Hybrid maximum principle for regional optimal control problems with nonsmooth interfaces

Térence Bayen* Anas Bouali† Florent Nacry‡

September 9, 2024

Abstract

In this paper, we consider a general Mayer optimal control problem whose dynamics is defined regionally, and, additionally, we suppose that the interface between two regions is nonsmooth in the sense that it is described by a locally Lipschitz continuous function. Our objective is to derive a hybrid maximum principle in this setting. Doing so, we consider a sequence of mollifiers which allows us to approximate uniformly the interface between two regions by a sequence of smooth functions. This makes possible to apply the hybrid maximum principle on a sequence of approximated optimal control problems involving the smooth interface in place of the nonsmooth one. By passing to the limit as the approximation parameter tends to zero, we obtain the desired necessary optimality conditions in the form of a nonsmooth hybrid maximum principle. Our approach relies on the obtention of subgradients of a locally Lipschitz continuous function via mollifiers.

1 Introduction

The Pontryagin Maximum Principle (PMP in short) developed in the late 1950's in [20] was a major breakthrough in Mathematics particularly in terms of its many applications. It provides first order necessary optimality conditions for optimal control problems governed by ordinary differential equations extending that way the theory of calculus of variations. It has now been developed in other frameworks such as for problems governed by partial differential equations, ordinary differential equations with delays, and hybrid control systems, which is the main subject of this paper.

By hybrid control system, we mean a dynamical control system that has the particularity to be discontinuous with respect to the state, *i.e.*, the dynamics may change over the time frame. This discontinuity may have several origins. It can be controlled by an automaton like for switched systems (see, *e.g.*, [19]). In that case, switchings of the dynamics occur at so-called *switching times* that can be controlled. Another important framework where discontinuities of the dynamics arise is when the dynamical system is defined over a partition of the state space into a countable number of open regions (see, *e.g.*, [9]). The theory of hybrid systems is very broad and we restrict our attention in this paper only to those hybrid control systems that enter into the previous setting. In this context, the dynamics is thus discontinuous w.r.t. the state each time the trajectory crosses an interface between two regions. In order to characterize optimal solutions to optimal control problems governed by a hybrid control system, the PMP was extended to the hybrid setting leading to the co-called *Hybrid Maximum Principle* (HMP in short), see, *e.g.*, [5, 12, 16, 18, 21, 24]. Note that several settings for hybrid systems are considered in the litterature leading to various approaches. In the case where the hybrid control system is defined over a partition of the state space (as what is considered in this paper), we will refer to *regional optimal control problems*. For the derivation of the corresponding HMP in this context, we can cite for instance [3, 4, 5, 7, 18].

As a consequence of discontinuities (w.r.t. the state) arising in hybrid control systems, the covector arising in the HMP is no longer absolutely continuous (in contrast with the covector arising in the PMP), but only

*Avignon Université, Laboratoire de Mathématiques d'Avignon (EA 2151) F-84018. terence.bayen@univ-avignon.fr

†MISTEA, Univ. Montpellier, INRAE, Institut Agro, 34060 Montpellier, France. anas.bouali@inrae.fr

‡Laboratoire de Mathématiques et Physique, Université de Perpignan Via Domitia, Perpignan, France. florent.nacry@univ-perp.fr

piecewise absolutely continuous. More precisely, at each crossing time (between two regions), the covector has a discontinuity that is colinear to the outward unit vector to the interface at the crossing point. This condition requires the boundary of the interface to be locally the graph of a C^1 function (as it is usually done in this setting, see, *e.g.*, [3, 4, 5, 18]) in order to properly define the discontinuity condition of the covector. The objective of this paper is to address the case of an interface between two regions that is nonsmooth in the sense that it coincides with the graph of a locally Lipschitz continuous function (see, *e.g.*, [14, 15] in which related questions were studied for state constraints optimal control problems). Our aim here is to provide the corresponding necessary optimality conditions in this case. Regions with nonsmooth interfaces arise typically when a partition of the state space consists of polyhedra or intersections of half spaces. In that case, a trajectory may leave a region at some corner point, so that the HMP as developed for instance in [5] no longer applies.

To handle a nonsmooth interface between two regions, our methodology relies on a regularization technique of the interface. We consider a sequence of C^1 functions converging uniformly to the Lipschitz continuous function describing a nonsmooth interface. This sequence is defined thanks to mollifiers following the approach in [23]. Besides the uniform convergence, it is shown in [23] that subgradients of the original function can be recovered as the convex envelope of the limit gradients of the approximated function (it is also called *gradient consistency*, see [25]). This property is of utmost importance in our approach since it will allow us to extend the discontinuity condition at some point of non-differentiability of the interface. Thanks to this sequence that approaches the nonsmooth interface, we can associate a sequence of approximated optimal control problems. We know that in this case, the HMP as in [5] can be applied on the approximated optimal control problems. By letting the approximation parameter go to zero, we obtain our main result (Theorem 4.1) which provides first-order necessary optimality conditions on an optimal control. This result is obtained using the gradient consistency property of the sequence of regularizing functions which gives us the desired discontinuity condition on the covector.

For proving Theorem 4.1, we use a strong transverse hypothesis on trajectories. Roughly speaking, this hypothesis means that locally, every admissible trajectory crosses the interface transversally. We refer to [6] for a detailed presentation of varied transverse hypotheses in the present hybrid setting. This assumption is needed in particular to ensure that the structure of an optimal trajectory remains unchanged whenever the nonsmooth interface is replaced by a smooth one (which is a non-evident issue as far as we know). There is indeed in general no reason that the structure of an optimal trajectory remains the same whenever considering (small) perturbations of the interface, but this property is kept under a strong transverse hypothesis as we introduce in this paper.

The paper is organized as follows. In Section 2, we introduce some notation and the hybrid optimal control problem. In Section 3, we introduce a sequence of smooth functions that approaches the nonsmooth interface and we also define a sequence of optimal control problems governed by a perturbed differential inclusion (defined thanks to the sequence of approximating functions). To simplify the exposition, we have developed our approach in this paper in the case of only one nonsmooth interface delimiting two open regions, but, an extension of our main result (Theorem 4.1) is also given at the end of the paper in the case of a finite number of regions (section 4.3). In order to prove Theorem 4.1, we start by proving a stability result concerning solutions to the perturbed differential inclusions based on graphical convergence of sets (without the use of a transverse hypothesis). Next, we prove that under a strong transverse hypothesis, optimal solutions to the approximated optimal control problems converge (up to a subsequence) to an optimal solution to the original optimal control problem. Finally, in Section 4, we derive the nonsmooth HMP[1] following [8] to address convergence of covectors (coming from the application of the HMP to the sequence of approximated optimal control problems). The paper ends up with a list of remarks on the employed methodology and perspectives.

2 Statement of the hybrid optimal control problem

2.1 Notation and preliminaries

In this section, we introduce some notation and also recall concepts of nonsmooth analysis related to set convergence. The letter $\mathbb{N}$ stands for the set of natural integers starting from 0, that is, $\mathbb{N} = \{0, 1, 2, \ldots\}$. Let

[1]By nonsmooth HMP, we mean that regions are with nonsmooth boundaries.

$m, n \in \mathbb{N}^\star := \mathbb{N} \setminus \{0\}$. For any extended-real number $q \in [1, \infty]$ and any real interval $I \subset \mathbb{R}$, we denote by $\mathrm{L}^q(I, \mathbb{R}^n)$ the usual Lebesgue space of q-integrable functions defined on I with values in $\mathbb{R}^n$, endowed with its usual norm $\|\cdot\|_{\mathrm{L}^q}$. Throughout the paper, $T > 0$ is given. A sequence (x_k) of absolutely continuous functions from $[0, T]$ into $\mathbb{R}^n$ is said to strongly-weakly converge (over the interval $[0, T]$) to an absolutely continuous function $x^\star : [0, T] \to \mathbb{R}^n$ whenever (x_k) converges to $x^\star$ in $\mathrm{L}^\infty(I, \mathbb{R}^n)$ and $(\dot{x}_k)$ weakly converges in $\mathrm{L}^2(I, \mathbb{R}^n)$ to $\dot{x}^\star$. It will be convenient to denote (as usual) $\dot{x}_k \rightharpoonup \dot{x}^\star$ the latter weak convergence. For convenience, we shall also note indifferently x or $x(\cdot)$ a function $x : [0, T] \to \mathbb{R}^n$.

Next, $B(x_0, r)$ (resp. $B[x_0, r]$) stands for the open (resp. closed) ball centered at $x_0 \in \mathbb{R}^n$ with radius $r > 0$. The closed unit ball $B[0, 1]$ of $\mathbb{R}^n$ is denoted $\mathbb{B}_{\mathbb{R}^n}$. We also write $\cdot$ (resp. $|\cdot|$) for the standard inner product (resp. Euclidean norm) of $\mathbb{R}^n$. For any subset $X \subset \mathbb{R}^n$, we denote by ∂X the boundary of X defined by $\partial X := \overline{X} \backslash \mathrm{int}(X)$, where $\overline{X}$ and $\mathrm{int}(X)$ stand respectively for the closure and the interior of X. Given a function $\rho : \mathbb{R}^n \to \mathbb{R}$, we define (as usual) its support as the closure of the set $\{\rho \neq 0\}$. The convex hull of the set X is denoted by $\mathrm{co}(X)$. We now recall the following definitions related to subgradients and set convergence:

