \\ &+ \underbrace{\int_\tau^T \int_0^1 \|\nabla_x g(y(s) + \eta(\tilde{y}_2^\alpha(s) - y(s)), \lambda, u(s), s) - \nabla_x g(\rho(s))\|_{\mathbb{R}^{n\times n}} \|w_{\tilde{u}}(s)\|_{\mathbb{R}^n} \, \mathrm{d}\eta \, \mathrm{d}s,}_{\Gamma_4(\alpha)} \end{aligned}$$

for all $t \in [\tau, T]$ and all $\alpha \in (0, \overline{\alpha}]$. The Grönwall lemma leads to

$$\|\chi_2^\alpha(t)\|_{\mathbb{R}^n} \leq (\|\chi_2^\alpha(\tau)\|_{\mathbb{R}^n} + \Gamma_4(\alpha)) \, e^{MT},$$

for all $t \in [\tau, T]$ and all $\alpha \in (0, \overline{\alpha}]$. From the uniform convergence of $\tilde{y}_2^\alpha$ to $y$ over $[0, T]$ when $\alpha \to 0$ and from the dominated convergence theorem, we prove that $\Gamma_4(\alpha)$ converges to zero when $\alpha \to 0$. It remains to prove that $\|\chi_2^\alpha(\tau)\|_{\mathbb{R}^n}$ converges to zero when $\alpha \to 0$. From integral representations it holds that

$$\chi_2^\alpha(\tau) = \int_{\tau-\alpha}^{\tau} \frac{g(\tilde{y}_2^\alpha(s), \lambda, v, s) - g(y(s), \lambda, v, s)}{\alpha} \, \mathrm{d}s + \int_{\tau-\alpha}^{\tau} \frac{g(y(s), \lambda, v, s) - g(y(s), \lambda, u(s), s)}{\alpha} \, \mathrm{d}s - w_{\tilde{u}}(\tau).$$

for all $\alpha \in (0, \overline{\alpha}]$. From the uniform convergence of $\tilde{y}_2^\alpha$ to $y$ over $[0, T]$ when $\alpha \to 0$, one can easily prove that the first term tends to $0_{\mathbb{R}^n}$ when $\alpha \to 0$. Finally, since $\tau$ is a Lebesgue point of the map $g(y(\cdot), \lambda, u(\cdot), \cdot)$ and from the value of $w_{\tilde{u}}(\tau)$, the second term tends to $0_{\mathbb{R}^n}$ when $\alpha \to 0$. The proof of Step 2 is complete.

*Step 3:* Let us prove that $\chi_3^\alpha$ uniformly converges to zero over $[0, T]$ when $\alpha \to 0$. From integral representations it holds that

$$\chi_3^\alpha(t) = \chi_3^\alpha(r) + \int_r^t \frac{g(\tilde{y}_1^\alpha(s), \lambda, \tilde{u}(\alpha)(s), s) - g(\tilde{y}_2^\alpha(s), \lambda, \tilde{u}(\alpha)(s), s)}{\alpha} - \nabla_x g(\rho(s)) w_{(\tilde{r}, \tilde{y_r})}(s) \, \mathrm{d}s,$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$. Using a Taylor expansion with integral rest, we obtain that

$$\begin{aligned}\|\chi_3^\alpha(t)\|_{\mathbb{R}^n} &\leq \|\chi_3^\alpha(r)\|_{\mathbb{R}^n} \\ &+ \left| \int_r^t \int_0^1 \|\nabla_x g(\tilde{y}_2^\alpha(s) + \eta(\tilde{y}_1^\alpha(s) - \tilde{y}_2^\alpha(s)), \lambda, \tilde{u}(\alpha)(s), s)\|_{\mathbb{R}^{n\times n}} \Bigg\| \underbrace{\frac{\tilde{y}_1^\alpha(s) - \tilde{y}_2^\alpha(s)}{\alpha} - w_{(\tilde{r},\tilde{y_r})}(s)}_{\chi_3^\alpha(s)} \Bigg\|_{\mathbb{R}^n} \mathrm{d}\eta \, \mathrm{d}s \right| \\ &+ \underbrace{\int_0^T \int_0^1 \|\nabla_x g(\tilde{y}_2^\alpha(s) + \eta(\tilde{y}_1^\alpha(s) - \tilde{y}_2^\alpha(s)), \lambda, \tilde{u}(\alpha)(s), s) - \nabla_x g(\rho(s))\|_{\mathbb{R}^{n\times n}} \|w_{(\tilde{r},\tilde{y_r})}(s)\|_{\mathbb{R}^n} \, \mathrm{d}\eta \, \mathrm{d}s}_{\Gamma_5(\alpha)},\end{aligned}$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$. The Grönwall lemma leads to

$$\|\chi_3^\alpha(t)\|_{\mathbb{R}^n} \leq (\|\chi_3^\alpha(r)\|_{\mathbb{R}^n} + \Gamma_5(\alpha))\, e^{MT},$$

for all $t \in [0,T]$ and all $\alpha \in (0,\overline{\alpha}]$. From the uniform convergences of $\tilde{y}_1^\alpha$ and $\tilde{y}_2^\alpha$ to $y$ over $[0,T]$ when $\alpha \to 0$ and from the dominated convergence theorem, we prove that $\Gamma_5(\alpha)$ converges to zero when $\alpha \to 0$. It remains to prove that $\|\chi_3^\alpha(r)\|_{\mathbb{R}^n}$ converges to zero when $\alpha \to 0$. From integral representations it holds that

$$\chi_3^\alpha(r) = \left( \frac{\tilde{y}_r(\alpha) - y_r}{\alpha} - \tilde{y}_r'(0) \right) + \left( \tilde{r}'(0) g(y(r), \lambda, u(r), r) + \frac{1}{\alpha} \int_{\tilde{r}(\alpha)}^r g(\tilde{y}_1^\alpha(s), \tilde{\lambda}(\alpha), \tilde{u}(\alpha)(s), s) \, \mathrm{d}s \right),$$

for all $\alpha \in (0,\overline{\alpha}]$. From differentiability of $\tilde{y}_r(\cdot)$ at 0, the first term converges to $0_{\mathbb{R}^n}$ when $\alpha \to 0$. Since $r \neq \tau$ and from the continuity of $\tilde{r}(\cdot)$ at 0, we know that the second term can be rewritten as

$$\tilde{r}'(0) g(y(r), \lambda, u(r), r) + \frac{1}{\alpha} \int_{\tilde{r}(\alpha)}^r g(y(s), \lambda, u(s), s) \, \mathrm{d}s - \frac{1}{\alpha} \int_{\tilde{r}(\alpha)}^r g(y(s), \lambda, u(s), s) - g(\tilde{y}_1^\alpha(s), \tilde{\lambda}(\alpha), u(s), s) \, \mathrm{d}s,$$

for sufficiently small $\alpha > 0$. Since $r$ is a Lebesgue point of the map $g(y(\cdot,g,\theta),\lambda,u(\cdot),\cdot)$ and from the differentiability of $\tilde{r}(\cdot)$ at 0, the sum of the two first terms in the above equation converges to $0_{\mathbb{R}^n}$ when $\alpha \to 0$. Finally the norm of the last term in the above equation can be bounded by

$$\left| \frac{1}{\alpha} \int_{\tilde{r}(\alpha)}^r \|g(y(s), \lambda, u(s), s) - g(\tilde{y}_1^\alpha(s), \tilde{\lambda}(\alpha), u(s), s)\|_{\mathbb{R}^n} \, \mathrm{d}s \right| \leq M \left| \frac{\tilde{r}(\alpha) - r}{\alpha} \right| (\|y - \tilde{y}_1^\alpha\|_{\mathrm{C}} + \|\lambda - \tilde{\lambda}(\alpha)\|_{\mathbb{R}^d}),$$

which tends to zero when $\alpha \to 0$, thanks to the differentiability of $\tilde{r}(\cdot)$ at 0, to the continuity of $\tilde{\lambda}(\alpha)$ at 0 and from the uniform convergence of $\tilde{y}_1^\alpha$ to $y$ over $[0,T]$ when $\alpha \to 0$. The proof of Step 3 is complete. This completes the proof of Proposition 4.1. □

# 5 Preliminaries: sensitivity analysis in the hybrid context

As explained in the overview of the proof of Theorem 2.1 developed in Section 2.4, a sensitivity analysis of the hybrid control system (CS) has to be performed to construct perturbed trajectories which visit exactly (and in the same order) the same regions than a given nominal trajectory. This is exactly the aim of the present section. To this aim we will use the results stated in the previous Section 4, but we will also invoke at several occasions the following conic implicit function theorem to prove the existence of perturbed crossing times (see Section 2.4 for more details).

**Lemma 5.1** (A conic implicit function theorem)**.** Let $\overline{\alpha} > 0$, $t^c \in (0,T)$ and $\delta > 0$. Consider a continuous map

$$\begin{array}{rccl} \mathcal{G}: & [0,\overline{\alpha}] \times [t^c - \delta, t^c + \delta] & \to & \mathbb{R} \\ & (\alpha, t) & \mapsto & \mathcal{G}(\alpha, t), \end{array}$$

satisfying $\mathcal{G}(0,t^c)=0$, such that $\nabla_\alpha \mathcal{G}(0,t^c)$ exists and such that $\nabla_t \mathcal{G}$ exists and is continuous over $[0,\overline{\alpha}]\times[t^c-\delta,t^c+\delta]$ with $\nabla_t\mathcal{G}(0,t^c)\neq 0$. Then there exist $0<\overline{\beta}\leq\overline{\alpha}$ and an implicit function $\tilde{t}\in \mathrm{C}([0,\overline{\beta}],[t^c-\delta,t^c+\delta])$, satisfying $\tilde{t}(0)=t^c$ and $\mathcal{G}(\alpha,\tilde{t}(\alpha))=0$ for all $\alpha\in[0,\overline{\beta}]$, that is differentiable at 0 with derivative $\tilde{t}'(0)=-\frac{\nabla_\alpha\mathcal{G}(0,t^c)}{\nabla_t\mathcal{G}(0,t^c)}$.

*Proof.* Consider the extension $\mathcal{G}_0:[-\overline{\alpha},\overline{\alpha}]\times[t^c-\delta,t^c+\delta]\to\mathbb{R}$ defined by

$$\forall(\alpha,t)\in[-\overline{\alpha},\overline{\alpha}]\times[t^c-\delta,t^c+\delta],\quad \mathcal{G}_0(\alpha,t):=\begin{cases}\mathcal{G}(\alpha,t) & \text{if}\quad \alpha\in[0,\overline{\alpha}],\\ 2\mathcal{G}(0,t)-\mathcal{G}(-\alpha,t) & \text{if}\quad \alpha\in[-\overline{\alpha},0].\end{cases}$$

From the assumptions of Lemma 5.1, one can easily derive that $\mathcal{G}_0(0,t^c)=0$, $\nabla_\alpha\mathcal{G}_0(0,t^c)$ exists and $\nabla_t\mathcal{G}_0$ exists and is continuous over $[-\overline{\alpha},\overline{\alpha}]\times[t^c-\delta,t^c+\delta]$ with $\nabla_t\mathcal{G}_0(0,t^c)\neq 0$. Using a classical version of the implicit function theorem (see [48, Theorem 9.3] and [43, Theorem E]), there exist $0<\overline{\beta}\leq\overline{\alpha}$ and an implicit function $\tilde{t}\in\mathrm{C}([-\overline{\beta},\overline{\beta}],[t^c-\delta,t^c+\delta])$, satisfying $\tilde{t}(0)=t^c$ and $\mathcal{G}_0(\alpha,\tilde{t}(\alpha))=0$ for all $\alpha\in[-\overline{\beta},\overline{\beta}]$, differentiable at 0 with derivative $\tilde{t}'(0)=-\frac{\nabla_\alpha\mathcal{G}_0(0,t^c)}{\nabla_t\mathcal{G}_0(0,t^c)}$. To conclude the proof, one has just to consider the restriction of the function $\tilde{t}$ to the interval $[0,\overline{\beta}]$ and to use the facts that $\nabla_\alpha\mathcal{G}_0(0,t^c)=\nabla_\alpha\mathcal{G}(0,t^c)$ and $\nabla_t\mathcal{G}_0(0,t^c)=\nabla_t\mathcal{G}(0,t^c)$. □

