The hybrid maximum principle for optimal control problems with spatially heterogeneous dynamics is a consequence of a Pontryagin maximum principle for “L1 square” local solutions

Terence Bayen, Anas Bouali, Loic Bourdin

THE HYBRID MAXIMUM PRINCIPLE FOR OPTIMAL CONTROL PROBLEMS WITH SPATIALLY HETEROGENEOUS DYNAMICS IS A CONSEQUENCE OF A PONTRYAGIN MAXIMUM PRINCIPLE FOR $\mathrm{L}^1_\square$–LOCAL SOLUTIONS

TÉRENCE BAYEN*, ANAS BOUALI†, AND LOÏC BOURDIN‡

Abstract. The title of the present work is a nod to the paper *The hybrid maximum principle is a consequence of Pontryagin maximum principle* by Dmitruk and Kaganovich (Systems and Control Letters, 2008). Here we investigate a similar but different framework of hybrid optimal control problems. Precisely we consider a general control system that is described by a differential equation involving a spatially heterogeneous dynamics. In that context the sequence of dynamics followed by the trajectory and the corresponding switching times are fully constrained by the state position. We prove with an explicit counterexample that the augmentation technique proposed by Dmitruk and Kaganovich cannot be fully applied to our setting, but we show that it can be adapted by introducing a new notion of local solution to classical optimal control problems and by establishing a corresponding Pontryagin maximum principle. Thanks to this method we derive a hybrid maximum principle adapted to our setting, with a simple proof that does not require any technical tool (such as implicit function arguments) to handle the dynamical discontinuities.

Key words. Optimal control, heterogeneous dynamics, hybrid maximum principle.

MSC codes. 34A38, 49K15.

1. Introduction.

1.1. General context. The *Pontryagin Maximum Principle* (in short, PMP), established at the end of the 1950s (see [27]), has originally been developed for optimal control problems where the control system is described by an ordinary differential equation (in short, ODE). It states the corresponding first-order necessary optimality conditions, in terms of an (absolutely continuous) costate function. As usual in optimization, the PMP remains valid for *local solutions* only (typically in uniform norm for the state and in L^1-norm for the control). Since then, the PMP has been adapted to many situations, in particular for control systems of different natures.

On the other hand, *hybrid systems* are, in a broad sense, dynamical systems that exhibit both continuous and discrete behaviors. They are particularly used in automation and robotics to describe complex systems in which, for example, logic decisions are combined with physical processes. We refer to [32] for an elementary introduction to hybrid systems. This theory is very large and it is commonly accepted that it includes ODEs with *heterogeneous dynamics*, that is, ODEs involving a family of different dynamics (used for example to describe evolutions in heterogeneous media) where the transitions from one dynamics to another are seen as discrete events.

The PMP has been extended to hybrid control systems, especially in the context of ODEs with heterogeneous dynamics (see, *e.g.*, [19, 22, 26, 29, 30, 31]), resulting in theorems often referred to as *Hybrid Maximum Principle* (in short, HMP). We emphasize that the frameworks are very varied. Indeed the rule that supervises the transitions between the different dynamics is usually described by additional variables

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

†Avignon Université, Laboratoire de Mathématiques d'Avignon (EA 2151) F-84018, France. (anas.bouali@univ-avignon.fr)

‡Université de Limoges, Institut de recherche XLIM. UMR CNRS 7252, France. (loic.bourdin@unilim.fr)

that can be free or constrained and, in that second case, the constraints can be of different natures. For example the *switching times* (*i.e.* the instants at which the control system moves from one dynamics to another) can be the resultant of a control decision or can be (fully or partially) determined by the time variable, the state position or both of them. Hence different versions of the HMP can be found in the literature, corresponding to different hybrid control systems that are presented under various names according to their nature (such as *multi-processes* [18], *switched systems* [28], *regional systems* [1], *systems on stratified domains* [14], *variable structure systems* [8]). In contrary to the classical PMP, the HMP is usually expressed in terms of an (only) piecewise absolutely continuous costate function that admits discontinuity jumps at the switching times. A common feature of most of the above references is that the mathematical framework somehow guarantees that local perturbations (typically in uniform norm for the state and in L^1-norm for the control) preserve the same hybrid structure (that is, the same sequence of dynamics) as the nominal one.

1.2. The augmentation technique of Dmitruk and Kaganovich. In the context of ODEs with heterogeneous dynamics, the difficult part of deriving a HMP lies in handling the dynamical discontinuities. To this end, an excellent strategy has been proposed in [19], in which the switching times are additional variables satisfying equality/inequality constraints involving the corresponding intermediate state values.

Roughly speaking, considering an optimal solution (associated with switching times denoted by τ_k^*), this technique consists in affine changes of time variable, mapping the intervals (τ_{k-1}^*, τ_k^*) into a common interval $(0,1)$. This procedure augments the dimensions of the variables and thus is categorized in the set of *augmentation techniques*. The authors prove that the augmented solution is a local solution to the augmented problem which is classical (that is, non-hybrid) by construction (since the discontinuities have been positioned at the endpoints of the interval $[0,1]$). Therefore the classical PMP can be applied to the augmented solution (expressed in terms of an augmented absolutely continuous costate function satisfying endpoint transversality conditions). Hence, by inverting the above affine changes of time variable, first-order necessary optimality conditions are derived for the original nonaugmented solution, expressed in terms of a nonaugmented (piecewise absolutely continuous) costate function satisfying discontinuity jumps at the switching times τ_k^* (whose expressions follow from the endpoint transversality conditions at 0 and 1 of the augmented costate function).