- the *Clarke subdifferential* of a locally-Lipschitz continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is given by (see [10]):
$$\partial_C f(x) = \mathrm{co}\,\{\lim_{k \to +\infty} \nabla f(x_k) \;;\; \Delta_f \ni x_k \to x\},$$
where Δ_f denotes the set of points $x \in \mathbb{R}^n$ where the (Fréchet) gradient $\nabla f(x)$ exists ;
- the *Clarke normal cone* of a set $S \subset \mathbb{R}^n$ at $x \in S$ is defined as the polar of the *Clarke tangent cone* $T^C(S; x)$, that is,
$$N^C(S; x) := \left\{\zeta \in \mathbb{R}^n \;;\; \langle \zeta, h\rangle \leq 0, \forall h \in T^C(S; x)\right\},$$
where
$$T^C(S; x) := \{h \in \mathbb{R}^n \;;\; \forall S \ni x_n \to x, \forall t_n \downarrow 0, \exists h_n \to h, \forall n \in \mathbb{N}, x_n + t_n h_n \in S\}\,;$$
- the *limit inferior* (resp., *limit superior*) of a sequence (S_k) of sets in $\mathbb{R}^n$ is defined as
$$\operatorname*{Lim\,inf}_{k \to +\infty} S_k := \{x \in \mathbb{R}^n \;;\; \forall \varepsilon > 0, \exists K \in \mathbb{N}, \forall k \geq K, S_k \cap B(x, \varepsilon) \neq \emptyset\},$$
(resp.,
$$\operatorname*{Lim\,sup}_{k \to +\infty} S_k := \{x \in \mathbb{R}^n \;;\; \forall \varepsilon > 0, \forall K \in \mathbb{N}, \exists k \geq K, S_k \cap B(x, \varepsilon) \neq \emptyset\}).$$

Whenever both sets coincides, one says that the sequence (S_k) *Painlevé-Peano-Kuratowski converges*. It is well-known (and not difficult to check) that $x \in \operatorname*{Lim\,sup}_{k \to +\infty} S_k$ if and only if there exist a sequence (x_k) converging to x and an increasing mapping $s : \mathbb{N} \to \mathbb{N}$ such that $x_k \in S_{s(k)}$ for all $k \in \mathbb{N}$. Such a sequential characterization of the upper-limit easily ensures ([23, Theorem 4.10]) that any (not necessarily closed) set $S \supset \operatorname*{Lim\,sup}_{k \to +\infty} S_k$ satisfies for all real $\rho > 0$ and all real $\varepsilon > 0$, the existence of $n_0 \in \mathbb{N}$ such that
$$S_n \cap \rho \mathbb{B}_{\mathbb{R}^n} \subset S + \varepsilon \mathbb{B}_{\mathbb{R}^n} \quad \text{for all } n \geq n_0. \tag{2.1}$$

Let $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a multimapping whose graph $\mathrm{gph}\, G$ is denoted by $\mathrm{gph}\, G := \{(x, y) \in \mathbb{R}^{2n} : y \in G(x)\}$. Recall that its graph is the set $\mathrm{gph}\, G := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \;;\; y \in G(x)\}$. A sequence of multimappings (G_k) from $\mathbb{R}^n$ into itself *graphically converges* to a multimapping $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ whenever the sequence of graphs $(\mathrm{gph}\, G_k)$ Peano-Painlevé-Kuratowski converges to $\mathrm{gph}\, G$. One says that G is *outer semicontinuous* at $x \in \mathbb{R}^n$ if for every sequence $x_k \to x$ and every sequence $y_k \to y$ such that $y_k \in F(x_k)$ for all $k \in \mathbb{N}$, then one has $y \in F(x)$. Obviously, the multimapping G is outer semicontinuous at each point of $\mathbb{R}^n$ if and only if its graph is closed in $\mathbb{R}^n \times \mathbb{R}^n$. One says that the multimapping M is *upper semicontinuous* at x whenever for each open set V of $\mathbb{R}^n$ with $G(x) \subset V$ there is a real $\varepsilon > 0$ such that
$$G\big(B(x, \varepsilon)\big) \subset V.$$

Recall that (see, e.g., [**?**, Proposition 1.39]) the outer semicontinuity of G at $x \in \mathbb{R}^n$ is equivalent to its upper semicontinuity at x provided that there are V a neighborhood of x in $\mathbb{R}^n$ and a compact $K \subset \mathbb{R}^n$ such that M is closed-valued on V and $M(V) \subset K$.

2.2 Problem statement

In this section, we introduce the optimal control problem studied throughout this paper. In the sequel, we consider two functions $f_1, f_2 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ of class C^1 (reprensenting the dynamics) and a nonempty compact subset U of $\mathbb{R}^m$. We suppose that for $i = 1, 2$ and that for every $x \in \mathbb{R}^n$, the velocity set

$$\widehat{F}_i(x) := \{f_i(x, u) \; ; \; u \in U\} \tag{2.2}$$

is a nonempty compact and convex subset in $\mathbb{R}^n$. For technical reasons in the proof of the HMP (based on an approximating sequence of the nominal control), we also suppose that for $i = 1, 2$ and that for every $(x, p) \in \mathbb{R}^n \times \mathbb{R}^n$, the set

$$\widehat{G}_i(x, p) := \{D_x f_i(x, u)^\top p \; ; \; u \in U\} \tag{2.3}$$

is a nonempty compact convex subset in $\mathbb{R}^n$ where as usual $D_x f_i(x, u)$ stands for the Jacobian matrix of $f_i(\cdot, u)$ at x, $i = 1, 2$, and where $^\top$ denotes the transpose operator. Note that the preceding convexity hypotheses are automatically verified by affine-control systems w.r.t. the control which already represent a large class of control systems in application models. Additionally, we suppose (in order to prevent blow-up phenomenon of trajectories) that for $i = 1, 2$, the function f_i has a linear growth, *i.e.*, there is $c \geq 0$ such that

$$|f_i(x, u)| \leq c(|x| + 1) \quad \text{for all } (x, u) \in \mathbb{R}^n \times U. \tag{2.4}$$

Let also Γ be a subset of $\mathbb{R}^n$ for which there is a locally Lipschitz continuous function $\psi : \mathbb{R}^n \to \mathbb{R}$ such that Γ is the zero level-set of ψ, *i.e.*,

$$\Gamma := \{x \in \mathbb{R}^n \; ; \; \psi(x) = 0\}.$$

In what follows, we set $X_1 := \{x \in \mathbb{R}^n \; ; \; \psi(x) < 0\}$ and $X_2 := \{x \in \mathbb{R}^n \; ; \; \psi(x) > 0\}$. We also consider a function $g : \mathbb{R}^n \to \mathbb{R}$ of class C^1 representing the terminal payoff. Finally, the admissible control set is the set $\mathcal{U}$ that contains all measurable functions $u : [0, T] \to \mathbb{R}^n$ such that $u(t) \in U$ for a.e. $t \in [0, T]$. Throughout this paper, we study the Mayer optimal control problem

$$\inf_{u \in \mathcal{U}} g(x(T)), \tag{2.5}$$

where x is a solution to the Cauchy problem

$$\left\{ \begin{array}{ll} \dot{x}(t) & = f(x(t), u(t)) \quad \text{a.e. } t \in [0, T], \\ x(0) & = x_0, \end{array} \right. \tag{2.6}$$

with $x_0 \in X_1$ and where the hybrid dynamics $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is defined regionally as follows:

$$\forall (x, u) \in \mathbb{R}^n \times \mathbb{R}^m, \quad f(x, u) := \left\{ \begin{array}{lll} f_1(x, u) & \text{if} & \psi(x) > 0, \\ f_2(x, u) & \text{if} & \psi(x) < 0. \end{array} \right. \tag{2.7}$$

Note that f is not defined on $\Gamma \times \mathbb{R}^m$. So, in absence of a transverse hypothesis on trajectories (see, *e.g.*, [6] for a presentation of weak and strong transverse hypotheses on trajectories), then, the differential inclusion formulation (as presented below) is more adapted in such a way to define a solution to (2.6). However, note that it is still possible to consider solutions to (2.6) provided that a transverse hypothesis on trajectories is introduced which is what we do starting from Section 3.3. Such transverse hypotheses exclude sliding modes (see [19]). At this step, we do not consider yet such a transverse hypothesis, so, we consider the Cauchy problem for the following differential inclusion

$$\left\{ \begin{array}{ll} \dot{x}(t) & \in F(x(t)) \quad \text{a.e. } t \in [0, T], \\ x(0) & = x_0, \end{array} \right. \tag{2.8}$$

where $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the Filippov multimapping associated with f which is defined for all $x \in \mathbb{R}^n$ by

$$F(x) := \left\{ \begin{array}{lll} \hat{F}_1(x) & \text{if} & \psi(x) > 0, \\ \hat{F}_2(x) & \text{if} & \psi(x) < 0, \\ \hat{F}_{1,2}(x) & \text{if} & \psi(x) = 0, \end{array} \right. \tag{2.9}$$

where for all $x \in \mathbb{R}^n$, the set $\widehat{F}_{1,2}(x)$ is the closed convex hull of $\widehat{F}_1(x)$ and $\widehat{F}_2(x)$. For more details on the Filippov multimapping, we refer to [13]. Having in mind that the multimappings $\widehat{F}_1$ and $\widehat{F}_2$ are convex-valued, one can easily check that for every $x \in \mathbb{R}^n$, one has

$$\widehat{F}_{1,2}(x) = \{\theta f_1(x,u) + (1-\theta) f_2(x,v) \; ; \; \theta \in [0,1] \; ; \; u, v \in U\}, \tag{2.10}$$

which is thus a non-empty closed compact convex subset of $\mathbb{R}^n$. It is readily seen that F is with nonempty convex compact values. It is also not difficult to check that F is outer semicontinuous, hence it is upper semicontinuous. This and the fact that the mapping F is with linear growth allows us to get the existence of an absolutely continuous function $x : [0,T] \to \mathbb{R}^n$ satisfying (2.8). We refer for instance to [1, Theorem 3, Chapter 2]) for more details on this point. In the sequel, we denote by $\mathcal{S}_T(x_0)$ all solutions to (2.8). We can now state an existence result related to (2.5).

Proposition 2.1. *There exists a solution to the optimal control problem*

$$\inf_{x(\cdot) \in \mathcal{S}_T(x_0)} g(x(T)). \tag{2.11}$$

Proof. The existence of a solution to (2.11) follows from the continuity of g and [2, Theorem 19.2.3] (see also [10, Theorem 1.11]) taking a minimizing sequence. $\square$

Our objective is to provide necessary optimality conditions for (2.5). Since the function ψ is only Lipschitz-continuous w.r.t. the state, we cannot directly apply the HMP as in [5] or in [18] (because if $x \in \Gamma$, $\partial_C \psi(x)$ may not be a singleton). To derive a nonsmooth HMP, we will proceed in two steps. First, we shall introduce a regularization of Γ and prove a stability result of optimal solutions to the auxiliary optimal control problems (Section 3). Next, we will apply the HMP on this sequence of problems for which the boundary is smooth, and necessary optimality conditions for (2.5) will be obtained by passing to the limit as the regularization parameter tends to zero (Section 4).