## 5.1 A regular solution to (CS) and auxiliary non-hybrid trajectories

Throughout Section 5 we fix $(x,\lambda,u)\in\mathrm{AC}([0,T],\mathbb{R}^n)\times\mathrm{PC}([0,T],\mathbb{R}^d)\times\mathrm{L}^\infty([0,T],\mathbb{R}^m)$ being a regular solution to (CS) and we will use the notations introduced in Definitions 2.1 and 2.2. For all $k\in\{1,...,N\}$, we introduce, following the notations from Section 4, the *auxiliary non-hybrid trajectory* $\tilde{z}_k:=y(\cdot,f_k,\theta_k)$ associated with the quadruplet $\theta_k:=(\lambda_k,\tilde{u}_k,t^c_{k-1},x(t^c_{k-1}))$, where the *auxiliary control* $\tilde{u}_k\in\mathrm{L}^\infty([0,T],\mathbb{R}^m)$ is defined by

$$\tilde{u}_k(t):=\begin{cases}u^+(t^c_{k-1}), & \text{for a.e. } t\in(t^c_0,t^c_{k-1}),\\ u(t), & \text{for a.e. } t\in(t^c_{k-1},t^c_k),\\ u^-(t^c_k), & \text{for a.e. } t\in(t^c_k,t^c_N).\end{cases}$$

We refer to Figure 6 in Section 2.4. Note that $\tilde{z}_k=x$ over $[t^c_{k-1},t^c_k]$ for all $k\in\{1,...,N\}$ (see Figure 7 in Section 2.4). As a consequence, from Cauchy-Lipschitz theorem and up to reducing $\delta>0$ provided in Definition 2.2, we will consider in the sequel that $[t^c_{k-1}-\delta,t^c_k+\delta]\cap[0,T]\subset I(f_k,\theta_k)$ for all $k\in\{1,...,N\}$. Furthermore, up to reducing $\delta>0$ again, we will consider that $\tilde{z}_k(t)\in\overline{\mathrm{B}}_{\mathbb{R}^n}(x(t^c_{k-1}),\frac{\nu}{2})$ for all $t\in[t^c_{k-1}-\delta,t^c_{k-1}+\delta]$ and all $k\in\{2,\dots,N\}$, and that $\tilde{z}_k(t)\in\overline{\mathrm{B}}_{\mathbb{R}^n}(x(t^c_k),\frac{\nu}{2})$ for all $t\in[t^c_k-\delta,t^c_k+\delta]$ and all $k\in\{1,\dots,N-1\}$.

Furthermore, from $(\mathcal{A}1)$ and for any $k\in\{1,\dots,N-1\}$, note that $\tilde{u}_k$ is continuous over $[t^c_k-\delta,T]$ and thus $\tilde{z}_k$ is of class $\mathrm{C}^1$ over $[t^c_k-\delta,t^c_k+\delta]$ with $\dot{z}_k(t)=f_k(\tilde{z}_k(t),\lambda_k,\tilde{u}_k(t),t)$ for all $t\in[t^c_k-\delta,t^c_k+\delta]$. In particular $t^c_k$ is a Lebesgue point of the map $f_k(\tilde{z}_k(\cdot),\lambda_k,\tilde{u}_k(\cdot),\cdot)$ and it holds that $\dot{z}_k(t^c_k)=(f_k)^-(t^c_k)$. Similarly, from $(\mathcal{A}1)$ and for any $k\in\{1,\dots,N-1\}$, note that $\tilde{u}_{k+1}$ is continuous over $[0,t^c_k+\delta]$ and thus $\tilde{z}_{k+1}$ is of class $\mathrm{C}^1$ over $[t^c_k-\delta,t^c_k+\delta]$ with $\dot{z}_{k+1}(t)=f_{k+1}(\tilde{z}_{k+1}(t),\lambda_{k+1},\tilde{u}_{k+1}(t),t)$ for all $t\in[t^c_k-\delta,t^c_k+\delta]$. In particular $t^c_k$ is a Lebesgue point of the map $f_{k+1}(\tilde{z}_{k+1}(\cdot),\lambda_{k+1},\tilde{u}_{k+1}(\cdot),\cdot)$ and it holds that $\dot{z}_{k+1}(t^c_k)=(f_{k+1})^+(t^c_k)$.

## 5.2 Convex perturbation of the regionally switching parameter

Consider the framework of Section 5.1. This entire Section 5.2 is dedicated to the proof of the next proposition which states a differentiability result at time $t=T$ for the trajectory $x$ with respect to a convex perturbation of the regionally switching parameter $\lambda$.

**Proposition 5.1.** Consider the framework of Section 5.1. Let $k\in\{1,...,N\}$ and let $\overline{\lambda}_k\in\mathbb{R}^d$. Then there exists $0<\overline{\alpha}\leq 1$ such that, for all $\alpha\in(0,\overline{\alpha}]$, there exists a perturbed solution $(x^\alpha,\lambda^\alpha,u^\alpha)\in\mathrm{AC}([0,T],\mathbb{R}^n)\times\mathrm{PC}([0,T],\mathbb{R}^d)\times\mathrm{L}^\infty([0,T],\mathbb{R}^m)$ to (CS) such that:

(i) The corresponding perturbed partition of $[0,T]$, denoted by $\{\tilde{t}_q(\alpha)\}_{q=0,\dots,N(\alpha)}$, satisfies $N(\alpha)=N$, with $\tilde{t}_q(\alpha)=t^c_q$ for all $q\in\{1,...,k-1\}$, and $\tilde{t}_q(\alpha)$ tends to $t^c_q$ when $\alpha\to 0$ for all $q\in\{k,...,N-1\}$.

(ii) The perturbed trajectory $x^\alpha$ follows the same regions than $x$, that is, $x^\alpha$ satisfies

$$x^\alpha(t) \in E_q(t) \text{ for all } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{1, \dots, N\},$$

with $x^\alpha(0) = x_{\text{init}} \in E_1(0)$ and $x^\alpha(T) \in E_N(T)$. Moreover $x^\alpha$ uniformly converges to $x$ over $[0, T]$ when $\alpha \to 0$.

(iii) The perturbed regionally switching parameter $\lambda^\alpha$ is given by the convex perturbation

$$\lambda^\alpha(t) = \begin{cases} \lambda_k + \alpha(\overline{\lambda}_k - \lambda_k) & \text{for a.e. } t \in (\tilde{t}_{k-1}(\alpha), \tilde{t}_k(\alpha)), \\ \lambda_q & \text{for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{1, \dots, N\} \backslash \{k\}. \end{cases}$$

(iv) The perturbed control $u^\alpha$ is given by

$$u^\alpha(t) = \tilde{u}_q(t) \text{ for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{1, \dots, N\},$$

where $\tilde{u}_q$ stands for the auxiliary control defined in Section 5.1 for all $q \in \{1, \dots, N\}$.

(v) The limit

$$\lim_{\alpha \to 0} \frac{x^\alpha(T) - x(T)}{\alpha} = w(T),$$

holds true, where

$$w(t) := \begin{cases} w_q(t) & \text{for all } t \in [t^c_{q-1}, t^c_q) \text{ and all } q \in \{k, ..., N-1\}, \\ w_N(t) & \text{for all } t \in [t^c_{N-1}, t^c_N], \end{cases}$$

where $w_k$ is the variation vector defined as the unique maximal solution (which is global) to the linearized Cauchy problem given by

$$\begin{cases} \dot{w}(t) = \nabla_x f_k(\tilde{z}_k(t), \lambda_k, \tilde{u}_k(t), t) w(t) + \nabla_\lambda f_k(\tilde{z}_k(t), \lambda_k, \tilde{u}_k(t), t)(\overline{\lambda}_k - \lambda_k), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{a.e. } t \in [t^c_{k-1} - \delta, t^c_k + \delta] \cap [0, T], \\ w(t^c_{k-1}) = 0_{\mathbb{R}^n}, \end{cases}$$

and $w_q$ is the variation vector defined by induction as the unique maximal solution (which is global) to the linearized Cauchy problem given by

$$\begin{cases} \dot{w}(t) = \nabla_x f_q(\tilde{z}_q(t), \lambda_q, \tilde{u}_q(t), t) w(t), \quad \text{a.e. } t \in [t^c_{q-1} - \delta, t^c_q + \delta] \cap [0, T], \\ w(t^c_{q-1}) = w_{q-1}(t^c_{q-1}) + \xi_{q-1}, \end{cases}$$

for all $q \in \{k+1, ..., N\}$, where $\xi_q \in \mathbb{R}^n$ stands for the jump vector defined by

$$\xi_q := \frac{\langle \nabla_x F_q(x(t^c_q), t^c_q), w_q(t^c_q) \rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t^c_q), t^c_q), (f_q)^-(t^c_q) \rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t^c_q), t^c_q)} ((f_{q+1})^+(t^c_q) - (f_q)^-(t^c_q)),$$

for all $q \in \{k, ..., N-1\}$.

(vi) The limit

$$\lim_{\alpha \to 0} \frac{\tilde{t}_q(\alpha) - t^c_q}{\alpha} = -\frac{\langle \nabla_x F_q(x(t^c_q), t^c_q), w_q(t^c_q) \rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t^c_q), t^c_q), (f_q)^-(t^c_q) \rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t^c_q), t^c_q)},$$

holds true for all $q \in \{k, \dots, N-1\}$.

### 5.2.1 Construction of perturbed auxiliary non-hybrid trajectories

**Lemma 5.2** (Construction of perturbed auxiliary non-hybrid trajectories)**.** Consider the frameworks of Section 5.1 and Proposition 5.1. Let $k \in \{1,...,N\}$ and let $\overline{\lambda}_k \in \mathbb{R}^d$. Then there exists $0 < \overline{\alpha} \leq 1$ and, for all $q \in \{k,...,N-1\}$, there exists a function $\tilde{t}_q \in \mathrm{C}([0,\overline{\alpha}],[t_q^c - \delta, t_q^c + \delta])$ differentiable at 0 with $\tilde{t}_q(0) = t_q^c$ and

$$\tilde{t}_q'(0) = -\frac{\langle \nabla_x F_q(x(t_q^c),t_q^c), w_q(t_q^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t_q^c),t_q^c), (f_q)^-(t_q^c)\rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t_q^c),t_q^c)},$$

such that the perturbed auxiliary non-hybrid trajectories $\tilde{z}_q^\alpha := y(\cdot, f_q, \theta_q^\alpha)$ associated with the perturbed quadruplets $\theta_q^\alpha$ defined by the induction

$$\theta_q^\alpha := \begin{cases} (\lambda_k + \alpha(\overline{\lambda}_k - \lambda_k), \tilde{u}_k, t_{k-1}^c, x(t_{k-1}^c)) & \text{if} \quad q = k, \\ (\lambda_q, \tilde{u}_q, \tilde{t}_{q-1}(\alpha), \tilde{z}_{q-1}^\alpha(\tilde{t}_{q-1}(\alpha))) & \text{if} \quad q \in \{k+1,\dots,N\}, \end{cases}$$

for all $\alpha \in [0,\overline{\alpha}]$ and all $q \in \{k,\dots,N\}$, satisfy:

- for all $q \in \{k,\dots,N\}$, it holds that $[t_{q-1}^c - \delta, t_q^c + \delta] \cap [0,T] \subset I(f_q, \theta_q^\alpha)$ for all $\alpha \in [0,\overline{\alpha}]$, that $\tilde{z}_q^\alpha$ uniformly converges to $\tilde{z}_q$ over $[t_{q-1}^c - \delta, t_q^c + \delta] \cap [0,T]$ when $\alpha \to 0$, and

$$\lim_{\alpha \to 0} \frac{\tilde{z}_q^\alpha(t_q^c) - \tilde{z}_q(t_q^c)}{\alpha} = w_q(t_q^c).$$