Hence Dmitruk and Kaganovich have entitled their paper [19] as *The hybrid maximum principle is a consequence of Pontryagin maximum principle*. The augmentation technique is particularly satisfactory because it allows to reduce the hybrid problem into a classical (non-hybrid) augmented problem, avoiding the use of technical arguments (such as implicit function theorems) to handle the dynamical discontinuities.

1.3. Framework and contributions of the present work. In the spirit of [1, 23], we consider a control system described by an ODE with *spatially* heterogeneous dynamics, in the sense that the state space is partitioned into several disjoint regions and each region has its own dynamics. In that context the sequence of dynamics followed by the trajectory and the corresponding switching times (called *crossing times* since they correspond to the instants at which the state goes from one region to another) are fully constrained by the state position.

A HMP corresponding to this setting has already been announced in [23] but with a sketch of proof which is, to our best knowledge, erroneous. Indeed the author invoke *needle-like perturbations* of the control, while they are not admissible in the present setting (see Appendix C for a counterexample). This issue has been corrected in our previous paper [2] by applying needle-like perturbations on *auxiliary controls*. Then, to handle the resulting perturbed crossing times, we used an inductive application of the implicit function theorem, which results into a technical and extended analysis. An attempt to derive a HMP corresponding to our setting, with the simpler approach of Dmitruk and Kaganovich, was also presented in [1]. Unfortunately, to our best knowledge, this proof is also incorrect. Indeed, in contrary to the framework of Dmitruk and Kaganovich in [19], our setting fails to guarantee that the augmented solution is a local solution to the classical augmented problem (see Section 3.4 for a counterexample) and, therefore, the classical PMP cannot be applied. We emphasize that our counterexample shows that, in our setting, a local perturbation (in uniform norm for the state and in L^1-norm for the control) does not preserve the hybrid structure of the nominal one in general.

Hence the main objective of this paper is to derive a HMP for our setting, with a correct proof that adapts the augmentation procedure of Dmitruk and Kaganovich. To this aim a new notion of local solution to classical optimal control problems (see the definition of $\mathrm{L}^1_\square$–local solution in Definition 2.2) and a corresponding version of the PMP (see Theorem 2.1) are required. Indeed we prove in Proposition 3.1 that, under appropriate assumptions (such as *transverse conditions* at the crossing times), the augmented solution is a $\mathrm{L}^1_\square$–local solution to the classical augmented problem and therefore the above new PMP can be applied. Finally, similarly to [19], by inverting the affine changes of time variable, a HMP for our setting is obtained (see Theorem 3.1).

1.4. Organization of the paper. In Section 2, a classical optimal control problem is considered (see Problem (P)), the new notion of $\mathrm{L}^1_\square$–local solution is introduced (see Definition 2.2) and a corresponding PMP is established (see Theorem 2.1). In Section 3, a hybrid optimal control problem with spatially heterogeneous dynamics is introduced (see Problem (HP)). Applying the augmentation procedure, Proposition 3.1 states that an augmented solution to Problem (HP) is a $\mathrm{L}^1_\square$–local solution to the corresponding classical augmented problem of the form of Problem (P). Hence, applying the above new PMP and inverting the affine changes of time variable, a HMP for Problem (HP) is obtained (see Theorem 3.1). An explicit counterexample showing that an augmented solution to Problem (HP) is not a local solution (in the usual sense) to the corresponding classical augmented problem is provided in Section 3.4. Finally the technical proofs of Proposition 3.1 and Theorem 3.1 are provided in Appendices A and B respectively. A counterexample showing that needle-like perturbations of the control are not admissible in our setting is provided in Appendix C.

2. Preliminaries and PMP for the new notion of $\mathrm{L}^1_{A\square}$–local solution.

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 $X \subset \mathbb{R}^d$, 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 (Lebesgue) measurable subset $A \subset \mathbb{R}$, we denote by $\mu(A)$ its (Lebesgue) measure. Furthermore, for any extended-real number $r \in [1, \infty]$ and any real interval $I \subset \mathbb{R}$, we denote by:

- $\mathrm{L}^r(I, \mathbb{R}^d)$ the usual 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 the 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.

Now take $I = [0, T]$ for some $T > 0$. Recall that a partition of the interval $[0, T]$ is a set $\mathbb{T} = \{\tau_k\}_{k=0,\dots,N}$ of real numbers such that $0 = \tau_0 < \tau_1 < \ldots < \tau_{N-1} < \tau_N = T$

for some $N \in \mathbb{N}^*$. In this paper a function $p : [0,T] \to \mathbb{R}^d$ is said to be *piecewise absolutely continuous*, with respect to a partition $\mathbb{T} = \{\tau_k\}_{k=0,\dots,N}$, if p is continuous at 0 and T and the restriction of p over each open interval (τ_{k-1},τ_k) admits an extension over $[\tau_{k-1},\tau_k]$ that is absolutely continuous. If so, p admits left and right limits at each $\tau_k \in (0,T)$, denoted respectively by $p^-(\tau_k)$ and $p^+(\tau_k)$. We denote by:

- $\mathrm{PAC}_{\mathbb{T}}([0,T],\mathbb{R}^d)$ the space of piecewise absolutely continuous functions, with respect to a partition $\mathbb{T}$ of the interval $[0,T]$, with values in $\mathbb{R}^d$.