3 Regularization of the interface and convergence analysis

3.1 Definition of approximated optimal control problems

In this section, we introduce a sequence of optimal control problems approaching (2.5) that are based on the regularization of the function ψ describing the interface between X_1 and X_2. Doing so, we shall next recall [23, Theorem 9.67] by Rockafellar and Wets. This theorem is about the convergence of the gradients of a sequence of smooth functions (defined by convolution) approximating uniformly a locally Lipschitz-continuous function. Let $\rho_k : \mathbb{R}^n \to \mathbb{R}_+$ be a sequence of bounded and continuous mollifiers[2] satisfying $\int_{\mathbb{R}^n} \rho_k(y)\,\mathrm{d}y = 1$ and $\mathrm{supp}(\rho_k) \subset B[0, \varepsilon_k]$ for any k, where $\varepsilon_k \downarrow 0$ as $k \to +\infty$ and let $\psi_k(x) := \int_{\mathbb{R}^n} \psi(x-y)\rho_k(y)\,\mathrm{d}y$ denotes the convolution of ψ with $(\rho_k)_k$. Following [23, Theorem 9.67] (see also [25]), the following properties are fulfilled:

- the sequence (ψ_k) uniformly converges to ψ over every compact subset of $\mathbb{R}^n$;
- the gradient consistency (allowing to obtain subgradients of ψ via mollifiers) holds true:

$$\partial_C \psi(x) = \mathrm{co} \left\{ \lim_{k \to +\infty} \nabla \psi_k(x_k) \; ; \; x_k \to x \right\}. \tag{3.1}$$

In order to approach (2.5) by a sequence of optimal control problems involving a smooth boundary, we proceed as follows. For $k \in \mathbb{N}$, define a smooth interface

$$\Gamma_k := \{x \in \mathbb{R}^n \; ; \; \psi_k(x) = 0\},$$

[2] *e.g.*, $\rho_k(y) := \frac{1}{\varepsilon_k^n}\rho(\frac{y}{\varepsilon_k})$ where $\rho(y) := C\exp(-\frac{1}{1-|y|^2})$ for $|y| < 1$ and $\rho(y) := 0$ for $|y| \geq 1$ with $C > 0$ such that $\int_{\mathbb{R}^n} \rho = 1$.

as well as the set-valued map $F_k : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$

$$F_k(x) := \begin{cases} \hat{F}_1(x) & \text{if} \quad \psi_k(x) > 0, \\ \hat{F}_2(x) & \text{if} \quad \psi_k(x) < 0, \\ \hat{F}_{1,2}(x) & \text{if} \quad \psi_k(x) = 0. \end{cases} \tag{3.2}$$

Similarly as for F, the set-valued map F_k is upper semi-continuous for every $k \in \mathbb{N}$. In addition, for every $k \in \mathbb{N}$ and for every $x \in \mathbb{R}^n$, $F_k(x)$ is a nonempty compact convex subset of $\mathbb{R}^n$. Since F_k is with linear growth (because of (2.4)), for every $k \in \mathbb{N}$, the differential inclusion

$$\begin{cases} \dot{x}(t) & \in F_k(x(t)) \quad \text{a.e. } t \in [0,T], \\ x(0) & = x_0, \end{cases} \tag{3.3}$$

has an absolutely continuous solution over $[0,T]$ denoted by x_k. Hence, using a similar argumentation as in the proof of Proposition 2.1, for every $k \in \mathbb{N}$, the optimal control problem

$$\inf_{x(\cdot) \in \mathcal{S}_T^k(x_0)} g(x_k(T)), \tag{3.4}$$

has a solution x_k^* where for any $k \in \mathbb{N}$, the set $\mathcal{S}_T^k(x_0)$ denotes all solutions to (3.3).

3.2 A stability result for hybrid differential inclusions

In this section, we address the following question : given a family of solutions to (3.3), does there exists a solution x to (2.8) such that, up to a subsequence, (x_k) strongly-weakly converges to x? Such a problem is developed in the literature through different approaches and under various assumptions. For the case of one Cauchy problem in $\mathbb{R}^n$, say $F_k \equiv F$, we refer the reader to the result [10, Chapter 4-Theorem 1.11] (which is based on Arzelà-Ascoli theorem). In [22], the strong-weak convergence is studied in the context of a general Banach space for a sequence of Cauchy problems under an appropriate Kuratowski-Mosco convergence (namely, $F_k(x_k) \to F(x)$ for every sequence (x_k) such that $x_k \rightharpoonup x$). As far as we know, the hypotheses required to apply [22] are not completely fulfilled by (F_k) and F as defined in (2.9) and (3.2). Another possible way is to follow the approach in [17] based on graphical convergence of sets.

Definition 3.1. *Let (G_k) be a sequence of multimappings from $\mathbb{R}^n$ into itself. The* (graphical) upper-limit *of the sequence (G_k) is the multimapping $G^\# : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by*

$$\mathrm{gph} G^\# := \operatorname*{Lim\,sup}_{k \to +\infty} \mathrm{gph} G_k.$$

According to the above sequential characterization of the upper-limit, it is readily seen that $(x,y) \in \mathrm{gph} G^\#$ if and only if there are two sequences (x_k) and (y_k) satisfying $x_k \to x$, $y_k \to y$ and an increasing function $s : \mathbb{N} \to \mathbb{N}$ such that $(x_k, y_k) \in \mathrm{gph} F_{s(k)}$ (or equivalently, $y_k \in F_{s(k)}(x_k)$) for all $k \in \mathbb{N}$.

Proposition 3.1. *Given a sequence (x_k) such that $x_k \in \mathcal{S}_T^k(x_0)$ for every $k \in \mathbb{N}$, there is $x \in \mathcal{S}_T(x_0)$ such that, up to a subsequence, (x_k) strongly-weakly converges to x.*

Proof. Let (x_k) be a sequence of functions from $[0,T]$ into $\mathbb{R}^n$ satisfying $x_k \in \mathcal{S}_T^k(x_0)$ for every $k \in \mathbb{N}$. Thanks to (2.4), there exists $c' \geq 0$ such that for all $k \in \mathbb{N}$, one has

$$|\dot{x}_k(t)| \leq c' \quad \text{a.e. } t \in [0,T].$$

This implies that (x_k) is uniformly bounded in $L^\infty([0,T], \mathbb{R}^n)$. Using the Arzelà-Ascoli theorem, we can find an absolutely continuous function $x : [0,T] \to \mathbb{R}^n$ such that (x_k) (up to a subsequence) uniformly converges to x and $(\dot{x}_k)$ weakly converges to $\dot{x}$ in $L^2([0,T], \mathbb{R}^n)$. Pick any real $\rho > 0$ such that for all $k \in \mathbb{N}$ and for almost every $t \in [0,T]$ one has

$$\big(x_k(t), \dot{x}_k(t)\big) \in \mathrm{gph}\, F_k \cap \rho \mathbb{B}_{\mathbb{R}^{2n}}.$$

Given any real $\varepsilon > 0$ we easily derive from the inclusion $\operatorname*{Lim\,sup}_{k \to +\infty} \mathrm{gph}\, F_k \subset \mathrm{gph}\, F$ (see (2.1)) that

$$\mathrm{gph}\, F_k \cap \rho \mathbb{B}_{\mathbb{R}^{2n}} \subset \mathrm{gph}\, F + \varepsilon \mathbb{B}_{\mathbb{R}^{2n}}.$$

It remains to apply [1, Theorem 1, Chapter 1] to obtain the inclusion $x \in \mathcal{S}_T(x_0)$. □

3.3 Convergence of optimal solutions to (3.4)

In this subsection, we introduce a hypothesis related to the behavior of trajectories at the interface. This will ensure convergence (up to a subsequence) of the sequence (x_k^*) to an optimal solution to (2.11). Doing so, we need to prove that every admissible solution x of (2.8) can be approached by a sequence (x_k) satisfying (3.3) (we refer to Proposition 3.2 for details). In order to prove such a result, we introduce the following transverse hypothesis (which is standard in a spatially hybrid setting [6, 18]). Let us recall that the initial condition x_0 and final time $T > 0$ are fixed and since both dynamics f_1 and f_2 satisfy a linear growth hypothesis, there is $R \geq 0$ such that for all $x(\cdot) \in \mathcal{S}_T(x_0)$, for every $t \in [0, T]$, one has $x(t) \in B[0, R]$.

Assumption 3.1. *There is $\kappa > 0$ such that*

$$\forall (x, u) \in (\Gamma \cap B[0, R]) \times U,\ \forall v \in \partial_C \psi(x), \quad v \cdot f_i(x, u) \geq \kappa, \tag{3.5}$$

where $i = 1, 2$.

Following [6], this assumption is called *strong transverse hypothesis*. Under this hypothesis, the differential inclusion in (2.8) can be replaced by the control system in (2.6) since (4.16) excludes sliding modes. In particular, Problem (2.5) and Problem (2.11) are equivalent. In the sequel, we suppose that this hypothesis is fulfilled.

Definition 3.2. *Given a solution x to the differential inclusion* (2.8)*, one says that a real $\tau \in (0, T)$ is a crossing time from X_1 to X_2 whenever there exists a real $\alpha > 0$ (small enough) such that*

$$\big(\psi(x(t)) - \psi(x(\tau))\big)(t - \tau) > 0 \quad \textit{for all } t \in [\tau - \alpha, \tau + \alpha] \backslash \{\tau\}.$$

A similar definition with Γ_k and (3.3) in place of Γ and (2.8) is omitted. Obviously, if τ is a crossing time of $x(\cdot) \in \mathcal{S}_T(x_0)$, then, one must have $\psi(x(\tau)) = 0$.