- for all $q \in \{k,\dots,N-1\}$, it holds that $\tilde{z}_q^\alpha(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_q^c),\nu)$ for all $(\alpha,t) \in [0,\overline{\alpha}] \times [t_q^c - \delta, t_q^c + \delta]$, that $F_q(\tilde{z}_q^\alpha(\tilde{t}_q(\alpha)), \tilde{t}_q(\alpha)) = 0$ for all $\alpha \in [0,\overline{\alpha}]$, and that the map $\alpha \in [0,\overline{\alpha}] \mapsto \tilde{z}_q^\alpha(\tilde{t}_q(\alpha)) \in \mathbb{R}^n$ is continuous over $[0,\overline{\alpha}]$ and differentiable at 0 with

$$\lim_{\alpha \to 0} \frac{\tilde{z}_q^\alpha(\tilde{t}_q(\alpha)) - \tilde{z}_q(t_q^c)}{\alpha} = w_q(t_q^c) + \tilde{t}_q'(0)(f_q)^-(t_q^c).$$

*Proof.* Let us fix $k \in \{1,...,N\}$ and $\overline{\lambda}_k \in \mathbb{R}^d$. The case $k = N$ follows directly from Proposition 4.1. In the sequel we deal with the case $k \in \{1,\dots,N-1\}$ and we will proceed by induction over $q \in \{k,\dots,N\}$. Note that we will construct $0 < \overline{\alpha} \leq 1$ in the base case and that it will be reduced a finite number of times at each step of the induction.

*Base case $q = k$.* We deduce from Proposition 4.1 that there exists $0 < \overline{\alpha} \leq 1$ such that $[t_{k-1}^c - \delta, t_k^c + \delta] \cap [0,T] \subset I(f_k, \theta_k^\alpha)$ for all $\alpha \in [0,\overline{\alpha}]$, that $\tilde{z}_k^\alpha$ uniformly converges to $\tilde{z}_k$ over $[t_{k-1}^c - \delta, t_k^c + \delta] \cap [0,T]$ when $\alpha \to 0$ (as illustrated in Figure 9(a) in Section 2.4), and that the map

$$(\alpha,t) \in [0,\overline{\alpha}] \times ([t_{k-1}^c - \delta, t_k^c + \delta] \cap [0,T]) \mapsto \tilde{z}_k^\alpha(t) \in \mathbb{R}^n, \tag{5.1}$$

is continuous. Since moreover $\tilde{z}_k(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_k^c), \frac{\nu}{2})$ for all $t \in [t_k^c - \delta, t_k^c + \delta]$, up to reducing $\overline{\alpha} > 0$, we have $\tilde{z}_k^\alpha(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_k^c),\nu)$ for all $(\alpha,t) \in [0,\overline{\alpha}] \times [t_k^c - \delta, t_k^c + \delta]$. We are now in a position to define the map

$$\begin{array}{rccl} G_k : & [0,\overline{\alpha}] \times [t_k^c - \delta, t_k^c + \delta] & \to & \mathbb{R} \\ & (\alpha,t) & \mapsto & F_k(\tilde{z}_k^\alpha(t),t), \end{array}$$

where $F_k : \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_k^c),\nu) \times [t_k^c - \delta, t_k^c + \delta] \to \mathbb{R}$ is the $\mathrm{C}^1$ function provided in Definition 2.2.

Let us check that $G_k$ satisfies all the assumptions of the conic implicit function theorem (Lemma 5.1). First, $G_k$ is continuous from the continuity of the map (5.1) and $G_k(0,t_k^c) = F_k(x(t_k^c),t_k^c) = 0$. Second, for any $\alpha \in [0,\overline{\alpha}]$, since $\tilde{u}_k$ is continuous over $[t_k^c - \delta, t_k^c + \delta]$ (see Figure 6 in Section 2.4), we know that $\tilde{z}_k^\alpha$ is of class $\mathrm{C}^1$ over $[t_k^c - \delta, t_k^c + \delta]$. This implies that $\nabla_t G_k(\alpha,t)$ exists with

$$\nabla_t G_k(\alpha,t) = \langle \nabla_x F_k(\tilde{z}_k^\alpha(t),t), f_k(\tilde{z}_k^\alpha(t), \lambda_k + \alpha(\overline{\lambda}_k - \lambda_k), \tilde{u}_k(t),t)\rangle_{\mathbb{R}^n} + \nabla_t F_k(\tilde{z}_k^\alpha(t),t),$$

for all $(\alpha,t) \in [0,\overline{\alpha}] \times [t_k^c - \delta, t_k^c + \delta]$. Furthermore, from the continuity of the map (5.1), one can see that $\nabla_t G_k$ is continuous over $[0,\overline{\alpha}] \times [t_k^c - \delta, t_k^c + \delta]$ and, from $(\mathcal{A}3)$, it holds that

$$\nabla_t G_k(0,t_k^c) = \langle \nabla_x F_k(x(t_k^c),t_k^c), (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t_k^c),t_k^c) \neq 0.$$

Finally, from the third item of Proposition 4.1, we get that

$$\lim_{\alpha\to 0} \frac{\tilde{z}_k^\alpha(t_k^c) - \tilde{z}_k(t_k^c)}{\alpha} = w_k(t_k^c),$$

which implies that $\nabla_\alpha G_k(0,t_k^c)$ exists with $\nabla_\alpha G_k(0,t_k^c) = \langle \nabla_x F_k(x(t_k^c),t_k^c), w_k(t_k^c)\rangle_{\mathbb{R}^n}$.

We deduce from the conic implicit function theorem (Lemma 5.1) that, up to reducing $\overline{\alpha} > 0$ (precisely, by taking $\overline{\alpha} = \overline{\beta}$), there exists a function $\tilde{t}_k \in \mathrm{C}([0,\overline{\alpha}],[t_k^c-\delta,t_k^c+\delta])$, such that $\tilde{t}_k(0) = t_k^c$ and $F_k(\tilde{z}_k^\alpha(\tilde{t}_k(\alpha)),\tilde{t}_k(\alpha)) = 0$ for all $\alpha \in [0,\overline{\alpha}]$ (see Figure 9(a) in Section 2.4), that is differentiable at 0 with

$$\tilde{t}_k'(0) = -\frac{\langle \nabla_x F_k(x(t_k^c),t_k^c), w_k(t_k^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_k(x(t_k^c),t_k^c), (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t_k^c),t_k^c)}.$$

From the continuities of the function $\tilde{t}_k$ and of the map (5.1), we deduce that the map $\alpha \in [0,\overline{\alpha}] \mapsto \tilde{z}_k^\alpha(\tilde{t}_k(\alpha)) \in \mathbb{R}^n$ is continuous over $[0,\overline{\alpha}]$. It remains to prove that

$$\lim_{\alpha\to 0} \frac{\tilde{z}_k^\alpha(\tilde{t}_k(\alpha)) - \tilde{z}_k(t_k^c)}{\alpha} = w_k(t_k^c) + \tilde{t}_k'(0)(f_k)^-(t_k^c).$$

To this aim, using integral representations, one can write

$$\begin{aligned}\frac{\tilde{z}_k^\alpha(\tilde{t}_k(\alpha)) - \tilde{z}_k(t_k^c)}{\alpha} = \frac{\tilde{z}_k^\alpha(t_k^c) - \tilde{z}_k(t_k^c)}{\alpha} + \frac{\tilde{t}_k(\alpha)-t_k^c}{\alpha}\frac{1}{\tilde{t}_k(\alpha)-t_k^c}\int_{t_k^c}^{\tilde{t}_k(\alpha)} f_k(\tilde{z}_k(s),\lambda_k,\tilde{u}_k(s),s)\,\mathrm{d}s \\ + \frac{1}{\alpha}\int_{t_k^c}^{\tilde{t}_k(\alpha)} f_k(\tilde{z}_k^\alpha(s),\lambda_k+\alpha(\overline{\lambda}_k-\lambda_k),\tilde{u}_k(s),s) - f_k(\tilde{z}_k(s),\lambda_k,\tilde{u}_k(s),s)\,\mathrm{d}s,\end{aligned}$$

for all $\alpha \in (0,\overline{\alpha}]$. We already proved that the first term tends to $w_k(t_k^c)$ when $\alpha \to 0$. Since $t_k^c$ is a Lebesgue point of the map $f_k(\tilde{z}_k(\cdot),\lambda_k,\tilde{u}_k(\cdot),\cdot)$ and since $\tilde{t}_k$ is differentiable at 0, the second term tends to $\tilde{t}_k'(0)(f_k)^-(t_k^c)$, finally the third term tends to zero when $\alpha \to 0$, since $\tilde{z}_k^\alpha$ uniformly converges to $\tilde{z}_k$ over $[t_k^c-\delta,t_k^c+\delta]$, $f_k$ is of class $\mathrm{C}^1$ and $\tilde{t}_k$ is differentiable at 0. Hence the proof for the base case is complete.

*Inductive step.* Let $q \in \{k+1,...,N\}$ and assume that the induction hypothesis holds true for all $\ell \in \{k,...,q-1\}$. The case $q = N$ follows directly from Proposition 4.1 and from the induction hypothesis (in particular from the differentiabilities at 0 of the function $\tilde{t}_{N-1}$ and of the map $\alpha \in [0,\overline{\alpha}] \mapsto \tilde{z}_{N-1}^\alpha(\tilde{t}_{N-1}(\alpha)) \in \mathbb{R}^n$). Therefore, in the sequel, we deal with the case $q \in \{k+1,\ldots,N-1\}$ and we will proceed similarly to the base case. Therefore some details will be omitted.

Thanks to the induction hypothesis ensuring the continuities of the function $\tilde{t}_{q-1}$ and of the map $\alpha \in [0,\overline{\alpha}] \mapsto \tilde{z}_{q-1}^\alpha(\tilde{t}_{q-1}(\alpha))$, we deduce from Proposition 4.1 that, up to reducing $\overline{\alpha}$, it holds that $[t_{q-1}^c-\delta,t_q^c+\delta] \subset I(f_q,\theta_q^\alpha)$ for all $\alpha \in [0,\overline{\alpha}]$, that $\tilde{z}_q^\alpha$ uniformly converges to $\tilde{z}_q$ over $[t_{q-1}^c-\delta,t_q^c+\delta]$ when $\alpha \to 0$ (see Figure 10 in Section 2.4 where $q = k+1$), and that the map

$$(\alpha,t) \in [0,\overline{\alpha}] \times [t_{q-1}^c-\delta,t_q^c+\delta] \mapsto \tilde{z}_q^\alpha(t) \in \mathbb{R}^n, \tag{5.2}$$

is continuous. Similarly to the base case, up to reducing $\overline{\alpha} > 0$, we get that $\tilde{z}_q^\alpha(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_q^c),\nu)$ for all $(\alpha,t) \in [0,\overline{\alpha}] \times [t_q^c-\delta,t_q^c+\delta]$ and thus we are in a position to define the map

$$\begin{array}{rccl} G_q: & [0,\overline{\alpha}] \times [t_q^c-\delta,t_q^c+\delta] & \to & \mathbb{R} \\ & (\alpha,t) & \mapsto & F_q(\tilde{z}_q^\alpha(t),t). \end{array}$$

Similarly to the base case, $G_q$ is continuous, $G_q(0,t_q^c) = F_q(x(t_q^c),t_q^c) = 0$ and $\nabla_t G_q(\alpha,t)$ exists and is continuous over $[0,\overline{\alpha}] \times [t_q^c-\delta,t_q^c+\delta]$ and

$$\nabla_t G_q(0,t_q^c) = \langle \nabla_x F_q(x(t_q^c),t_q^c), (f_q)^-(t_q^c)\rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t_q^c),t_q^c) \neq 0.$$

Finally, from the third item of Proposition 4.1 and from the induction hypothesis (in particular from the differentiabilities at 0 of the function $\tilde{t}_{q-1}$ and of the map $\alpha \in [0,\overline{\alpha}] \mapsto \tilde{z}_{q-1}^\alpha(\tilde{t}_{q-1}(\alpha)) \in \mathbb{R}^n$), we get that

$$\lim_{\alpha\to 0} \frac{\tilde{z}_q^\alpha(t_q^c) - \tilde{z}_q(t_q^c)}{\alpha} = w_q(t_q^c),$$

which implies that $\nabla_\alpha G_q(0,t_q^c)$ exists with $\nabla_\alpha G_q(0,t_q^c) = \langle \nabla_x F_q(x(t_q^c),t_q^c), w_q(t_q^c)\rangle_{\mathbb{R}^n}$.