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

2.1. A classical optimal control problem and $\mathrm{L}^1_{A\square}$–local solution. Let n, m, d and $\ell \in \mathbb{N}^*$ be four fixed positive integers and $T > 0$ be a fixed positive real number. In the present section we consider a classical Mayer optimal control problem with parameter and mixed terminal state constraints given by

$$
(\mathrm{P}) \qquad \begin{array}{ll}
\text{minimize} & \phi(x(0),x(T),\lambda), \\
\text{subject to} & (x,u,\lambda) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m) \times \mathbb{R}^d, \\
& \dot{x}(t) = f(x(t),u(t),\lambda), \quad \text{a.e. } t \in [0,T], \\
& g(x(0),x(T),\lambda) \in \mathrm{S}, \\
& u(t) \in \mathrm{U}, \quad \text{a.e. } t \in [0,T],
\end{array}
$$

where the Mayer cost function $\phi : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$, the dynamics $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^n$ and the constraint function $g : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^\ell$ are of class C^1, and where $\mathrm{S} \subset \mathbb{R}^\ell$ is a nonempty closed convex subset and $\mathrm{U} \subset \mathbb{R}^m$ is a nonempty subset. As usual in the literature, $x \in \mathrm{AC}([0,T],\mathbb{R}^n)$ is called the *state* (or the *trajectory*), $u \in \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ is called the *control* and $\lambda \in \mathbb{R}^d$ is called the *parameter*. A triplet $(x,u,\lambda) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m) \times \mathbb{R}^d$ is said to be *admissible* for Problem (P) if it satisfies all the constraints of Problem (P). Finally, such an admissible triplet is said to be a *global solution* to Problem (P) if it minimizes the Mayer cost $\phi(x(0),x(T),\lambda)$ among all admissible triplets.

Remark 2.1. (i) All along this paper (not only for Problem (P)), we have chosen to deal with optimal control problems with (only) Mayer cost, fixed final time and autonomous dynamics. It is well known in the literature (see, *e.g.*, [9, 15, 16]) that standard techniques (such as augmentation or changes of variables) allow to deal with more general Bolza cost, free final time and time-dependent dynamics. Similarly, in Problem (P), we assume for simplicity that ϕ, f and g are of class C^1 and also some topological properties for S. However the results that we will present in this section can be extended to weaker assumptions (see, *e.g.*, [17, 33]). Overall, our aim in this paper is not to address the most general framework possible. We keep our setting as simple as possible to stay focused on the novel aspects of our work.

(ii) The presence of a parameter $\lambda \in \mathbb{R}^d$ in Problem (P) can also be treated thanks to an augmentation (see, *e.g.*, [9]). It is noteworthy that the main problem considered in the present work (see Problem (HP) in the next Section 3) is a hybrid optimal control problem which does not involve any parameter. However the proof of our main result (Theorem 3.1) is based on a reduction of Problem (HP) into a classical optimal control problem of the form of Problem (P) that involves parameters. This is the only reason why we need to consider the presence of a parameter $\lambda \in \mathbb{R}^d$ in Problem (P).

The classical PMP [27] has originally been developed for global solutions but, as usual in optimization, it remains valid for *local solutions*. As a consequence, several notions of local solution to classical optimal control problems, and the corresponding versions of the PMP, have been developed in the literature (see, *e.g.*, [11, 25]). Let us introduce two new notions of local solution which will play central roles in our work.

Definition 2.1 (L^1_A–local solution). *An admissible triplet (x^*, u^*, λ^*) is said to be a L^1_A–local solution to Problem* (P)*, for a measurable subset $A \subset [0,T]$, if, for all $R \geq \|u^*\|_{\mathrm{L}^\infty}$, there exists $\eta > 0$ such that $\phi(x^*(0), x^*(T), \lambda^*) \leq \phi(x(0), x(T), \lambda)$ for all admissible triplets (x, u, λ) satisfying*

$$\begin{cases} \|x - x^*\|_{\mathrm{C}} + \|u - u^*\|_{\mathrm{L}^1} + \|\lambda - \lambda^*\|_{\mathbb{R}^d} \leq \eta, \\ \|u\|_{\mathrm{L}^\infty} \leq R, \\ u(t) = u^*(t) \ \text{a.e. } t \in [0,T] \backslash A. \end{cases}$$

Definition 2.2 ($\mathrm{L}^1_{A\square}$–local solution). *An admissible triplet (x^*, u^*, λ^*) is said to be a $\mathrm{L}^1_{A\square}$–local solution to Problem* (P)*, for a measurable subset $A \subset [0,T]$, if there exists an increasing family $(A_\varepsilon)_{\varepsilon>0}$ of measurable subsets of A, satisfying $\mu(A_\varepsilon) \to \mu(A)$ as $\varepsilon \to 0$, such that (x^*, u^*, λ^*) is a $\mathrm{L}^1_{A_\varepsilon}$–local solution to Problem* (P) *for all $\varepsilon > 0$.*

Remark 2.2. (i) The notations L^1_A and $\mathrm{L}^1_{A\square}$ are very close, while the corresponding definitions are (slightly) different. Therefore the reader needs to be careful with these two different concepts, for which we will give each one a version of the PMP (see Lemma 2.1 for L^1_A–local solutions and Theorem 2.1 for $\mathrm{L}^1_{A\square}$–local solutions).