Lemma 3.1. *Every admissible trajectory of* (2.5) *has at most one crossing time on Γ from X_1 to X_2.*

Proof. First, every crossing time of an admissible solution $x \in \mathcal{S}_T(x_0)$ is isolated. Indeed, let $\tau \in (0, T)$ be a crossing time of $x(\cdot)$, it follows that $\psi(x(\tau)) = 0$. In addition, by composition, the mapping $t \mapsto \psi(x(t))$ is absolutely continuous. Moreover, one has $\frac{\mathrm{d}}{\mathrm{d}t}\psi(x(t)) = \nabla\psi(x(t)) \cdot \dot{x}(t)$ for almost every $t \in [0, T]$. Now, from (4.16), we get the existence of $\eta > 0$ such that for a.e. $t \in [\tau, \tau + \eta]$,

$$\nabla\psi(x(t)) \cdot f_2(x(t), u(t)) \geq \frac{\kappa}{2}.$$

Hence, $t \mapsto \psi(x(t))$ is increasing in a right neighborhood of τ, whence the result. Suppose now that the mapping x has a crossing time and set

$$\tau := \min\{t \in [0, T]\ ;\ \psi\big(x(t)\big) = 0\}.$$

If there is $\tau' \in (\tau, T)$ such that $\psi(x(\tau')) = 0$ and $x(t) \in X_2$ for $t \in (\tau, \tau')$, then, one has $\psi(x(t)) > 0$ for every $t \in (\tau, \tau')$. Take $r > 0$ small enough. We have (keeping in mind that $\psi \circ x$ is absolutely continuous on $[0, T]$)

$$\psi\big(x(\tau')\big) - \psi\big(x(\tau' - r)\big) = \int_{\tau'-r}^{\tau'} \nabla\psi\big(x(t)\big) \cdot \dot{x}(t)\, \mathrm{d}t = \int_{\tau'-r}^{\tau'} \nabla\psi\big(x(t)\big) \cdot f_2(x(t), u(t))\, \mathrm{d}t.$$

On the one hand, for r small enough, the integral above is bounded below by a positive constant (using the transversality condition and a continuity argumentation). On the other hand,

$$\psi(x(\tau')) - \psi(x(\tau' - r)) = -\psi(x(\tau' - r)) < 0,$$

for every $r > 0$ small enough. This is thus a contradiction which ends the proof. □

The preceding property can be transferred to admissible trajectories of (3.4) for k large enough as we show in the next lemma.

Lemma 3.2. *For k large enough, every admissible solution of* (3.3) *has zero or one crossing time on Γ_k from X_1 to X_2.*

Proof. Suppose that for every $k \in \mathbb{N}$, there are $n_k \in \mathbb{N}$, $x_{n_k} \in \Gamma_{n_k} \cap B[0,R]$, $u_{n_k} \in U$ such that

$$\nabla\psi_{n_k}(x_{n_k}) \cdot f_i(x_{n_k}, u_{n_k}) < \frac{\kappa}{2},$$

where $i = 1,2$ is fixed. Since U is compact, we may assume that (u_{n_k}) converges to some $u \in U$ and that (x_{n_k}) also converges to some vector $\tilde{x} \in \Gamma$. By using the gradient consistency property, we deduce that (up to a subsequence), there exists $v \in \partial_C\psi(\tilde{x})$ such that $v = \lim_{k\to+\infty} \nabla\psi_{n_k}(x_{n_k})$. It follows that $v \cdot f_i(\tilde{x}, u) \leq \frac{\kappa}{2}$ which contradicts assumption 4.16. Hence, we deduce that there is $k_0 \in \mathbb{N}$ such that for every $k \geq k_0$, one has:

$$\forall x \in \Gamma_k \cap B[0,R],\ \forall u \in U,\ \nabla\psi_k(x) \cdot f_i(x,u) \geq \frac{\kappa}{2}, \tag{3.6}$$

for $i = 1,2$. Let $k \geq k_0$ and $x \in \mathcal{S}_T^k(x_0)$. Suppose that x has at least two crossing times $\tau < \tau'$ (isolated because of (3.6)) such that $\psi_k(x(\tau)) = \psi_k(x(\tau')) = 0$ so that $x(t) \in X_2$ for $t \in (\tau, \tau')$. We thus have $\psi(x(t)) > 0$ for every $t \in (\tau, \tau')$. By taking r small enough, we find that

$$\forall k \geq k_0,\ \psi_k(x(\tau')) - \psi_k(x(\tau' - r)) = \int_{\tau'-r}^{\tau'} \nabla\psi_k(x(t)) \cdot f_2(x(t), u(t))\,\mathrm{d}t \geq \frac{r\kappa}{4} > 0\,.$$

But, in the preceding equality, one has $\psi_k(x(\tau')) - \psi_k(x(\tau' - r)) = -\psi_k(x(\tau' - r)) < 0$ which is a contradiction. This ends the proof. $\square$

Proposition 3.2. *Up to a subsequence, (x_k^*) strongly-weakly converges to an optimal solution x^* of* (2.5).

Proof. We know from Proposition 3.1 that there is an admissible solution x^* to (2.5) such that, up to a subsequence, (x_k^*) strongly-weakly converges to x^*. Now, let us consider an admissible solution x of (2.5) associated with some control $u \in \mathcal{U}$. We define $x_k : [0,T] \to \mathbb{R}^n$ as the unique solution to the system

$$\left\{ \begin{array}{lll} \dot{x}_k(t) = f_1(x_k(t), u(t)) & \text{if} & \psi_k(x_k(t)) < 0, \\ \dot{x}_k(t) = f_2(x_k(t), u(t)) & \text{if} & \psi_k(x_k(t)) > 0, \end{array} \right. \tag{3.7}$$

such that $x_k(0) = x_0$. Let us prove that (x_k) strongly-weakly converges to x over $[0,T]$. If x has no crossing time, the result is obvious. Suppose then that x has a (single) crossing time $\tau \in (0,T)$. It follows from Lemma 3.2 that for k large enough, x_k also has a unique crossing time over $(0,T)$. Up to a subsequence, we may assume that (x_k) strongly-weakly converges to some absolutely continuous function $\hat{x}$ over $[0,T]$ where $\hat{x}$ satisfies (2.8) (Proposition 3.1). From the control system (3.7), we easily get that $\hat{x}(t) = x(t)$ for every $t \in [0, \tau - \alpha]$ for every $\alpha > 0$ small enough. It follows that $\hat{x} = x$ over $[0,\tau]$ (using the continuity of x and $\hat{x}$ at $t = \tau^-$). Next, $\hat{x}$ and x satisfy the same ordinary differential equation over $[\tau, T]$. Since $\hat{x}(\tau) = x(\tau)$, we get that $\hat{x} = x$ over $[\tau, T]$. This proves that (x_k) strongly-weakly converges to x over $[0,T]$.

To conclude the proof, take any admissible trajectory x of (2.5) and construct an admissible sequence (x_k) for (3.4) by (3.7). Since we have

$$g(x_k^*(T)) \leq g(x_k(T))$$

for every $k \in \mathbb{N}$, we deduce that $g(x^*(T)) \leq g(x(T))$ by letting $k \to +\infty$ and using the continuity of g. Since x is any admissible trajectory of (2.5), the result follows. $\square$

Remark 3.1. *In presence of mixed initial-terminal constraints, Proposition 3.2 may no longer hold true unless additional controllability hypotheses on the dynamics are imposed (in such a way that every admissible solution to the original problem can be approached by admissible solutions to the approximated problems).*

4 Nonsmooth hybrid maximum principle

Our goal now is to exploit optimality conditions on (3.4) to deduce optimality conditions on (2.5). Let us then consider $x^* \in \mathcal{S}_T(x_0)$ an optimal trajectory of (2.5). We suppose that x^* has a single crossing time $\tau^* \in (0,T)$ from X_1 to X_2 (remind Lemma 3.1). By using the results of the previous section, for every $k \in \mathbb{N}$, there exists a solution x_k^* to (3.4) such that (up to a subsequence), (x_k^*) strongly-weakly converges to x^*. So, for every $k \in \mathbb{N}$, there is an addmissible control $u_k^* \in \mathcal{U}$ associated with x_k^*. Our next step is to apply the HMP from [5] on x_k^* (for k large enough according to Lemma 3.2) which is possible since Γ_k is smooth.

4.1 Application of the HMP on (3.4) and consequences

Since x^* has a single crossing time, for every k large enough, x_k^* also has a single crossing time τ_k and by uniform convergence of (x_k^*), we get that $\tau_k \to \tau^*$ as $k \to +\infty$. We can now straightforwardly apply the HMP from [5] on (3.4). For every k large enough, there is a piecewise absolutely continuous function $p_k : [0,T] \to \mathbb{R}^n$ (a covector) such that the following properties are fulfilled:

(i) Adjoint equation :

$$\dot{p}_k(t) = -D_x f(x_k^*(t), u_k^*(t))^\top p_k(t) \quad \text{a.e. } t \in [0,T], \tag{4.1}$$

(ii) Terminal condition :

$$p_k(T) = -\nabla g(x_k^*(T)), \tag{4.2}$$

(iii) Hamiltonian maximization condition :

$$u_k^*(t) \in \operatorname{argmax}_{\omega \in \mathrm{U}} p_k(t) \cdot f(x_k^*(t), \omega) \quad \text{a.e. } t \in [0,T], \tag{4.3}$$

(iv) Discontinuity condition of the covector : there is $\nu_k \in \mathbb{R}$ such that

$$p_k^+(\tau_k) - p_k^-(\tau_k) = \nu_k \nabla \psi_k(x_k^*(\tau_k)). \tag{4.4}$$

Observe that the adjoint equation (4.1) can be rewritten as

$$\left\{ \begin{array}{lll} \dot{p}_k(t) = -D_x f_1(x_k^*(t), u_k^*(t))^\top p_k(t) & \text{for} & \text{a.e. } t \in (0,\tau_k), \\ \dot{p}_k(t) = -D_x f_2(x_k^*(t), u_k^*(t))^\top p_k(t) & \text{for} & \text{a.e. } t \in (\tau_k, T). \end{array} \right. \tag{4.5}$$

Finally, the Hamiltonian function $t \mapsto p_k(t) \cdot f(x_k^*(t), u_k^*(t))$ is constant almost everywhere from [5] so that there is $c_k \in \mathbb{R}$ such that

$$p_k(t) \cdot f_1(x_k^*(t), u_k^*(t)) = p_k(t') \cdot f_2(x_k^*(t'), u_k^*(t')) = c_k, \tag{4.6}$$

for a.e. $(t,t') \in (0,\tau_k) \times (\tau_k, T)$. The main issue now is to prove the uniform boundedness of (p_k) in order to prove the existence of a covector associated with x^* that satisfies the adjoint equation together with a discontinuity condition at the crossing time.