From the conic implicit function theorem (Lemma 5.1), up to reducing $\overline{\alpha} > 0$, there exists a function $\tilde{t}_q \in \mathrm{C}([0,\overline{\alpha}],[t_q^c-\delta,t_q^c+\delta])$, such that $\tilde{t}_q(0) = t_q^c$ and $F_q(\tilde{z}_q^\alpha(\tilde{t}_q(\alpha)),\tilde{t}_q(\alpha)) = 0$ for all $\alpha \in [0,\overline{\alpha}]$, that is differentiable at 0 with

$$\tilde{t}_q'(0) = -\frac{\langle \nabla_x F_q(x(t_q^c),t_q^c), w_q(t_q^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t_q^c),t_q^c), (f_q)^-(t_q^c)\rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t_q^c),t_q^c)}.$$

From the continuities of the function $\tilde{t}_q$ and of the map (5.2), we deduce that the map $\alpha \in [0,\overline{\alpha}] \mapsto \tilde{z}_q^\alpha(\tilde{t}_q(\alpha)) \in \mathbb{R}^n$ is continuous over $[0,\overline{\alpha}]$. Similarly to the base case, one can easily prove that

$$\lim_{\alpha\to 0} \frac{\tilde{z}_q^\alpha(\tilde{t}_q(\alpha)) - \tilde{z}_q(t_q^c)}{\alpha} = w_q(t_q^c) + \tilde{t}_q'(0)(f_q)^-(t_q^c),$$

which completes the proof for the inductive step. □

### 5.2.2 Admissibility of the perturbed auxiliary non-hybrid trajectories

**Lemma 5.3.** Consider the framework of Lemma 5.2. Then, up to reducing $\overline{\alpha}$, the following properties are satisfied:

1. There exists $s'_{k-1} \in (t_{k-1}^c, t_{k-1}^c+\delta]$ such that $\tilde{z}_k^\alpha(t) \in E_k(t)$ for all $(\alpha,t) \in [0,\overline{\alpha}] \times (t_{k-1}^c, s'_{k-1}]$ (and for all $(\alpha,t) \in [0,\overline{\alpha}] \times [t_0^c, s'_{k-1}]$ if $k=1$).
2. For all $q \in \{k,...,N-1\}$, there exists $s_q \in [t_q^c-\delta, t_q^c)$ such that $\tilde{z}_q^\alpha(t) \in E_q(t)$ for all $(\alpha,t) \in [0,\overline{\alpha}] \times [s_q, \tilde{t}_q(\alpha))$.
3. For all $q \in \{k,...,N-1\}$, there exists $s'_q \in (t_q^c, t_q^c+\delta]$ such that $\tilde{z}_{q+1}^\alpha(t) \in E_{q+1}(t)$ for all $(\alpha,t) \in [0,\overline{\alpha}] \times (\tilde{t}_q(\alpha), s'_q]$.
4. There exists $s_N \in [t_N^c-\delta, t_N^c)$ such that $\tilde{z}_N^\alpha(t) \in E_N(t)$ for all $(\alpha,t) \in [0,\overline{\alpha}] \times [s_N, t_N^c]$.

*Proof.* This proof does not require induction. We will prove each item separately. Note that we will reduce $\overline{\alpha}$ in each item.

*Proof of the fourth item.* Recall that $\tilde{z}_N = x$ over $[t_{N-1}^c, t_N^c]$ and that $x(t) \in E_N(t)$ for all $t \in [t_N^c-\delta, t_N^c]$. From ($\mathcal{C}1$) and since $\tilde{z}_N^\alpha$ converges uniformly to $\tilde{z}_N$ over $[t_{N-1}^c-\delta, t_N^c] \cap [0,T]$ when $\alpha \to 0$, one can easily conclude the proof of the fourth item by reducing $\overline{\alpha} > 0$ and by taking $s_N = t_N^c - \delta$.

*Proof of the first item.* If $k=1$, then the proof is similar to the above fourth item. Therefore let us deal with the case $k \in \{2,\dots,N\}$. Recall that $\tilde{z}_k(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_{k-1}^c), \frac{\nu}{2})$ for all $t \in [t_{k-1}^c-\delta, t_{k-1}^c+\delta]$. Since $\tilde{z}_k^\alpha$ uniformly converges to $\tilde{z}_k$ over $[t_{k-1}^c-\delta, t_k^c+\delta] \cap [0,T]$ when $\alpha \to 0$, up to reducing $\overline{\alpha} > 0$, we get that $\tilde{z}_k^\alpha(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t_{k-1}^c), \nu)$, and therefore $\tilde{z}_k^\alpha(t) \in E_k(t)$ if and only if $F_{k-1}(\tilde{z}_k^\alpha(t),t) > 0$, for all $(\alpha,t) \in [0,\overline{\alpha}] \times [t_{k-1}^c-\delta, t_{k-1}^c+\delta]$. By contradiction let us assume that

$$\forall s'_{k-1} \in (t_{k-1}^c, t_{k-1}^c+\delta], \quad \forall 0<\beta\leq\overline{\alpha}, \quad \exists \alpha \in [0,\beta], \quad \exists t \in (t_{k-1}^c, s'_{k-1}], \quad F_{k-1}(\tilde{z}_k^\alpha(t),t) \leq 0. \tag{5.3}$$

Let $s'_{k-1} \in (t_{k-1}^c, t_{k-1}^c+\delta]$ and $0<\beta\leq\overline{\alpha}$ and consider $(\alpha,t)$ given in (5.3). Since $F_{k-1}(\tilde{z}_k^\alpha(t_{k-1}^c), t_{k-1}^c) = F_{k-1}(x(t_{k-1}^c), t_{k-1}^c) = 0$, we obtain that

$$F_{k-1}(\tilde{z}_k^\alpha(t),t) - F_{k-1}(\tilde{z}_k^\alpha(t_{k-1}^c), t_{k-1}^c) \leq 0.$$

Since $\tilde{z}_k^\alpha$ is of class $\mathrm{C}^1$ over $[t_{k-1}^c-\delta, t_{k-1}^c+\delta]$, note that the above inequality can be rewritten as

$$\frac{1}{t-t_{k-1}^c}\int_{t_{k-1}^c}^t \Psi_{k-1}(s)\,\mathrm{d}s \leq \frac{1}{t-t_{k-1}^c}\int_{t_{k-1}^c}^t \Psi_{k-1}(s) - \Psi_{k-1}^\alpha(s)\,\mathrm{d}s, \tag{5.4}$$

where

$$\Psi_{k-1}(s) := \langle \nabla_x F_{k-1}(\tilde{z}_k(s),s), f_k(\tilde{z}_k(s),\lambda_k,\tilde{u}_k(s),s)\rangle_{\mathbb{R}^n} + \nabla_t F_{k-1}(\tilde{z}_k(s),s),$$

and

$$\Psi^\alpha_{k-1}(s) := \langle \nabla_x F_{k-1}(\tilde{z}^\alpha_k(s),s), f_k(\tilde{z}^\alpha_k(s), \lambda_k + \alpha(\overline{\lambda}_k - \lambda_k), \tilde{u}_k(s), s)\rangle_{\mathbb{R}^n} + \nabla_t F_{k-1}(\tilde{z}^\alpha_k(s),s),$$

for all $s \in [t^c_{k-1} - \delta, t^c_{k-1} + \delta]$. Since $\tilde{u}_k$ is continuous at $t^c_{k-1}$, note that $t^c_{k-1}$ is a Lebesgue point of $\Psi_{k-1}$. Therefore, when making tend $s'_{k-1} \to t^c_{k-1}$ and $\beta \to 0$, we make tend $\alpha \to 0$ and $t \to t^c_{k-1}$ and thus the left term of (5.4) tends to

$$\langle \nabla_x F_{k-1}(x(t^c_{k-1}), t^c_{k-1}), (f_k)^+(t^c_{k-1})\rangle_{\mathbb{R}^n} + \nabla_t F_{k-1}(x(t^c_{k-1}), t^c_{k-1}).$$

It remains to prove that the right term of (5.4) converges to zero when $\alpha \to 0$ and $t \to t^c_k$. To this aim recall that $\tilde{z}^\alpha_k(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t^c_{k-1}), \nu)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times [t^c_{k-1} - \delta, t^c_{k-1} + \delta]$ and that $\nabla_x F_{k-1}$ and $\nabla_t F_{k-1}$ are uniformly continuous over the compact set $\overline{\mathrm{B}}_{\mathbb{R}^n}(x(t^c_{k-1}), \nu) \times [t^c_{k-1} - \delta, t^c_{k-1} + \delta]$ (since $F_{k-1}$ is of class $\mathrm{C}^1$). Therefore, since $\tilde{z}^\alpha_k$ uniformly converges to $\tilde{z}_k$ over $[t^c_{k-1} - \delta, t^c_{k-1} + \delta]$ when $\alpha \to 0$, one can easily prove that the right term of (5.4) tends to zero when $\alpha \to 0$ and $t \to t^c_k$.

Hence we have obtained that

$$\langle \nabla_x F_{k-1}(x(t^c_{k-1}), t^c_{k-1}), (f_k)^+(t^c_{k-1})\rangle_{\mathbb{R}^n} + \nabla_t F_{k-1}(x(t^c_{k-1}), t^c_{k-1}) \leq 0,$$

which raises a contradiction with $(\mathcal{A}3)$. Therefore we have proved the negation of (5.3) which is given by

$$\exists s'_{k-1} \in (t^c_{k-1}, t^c_{k-1} + \delta], \quad \exists 0 < \beta \leq \overline{\alpha}, \quad \forall \alpha \in [0, \beta], \quad \forall t \in (t^c_{k-1}, s'_{k-1}], \quad F_{k-1}(\tilde{z}^\alpha_k(t), t) > 0,$$

which concludes the proof of the first item by reducing $\overline{\alpha} > 0$ to $\beta$.