(ii) The concept of $\mathrm{L}^1_{[0,T]}$–local solution coincides with the classical notion of L^1–local solution well established in the literature (see, *e.g.*, [11, 25]). Therefore, in the sequel, we simply write L^1–local solution instead of $\mathrm{L}^1_{[0,T]}$–local solution. To be consistent we simply write $\mathrm{L}^1_\square$–local solution instead of $\mathrm{L}^1_{[0,T]\square}$–local solution.

(iii) With respect to the classical concept of L^1–local solution, the refined notion of L^1_A–local solution imposes on admissible controls to match the nominal one almost everywhere outside the measurable subset $A \subset [0,T]$. This feature is crucial to reduce the hybrid optimal control problem considered in the next Section 3 into a classical optimal control problem. This is not possible with the classical concept of L^1–local solution, as shown by a counterexample in Section 3.4.

(iv) For a measurable subset $A \subset [0,T]$, it is clear that a L^1_A–local solution is automatically a $\mathrm{L}^1_{A\square}$–local solution. However the converse is not true in general (see the counterexample in Section 3.4). From a general point of view, the implications

$$\begin{tikzcd}[arrows=Rightarrow] & & \mathrm{L}^1_{A'}\text{–local solution} \arrow[dr] & \\ \text{global solution} \arrow[r] & \mathrm{L}^1_A\text{–local solution} \arrow[ur] \arrow[dr] & & \mathrm{L}^1_{A'\square}\text{–local solution} \\ & & \mathrm{L}^1_{A\square}\text{–local solution} \arrow[ur] & \end{tikzcd}$$

hold true for any measurable subsets $A' \subset A \subset [0,T]$, but not the converses in general.

2.2. PMP for $\mathrm{L}^1_{A\square}$–local solutions and comments. Recall first that the *normal cone* to S at some point $z \in \mathrm{S}$ is defined by

$$\mathrm{N_S}[z] := \{z'' \in \mathbb{R}^\ell \mid \forall z' \in \mathrm{S},\ \langle z'', z' - z\rangle_{\mathbb{R}^\ell} \leq 0\},$$

and that g is said to be *submersive* at a point of $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ if the differential of g at this point is surjective. Finally recall that the *Hamiltonian* $\mathcal{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ associated with Problem (P) is defined by $\mathcal{H}(x, u, \lambda, p) := \langle p, f(x, u, \lambda)\rangle_{\mathbb{R}^n}$ for all $(x, u, \lambda, p) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^d \times \mathbb{R}^n$. We are now in a position to establish a new version of the PMP that is dedicated to $\mathrm{L}^1_{A\square}$–local solutions to Problem (P).

Theorem 2.1 (PMP for $\mathrm{L}^1_{A\square}$–local solutions). *If (x^*, u^*, λ^*) is a $\mathrm{L}^1_{A\square}$–local solution to Problem* (P)*, for a measurable subset $A \subset [0,T]$, such that g is submersive at $(x^*(0), x^*(T), \lambda^*)$, then there exists a nontrivial pair $(p, p^0) \in \mathrm{AC}([0,T], \mathbb{R}^n) \times \mathbb{R}_+$ satisfying:*

(i) the Hamiltonian system $\dot{x}^*(t) = \nabla_p \mathcal{H}(x^*(t), u^*(t), \lambda^*, p(t))$ and $-\dot{p}(t) = \nabla_x \mathcal{H}(x^*(t), u^*(t), \lambda^*, p(t))$ for almost every $t \in [0,T]$;

(ii) the endpoint transversality condition

$$\begin{pmatrix} p(0) \\ -p(T) \\ \displaystyle\int_0^T \nabla_\lambda \mathcal{H}(x^*(s), u^*(s), \lambda^*, p(s))\, ds \end{pmatrix} = p^0 \nabla \phi(x^*(0), x^*(T), \lambda^*) + \nabla g(x^*(0), x^*(T), \lambda^*)\xi,$$

for some $\xi \in \mathrm{N}_{\mathrm{S}}[g(x^*(0), x^*(T), \lambda^*)]$;

(iii) the Hamiltonian maximization condition $u^*(t) \in \arg\max_{\omega \in \mathrm{U}} \mathcal{H}(x^*(t), \omega, \lambda^*, p(t))$ for almost every $t \in A$.

The proof of Theorem 2.1 is quite simple and will be developed in a few lines. It is based on the next preliminary PMP for L^1_A–local solutions to Problem (P).