Proposition 4.1. *The sequence (ν_k) is bounded.*

Proof. By contradiction, suppose that $|\nu_k| \to +\infty$. Observe first that the sequences $(g(x_k(T)))$ is bounded. It follows using (4.1) backward in time together with (4.2) that $(p_k(\tau_k^+))$ is also bounded. We deduce that there is a constant $M \geq 0$ such that

$$\forall k \in \mathbb{N},\ \forall t \in (\tau_k, T], \quad |p_k(t)| \leq M. \tag{4.7}$$

Now, from the uniform convergence of (x_k^*) to x^* over $[0,T]$, we get that $\psi(x_k^*(\tau_k)) \to \psi(x^*(\tau^*))$ as $k \to +\infty$. Using the gradient consistency property (3.1), we may assume that there exists $v \in \partial_C \psi(x^*(\tau^*))$ (with $v \neq 0$ because of (4.16)) such that

$$\nabla \psi_k(x_k^*(\tau_k)) \to v,$$

as $k \to +\infty$ (extracting a subsequence if necessary). Let us now prove that the sequence $(p_k(\tau_k^-))$ is unbounded. By contradiction, if $(p_k(\tau_k^-))$ is bounded, then, dividing (4.4) by ν_k and letting $k \to +\infty$ yields that $v = 0$. Since $x^*(\tau^*) \in \Gamma$, this contradicts (4.16), hence, $(p_k(\tau_k^-))$ is unbounded so that we may assume that $|p_k(\tau_k^-)| \to +\infty$ as $k \to +\infty$ (extracting a subsequence if necessary). For every $k \in \mathbb{N}$, the sequence $(p_k(\tau_k^-)/|p_k(\tau_k^-)|)$ belongs to the unit sphere of $\mathbb{R}^n$ so that (extracting a subsequence if necessary), we may assume that there is $\xi \in \mathbb{R}^n$ such that $|\xi| = 1$ and $p_k(\tau_k^-)/|p_k(\tau_k^-)| \to \xi$ as $k \to +\infty$. Dividing now (4.4) by $|p_k(\tau_k^-)|$ yields

$$\frac{p_k(\tau_k^+)}{|p_k(\tau_k^-)|} - \frac{p_k(\tau_k^-)}{|p_k(\tau_k^-)|} = \frac{\nu_k}{|p_k(\tau_k^-)|}\nabla\psi_k(x_k^*(\tau_k))$$

Since the left member of the preceding equality goes to $-\xi$ and $\nabla\psi_k(x_k^*(\tau_k)) \to v$ as $k \to +\infty$, we obtain that the sequence $(\nu_k/|p_k(\tau_k^-)|)$ is bounded. Hence, extracting a subsequence if necessary, we may suppose that there is $r \in \mathbb{R}$ such that $(\nu_k/|p_k(\tau_k^-)|) \to r$ as $k \to +\infty$. It follows that

$$-\xi = rv. \tag{4.8}$$

On the other hand, for all $k \in \mathbb{N}$, $t \mapsto u_k^*(t)$ is a Lebesgue measurable function such that $u_k^*(t) \in U$ for a.e. $t \in [0, T]$. Since U is bounded, $u_k^* \in L^1([0, 1], \mathbb{R}^m)$, thus, almost every point of $[0, T]$ is a Lebesgue point of u_k^*. It follows that for every $k \in \mathbb{N}$, there exists $t_k \in (0, \tau_k)$ such that t_k is a Lebesgue point of u_k^*, $u_k(t_k) \in U$, and moreover, we may suppose that $|t_k - \tau_k| \to 0$ as $k \to +\infty$. Now, observe that one has:

$$\begin{aligned}\frac{c_k}{|p_k(\tau_k^-)|} &= \frac{p_k(t_k)}{|p_k(\tau_k^-)|} \cdot f_1(x_k^*(t_k), u_k^*(t_k)) \\ &= \frac{(p_k(t_k) - p_k(\tau_k^-))}{|p_k(\tau_k^-)|} \cdot f_1(x_k^*(t_k), u_k^*(t_k)) + \frac{p_k(\tau_k^-)}{|p_k(\tau_k^-)|} \cdot f_1(x_k^*(t_k), u_k^*(t_k)).\end{aligned} \tag{4.9}$$

Next, $c_k = p_k(t) \cdot f_2(x_k(t), u_k^*(t))$ for a.e. $t \in (\tau_k, T)$. Hence, taking for every $k \in \mathbb{N}$ a Lebesgue point of u_k^* in (τ_k, T) such that $u_k^*(t) \in U$ and using (4.7), we deduce that the sequence (c_k) is bounded. Additionally, by using (4.1) backward in time and Gronwall's Lemma, it follows that

$$\forall t \in [0, \tau_k],\ |p_k(t)| \leq c_1|p_k(\tau_k^-)|,$$

where $c_1 \geq 0$ is a constant. We deduce from (4.1) that $|\dot{p}_k(t)| \leq c_1'|p_k(\tau_k^-)|$ for every $t \in [0, \tau_k]$ where $c_1' \geq 0$ is a constant. Finally, we get that

$$|p_k(t_k) - p_k(\tau_k^-)| \leq c_1'|p_k(\tau_k^-)||t_k - \tau_k|.$$

Combining the preceding inequality with (4.9), we obtain that

$$\frac{p_k(\tau_k^-)}{|p_k(\tau_k^-)|} \cdot f_1(x_k^*(t_k), u_k^*(t_k)) = o(1),$$

where $o(1) \to 0$ as $k \to +\infty$. Extracting a subsequence if necessary (remind that U is compact), we may suppose that there is $\omega \in U$ such that $u_k^*(t_k) \to \omega$ as $k \to +\infty$. Hence, we get that

$$\xi \cdot f_1(x^*(\tau^*), \omega) = 0. \tag{4.10}$$

Combining (4.8) and (4.10), we find that $v \cdot f_1(x^*(\tau^*), \omega) = 0$. This equality is a contradiction with the transverse hypothesis (4.16) since $v \in \partial_C\psi(x^*(\tau^*))$. This ends the proof. □

Remark 4.1. *This proof mainly relies on the (strong) transverse hypothesis* (4.16) *made on the dynamics at the interface* Γ. *Note that we could not directly handle* $f_i(x_k^*(\tau_k), u_k^*(\tau_k))$ *for* $i = 1, 2$ *since* u_k^* *is only defined almost everywhere. Our methodology thus consisted in considering an optimal solution to* (3.4) *for which we have no information on the regularity of the control* u_k^*. *This made our approach more delicate at this step of the proof. But, note that if* u_k^* *is right and left continuous at* $t = \tau_k$, *then, following* [5], *we can show that*

$$\nu_k = \frac{p_k^+(\tau_k) \cdot (f(x_k^*(\tau_k), u_k^*(\tau_k^-)) - f(x_k^*(\tau_k), u_k^*(\tau_k^+)))}{\nabla\psi_k(x_k(\tau_k)) \cdot f(x_k^*(\tau_k), u_k(\tau_k^-))},$$

using the Hamiltonian constancy. Combining this expression together with the transverse hypothesis, we would more easily deduce that (ν_k) *is bounded. Unfortunately, our approach does not allow us a priori to use this expression since we do not know in advance if* u_k^* *is defined at* $t = \tau_k^{\pm}$.

We can now prove the following convergence result on the sequence (p_k).

Proposition 4.2. *There exists a function $p : [0,T]$ such that its restriction to $[0,\tau^*)$, resp. to $(\tau^*,T]$ can be extended to an absolutely continuous function over $[0,\tau^*]$, resp. over $[\tau^*,T]$. Additionally, up to a subsequence, (p_k) strongly-weakly converges to p over every set $[0,\tau^*-\alpha]\cup[\tau^*+\alpha,T]$ (for α small enough). Finally, p satisfies*

$$\begin{cases} \dot p(t) = -D_x f_1(x^*(t),u^*(t))^\top p(t) \quad \text{for} \quad \text{a.e. } t\in[0,\tau^*], \\ \dot p(t) = -D_x f_2(x^*(t),u^*(t))^\top p(t) \quad \text{for} \quad \text{a.e. } t\in[\tau^*,T]. \end{cases} \tag{4.11}$$

Proof. The proof of this result is done in four steps:

Step 1. Definition of auxiliary trajectories and covectors. First, let us denote by $\tilde u^*$ the optimal control corresponding to the optimal trajectory x^* of Problem (2.5). Then, given $\eta>0$, let us define an auxiliary trajectory x_1^* as the unique solution to the following Cauchy problem:

$$\begin{cases} \dot x_1^*(t) &= f_1(x_1^*(t),\tilde u^*(t)), \quad \text{for a.e. } t\in[0,\tau^*+\eta], \\ x_1^*(0) &= x_0, \end{cases}$$

As well, we can define an auxiliary trajectory x_2^* as the unique solution to the following Cauchy problem:

$$\begin{cases} \dot x_2^*(t) &= f_2(x_2^*(t),\tilde u^*(t)), \quad \text{for a.e. } t\in[\tau^*-\eta,T], \\ x_2^*(T) &= x^*(T). \end{cases}$$

Note that the existence of $\eta>0$ (such that $\tau^*-\eta\ge 0$ and $\tau^*+\eta\le T$) is guaranteed by Cauchy-Lipschitz's Theorem. Furthermore, we can define p_2 as the unique solution to the following Cauchy problem:

$$\begin{cases} \dot p_2(t) &= -D_x f_2(x_2^*(t),\tilde u^*(t))^\top p_2(t), \quad \text{for a.e. } t\in[\tau^*-\eta,T], \\ p_2(T) &= -\nabla g(x^*(T)), \end{cases}$$

and, similarly p_1 is defined as the unique solution to the Cauchy problem:

$$\begin{cases} \dot p_1(t) &= -D_x f_1(x_1^*(t),\tilde u^*(t))^\top p_1(t), \quad \text{for a.e. } t\in[0,\tau^*+\eta], \\ p_1(\tau^*) &= p_2(\tau^*)-\nu v, \end{cases}$$

where $\nu\in\mathbb{R}$ and $v\in\partial_C\psi(x^*(T))$ are limits (up to a subsequence) of the sequences (ν_k) and $(\nabla\psi_k(x_k^*(\tau_k)))$ respectively. Note that the restriction of x_1^* and p_1, resp. x_2^* and p_2, coincides with x^* and p over $[0,\tau^*]$, resp. over $[\tau^*,T]$. Now, let us define $x_{k,1}^*$ as the unique solution to the following Cauchy problem:

$$\begin{cases} \dot x_{k,1}^*(t) &= f_1(x_{k,1}^*(t),u_k^*(t)), \quad \text{for a.e. } t\in[0,\tau^*+\eta], \\ x_{k,1}^*(0) &= x_0, \end{cases}$$

and similarly, $x_{k,2}^*$ is defined as the solution to the following Cauchy problem:

$$\begin{cases} \dot x_{k,2}^*(t) &= f_2(x_{k,2}^*(t),u_k^*(t)), \quad \text{for a.e. } t\in[\tau^*-\eta,T], \\ x_{k,2}^*(T) &= x_k^*(T). \end{cases}$$