*Proof of the second item.* Let $q \in \{k, ..., N-1\}$ be fixed. This proof is similar to the above one, with an additional difficulty due to the presence of the implicit function $\tilde{t}_q$. Recall that $\tilde{z}^\alpha_q(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t^c_q), \nu)$, and therefore $\tilde{z}^\alpha_q(t) \in E_q(t)$ if and only if $F_q(\tilde{z}^\alpha_q(t), t) < 0$, for all $(\alpha, t) \in [0, \overline{\alpha}] \times [t^c_q - \delta, t^c_q + \delta]$. Also recall that $\tilde{t}_q(\alpha)$ tends to $t^c_q$ when $\alpha \to 0$. Therefore, for any $s_q \in [t^c_q - \delta, t^c_q)$, there exists $0 < \overline{\beta}(s_q) \leq \overline{\alpha}$ such that $s_q < \tilde{t}_q(\alpha) \leq t^c_q + \delta$ for all $\alpha \in [0, \overline{\beta}(s_q)]$. By contradiction let us assume that

$$\forall s_q \in [t^c_q - \delta, t^c_q), \quad \forall 0 < \beta \leq \overline{\beta}(s_q), \quad \exists \alpha \in [0, \beta], \quad \exists t \in [s_q, \tilde{t}_q(\alpha)), \quad F_q(\tilde{z}^\alpha_q(t), t) \geq 0. \tag{5.5}$$

Let $s_q \in [t^c_q - \delta, t^c_q)$ and $0 < \beta \leq \overline{\beta}(s_q)$ and consider $(\alpha, t)$ given in (5.5). Since $F_q(\tilde{z}^\alpha_q(\tilde{t}_q(\alpha)), \tilde{t}_q(\alpha)) = 0$ (see Lemma 5.2), we obtain that

$$F_q(\tilde{z}^\alpha_q(\tilde{t}_q(\alpha)), \tilde{t}_q(\alpha)) - F_q(\tilde{z}^\alpha_q(t), t) \leq 0.$$

Since $\tilde{z}^\alpha_q$ is of class $\mathrm{C}^1$ over $[t^c_q - \delta, t^c_q + \delta]$, note that the above inequality can be rewritten as

$$\frac{1}{\tilde{t}_q(\alpha) - t} \int_t^{\tilde{t}_q(\alpha)} \Psi_q(s)\, \mathrm{d}s \leq \frac{1}{\tilde{t}_q(\alpha) - t} \int_t^{\tilde{t}_q(\alpha)} \Psi_q(s) - \Psi^\alpha_q(s)\, \mathrm{d}s, \tag{5.6}$$

where

$$\Psi_q(s) := \begin{cases} \langle \nabla_x F_k(\tilde{z}_k(s), s), f_k(\tilde{z}_k(s), \lambda_k, \tilde{u}_k(s), s)\rangle_{\mathbb{R}^n} + \nabla_t F_k(\tilde{z}_k(s), s), & \text{if } q = k, \\ \langle \nabla_x F_q(\tilde{z}_q(s), s), f_q(\tilde{z}_q(s), \lambda_q, \tilde{u}_q(s), s)\rangle_{\mathbb{R}^n} + \nabla_t F_k(\tilde{z}_q(s), s), & \text{if } q \in \{k+1, ..., N-1\}, \end{cases}$$

and

$$\Psi^\alpha_q(s) := \begin{cases} \langle \nabla_x F_k(\tilde{z}^\alpha_k(s), s), f_k(\tilde{z}^\alpha_k(s), \lambda_k + \alpha(\overline{\lambda}_k - \lambda_k), \tilde{u}_k(s), s)\rangle_{\mathbb{R}^n} + \nabla_t F_k(\tilde{z}^\alpha_k(s), s), & \text{if } q = k, \\ \langle \nabla_x F_q(\tilde{z}^\alpha_q(s), s), f_q(\tilde{z}^\alpha_q(s), \lambda_q, \tilde{u}_q(s), s)\rangle_{\mathbb{R}^n} + \nabla_t F_q(\tilde{z}^\alpha_q(s), s), & \text{if } q \in \{k+1, ..., N-1\}, \end{cases}$$

for all $s \in [t^c_q - \delta, t^c_q + \delta]$. Since $\tilde{u}_q$ is continuous at $t^c_q$, note that $t^c_q$ is a Lebesgue point of $\Psi_q$. Therefore, when making tend $s_q \to t^c_q$ and $\beta \to 0$, we make tend $\alpha \to 0$, $\tilde{t}_q(\alpha) \to t^c_q$ and $t \to t^c_q$ and thus the left term of (5.6) tends to

$$\langle \nabla_x F_q(x(t^c_q), t^c_q), (f_q)^-(t^c_q)\rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t^c_q), t^c_q),$$

and, using similar arguments as in the proof of the first item, we obtain that the right term of (5.6) tends to zero when $\alpha \to 0$, $\tilde{t}_q(\alpha) \to t^c_q$ and $t \to t^c_q$.

Hence we have obtained that

$$\langle \nabla_x F_q(x(t_q^c), t_q^c), (f_q)^-(t_q^c) \rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t_q^c), t_q^c) \leq 0,$$

which raises a contradiction with ($\mathcal{A}3$). Therefore we have proved the negation of (5.5) which is given by

$$\exists s_q \in [t_q^c - \delta, t_q^c), \quad \exists 0 < \beta \leq \overline{\beta}(s_q), \quad \forall \alpha \in [0, \beta], \quad \forall t \in [s_q, \tilde{t}_q(\alpha)), \quad F_q(\tilde{z}_q^\alpha(t), t) > 0.$$

which concludes the proof of the second item by reducing $\overline{\alpha} > 0$ to $\beta$.

*Proof of the third item.* The proof is similar to the above one. □

**Lemma 5.4** (Admissibility of the perturbed auxiliary non-hybrid trajectories)**.** Consider the framework of Lemma 5.2. Then, up to reducing $\overline{\alpha} > 0$, it holds that:

1. $\tilde{z}_k^\alpha(t) \in E_k(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (t_{k-1}^c, \tilde{t}_k(\alpha))$ (and for all $(\alpha, t) \in [0, \overline{\alpha}] \times [t_0^c, \tilde{t}_k(\alpha))$ if $k = 1$).
2. $\tilde{z}_q^\alpha(t) \in E_q(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha))$ and all $q \in \{k+1, ..., N-1\}$.
3. $\tilde{z}_N^\alpha(t) \in E_N(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (\tilde{t}_{N-1}(\alpha), t_N^c]$.

*Proof.* This proof does not require induction. Let us prove the second item only. The other items can be proved similarly (and note that $\overline{\alpha} > 0$ is reduced in each item). Let $q \in \{k+1, \ldots, N-1\}$. From Lemma 5.3, we know that:

- there exists $s'_{q-1} \in (t_{q-1}^c, t_{q-1}^c + \delta]$ such that $\tilde{z}_q^\alpha(t) \in E_q(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (\tilde{t}_{q-1}(\alpha), s'_{q-1}]$.
- there exists $s_q \in [t_q^c - \delta, t_q^c)$ such that $\tilde{z}_q^\alpha(t) \in E_q(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times [s_q, \tilde{t}_q(\alpha))$.

Now recall that $\tilde{z}_q = x$ over $[t_{q-1}^c, t_q^c]$ and that $x(t) \in E_q(t)$ for all $t \in (t_{q-1}^c, t_q^c)$ and thus for all $t \in [s'_{q-1}, s_q]$. From ($\mathcal{C}1$) and since $\tilde{z}_q^\alpha$ converges uniformly to $\tilde{z}_q$ over $[t_{q-1}^c - \delta, t_q^c + \delta]$ when $\alpha \to 0$, one can easily see that, up to reducing $\overline{\alpha} > 0$, one has $\tilde{z}_q^\alpha(t) \in E_q(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times [s'_{q-1}, s_q]$. We finally deduce that $\tilde{z}_q^\alpha(t) \in E_q(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha))$. The proof of the second item is complete. □

#### 5.2.3 Proof of Proposition 5.1

Let us fix $k \in \{1, ..., N\}$ and $\overline{\lambda}_k \in \mathbb{R}^d$. Consider the perturbed auxiliary non-hybrid trajectories $\tilde{z}_q^\alpha = y(\cdot, f_q, \theta_q^\alpha)$ over $[t_{q-1}^c - \delta, t_q^c + \delta] \cap [0, T]$ for all $q \in \{k, \ldots, N\}$ and all $\alpha \in (0, \overline{\alpha}]$ constructed in Lemma 5.2, together with the corresponding implicit functions $\tilde{t}_q$ for all $q \in \{k, \ldots, N-1\}$. As explained in Section 2.4, we define by concatenation

$$x^\alpha(t) := \begin{cases} x(t) & \text{for all } t \in [t_0^c, t_{k-1}^c], \\ \tilde{z}_k^\alpha(t) & \text{for all } t \in [t_{k-1}^c, \tilde{t}_k(\alpha)], \\ \tilde{z}_q^\alpha(t) & \text{for all } t \in [\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)] \text{ and all } q \in \{k+1, \ldots, N-1\}, \\ \tilde{z}_N^\alpha(t) & \text{for all } t \in [\tilde{t}_{N-1}(\alpha), t_N^c], \end{cases}$$

and

$$\lambda^\alpha(t) := \begin{cases} \lambda(t) & \text{for a.e. } t \in (t_0^c, t_{k-1}^c), \\ \lambda_k + \alpha(\overline{\lambda}_k - \lambda_k) & \text{for a.e. } t \in (t_{k-1}^c, \tilde{t}_k(\alpha)), \\ \lambda_q & \text{for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{k+1, \ldots, N-1\}, \\ \lambda_N & \text{for a.e. } t \in (\tilde{t}_{N-1}(\alpha), t_N^c), \end{cases}$$

and

$$u^\alpha(t) := \begin{cases} u(t) & \text{for a.e. } t \in (t_0^c, t_{k-1}^c), \\ \tilde{u}_k(t) & \text{for a.e. } t \in (t_{k-1}^c, \tilde{t}_k(\alpha)), \\ \tilde{u}_q(t) & \text{for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{k+1, \ldots, N-1\}, \\ \tilde{u}_N(t) & \text{for a.e. } t \in (\tilde{t}_{N-1}(\alpha), t_N^c), \end{cases}$$

for all $\alpha \in (0, \overline{\alpha}]$. From the construction and the results developed in Lemmas 5.2 and 5.4, one can easily see that $(x^\alpha, \lambda^\alpha, u^\alpha) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{PC}([0,T],\mathbb{R}^d) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ is a (perturbed) solution to (CS), admitting the $\tilde{t}_q(\alpha)$ as crossing times, where we have introduced $\tilde{t}_q(\alpha) := t_q^c$ for all $q \in \{1,\dots,k-1\}$ and all $\alpha \in (0,\overline{\alpha}]$. The first, third, fourth and sixth items of Proposition 5.1 also directly follow, as well as the first assertion of the second item. The second assertion of the second item follows from the uniform convergence of $\tilde{z}_q^\alpha$ to $\tilde{z}_q$ over $[t_{q-1}^c - \delta, t_q^c + \delta] \cap [0,T]$ for all $q \in \{k,\dots,N\}$ when $\alpha \to 0$, from the convergence of $\tilde{t}_q(\alpha)$ to $t_q^c$ for all $q \in \{k,\dots,N-1\}$ when $\alpha \to 0$, and from the equality $\tilde{z}_q = x$ over $[t_{q-1}^c, t_q^c]$ for all $q \in \{k,\dots,N\}$. Finally the fifth item follows from Lemma 5.2 since it holds that

$$\lim_{\alpha \to 0} \frac{x^\alpha(T) - x(T)}{\alpha} = \lim_{\alpha \to 0} \frac{\tilde{z}_N^\alpha(t_N^c) - \tilde{z}_N(t_N^c)}{\alpha} = w_N(t_N^c) = w(T),$$

which concludes the proof of Proposition 5.1.

## 5.3 Needle-like perturbation of the control

Consider the framework of Section 5.1. This entire Section 5.3 is dedicated to the proof of the next proposition which states a differentiability result at time $t = T$ for the trajectory $x$ with respect to a needle-like perturbation of the control $u$. Since the proofs of this section are very similar to the ones of the previous Section 5.2, they are omitted.

**Proposition 5.2.** Consider the framework of Section 5.1. Let $k \in \{1, ..., N\}$, let $v \in \mathbb{R}^m$ and let $\tau \in (t_{k-1}^c, t_k^c)$ be a Lebesgue point of the map $h(x(\cdot), \lambda(\cdot), u(\cdot), \cdot)$. Then there exists $0 < \overline{\alpha} < \min\{1, \tau - t_{k-1}^c\}$ such that, for all $\alpha \in (0, \overline{\alpha}]$, there exists a perturbed solution $(x^\alpha, \lambda^\alpha, u^\alpha) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{PC}([0,T],\mathbb{R}^d) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ to (CS) such that:

(i) The corresponding perturbed partition of $[0,T]$, denoted by $\{\tilde{t}_q(\alpha)\}_{q=0,\dots,N(\alpha)}$, satisfies $N(\alpha) = N$, with $\tilde{t}_q(\alpha) = t_q^c$ for all $q \in \{1, ..., k-1\}$, and $\tilde{t}_q(\alpha)$ tends to $t_q^c$ when $\alpha \to 0$ for all $q \in \{k, ..., N-1\}$.