Lemma 2.1 (PMP for L^1_A–local solutions). *If (x^*, u^*, λ^*) is a L^1_A–local solution to Problem* (P)*, for a measurable subset $A \subset [0,T]$, such that g is submersive at $(x^*(0), x^*(T), \lambda^*)$, then the conclusion of Theorem 2.1 holds true.*

About the proof of Lemma 2.1. A PMP for L^1_A–local solutions to classical optimal control problems can be established via many different methods known in the literature. In our context, since the measurable subset A can be of complex nature (such as a Cantor set of positive measure), the classical needle-like perturbations of the control (see, *e.g.*, [15, 27]) may not be suitable for the sensitivity analysis of the control system and, therefore, one may prefer to use *implicit spike variations* (see, *e.g.*, [10, 12, 24]). To deal with the parameter $\lambda \in \mathbb{R}^d$ in Problem (P), one can simply augment the state variable from x to (x, λ) by adding the state equation $\dot{\lambda}(t) = 0_{\mathbb{R}^d}$ (see, *e.g.*, [9]). Finally, to deal with the general mixed terminal state constraints $g(x(0), x(T), \lambda) \in \mathrm{S}$ in Problem (P), one may use the Ekeland variational principle on a penalized functional involving the square of the distance function to S (see, *e.g.*, [12, 20]). Since all these techniques are very well known in the literature, the proof of Lemma 2.1 is omitted. □

Proof of Theorem 2.1. Consider an increasing family $(A_\varepsilon)_{\varepsilon>0}$ of measurable subsets of A associated with (x^*, u^*, λ^*) and a decreasing positive sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ such that $\varepsilon_k \to 0$. In the sequel we denote by $A_k := A_{\varepsilon_k}$ and by $(p_k, p_k^0) \in \mathrm{AC}([0,T], \mathbb{R}^n) \times \mathbb{R}_+$ the nontrivial pair provided by Lemma 2.1 (with $\xi_k \in \mathrm{N}_{\mathrm{S}}[g(x^*(0), x^*(T), \lambda^*)]$) for all $k \in \mathbb{N}$. From linearity and submersiveness, the pair (ξ_k, p_k^0) is nontrivial and can be renormalized so that $\|(\xi_k, p_k^0)\|_{\mathbb{R}^\ell \times \mathbb{R}} = 1$ for all $k \in \mathbb{N}$. Therefore, up to a subsequence that we do not relabel, the sequence $(\xi_k, p_k^0)_{k\in\mathbb{N}}$ converges to some nontrivial pair (ξ, p^0) satisfying $(\xi, p^0) \in \mathrm{N}_{\mathrm{S}}[g(x^*(0), x^*(T), \lambda^*)] \times \mathbb{R}_+$ from closure of the normal cone. Define $p \in \mathrm{AC}([0,T], \mathbb{R}^n)$ as the unique global solution to

$$\begin{cases} \dot{p}(t) = -\nabla_x f(x^*(t), u^*(t), \lambda^*)^\top p(t), \quad \text{a.e. } t \in [0,T], \\ p(T) = -p^0 \nabla_2 \phi(x^*(0), x^*(T), \lambda^*) - \nabla_2 g(x^*(0), x^*(T), \lambda^*)\xi. \end{cases}$$

The Hamiltonian system and the second component of the endpoint transversality condition are satisfied. Since p and p_k satisfy the same linear differential equation and $p_k(T) \to p(T)$, the sequence $(p_k)_{k\in\mathbb{N}}$ uniformly converges to p over $[0,T]$. We deduce the first and third components of the endpoint transversality condition and, from submersiveness, that the pair (p,p^0) is nontrivial. Still from Lemma 2.1, there exists a null set $N_k \subset A_k$ such that $\mathcal{H}(x^*(t),u^*(t),\lambda^*,p_k(t)) \geq \mathcal{H}(x^*(t),\omega,\lambda^*,p_k(t))$ for all $\omega \in \mathrm{U}$ and all $t \in A_k \backslash N_k$, for all $k \in \mathbb{N}$. Now let us prove that the Hamiltonian maximization condition holds true at any $t \in \tilde{A} := (\cup_{k\in\mathbb{N}} A_k)\backslash(\cup_{k\in\mathbb{N}} N_k)$ which is a measurable subset of A with full measure. Let $t \in \tilde{A}$ and take $k_0 \in \mathbb{N}$ such that $t \in A_k\backslash N_k$ for all $k \geq k_0$. Therefore the previous inequality holds true at t for all $\omega \in \mathrm{U}$ and all $k \geq k_0$. From convergence of $p_k(t)$ to $p(t)$, we get that $\mathcal{H}(x^*(t),u^*(t),\lambda^*,p(t)) \geq \mathcal{H}(x^*(t),\omega,\lambda^*,p(t))$ for all $\omega \in \mathrm{U}$, which ends the proof. □

Remark 2.3. (i) First of all we bring the reader's attention to the fact that the Hamiltonian maximization condition in Lemma 2.1 and Theorem 2.1 holds true only almost everywhere over A (and not over the entire interval $[0,T]$). This is the only difference with the classical PMP and this is due, of course, to the restrictions to L^1_A– and $\mathrm{L}^1_{A\square}$–local solutions (see Definitions 2.1 and 2.2 and Item (iii) of Remark 2.2).

(ii) Even if the conclusions of Lemma 2.1 and Theorem 2.1 are exactly the same, we recall that a $\mathrm{L}^1_{A\square}$–local solution is not a L^1_A–local solution in general (see Item (iv) of Remark 2.2). Therefore Theorem 2.1 is not only a consequence of Lemma 2.1 but also a strict extension. From the diagram in Remark 2.2, it is also clear that the classical PMP (for global solutions or for L^1–local solutions) is a particular case of both Lemma 2.1 and Theorem 2.1 (by taking $A = [0,T]$).