Furthermore, we define $p_{k,2}$ as the unique solution to the following Cauchy problem:

$$\begin{cases} \dot p_{k,2}(t) &= -D_x f_2(x_{k,2}^*(t),u_{k,2}^*(t))^\top p_{k,2}(t), \quad \text{for a.e. } t\in[\tau^*-\eta,T], \\ p_{k,2}(T) &= -\nabla g(x_{k,2}^*(T)), \end{cases}$$

and $p_{k,1}$ as the unique solution to the following Cauchy problem:

$$\begin{cases} \dot p_{k,1}(t) &= -D_x f_1(x_{k,1}^*(t),u_{k,1}^*(t))^\top p_{k,1}(t), \quad \text{for a.e. } t\in[0,\tau^*+\eta], \\ p_{k,1}(\tau_k) &= p_{k,2}(\tau_k)-\nu_k\nabla\psi_k(x_{k,2}(\tau_k)). \end{cases}$$

By using a standard argumentation (based on the boundedness of U, the boundedness of the sequence (x_k) for the L^∞-norm, and the continuity of f_i and $D_x f_i$), we deduce that there exists $k_0 \in \mathbb{N}^*$ such that $x^*_{k,1}$ and $p_{k,1}$, resp. $x^*_{k,2}$ and $p_{k,2}$, are well defined over $[0, \tau^* + \eta]$, resp. over $[\tau^* - \eta, T]$, for all $k \geq k_0$.

Step 2. Convergence over $[\tau^* - \eta, T]$**.** One can see that the pair $z_{k,2} := (x^*_{k,2}, p_{k,2})$ satisfies the following differential inclusion

$$\dot{z}_{k,2}(t) \in H_2(z_{k,2}(t)) \quad \text{a.e. } t \in [\tau^* - \eta, T],$$

where $H_2 : B(0,R) \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n \times \mathbb{R}^n$ is defined as $H_2(z) := \widehat{F}_2(\tilde{x}) \times \widehat{G}_2(z)$ for every $z := (\tilde{x}, p) \in \mathbb{R}^n \times \mathbb{R}^n$. From the hypotheses on the velocity sets, we check that for each $z \in B(0,R) \times \mathbb{R}^n$, $H_2(z)$ is with nonempty compact and convex values. In addition, H_2 is also with linear growth (using the linearity of the adjoint equation, the fact that $x \in B(0,R)$ and the compactness of U). Moreover, from the compactness of U and the uniform continuity of the mapping $(x,u,p) \mapsto \big(f_2(x,u), -D_x f_2(x,u)^\top p\big)$ over compact sets one can deduce that H_2 upper semi-continuous at each $(x,p) \in B(0,R) \times \mathbb{R}^n$. Hence, we are in a position to apply [10, Theorem 1.11]. It follows that there is an absolutely continuous function $z_2 := (x^*_2, p_2) : [\tau^* - \eta, T] \to \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$\dot{z}_2(t) \in H_2(z_2(t)) \quad \text{a.e. } t \in [\tau^* - \eta, T],$$

such that $(z_{k,2})$ strongly-weakly converges to z_2 over $[\tau^* - \eta, T]$. Moreover, one also gets $x_2(T) = x^*(T)$ and $p_2(T) = -\nabla g(x^*(T))$. From the uniform convergence of $x^*_{k,2}$ towards x^*_2 and the definition of auxiliary trajectories we get $x^*_2 = x^*$ over $[\tau^*, T]$. Furthermore, from the Filippov selection Lemma, there exists an admissible control u^*_2 such that

$$\dot{x}^*_2(t) = f_2(x^*_2(t), u^*_2(t)), \quad \dot{p}_2(t) = -D_x f_2(x^*_2(t), u^*_2(t))^\top p_2(t),$$

for almost every $t \in [\tau^* - \eta, T]$.

Step 3. Convergence over $[0, \tau^* + \eta]$**.** Similarly, one can see that the pair $z_{k,1} := (x^*_{k,1}, p_{k,1})$ satisfies the following differential inclusion

$$\dot{z}_{k,1}(t) \in H_1(z_{k,1}(t)) \quad \text{a.e. } t \in [0, \tau^* + \eta],$$

where $H_1 : B(0,R) \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n \times \mathbb{R}^n$ is defined as $H_1(z) := \widehat{F}_1(\tilde{x}) \times \widehat{G}_1(z)$ for every $z := (\tilde{x}, p) \in \mathbb{R}^n \times \mathbb{R}^n$. Using similar arguments as in Step 2, it follows that there is an absolutely continuous function $z_1 = (x^*_1, p_1) : [0, \tau^* + \eta] \to \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$\dot{z}_1(t) \in H_1(z_1(t)) \quad \text{a.e. } t \in [0, \tau^* + \eta],$$

such that $(z_{k,1})$ strongly-weakly converges to z_1 over $[0, \tau^* + \eta]$. Moreover, one also gets $x_1(0) = x^*(0) = x_0$ and $p_1(\tau^*) = p_2(\tau^*) - \nu v$. From the uniform convergence of $x^*_{k,1}$ towards x^*_1 and the definition of auxiliary trajectories we get that $x^*_1 = x^*$ over $[0, \tau^*]$. Furthermore, from the Filippov selection lemma, we get the existence of an admissible control u^*_1 such that

$$\dot{x}^*_1(t) = f_1(x^*_1(t), u^*_1(t)), \quad \dot{p}_1(t) = -D_x f_1(x^*_1(t), u^*_1(t))^\top p_1(t),$$

for almost every $t \in [0, \tau^* + \eta]$.

Step 4. Definition of the covector p**.** Finally, we define u^* and p as follows

$$u^*(t) := \begin{cases} u^*_1(t) & \text{a.e. } t \in [0, \tau^*], \\ u^*_2(t) & \text{a.e. } t \in [\tau^*, T]. \end{cases}$$

and

$$p(t) := \begin{cases} p_1(t) & \text{for all } t \in [0, \tau^*), \\ p_2(t) & \text{for all } t \in (\tau^*, T]. \end{cases}$$

From the preceding steps 2-3, one can see that u^* is a control function corresponding to the optimal trajectory x^*. Moreover, we get that the covector p (that is piecewise absolutely continuous) satisfies (4.11) which completes the proof.

$\square$

4.2 Derivation of a nonsmooth HMP in the case of one crossing time

In this section, we provide necessary optimality conditions for an optimal solution to (2.5) in terms of a nonsmooth hybrid maximum principle. Our main result is as follows.

Theorem 4.1. *Suppose that the hypotheses in Section 2 and that Assumption 3.1 are fulfilled. Let $x^* \in \mathcal{S}_T(x_0)$ be an optimal solution to (2.5) and let $u^* \in \mathcal{U}$ an associated optimal control. Then, x^* has zero or one crossing time and there is a piecewise absolutely continuous function $p : [0,T] \to \mathbb{R}^n$ satisfying:*

(i) *the adjoint equation*

$$\dot{p}(t) = -D_x f(x^*(t), u^*(t))^\top p(t) \quad \text{a.e. } t \in [0,T],$$

(ii) *the terminal condition*

$$p(T) = -\nabla g(x^*(T)),$$

(iii) *the Hamiltonian maximization condition*

$$u^*(t) \in \operatorname{argmax}_{\omega \in \mathrm{U}}\, p(t) \cdot f(x^*(t), \omega) \quad \text{a.e. } t \in [0,T],$$

(iv) *the discontinuity condition*

$$p^+(\tau) - p^-(\tau) = \nu v, \tag{4.12}$$

where $\nu \in \mathbb{R}$ and $v \in \partial_C \psi(x^*(\tau^*))$ if x^* has a crossing time $\tau^* \in (0,T)$ (otherwise p has no discontinuity),

(v) *the Hamiltonian constancy : there is $\tilde{c} \in \mathbb{R}$ such that for a.e. $(t,t') \in (0,\tau^*) \times (\tau^*, T)$*

$$p(t) \cdot f_1(x^*(t), u^*(t)) = p(t') \cdot f_2(x^*(t'), u^*(t')) = \tilde{c}.$$

Proof. We prove each item step by step.

(i) Adjoint equation : the existence of a piecewise absolutely continuous function $p : [0,T] \to \mathbb{R}^n$ satisfying the adjoint equation follows from Proposition 4.2.

(ii) Terminal condition : since for all $k \in \mathbb{N}$, $p_k(T) = -\nabla g(x_k(T))$, the result follows from Proposition 4.2 letting $k \to +\infty$.

(iii) Hamiltonian maximization condition : from (4.3), we get that for every $k \in \mathbb{N}$,

$$p_k(t) \cdot f(x_k^*(t), \omega) \le p_k(t) \cdot f(x_k^*(t), u_k^*(t)) \quad \text{a.e. } t \in [0,T].$$

Now, let $t_0 \in (0,T)\backslash\{\tau^*\}$ be a Lebesgue point of $t \mapsto p_k(t) \cdot f(x_k^*(t), u_k^*(t))$ and let $\varepsilon > 0$. It follows that

$$\frac{1}{\varepsilon}\int_{t_0}^{t_0+\varepsilon} p_k(t) \cdot f(x_k^*(t), \omega)\,\mathrm{d}t \le \frac{1}{\varepsilon}\int_{t_0}^{t_0+\varepsilon} p_k(t) \cdot \dot{x}_k^*(t)\,\mathrm{d}t.$$

By using the uniform convergence of (x_k) to x^* over $[0,T]$, the uniform convergence of (p_k) to p over every compact subset of $[0,\tau^*) \cup (\tau^*, T]$, and the weak convergence of $\dot{x}_k^*$ to $\dot{x}^*$, we find that

$$\frac{1}{\varepsilon}\int_{t_0}^{t_0+\varepsilon} p(t) \cdot f(x^*(t), \omega)\,\mathrm{d}t \le \frac{1}{\varepsilon}\int_{t_0}^{t_0+\varepsilon} p(t) \cdot \dot{x}^*(t)\,\mathrm{d}t.$$

Letting $\varepsilon \downarrow 0$, we obtain the Hamiltonian maximization condition (iii) since almost every point of $[0,T]$ is a Lebesgue point of u^*.