(ii) The perturbed trajectory $x^\alpha$ follows the same regions than $x$, that is, $x^\alpha$ satisfies

$$x^\alpha(t) \in E_q(t) \text{ for all } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{1,\dots,N\},$$

with $x^\alpha(0) = x_{\mathrm{init}} \in E_1(0)$ and $x^\alpha(T) \in E_N(T)$. Moreover $x^\alpha$ uniformly converges to $x$ over $[0,T]$ when $\alpha \to 0$.

(iii) The perturbed regionally switching parameter $\lambda^\alpha$ is given by

$$\lambda^\alpha(t) = \lambda_q \text{ for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{1,\dots,N\}.$$

(iv) The perturbed control $u^\alpha$ is given by

$$u^\alpha(t) = \begin{cases} v & \text{for a.e. } t \in (\tau - \alpha, \tau), \\ \tilde{u}_k(t) & \text{for a.e. } t \in (t_{k-1}^c, \tau - \alpha) \cup (\tau, \tilde{t}_k(\alpha)), \\ \tilde{u}_q(t) & \text{for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{1,\dots,N\} \backslash \{k\}, \end{cases}$$

where $\tilde{u}_q$ stands for the auxiliary control defined in Section 5.1 for all $q \in \{1,\dots,N\}$.

(v) The limit

$$\lim_{\alpha \to 0} \frac{x^\alpha(T) - x(T)}{\alpha} = w(T),$$

holds true, where

$$w(t) := \begin{cases} w_q(t), & \text{for all } t \in [t_{q-1}^c, t_q^c) \text{ and all } q \in \{k, ..., N-1\}, \\ w_N(t), & \text{for all } t \in [t_{N-1}^c, t_N^c], \end{cases}$$

where $w_k$ is the variation vector defined as the unique maximal solution (which is global) to the linearized Cauchy problem given by

$$\left\{ \begin{array}{l} \dot{w}(t) = \nabla_x f_k(\tilde{z}_k(t), \lambda_k, \tilde{u}_k(t), t) w(t), \quad \text{a.e. } t \in [t^c_{k-1} - \delta, t^c_k + \delta] \cap [0, T], \\ w(\tau) = f_k(\tilde{z}_k(\tau), \lambda_k, v, \tau) - f_k(\tilde{z}_k(\tau), \lambda_k, \tilde{u}_k(\tau), \tau), \end{array} \right.$$

and $w_q$ is the variation vector defined by induction as the unique maximal solution (which is global) to the linearized Cauchy problem given by

$$\left\{ \begin{array}{l} \dot{w}(t) = \nabla_x f_q(\tilde{z}_q(t), \lambda_q, \tilde{u}_q(t), t) w(t), \quad \text{a.e. } t \in [t^c_{q-1} - \delta, t^c_q + \delta] \cap [0, T], \\ w(t^c_{q-1}) = w_{q-1}(t^c_{q-1}) + \xi_{q-1}, \end{array} \right.$$

for all $q \in \{k+1, ..., N\}$, where $\xi_q \in \mathbb{R}^n$ stands for the jump vector defined by

$$\xi_q := \frac{\langle \nabla_x F_q(x(t^c_q), t^c_q), w_q(t^c_q) \rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t^c_q), t^c_q), (f_q)^-(t^c_q) \rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t^c_q), t^c_q)} ((f_{q+1})^+(t^c_q) - (f_q)^-(t^c_q)),$$

for all $q \in \{k, ..., N-1\}$.

(vi) The limit

$$\lim_{\alpha \to 0} \frac{\tilde{t}_q(\alpha) - t^c_q}{\alpha} = - \frac{\langle \nabla_x F_q(x(t^c_q), t^c_q), w_q(t^c_q) \rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t^c_q), t^c_q), (f_q)^-(t^c_q) \rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t^c_q), t^c_q)},$$

holds true for all $q \in \{k, \dots, N-1\}$.

### 5.3.1 Construction of perturbed auxiliary non-hybrid trajectories

**Lemma 5.5** (Construction of perturbed auxiliary non-hybrid trajectories)**.** Consider the frameworks of Section 5.1 and Proposition 5.2. Let $k \in \{1, ..., N\}$, let $v \in \mathbb{R}^m$ and let $\tau \in (t^c_{k-1}, t^c_k)$ be a Lebesgue point of the map $h(x(\cdot), \lambda(\cdot), u(\cdot), \cdot)$. Then there exists $0 < \overline{\alpha} < \min\{1, \tau - t^c_{k-1}\}$ and, for all $q \in \{k, ..., N-1\}$, there exists a function $\tilde{t}_q \in \mathrm{C}([0, \overline{\alpha}], [t^c_q - \delta, t^c_q + \delta])$ differentiable at 0 with $\tilde{t}_q(0) = t^c_q$ and

$$\tilde{t}'_q(0) = - \frac{\langle \nabla_x F_q(x(t^c_q), t^c_q), w_q(t^c_q) \rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t^c_q), t^c_q), (f_q)^-(t^c_q) \rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t^c_q), t^c_q)},$$

such that the perturbed auxiliary non-hybrid trajectories $\tilde{z}^\alpha_q := y(\cdot, f_q, \theta^\alpha_q)$ associated with the perturbed quadruplets $\theta^\alpha_q$ defined by the induction

$$\theta^\alpha_q := \left\{ \begin{array}{ll} (\lambda_k, \tilde{u}^\alpha_k, t^c_{k-1}, x(t^c_{k-1})) & \text{if} \quad q = k, \\ (\lambda_q, \tilde{u}_q, \tilde{t}_{q-1}(\alpha), \tilde{z}^\alpha_{q-1}(\tilde{t}_{q-1}(\alpha)) & \text{if} \quad q \in \{k+1, \dots, N\}, \end{array} \right.$$

for all $\alpha \in [0, \overline{\alpha}]$ and all $q \in \{k, \dots, N\}$, where $\tilde{u}^\alpha_k$ is the needle-like perturbation of $\tilde{u}_k$ (see Figure 8 in Section 2.4) given by

$$\tilde{u}^\alpha_k(t) := \left\{ \begin{array}{ll} v & \text{if } t \in [\tau - \alpha, \tau), \\ \tilde{u}_k(t) & \text{if } t \notin [\tau - \alpha, \tau), \end{array} \right.$$

for almost every $t \in [0, T]$, satisfy:

- for all $q \in \{k, \dots, N\}$, it holds that $[t^c_{q-1} - \delta, t^c_q + \delta] \cap [0, T] \subset I(f_q, \theta^\alpha_q)$ for all $\alpha \in [0, \overline{\alpha}]$, that $\tilde{z}^\alpha_q$ uniformly converges to $\tilde{z}_q$ over $[t^c_{q-1} - \delta, t^c_q + \delta] \cap [0, T]$ when $\alpha \to 0$, and

$$\lim_{\alpha \to 0} \frac{\tilde{z}^\alpha_q(t^c_q) - \tilde{z}_q(t^c_q)}{\alpha} = w_q(t^c_q).$$

- for all $q \in \{k, \dots, N-1\}$, it holds that $\tilde{z}^\alpha_q(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(t^c_q), \nu)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times [t^c_q - \delta, t^c_q + \delta]$, that $F_q(\tilde{z}^\alpha_q(\tilde{t}_q(\alpha)), \tilde{t}_q(\alpha)) = 0$ for all $\alpha \in [0, \overline{\alpha}]$, and that the map $\alpha \in [0, \overline{\alpha}] \mapsto \tilde{z}^\alpha_q(\tilde{t}_q(\alpha)) \in \mathbb{R}^n$ is continuous over $[0, \overline{\alpha}]$ and differentiable at 0 with

$$\lim_{\alpha \to 0} \frac{\tilde{z}^\alpha_q(\tilde{t}_q(\alpha)) - \tilde{z}_q(t^c_q)}{\alpha} = w_q(t^c_q) + \tilde{t}'_q(0)(f_q)^-(t^c_q).$$

*Proof.* The proof is very similar to the one of Lemma 5.2 and thus is omitted. The only difference is that, for the base case, one must note that we fix $\delta_0 \in [0, \delta]$ such that $\tau < t^c_k - \delta_0$ in order to have $\tilde{u}^\alpha_k = \tilde{u}_k$ for almost every $t \in [t^c_k - \delta_0, t^c_k + \delta_0]$ where $\tau \in (t^c_{k-1}, t^c_k)$ stands for a Lebesgue point of $h(x(\cdot), \lambda(\cdot), u(\cdot), \cdot)$. □

#### 5.3.2 Admissibility of the perturbed auxiliary non-hybrid trajectories

**Lemma 5.6.** Consider the framework of Lemma 5.5. Then, up to reducing $\overline{\alpha} > 0$, the following properties are satisfied:

1. There exists $s'_{k-1} \in (t^c_{k-1}, t^c_{k-1} + \delta]$ such that $\tilde{z}^\alpha_k(t) \in E_k(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (t^c_{k-1}, s'_{k-1}]$ (and for all $(\alpha, t) \in [0, \overline{\alpha}] \times [t^c_0, s'_{k-1}]$ if $k = 1$).
2. For all $q \in \{k, ..., N-1\}$, there exists $s_q \in [t^c_q - \delta, t^c_q)$ such that $\tilde{z}^\alpha_q(t) \in E_q(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times [s_q, \tilde{t}_q(\alpha))$.
3. For all $q \in \{k, ..., N-1\}$, there exists $s'_q \in (t^c_q, t^c_q + \delta]$ such that $\tilde{z}^\alpha_{q+1}(t) \in E_{q+1}(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (\tilde{t}_q(\alpha), s'_q]$.
4. There exists $s_N \in [t^c_N - \delta, t^c_N)$ such that $\tilde{z}^\alpha_N(t) \in E_N(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times [s_N, t^c_N]$.

*Proof.* The proof is very similar to the one of Lemma 5.3 and thus is omitted. □

**Lemma 5.7** (Admissibility of the perturbed auxiliary non-hybrid trajectories)**.** Consider the framework of Lemma 5.5. Then, up to reducing $\overline{\alpha} > 0$, it holds that:

1. $\tilde{z}^\alpha_k(t) \in E_k(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (t^c_{k-1}, \tilde{t}_k(\alpha))$ (and for all $(\alpha, t) \in [0, \overline{\alpha}] \times [t^c_0, \tilde{t}_k(\alpha))$ if $k = 1$).
2. $\tilde{z}^\alpha_q(t) \in E_q(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha))$ and all $q \in \{k+1, ..., N-1\}$.
3. $\tilde{z}^\alpha_N(t) \in E_N(t)$ for all $(\alpha, t) \in [0, \overline{\alpha}] \times (\tilde{t}_{N-1}(\alpha), t^c_N]$.