(iii) As explained in [6, 7], the submersiveness hypothesis can be removed but, in that case, all items of Lemma 2.1 and Theorem 2.1 remain valid, except Item (ii).

(iv) Consider the framework of Theorem 2.1 for a $\mathrm{L}^1_\square$–local solution (x^*,u^*,λ^*). Using the Hamiltonian system and the Hamiltonian maximization condition over $[0,T]$ and applying [21, Theorem 2.6.1], we obtain the *Hamiltonian constancy condition* $\mathcal{H}(x^*(t),u^*(t),\lambda^*,p(t)) = c$ for almost every $t \in [0,T]$, for some $c \in \mathbb{R}$.

3. Derivation of a HMP for spatially heterogeneous dynamics. In this section we consider a partition of the state space $\mathbb{R}^n = \cup_{j\in\mathcal{J}} \overline{X_j}$, where $\mathcal{J}$ is a family of indexes and the nonempty open subsets $X_j \subset \mathbb{R}^n$, called *regions*, are disjoint. Our aim is to derive first-order necessary optimality conditions in a Pontryagin form for the *hybrid* optimal control problem with mixed terminal state constraints given by

$$\text{(HP)} \qquad \begin{array}{ll} \text{minimize} & \phi(x(0),x(T)), \\ \text{subject to} & (x,u) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m), \\ & \dot{x}(t) = h(x(t),u(t)), \quad \text{a.e. } t \in [0,T], \\ & g(x(0),x(T)) \in \mathrm{S}, \\ & u(t) \in \mathrm{U}, \quad \text{a.e. } t \in [0,T], \end{array}$$

where the data assumptions and the terminology for Problem (HP) are the same as those for Problem (P), except that the dynamics $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is *spatially heterogeneous*, in the sense that it is defined *regionally* by

$$\forall (x,u) \in \mathbb{R}^n \times \mathbb{R}^m, \quad h(x,u) := h_j(x,u) \quad \text{when } x \in X_j,$$

where the *subdynamics* $h_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are of class C^1. Note that $h(x,u)$ is not defined when $x \notin \cup_{j\in\mathcal{J}} X_j$ but this fact will have no impact on the rest of this work (see Item (i) in Remark 3.1). Finally, in contrary to Problem (P) and as explained in Item (ii) of Remark 2.1, note that Problem (HP) does not involve any parameter.

3.1. Regular solutions to the hybrid control system. Due to the discontinuities of the spatially heterogeneous dynamics h, we need to precise the definition of a solution to the hybrid control system

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

associated with Problem (HP).

DEFINITION 3.1 (Solution to (HS)). *A pair* $(x, u) \in \mathrm{AC}([0, T], \mathbb{R}^n) \times \mathrm{L}^\infty([0, T], \mathbb{R}^m)$ *is said to be a solution to* (HS) *if there exists a partition* $\mathbb{T} = \{\tau_k\}_{k=0,\dots,N}$ *such that:*

(i) For all $k \in \{1, \dots, N\}$, *there exists* $j(k) \in \mathcal{J}$ *(with* $j(k) \neq j(k-1)$*) such that* $x(t) \in X_{j(k)}$ *for almost every* $t \in (\tau_{k-1}, \tau_k)$.

(ii) $x(0) \in X_{j(1)}$ *and* $x(T) \in X_{j(N)}$.

(iii) $\dot{x}(t) = h_{j(k)}(x(t), u(t))$ *for almost every* $t \in (\tau_{k-1}, \tau_k)$ *and all* $k \in \{1, \dots, N\}$.

In that case, to ease notation, we set $f_k := h_{j(k)}$ *and* $E_k := X_{j(k)}$ *for all* $k \in \{1, \dots, N\}$. *With this system of notations, we have*

$$\begin{cases} x(t) \in E_1, & \forall t \in [\tau_0, \tau_1), \\ x(t) \in E_k, & \forall t \in (\tau_{k-1}, \tau_k), \quad \forall k \in \{2, \dots, N-1\}, \\ x(t) \in E_N, & \forall t \in (\tau_{N-1}, \tau_N], \\ \dot{x}(t) = f_k(x(t), u(t)), & \text{a.e. } t \in (\tau_{k-1}, \tau_k), \quad \forall k \in \{1, \dots, N\}. \end{cases}$$

Finally the times τ_k, *for* $k \in \{1, \dots, N-1\}$, *are called crossing times since they correspond to the instants at which the trajectory* x *goes from the region* E_k *to the region* E_{k+1}, *and thus* $x(\tau_k) \in \partial E_k \cap \partial E_{k+1}$.

Our main result (Theorem 3.1 stated in Section 3.3) is based on some regularity assumptions made on the behavior of the optimal pair of Problem (HP) at each crossing time. These hypotheses are precised in the next definition.

DEFINITION 3.2 (Regular solution to (HS)). *Following the notations introduced in Definition 3.1, a solution* $(x, u) \in \mathrm{AC}([0, T], \mathbb{R}^n) \times \mathrm{L}^\infty([0, T], \mathbb{R}^m)$ *to* (HS) *is said to be regular if the following conditions are both satisfied:*

(i) At each crossing time τ_k, *there exists a* C^1 *function* $F_k : \mathbb{R}^n \to \mathbb{R}$ *such that*

$$\exists \nu_k > 0, \quad \forall z \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x(\tau_k), \nu_k), \quad \begin{cases} z \in E_k & \Leftrightarrow & F_k(z) < 0, \\ z \in \partial E_k \cap \partial E_{k+1} & \Leftrightarrow & F_k(z) = 0, \\ z \in E_{k+1} & \Leftrightarrow & F_k(z) > 0. \end{cases}$$

In particular it holds that $F_k(x(\tau_k)) = 0$.