(iv) Discontinuity condition. Suppose that x^* has a crossing point at some time τ^*. From Lemma 3.1, we know that τ^* is unique. Now, for every $k \in \mathbb{N}$ large enough, one has $p_k(\tau_k^*) - p_k(\tau_k^-) = \nu_k \nabla \psi_k(x_k(\tau_k))$. Remind that the sequence (ν_k) is bounded, so, extracting a subsequence if necessary, we may assume that $\nu_k \to \nu$ as $k \to +\infty$ where $\nu \in \mathbb{R}$. From the gradient consistency, there is $v \in \partial_C \psi(x^*(\tau^*))$ such that (up to a subsequence), $\psi_k(x_k(\tau_k)) \to +\infty$ as $k \to +\infty$. Hence, by passing to the limit as $k \to +\infty$, we find that $p^+(\tau^*) - p^-(\tau^*) = \nu v$, whence the result.

(v) Since (c_k) is bounded, we may assume that there is $\tilde{c} \in \mathbb{R}$ such that, up to a subsequence, one has $c_k \to \tilde{c}$ as $k \to +\infty$. Next, we proceed as for proving (iii). Let $t_0 \in (0,\tau)\backslash\{\tau^*\}$ be a Lebesgue point of u^* such that $u^*(t_0) \in U$ (we know that a.e. point of $(0,T)$ satisfies this property since u^* is a measurable (bounded) function such that $u^*(t) \in U$ for a.e. $t \in [0,T]$). We get for $\varepsilon > 0$ small enough

$$\frac{1}{\varepsilon}\int_{t_0}^{t_0+\varepsilon} p_k(t)\cdot \dot{x}_k^*(t)\,\mathrm{d}t = \tilde{c}_k.$$

As for proving (iii), we let $k \to +\infty$ (using weak convergence of (x_k^*)) and then $\varepsilon \downarrow 0$, and we obtain the result over $(0,\tau^*)$. Next, the same reasoning is done over (τ^*,T) which ends the proof. $\square$

Remark 4.2. *Whenever $\nu \geq 0$,* (4.17) *becomes*

$$p^+(\tau^*) - p^-(\tau^*) \in N^C(\overline{X}_1; x^*(\tau^*)).$$

Indeed, in that case, one has $N^C(\overline{X}_1; x^*(\tau^*)) = \mathbb{R}_+\partial_C\psi(x^*(\tau^*))$ since $0 \notin \partial_C\psi(x^*(\tau^*))$ because of the transversality condition (4.16).

We give now a simple example highlighting the application of the nonsmooth HMP. Consider the Lipschitz-continuous function $\psi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\psi(x_1,x_2) := \max(x_1, x_1+x_2-1),$$

so that $\Gamma = \{(x_1,x_2)\in\mathbb{R}^2\ ;\ \psi(x_1,x_2)=0\}$. It is easily seen that at the point $(0,1)$ the sub-differential of ψ is the convex combination of the gradients of the maps $(x_1,x_2)\mapsto x_1$ and $(x_1,x_2)\mapsto x_1+x_2-1$, hence

$$\partial_C\psi(0,1) = \{(1,\theta)\ ;\ \theta\in[0,1]\}.$$

Now, we consider the hybrid control system

$$\left\{\begin{array}{ll}\dot{x}_1 & = \frac{\sqrt{3}}{2}(1-u)\\ \dot{x}_2 & = \frac{1}{2}(1-u)\end{array}\right. \quad \text{if } \psi(x_1,x_2)<0 \quad \text{and} \quad \left\{\begin{array}{ll}\dot{x}_1 & = x_2\\ \dot{x}_2 & = \frac{u}{2}\end{array}\right. \quad \text{if } \psi(x_1,x_2)>0,$$

together with the initial condition $x(0) = (-\sqrt{3},0)$ and the optimal control problem

$$-x_1(2) \to \min$$

where the minimimum is taken w.r.t. measurable control functions $u : [0,2]\to[-1,\frac{1}{2}]$. We see that every admissible trajectory crosses the interface at the point $(0,1)$. Hence, the transverse condition needs only to be verified at the point $(x_1,x_2)=(0,1)$. For all $\theta\in[0,1]$ and all $u\in[-1,\frac{1}{2}]$, one has

$$\Big(\frac{\sqrt{3}(1-u)}{2},\frac{1-u}{2}\Big)\cdot(1,\theta) = \frac{(1-u)}{2}(\sqrt{3}+\theta) \geq \frac{\sqrt{3}}{4}\ ;\ \Big(1,\frac{u}{2}\Big)\cdot(1,\theta) = 1+\frac{u\theta}{2}\geq\frac{1}{2},$$

so that the transverse condition is verified. Now, since the objective is equivalent to maximize $x_1(2)$, an optimal control will necessarily be such that $u(t)=-1$ as long $x(t)\in X_1$ (where $X_1 = \{x\in\mathbb{R}^2\ ;\ \psi(x)<0\}$). We thus get that an optimal trajectory has a single crossing time at $\tau := +1$. If we write the HMP, we introduce the (conserved) Hamiltonian

$$H = \frac{\sqrt{3}}{2}p_1(1-u) + \frac{1}{2}p_2(1-u) = q_1x_2 + q_2\frac{u}{2},$$

where (p_1,p_2) is the covector over $[0,1]$ such that $\dot{p}_1 = \dot{p}_2 = 0$ over $[0,1]$ (hence, (p_1,p_2) is constant) and where (q_1,q_2) is the covector over $[1,2]$ such that $\dot{q}_1 = 0$ and $\dot{q}_2 = -1$. Additionally, the terminal condition gives us $q_1(2)=1$ and $q_2(2)=0$ so that $q_1(t)=+1$ and $q_2(t)=2-t$ for $t\in[1,2]$ and consequently $u(t)=\frac{1}{2}$ a.e. over $[1,2]$ (using the maximization condition). If we write the jump condition and the conservation of the Hamiltonian, we get the three equalities

$$\begin{array}{rl} q_1 - p_1 & = \nu\\ q_2(\tau)-p_2 & = \nu\theta\\ \sqrt{3}p_1+p_2 & = 2\end{array} \quad\Rightarrow\quad \begin{array}{rl} 1-p_1 & = \nu\\ 1-p_2 & = \nu\theta\\ \sqrt{3}p_1+p_2 & = 2\end{array}$$

where unknown parameters are $p_1, p_2, \nu \in \mathbb{R}$ and $\theta \in [0,1]$. This gives the relations $p_2 = 2 - \sqrt{3}p_1$ and $p_1 = 1 - \nu$ so that $p_2 = 2 - \sqrt{3}(1-\nu) = 1 - \nu\theta$. Finally, we get $(\sqrt{3}+\theta)\nu = \sqrt{3} - 1$. We thus deduce from the preceding equality that ν cannot be null. Hence, ν is necessarily such that $\nu \neq 0$ so that the covector is necessarily discontinuous at the crossing time (it has a jump). In addition, to ensure that $u = -1$ over $[0,1]$, one must have $\frac{\sqrt{3}}{2}p_1 + \frac{1}{2}p_2 > 0$ from the maximization condition. We see that it is possible to choose adequately the pair $(\nu, \theta) \in \mathbb{R}^* \times [0,1]$ to fulfill this constraint. For instance, taking $\theta = 0$ gives us $\nu = 1 - 1/\sqrt{3}$, $p_1 = 1/\sqrt{3}$ and $p_2 = 1$ so that $(\sqrt{3}/2)p_1 + p_2/2 > 0$ as wanted.

In conclusion, this example highlights the fact that (depending on the problem) the covector must have a jump at the crossing time but the direction of the jump in the subdifferential of ψ may be non-unique.

4.3 Nonsmooth HMP in the case of multiple crossing times

Our approach for proving Theorem 4.1 can be re-employed in order to obtain the HMP in the more general case of a partition of the state space into a finite number of regions meaning that

$$\mathbb{R}^n = \bigcup_{i \in \mathcal{I}} \overline{X}_i,$$

where $\mathcal{I} := \{1, ..., N\}$, $N \in \mathbb{N}^*$, and where for every $i \in \mathcal{I}$, X_i is an open subset of $\mathbb{R}^n$ with non-empty interior and with a locally Lipschitz continuous boundary. To this partition, we associate a family of dynamics $(f_i)_{i \in \mathcal{I}}$ where each $f_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is of class C^1 and we define a hybrid dynamics as

$$\forall (x,u) \in X_i \times \mathbb{R}^m, \quad f(x,u) := f_i(x,u). \tag{4.13}$$

Note that f is not properly defined at the boundary of the sets X_i, but this fact will have no importance since our approach relies on a similar transverse hypothesis as (3.1). The hypotheses satisfied by f as given by (4.13) are as follows. We suppose that for every $i \in \mathcal{I}$ and that for every $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$, the sets $\widehat{F}_i(x)$ and $\widehat{G}_i(x,p)$ (remind (2.2)-(2.3)) are non-empty compact convex subsets of $\mathbb{R}^n$ and that each f_i satisfies (2.4). The Mayer optimal control problem now becomes

$$\inf_{u \in \mathcal{U}} g(x(T)), \tag{4.14}$$

where x is a solution to the Cauchy problem

$$\left\{ \begin{array}{ll} \dot{x}(t) & = f(x(t), u(t)) \quad \text{a.e. } t \in [0,T], \\ x(0) & = x_0, \end{array} \right. \tag{4.15}$$

and where x_0 is supposed to belong to X_1 and f is defined by (4.13). Additionally, we suppose that each interface[3] $\overline{X}_i \cap \overline{X}_{i+1}$ is described by a locally Lipschitz continuous function $\psi_i : \mathbb{R}^n \to \mathbb{R}$ which allows us to introduce the following transverse hypothesis (in the spirit of (3.1)). Remind that there is $R \geq 0$ such that every solution to (4.15) fulfills $x(t) \in B[0,R]$ for every $t \in [0,T]$ (see Section 3.3).