*Proof.* The proof is very similar to the one of Lemma 5.4 and thus is omitted. □

#### 5.3.3 Proof of Proposition 5.2

Let us fix $k \in \{1, ..., N\}$, $v \in \mathbb{R}^m$ and $\tau \in (t^c_{k-1}, t^c_k)$ being a Lebesgue point of the map $h(x(\cdot), \lambda(\cdot), u(\cdot), \cdot)$. Consider the perturbed auxiliary non-hybrid trajectories $\tilde{z}^\alpha_q = y(\cdot, f_q, \theta^\alpha_q)$ over $[t^c_{q-1} - \delta, t^c_q + \delta] \cap [0, T]$ for all $q \in \{k, \ldots, N\}$ and all $\alpha \in (0, \overline{\alpha}]$ constructed in Lemma 5.5, together with the corresponding implicit functions $\tilde{t}_q$ for all $q \in \{k, \ldots, N-1\}$. As explained in Section 2.4, we define by concatenation

$$x^\alpha(t) := \begin{cases} x(t) & \text{for all } t \in [t^c_0, t^c_{k-1}], \\ \tilde{z}^\alpha_k(t) & \text{for all } t \in [t^c_{k-1}, \tilde{t}_k(\alpha)], \\ \tilde{z}^\alpha_q(t) & \text{for all } t \in [\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)] \text{ and all } q \in \{k+1, \ldots, N-1\}, \\ \tilde{z}^\alpha_N(t) & \text{for all } t \in [\tilde{t}_{N-1}(\alpha), t^c_N], \end{cases}$$

and

$$\lambda^\alpha(t) := \begin{cases} \lambda(t) & \text{for a.e. } t \in (t^c_0, t^c_{k-1}), \\ \lambda_k & \text{for a.e. } t \in (t^c_{k-1}, \tilde{t}_k(\alpha)), \\ \lambda_q & \text{for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{k+1, \ldots, N-1\}, \\ \lambda_N & \text{for a.e. } t \in (\tilde{t}_{N-1}(\alpha), t^c_N), \end{cases}$$

and

$$u^\alpha(t) = \begin{cases} u(t) & \text{for a.e. } t \in (t^c_0, t^c_{k-1}), \\ v & \text{for a.e. } t \in (\tau - \alpha, \tau), \\ \tilde{u}_k(t) & \text{for a.e. } t \in (t^c_{k-1}, \tau - \alpha) \cup (\tau, \tilde{t}_k(\alpha)), \\ \tilde{u}_q(t) & \text{for a.e. } t \in (\tilde{t}_{q-1}(\alpha), \tilde{t}_q(\alpha)) \text{ and all } q \in \{k+1, \ldots, N-1\}, \\ \tilde{u}_N(t) & \text{for a.e. } t \in (\tilde{t}_{N-1}(\alpha), t^c_N), \end{cases}$$

for all $\alpha \in (0, \overline{\alpha}]$. The remaining details of the proof are omitted, since we use similar arguments than the proof of Proposition 5.1.

# 6 Proof of Theorem 2.1

Let $(x,\lambda,u) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{PC}([0,T],\mathbb{R}^d) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ be a solution to (OCP), that is moreover a regular solution to (CS). In the sequel we will use the notations introduced in Definitions 2.1 and 2.2 and the results obtained in the previous Section 5.

**Introduction of the adjoint vector.** We define the adjoint vector $p \in \mathrm{PAC}^{\mathbb{T}}([0,T],\mathbb{R}^n)$ as

$$p(t) := \begin{cases} p_1(t) & \text{for all } t \in [t_0^c, t_1^c), \\ p_k(t) & \text{for all } t \in (t_{k-1}^c, t_k^c) \text{ and all } k \in \{2,...,N-1\}, \\ p_N(t) & \text{for all } t \in (t_{N-1}^c, t_N^c], \end{cases}$$

where $p_N$ is defined as the unique maximal solution (which is global) to the linear Cauchy problem given by

$$\begin{cases} \dot{p}(t) = -\nabla_x f_N(\tilde{z}_N(t), \lambda_N, \tilde{u}_N(t), t)^\top p(t), \quad \text{a.e. } t \in [t_{N-1}^c - \delta, T] \cap [0,T], \\ p(T) = -\nabla \phi(x(T)), \end{cases}$$

and $p_k$ is defined by backward induction as the unique maximal solution (which is global) to the linear Cauchy problem given by

$$\begin{cases} \dot{p}(t) = -\nabla_x f_k(\tilde{z}_k(t), \lambda_k, \tilde{u}_k(t), t)^\top p(t), \quad \text{a.e. } t \in [t_{k-1}^c - \delta, t_k^c + \delta] \cap [0,T], \\ p^-(t_k^c) = p_{k+1}^+(t_k^c) - \chi_k, \end{cases}$$

for all $k \in \{1,...,N-1\}$, where $\chi_k \in \mathbb{R}^n$ stands for the jump vector defined by

$$\chi_k := -\frac{\langle p_{k+1}^+(t_k^c), (f_{k+1})^+(t_k^c) - (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_k(x(t_k^c), t_k^c), (f_k)^-(t_k^c)\rangle_{\mathbb{R}^n} + \nabla_t F_k(x(t_k^c), t_k^c)} \nabla_x F_k(x(t_k^c), t_k^c),$$

for all $k \in \{1,...,N-1\}$.

Recall that $\tilde{z}_k(t) = x(t)$ for all $t \in [t_{k-1}^c, t_k^c]$ and all $k \in \{1,...,N\}$. Through concatenation of the above linear Cauchy problems, one can easily see that the first item of Theorem 2.1 is fullfilled. Furthermore, from the above Cauchy conditions, the second and third items of Theorem 2.1 also trivially follow.

**The Hamiltonian maximization condition.** Let us fix $k \in \{1,...,N\}$, $v \in \mathrm{U}$ and $\tau \in (t_{k-1}^c, t_k^c)$ being a Lebesgue point of $h(x(\cdot),\lambda(\cdot),u(\cdot),\cdot)$. Consider $0 < \overline{\alpha} < \min\{1, \tau - t_{k-1}^c\}$ given in Proposition 5.2. From the construction detailed in Proposition 5.2 and explained in detail in Section 2.4, and since U is assumed to be closed (which guarantees that the limits $u^-(t_k^c)$ and $u^+(t_k^c)$ belongs to U), one can easily see that the perturbed solution $(x^\alpha, \lambda^\alpha, u^\alpha)$ to (CS) satisfies all the constraints of Problem (OCP) for all $\alpha \in (0,\overline{\alpha}]$. Thus, from optimality of the triplet $(x,\lambda,u)$, we get that

$$\frac{\phi(x^\alpha(T)) - \phi(x(T))}{\alpha} \geq 0,$$

for all $\alpha \in (0,\overline{\alpha}]$ and, taking the limit $\alpha \to 0$, we get from Proposition 5.2 that $\langle \nabla\phi(x(T)), w(T)\rangle_{\mathbb{R}^n} \geq 0$ which can be rewritten as $\langle p(T), w(T)\rangle_{\mathbb{R}^n} \leq 0$.

From the linear Cauchy problems satisfied by $p$ and $w$ over each open interval $(t_{q-1}^c, t_q^c)$ for $q \in \{k,\dots,N-1\}$ (and over $(t_{N-1}^c, t_N^c]$ for $q = N$), one can easily see that the scalar product $\langle p(\cdot), w(\cdot)\rangle_{\mathbb{R}^n}$ is constant over each of these intervals.

Now let us prove that $\langle p^+(t_q^c), w^+(t_q^c)\rangle_{\mathbb{R}^n} = \langle p^-(t_q^c), w^-(t_q^c)\rangle_{\mathbb{R}^n}$ at each crossing time $t_q$ for $q \in \{k,\dots,N-1\}$. To this aim note that the definition of $\chi_q$ has been selected to get that

$$\frac{\langle p^+(t_q^c), (f_{q+1})^+(t_q^c) - (f_q)^-(t_q^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t_q^c), t_q^c), (f_q)^-(t_q^c)\rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t_q^c), t_q^c)} = \frac{\langle p^-(t_q^c), (f_{q+1})^+(t_q^c) - (f_q)^-(t_q^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t_q^c), t_q^c), (f_{q+1})^+(t_q^c)\rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t_q^c), t_q^c)},$$

by replacing $p^-(t_q^c)$ in the above right-hand term by $p^-(t_q^c) = p^+(t_q^c) - \chi_q$. In particular $\chi_q$ can thus be rewritten as

$$\chi_q = -\frac{\langle p^-(t_q^c), (f_{q+1})^+(t_q^c) - (f_q)^-(t_q^c)\rangle_{\mathbb{R}^n}}{\langle \nabla_x F_q(x(t_q^c), t_q^c), (f_{q+1})^+(t_q^c)\rangle_{\mathbb{R}^n} + \nabla_t F_q(x(t_q^c), t_q^c)} \nabla_x F_q(x(t_q^c), t_q^c),$$

which leads to $\langle p^-(t_q^c), \xi_q\rangle_{\mathbb{R}^n} + \langle \chi_q, w^-(t_q^c)\rangle_{\mathbb{R}^n} + \langle \chi_q, \xi_q\rangle_{\mathbb{R}^n} = 0$. Therefore, from the equality $\langle p^+(t_q^c), w^+(t_q^c)\rangle_{\mathbb{R}^n} = \langle p^-(t_q^c) + \chi_q, w^-(t_q^c) + \xi_q\rangle_{\mathbb{R}^n}$, we get that $\langle p^+(t_q^c), w^+(t_q^c)\rangle_{\mathbb{R}^n} = \langle p^-(t_q^c), w^-(t_q^c)\rangle_{\mathbb{R}^n}$.

Finally, by simple backward induction, we have obtained that $\langle p(\tau), w(\tau)\rangle_{\mathbb{R}^n} \leq 0$. From the value of $w(\tau)$ given in Proposition 5.2, this inequality gives

$$H(\tilde{z}_k(\tau), \lambda_k, v, p_k(\tau), \tau) \leq H(\tilde{z}_k(\tau), \lambda_k, \tilde{u}_k(\tau), p_k(\tau), \tau),$$

which can be rewritten as

$$H(x(\tau), \lambda(\tau), v, p(\tau), \tau) \leq H(x(\tau), \lambda(\tau), u(\tau), p(\tau), \tau),$$

which concludes this paragraph.

**The averaged Hamiltonian gradient condition.** Let us fix $k \in \{1, ..., N\}$. Consider some $\overline{\lambda}_k \in \Lambda$ and $0 < \overline{\alpha} \leq 1$ given in Proposition 5.1. From the convexity of $\Lambda$ and the construction detailed in Proposition 5.1 and explained in detail in Section 2.4, one can easily see that the perturbed solution $(x^\alpha, \lambda^\alpha, u^\alpha)$ to (CS) satisfies all the constraints of Problem (OCP) for all $\alpha \in (0, \overline{\alpha}]$. Thus, from optimality of the triplet $(x, \lambda, u)$, we get that

$$\frac{\phi(x^\alpha(T)) - \phi(x(T))}{\alpha} \geq 0,$$

for all $\alpha \in (0, \overline{\alpha}]$ and, taking the limit $\alpha \to 0$, we get from Proposition 5.1 that $\langle \nabla\phi(x(T)), w(T)\rangle_{\mathbb{R}^n} \geq 0$ which can be rewritten as $\langle p(T), w(T)\rangle_{\mathbb{R}^n} \leq 0$.

Using similar arguments than in the previous paragraph, one can derive that $\langle p^-(t_k^c), w^-(t_k^c)\rangle_{\mathbb{R}^n} \leq 0$. Now recall that the classical Duhamel formula leads to $p(s) = \Phi(t_k^c, s)^\top p^-(t_k^c)$ for all $s \in (t_{k-1}^c, t_k^c)$ and

$$w^-(t_k^c) = \int_{t_{k-1}^c}^{t_k^c} \Phi(t_k^c, s) \nabla_\lambda f_k(\tilde{z}_k(s), \lambda_k, \tilde{u}_k(s), s)(\overline{\lambda}_k - \lambda_k)\, \mathrm{d}s,$$

where $\Phi$ stands for the state transition matrix associated with the matrix function $\nabla_x f_k(\tilde{z}_k(\cdot), \lambda_k, \tilde{u}_k(\cdot), \cdot)$. Therefore the inequality $\langle p^-(t_k^c), w^-(t_k^c)\rangle_{\mathbb{R}^n} \leq 0$ gives

$$\left\langle \int_{t_{k-1}^c}^{t_k^c} \nabla_\lambda f_k(\tilde{z}_k(s), \lambda_k, \tilde{u}_k(s), s)^\top p(s)\, \mathrm{d}s, \overline{\lambda}_k - \lambda_k \right\rangle_{\mathbb{R}^d} \leq 0,$$

which can be rewritten as

$$\left\langle \int_{t_{k-1}^c}^{t_k^c} \nabla_\lambda H(x(s), \lambda_k, u(s), p(s), s)\, \mathrm{d}s, \overline{\lambda}_k - \lambda_k \right\rangle_{\mathbb{R}^d} \leq 0.$$

Since the above inequality is satisfied for any $\overline{\lambda}_k \in \Lambda$, this paragraph is complete, and so is the proof of Theorem 2.1.