(ii) At each crossing time τ_k, *there exists* $\alpha_k > 0$ *and* $\beta_k > 0$ *such that the transverse conditions*

$$\text{(TC)} \quad \begin{cases} \langle \nabla F_k(x(\tau_k)), f_k(x(\tau_k), u(t)) \rangle_{\mathbb{R}^n} \geq \beta_k, & \text{a.e. } t \in [\tau_k - \alpha_k, \tau_k), \\ \langle \nabla F_k(x(\tau_k)), f_{k+1}(x(\tau_k), u(t)) \rangle_{\mathbb{R}^n} \geq \beta_k, & \text{a.e. } t \in (\tau_k, \tau_k + \alpha_k], \end{cases}$$

are both satisfied.

Remark 3.1. (i) Definition 3.1 does not include the possibility of an infinite number of crossing times (excluding the Zeno phenomenon [34]). Also it does not allow trajectories bouncing against a boundary of a region, or moving along a boundary (excluding situations as described in [1]). This last restriction is the reason why the fact that $h(x, u)$ is not defined when $x \notin \cup_{j \in \mathcal{J}} X_j$ has no impact on the present work. Finally Definition 3.1 allows terminal states $x(0)$ and $x(T)$ that belong to regions only (and not to their boundaries). Possible relaxations are presented in Remark 3.4.

(ii) The transverse conditions (TC) have a geometrical interpretation, meaning that x does not cross the boundary $\partial E_k \cap \partial E_{k+1}$ tangentially. At a crossing time τ_k, the transverse conditions

$$\text{(TC')} \quad \begin{cases} u \text{ admits left and right limits at } \tau_k \text{ denoted by } u^-(\tau_k) \text{ and } u^+(\tau_k), \\ \langle \nabla F_k(x(\tau_k)), f_k(x(\tau_k), u^-(\tau_k)) \rangle_{\mathbb{R}^n} > 0 \\ \langle \nabla F_k(x(\tau_k)), f_{k+1}(x(\tau_k), u^+(\tau_k)) \rangle_{\mathbb{R}^n} > 0, \end{cases}$$

considered in the papers [1, 23], are (slightly) stronger than (TC).

3.2. Reduction into a classical optimal control problem with parameter. To establish a correspondence from the hybrid optimal control problem (HP) to a classical optimal control problem with parameter of the form of Problem (P), we will proceed as in [19] to affine changes of time variable. Precisely let $(x^*, u^*) \in \mathrm{AC}([0,T], \mathbb{R}^n) \times \mathrm{L}^\infty([0,T], \mathbb{R}^m)$ be a solution to (HS), associated with a partition $\mathbb{T}^* = \{\tau_k^*\}_{k=0,\dots,N}$, and let E_k^* and f_k^* stand for the corresponding regions and functions (see Definition 3.1). We introduce $(y^*, v^*) \in \mathrm{AC}([0,1], \mathbb{R}^{nN}) \times \mathrm{L}^\infty([0,1], \mathbb{R}^{mN})$ defined by

$$\text{(3.1)} \quad y_k^*(s) := x^*(\tau_{k-1}^* + (\tau_k^* - \tau_{k-1}^*)s) \quad \text{and} \quad v_k^*(s) := u^*(\tau_{k-1}^* + (\tau_k^* - \tau_{k-1}^*)s),$$

for all $s \in [0,1]$ and all $k \in \{1, \dots, N\}$. To invert the changes of time variable, it holds

$$\text{(3.2)} \qquad x^*(t) = y_k^* \left(\frac{t - \tau_{k-1}^*}{\tau_k^* - \tau_{k-1}^*} \right) \quad \text{and} \quad u^*(t) = v_k^* \left(\frac{t - \tau_{k-1}^*}{\tau_k^* - \tau_{k-1}^*} \right),$$

for all $t \in [\tau_{k-1}^*, \tau_k^*]$ and all $k \in \{1, \dots, N\}$. In particular note that $(x^*(0), x^*(T)) = (y_1^*(0), y_N^*(1))$. From a more general point of view, it holds that $x^*(\tau_k^*) = y_{k+1}^*(0)$ for all $k \in \{0, \dots, N-1\}$ and $x^*(\tau_k^*) = y_k^*(1)$ for all $k \in \{1, \dots, N\}$. Note that the triplet $(y^*, v^*, \mathbb{T}^*)$ satisfies

$$\dot{y}^*(s) = f^*(y^*(s), v^*(s), \mathbb{T}^*), \quad \text{a.e. } s \in [0,1],$$

where $f^* : \mathbb{R}^{nN} \times \mathbb{R}^{mN} \times \mathbb{R}^{N+1} \to \mathbb{R}^{nN}$ is the C^1 function defined by

$$f^*(y, v, \mathbb{T}) := \Big((\tau_1 - \tau_0) f_1^*(y_1, v_1), \dots, (\tau_N - \tau_{N-1}) f_N^*(y_N, v_N) \Big),$$