Assumption 4.1. *For every $i \in \mathcal{I}$, there is $\kappa_i > 0$ such that*

$$\forall (x,u) \in (\overline{X}_i \cap \overline{X}_{i+1} \cap B[0,R]) \times U, \ \forall v \in \partial_C \psi_i(x), \quad v \cdot f_j(x,u) \geq \kappa_i, \tag{4.16}$$

where $j \in \{i, i+1\}$.

The next result is an adaptation of Theorem 4.1 to the case of $N \geq 1$ regions and multiple crossing times.

Theorem 4.2. *Suppose that the dynamics f as given by (4.13) satisfies the preceding hypotheses and that Assumption 4.1 is fulfilled. Let x^* be an optimal solution to (4.14) with N crossing times $(\tau_i^*)_{1 \leq i \leq N}$ from X_i into X_{i+1} and let $u^* \in \mathcal{U}$ be an associated optimal control. Then, there is a piecewise absolutely continuous function $p : [0,T] \to \mathbb{R}^n$ satisfying:*

[3]Actually, a description of $\overline{X}_i \cap \overline{X}_j$ for an arbitrary pair $(i,j) \in \mathcal{I} \times \mathcal{I}$ such that $i \neq j$ is not needed since we shall assume (up to a re-indexation of the sets X_i) that an optimal solution to (4.14) will visit successively the sets X_1, X_2,...,X_N.

(i) *the adjoint equation*
$$\dot{p}(t) = -D_x f(x^*(t), u^*(t))^\top p(t) \quad \text{a.e. } t \in [0, T],$$
(ii) *the terminal condition*
$$p(T) = -\nabla g(x^*(T)),$$
(iii) *the Hamiltonian maximization condition*
$$u^*(t) \in \operatorname{argmax}_{\omega \in \mathrm{U}} \, p(t) \cdot f(x^*(t), \omega) \quad \text{a.e. } t \in [0, T],$$
(iv) *the discontinuity condition*
$$p^+(\tau_i^*) - p^-(\tau_i^*) = \nu_i v_i, \tag{4.17}$$
where $\nu_i \in \mathbb{R}$ *and* $v_i \in \partial_C \psi_i(x^*(\tau_i^*))$.
(v) *the Hamiltonian constancy : there is* $\tilde{c} \in \mathbb{R}$ *such that for a.e.* $t \in (0, T)$,
$$p(t) \cdot f(x^*(t), u^*(t)) = \tilde{c}.$$

The proof of this theorem is based on a regularization via mollifiers of each nonsmooth interface and it follows exactly the same scheme as for proving Theorem 4.1, that is why, it is omitted.

5 Discussions and conclusion

In this paper, we have given a so-called "nonsmooth" HMP for a Mayer optimal control problem governed by a hybrid control system defined regionally. By nonsmooth, we mean that the interface between two regions is described by a locally Lipschitz continuous function (and not of class C^1). The originality of our approach consists in regularizing the interface and in using the gradient consistency property arising from the regularization of a locally Lipschitz continuous function via mollifiers as given in [23, Theorem 9.67] by Rockafellar and Wets. We are then in a position to apply the "classical" HMP (valid for regions with smooth boundaries) on the regularized problems. By passing to the limit as the regularization parameter goes to zero, we have obtained a discontinuity condition for the covector that naturally generalizes the one encountered in the hybrid setting when interfaces between regions are smooth, see, *e.g.*, [5, 18]. It is also in line with the discontinuity condition in [11, Theorem 22.20] which (as far as we know) covers temporally hybrid optimal control problems (in the sense that no partition of the state space is considered and the change of dynamics occurs at free instants).

When the system involves mixed initial terminal constraints $\phi(x(0), x(T)) \in C$, then our approach may not be directly applicable. Indeed, it may be more involved to prove that any admissible trajectory to the original optimal control problem can be approached by an admissible trajectory solution to the approximated optimal control problems. However, this issue may be overcome considering additional (and reasonable) controllability hypotheses on the dynamics.

Our methodology also differs from other approaches like the augmentation technique as in [5] or the sensitivity analysis as it is done in [4, 18] (which follows the standard techniques based on the variation vector for proving the PMP, see [20]). At first glance, our approach seems rather straightforward, however, when introducing an approximated optimal control problem, we have in general no information on the regularity of the approximated optimal control (such as right and left continuity at a crossing time). This is a difficulty inherent to our approach. As far as we know, we can point out a similar difficulty if one wants to apply Ekeland's variational principle to prove the HMP with mixed initial-terminal constraints. Indeed, in that case, there are no information on the regularity of an approximated control (like right and left continuity at a crossing time) arising from this principle after a penalization of the mixed initial-terminal constraint. Finding a way to bypass this issue is an interesting question that could be the matter of future works.

Finally, another possible way to derive a nonsmooth HMP in our setting could be to follow [5] and to use a nonsmooth PMP (in place of the usual one) on some augmented optimal control problem involving nonsmooth data (since the interface is nonsmooth). This could also be an interesting perspective for future works on the subject. As well, it could be interesting to pursue the analysis made in this work, but under a weak transverse hypothesis at the interface (see [6]).

Acknowledgements

We thank Abderrahim Jourani for fruitful exchanges concerning our approach. We would also like to thank Edouard Pauwels for fruitful exchanges on the chain rule formula. We are also grateful to Alberto Seeger for helpful discussions on the gradient consistency property.

References

[1] J.-P. Aubin, A. Cellina, *Differential Inclusions: Set-Valued Maps and Viability Theory*, Springer-Verlag Berlin Heidelberg, 1984.

[2] J.-P. Aubin, A.M. Bayen, P. Saint-Pierre, *Viability Theory, New Directions*, Second Editions, Springer, 2011.

[3] G. Barles, A. Briani, E. Trélat, *Value function for regional control problems via dynamic programming and Pontryagin maximum principle*, Math. Control Relat. Fields, vol. 8, 3-4, pp. 509–533, 2018.

[4] T. Bayen, A. Bouali, L. Bourdin, *Hybrid maximum principle with regionally switching parameter*, preprint HAL, 2023, `hal-03638701`.

[5] T. Bayen, A. Bouali, L. Bourdin, *The hybrid maximum principle for optimal control problems with spatially heterogeneous dynamics is a consequence of a Pontryagin maximum principle for* $L^1_\square$ *local solutions*, SIAM J. Control Optim., vol. 62, 4, 2024.

[6] T. Bayen, A. Bouali, L. Bourdin, O. Cots, *On the reduction of a spatially hybrid optimal control problem into a temporally hybrid optimal control problem*, AIMS on Applied Mathematics. Ivan Kupka legacy: A Tour Through Controlled Dynamics, 12, 2024.

[7] T. Bayen, A. Bouali, L. Bourdin, O. Cots, *Loss control regions in optimal control problems*, J. Differential Equations, vol. 405, pp. 359–397, 2024.

[8] T. Bayen, K. Boumaza, A. Rapaport, *Necessary optimality condition for the minimal time crisis relaxing transverse condition via regularization*, ESAIM Control Optim. Calc. Var., vol. 27, pp. 105, 2021.

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

[10] F.H. Clarke, *Optimization and nonsmooth analysis*, Society for Industrial and Applied Mathematics, 1990.

[11] F.H. Clarke, *Functional analysis, calculus of variations and optimal control*, London, Springer, 2013.

[12] F.H. Clarke, R.B. Vinter, *Optimal multiprocesses*, SIAM J. Control Optim., vol.27, pp. 1072–1091, 1989.

[13] A.F. Filippov, *Classical solutions of differential equations with multi-valued right-hand side*, SIAM J. Control, vol. 5, 4, pp. 609–621, 1967.

[14] H. Frankowska, *Regularity of minimizers and of adjoint states in optimal control under state constraints*, J. of Convex Anal., vol. 13, 2, pp. 299–328, 2006.

[15] H. Frankowska, F. Rampazzo, *Filippov's and Filippov-Wazewski's theorems on closed domains*, J. Diff. Eqs., vol. 161, pp. 449–478, 2000.

[16] M. Garavello, B. Piccoli, *Hybrid necessary principle*, SIAM J. Control Optim., vol. 43, pp. 1867–1887, 2005.

[17] R. Goebel, A.R. Teel, *Solutions to hybrid inclusions via set and graphical convergence with stability theory applications*, Automatica, vol. 42, 4, pp. 573–587, 2006.

[18] T. Haberkorn, E. Trélat, *Convergence results for smooth regularizations of hybrid nonlinear optimal control problems*, SIAM J. Control Optim., vol. 49, 4, pp. 1498–1522, 2011.

[19] D. Liberzon, *Switching in systems and control*, Boston, Birkhäuser, 2003.

[20] L.S. Pontryagin, V.G. Boltyanskiy, R.V. Gamkrelidze, E.F. Mishchenko, *Mathematical theory of optimal processes*, The Macmillan Company, 1964.

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

[22] N.S. Papageorgiou, *A stability result for differential inclusions in Banach spaces*, J. Math. Anal. Appl., vol. 118, 1, pp. 232–246, 1986.

[23] R.T. Rockafellar, R.J.-B. Wets, *Variational Analysis*, Springer-Verlag, 2009.

[24] H.J. Sussmann, *A maximum principle for hybrid optimal control problems*, Proc. 38th IEEE Conf. Decis. Control, pp. 425–430, 1999.

[25] C. Xiaojun, *Smoothing methods for nonsmooth, nonconvex minimization*, Math. Program., vol. 134, pp. 71–99, 2012.

[26] L. Thibault, *Unilateral Variational Analysis in Banach Spaces. Part I, General Theory*, World Scientific (2023).

[27] L. Thibault, *Unilateral Variational Analysis in Banach Spaces. Part II, Special Classes of Functions and Sets*, World Scientific (2023).