# References


[1] S. Adly, L. Bourdin, and G. Dhar, *Application of a universal separating vector lemma to optimal sampled-data control problems with nonsmooth Mayer cost function*, Submitted, (2022).

[2] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*, vol. 2, Springer Science & Business Media, 2004.

[3] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability theory*, Springer, Heidelberg, second ed., 2011.

[4] J.-P. Aubin and A. Cellina, *Differential inclusions: set-valued maps and viability theory*, vol. 264, Springer Science & Business Media, 2012.

[5] J.-P. Aubin and H. Frankowska, *Set-valued analysis*, Birkhäuser, 1990.

[6] A. Balluchi, M. D. Benedetto, C. Pinello, C. Rossi, and A. Sangiovanni-Vincentelli, *Hybrid control for automotive engine management: the cut-off case*, in International workshop on hybrid systems: computation and control, Springer, Berlin, Heidelberg, 1998, pp. 13–32.

[7] G. Barles, A. Briani, and E. Trélat, *Value function for regional control problems via dynamic programming and pontryagin maximum principle*, Math. Control Relat. Fields, 8 (2018), pp. 509–533.

[8] T. Bayen, A. Bouali, and L. Bourdin, *Optimal control problems with non-control regions: necessary optimality conditions*, Submitted to proceedings of the IFAC workshop CAO, (2022).

[9] T. Bayen, K. Boumaza, and A. Rapaport, *Necessary optimality condition for the minimal time crisis relaxing transverse condition via regularization*, ESAIM Control Optim. Calc. Var., 27 (2021).

[10] T. Bayen and L. Pfeiffer, *Second-order analysis for the time crisis problem*, J. Convex Anal., 27 (2020), pp. 139–163.

[11] T. Bayen and A. Rapaport, *Minimal time crisis versus minimum time to reach a viability kernel: a case study in the prey-predator model*, Optim. Control Appl., 40 (2019), pp. 330–350.

[12] R. Bellman, *Dynamic programming*, Princeton University Press, Princeton, N. J., 1957.

[13] S. C. Bengea and R. A. DeCarlo, *Optimal control of switching systems*, Automatica, 41 (2005), pp. 11–27.

[14] M. Bergounioux and L. Bourdin, *Pontryagin maximum principle for general Caputo fractional optimal control problems with Bolza cost and terminal constraints*, ESAIM Control Optim. Calc. Var., 26 (2020), p. 35.

[15] A. Beydoun, L. Y. Wang, J. Sun, and S. Sivashankar, *Hybrid control of automotive powertrain systems: a case study*, in International workshop on hybrid systems: computation and control, 1998, pp. 33–48.

[16] L. Bolzoni, E. Bonacini, C. Soresina, and M. Groppi, *Time-optimal control strategies in sir epidemic models*, Mathematical biosciences, 292 (2017), pp. 86–96.

[17] J. F. Bonnans, *Course on optimal control*, OROC Ensta Paris-Tech and optimization master, U. Paris-Saclay, (2017).

[18] L. Bourdin, *Note on Pontryagin maximum principle with running state constraints and smooth dynamics–proof based on the Ekeland variational principle*, Research notes, hal-01302222, (2016).

[19] L. Bourdin and G. Dhar, *Continuity/constancy of the Hamiltonian function in a Pontryagin maximum principle for optimal sampled-data control problems with free sampling times*, Math. Control Signals Systems, 31 (2019), pp. 503–544.

[20] ——, *Optimal sampled-data controls with running inequality state constraints: Pontryagin maximum principle and bouncing trajectory phenomenon*, Math. Program., (2020), pp. 1–45.

[21] L. Bourdin and E. Trélat, *Pontryagin maximum principle for optimal sampled-data control problems*, in Proceedings of the IFAC workshop CAO, vol. 48, 2015, pp. 80–84.

[22] ——, *Optimal sampled-data control, and generalizations on time scales*, Math. Control Relat. Fields, 6 (2016), pp. 53–94.

[23] ——, *Linear–quadratic optimal sampled-data control problems: convergence result and Riccati theory*, Automatica, 79 (2017), pp. 273–281.

[24] ——, *Robustness under control sampling of reachability in fixed time for nonlinear control systems*, Math. Control Signals Systems, 33 (2021), pp. 515–551.

[25] A. Bressan and Y. Hong, *Optimal control problems on stratified domains*, Netw. Heterog. Media, 2 (2007), pp. 313–331.

[26] P. E. Caines, F. H. Clarke, X. Liu, and R. B. Vinter, *A maximum principle for hybrid optimal control problems with pathwise state constraints*, in Proc. IEEE Conf. Decis. Control, 2006, pp. 4821–4825.

[27] F. Clarke, *Functional analysis, calculus of variations and optimal control*, vol. 264 of Graduate Texts in Mathematics, Springer, London, 2013.

[28] F. H. Clarke and R. B. Vinter, *Applications of optimal multiprocesses*, SIAM J. Control Optim., 27 (2006), pp. 1048–1071.

[29] L. Daniel, *Switching in systems and control*, vol. 190, Springer, 2003.

[30] A. Dmitruk, *On the development of Pontryagin's maximum principle in the works of a. Ya. Dubovitskii and AA Milyutin*, Control Cybernet., 38 (2009), pp. 923–957.

[31] A. Dmitruk and A. Kaganovich, *The hybrid maximum principle is a consequence of Pontryagin maximum principle*, Systems Control Lett., 57 (2008), pp. 964–970.

[32] A. V. Dmitruk and N. P. Osmolovskiĭ, *On the proof of the Pontryagin maximum principle by means of needle variations*, Fundam. Prikl. Mat., 19 (2014), pp. 49–73.

[33] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl., 47 (1974), pp. 324–353.

[34] H. O. Fattorini, *Infinite dimensional optimization and control theory*, Cambridge University Press, 1999.

[35] C. Ferrier and R. Epenoy, *Optimal control for engines with electro-ionic propulsion under constraint of eclipse*, Acta Astronautica, 48 (2001), pp. 181–192.

[36] A. F. Filippov, *On certain questions in the theory of optimal control*, Journal of the Society for Industrial and Applied Mathematics, Series A: Control, 1 (1962), pp. 76–84.

[37] E. Frazzoli, M. A. Dahleh, and E. Feron, *Robust hybrid control for autonomous vehicle motion planning*, in Proc. IEEE Conf. Decis. Control, vol. 1, 2000, pp. 821–826.

[38] M. Garavello and B. Piccoli, *Hybrid necessary principle*, SIAM J. Control Optim., 43 (2005), pp. 1867–1887.

[39] S. Geffroy and R. Epenoy, *Optimal low-thrust transfers with constraints—generalization of averaging techniques*, Acta Astronautica, 41 (1997), pp. 133–149.

[40] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid dynamical systems*, IEEE Control Syst., 29 (2009), pp. 28–93.

[41] K. F. Graham and A. V. Rao, *Minimum-time trajectory optimization of low-thrust earth-orbit transfers with eclipsing*, Journal of Spacecraft and Rockets, 53 (2016), pp. 289–303.

[42] T. Haberkorn and E. Trélat, *Convergence results for smooth regularizations of hybrid nonlinear optimal control problems*, SIAM J. Control Optim., 49 (2011), pp. 1498–1522.

[43] H. Halkin, *Implicit functions and optimization problems without continuous differentiability of the data*, SIAM J. Control Optim., 12 (1974), pp. 229–236.

[44] J. M. Holtzman and H. Halkin, *Discretional convexity and the maximum principle for discrete systems*, SIAM J. Control Optim., 4 (1966), pp. 263–275.

[45] J. A. Kechichian, *Low-thrust inclination control in presence of earth shadow*, Journal of Spacecraft and Rockets, 35 (1998), pp. 526–532.

[46] A. Korytowski, *A simple proof of the maximum principle with endpoint constraints*, Control Cybernet., 43 (2014), pp. 5–13.

[47] X. Liu and P. Stechlinkski, *Infectious disease modeling*, vol. 965, Springer, 2017.

[48] L. H. Loomis and S. Sternberg, *Advanced calculus*, Addison-Wesley, Reading, Mass., 1968.

[49] J. Lygeros, N. G. Datta, and S. Shankar, *Verified hybrid controllers for automated vehicles*, IEEE Trans. Automat. Contr., 43 (1998), pp. 522–539.

[50] A. Pakniyat and P. E. Caines, *On the hybrid minimum principle: the hamiltonian and adjoint boundary conditions*, IEEE Trans. Automat. Contr., 66 (2020), pp. 1246–1253.

[51] M. Pichan, H. Rastegar, and M. Monfared, *A new digital control of four-leg inverters in the natural reference frame for renewable energy–based distributed generation*, Int. Trans. Electr. Energy Syst., 29 (2019).

[52] L. S. Pontryagin, *Mathematical theory of optimal processes*, John Wiley & Sons, Inc. New York-London, 1962.

[53] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, Translated by D. E. Brown, A Pergamon Press Book. The Macmillan Co., New York, 1964.

[54] A. Rantzer and M. Johansson, *Piecewise linear quadratic optimal control*, IEEE Trans. Automat. Contr., 45 (2000), pp. 629–637.

[55] P. Riedinger, F. Kratz, C. Iung, and C. Zanne, *Linear quadratic optimization for hybrid systems*, in Proc. IEEE Conf. Decis. Control, vol. 3, 1999, pp. 3059–3064.

[56] P. Riedinger, C. lung, and F. Kratz, *An optimal control approach for hybrid systems*, Eur. J. Control, 9 (2003), pp. 449–458.

[57] R. T. Rockafellar and R. J. B. Wets, *Variational analysis*, vol. 317, Springer Science & Business Media, 2009.

[58] C. Silva and E. Trélat, *Smooth regularization of bang-bang optimal control problems*, IEEE Trans. Automat. Control, 55 (2010), pp. 2488–2499.

[59] M. Soler, A. Olivares, and E. Staffetti, *Hybrid optimal control approach to commercial aircraft trajectory planning*, J. Guid. Control Dyn., 33 (2010), pp. 985–991.

[60] H. J. Sussmann, *A nonsmooth hybrid maximum principle*, in Stability and stabilization of nonlinear systems, vol. 246, Springer London, 1999, pp. 325–354.

[61] C. Tomlin, G. Pappas, and S. Sastry, *Conflict resolution for air traffic management: a study in multiagent hybrid systems*, IEEE Trans. Automat. Contr., 43 (1998), pp. 509–521.

[62] V. I. Utkin, *Sliding modes in control and optimization*, in Comm. Control Engrg. Ser., 1992.

[63] A. J. Van Der Schaft and J. M. Schumacher, *An introduction to hybrid dynamical systems*, vol. 251, Springer London, 2000.

[64] R. Woollands and E. Taheri, *Optimal low-thrust gravity perturbed orbit transfers with shadow constraints*, in The 2019 AAS/AIAA astrodynamics specialist conference, 2019.

[65] X. Xu and P. J. Antsaklis, *An approach for solving general switched linear quadratic optimal control problems*, in Proc. IEEE Conf. Decis., vol. 3, 2001, pp. 2478–2483.

[66] F. Zhu and P. J. Antsaklis, *Optimal control of switched hybrid systems: a brief survey*, Technical Report of the ISIS Group at the University of Notre Dame, 3 (2011).