for all $y = (y_1, \dots, y_N) \in \mathbb{R}^{nN}$, $v = (v_1, \dots, v_N) \in \mathbb{R}^{mN}$ and $\mathbb{T} = \{\tau_0, \dots, \tau_N\} \in \mathbb{R}^{N+1}$. Furthermore it holds that

$$\text{(3.3)} \qquad \begin{cases} y_1^*(s) \in E_1, & \forall s \in [0,1), \\ y_k^*(s) \in E_k, & \forall s \in (0,1), \quad \forall k \in \{2, \dots, N-1\}, \\ y_N^*(s) \in E_N, & \forall s \in (0,1], \end{cases}$$

and $y_{k+1}^*(0) = y_k^*(1) \in \partial E_k^* \cap \partial E_{k+1}^*$ for all $k \in \{1, \dots, N-1\}$. Also note that $\mathbb{T}^* \in \Delta$ where $\Delta \subset \mathbb{R}^{N+1}$ is the nonempty closed convex set defined by

$$\Delta := \{ \mathbb{T} = \{\tau_k\}_{k=0,\dots,N} \in \mathbb{R}^{N+1} \mid 0 = \tau_0 \leq \tau_1 \leq \ldots \leq \tau_{N-1} \leq \tau_N = T \}.$$

Now assume that the pair (x^*, u^*) is moreover a regular solution to (HS) and denote by F_k^* and $\nu_k^* > 0$ the corresponding functions and positive radii (see Definition 3.2). In that context note that $F_k^*(x(\tau_k^*)) = F_k^*(y_k^*(1)) = 0$ for all $k \in \{1, \dots, N-1\}$. Finally it is clear that, if the pair (x^*, u^*) is furthermore admissible for Problem (HP), then the triplet $(y^*, v^*, \mathbb{T}^*)$ is admissible for the classical optimal control problem with parameter given by

$$
(\mathrm{CP}^*) \qquad \begin{array}{ll} \text{minimize} & \phi^*(y(0),y(1),\mathbb{T}), \\ \text{subject to} & (y,v,\mathbb{T}) \in \mathrm{AC}([0,1],\mathbb{R}^{nN}) \times \mathrm{L}^\infty([0,1],\mathbb{R}^{mN}) \times \mathbb{R}^{N+1}, \\ & \dot{y}(s) = f^*(y(s),v(s),\mathbb{T}), \quad \text{a.e. } s \in [0,1], \\ & g^*(y(0),y(1),\mathbb{T}) \in \mathrm{S}^*, \\ & v(s) \in \mathrm{U}^N, \quad \text{a.e. } s \in [0,1], \end{array}
$$

where $\phi^* : \mathbb{R}^{nN} \times \mathbb{R}^{nN} \times \mathbb{R}^{N+1} \to \mathbb{R}$ and $g^* : \mathbb{R}^{nN} \times \mathbb{R}^{nN} \times \mathbb{R}^{N+1} \to \mathbb{R}^{\ell^*}$ are the C^1 functions defined by $\phi^*(y^0,y^1,\mathbb{T}) := \phi(y_1^0,y_N^1)$ and

$$
g^*(y^0,y^1,\mathbb{T}) := (g(y_1^0,y_N^1), y_2^0 - y_1^1, \ldots, y_N^0 - y_{N-1}^1, F_1^*(y_1^1), \ldots, F_{N-1}^*(y_{N-1}^1), \mathbb{T}),
$$

for all $y^0 = (y_1^0,\ldots,y_N^0)$, $y^1 = (y_1^1,\ldots,y_N^1) \in \mathbb{R}^{nN}$ and $\mathbb{T} = \{\tau_0,\ldots,\tau_N\} \in \mathbb{R}^{N+1}$, where $\ell^* := \ell + n(N-1) + (N-1) + (N+1)$, and where $\mathrm{S}^* \subset \mathbb{R}^{\ell^*}$ stands for the nonempty closed convex set defined by

$$
\mathrm{S}^* := \mathrm{S} \times \{0_{\mathbb{R}^n}\}^{N-1} \times \{0\}^{N-1} \times \Delta.
$$

Proposition 3.1. *If $(x^*,u^*) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ is a global solution to Problem* (HP)*, that is moreover a regular solution to* (HS)*, associated with a partition $\mathbb{T}^* = \{\tau_k^*\}_{k=0,\ldots,N}$, then the triplet $(y^*,v^*,\mathbb{T}^*)$ constructed above is a $\mathrm{L}^1_\square$–local solution to Problem* (CP*)*.*

Proof. The proof of Proposition 3.1 is postponed to Appendix A. We prove that the triplet $(y^*,v^*,\mathbb{T}^*)$ is a $\mathrm{L}^1_{[\varepsilon,1-\varepsilon]}$–local solution to Problem (CP*) for any $0 < \varepsilon < 1/2$. □

Remark 3.2. (i) Consider the framework of Proposition 3.1. In Section 3.4 we will provide a counterexample showing that the triplet $(y^*,v^*,\mathbb{T}^*)$ is not a L^1–local solution to Problem (CP*) in general. This highlights the fact that the classical PMP cannot be applied to the triplet $(y^*,v^*,\mathbb{T}^*)$. However, thanks to Proposition 3.1, we can apply the new PMP for $\mathrm{L}^1_\square$–local solution obtained in Theorem 2.1. This allows us to derive a HMP for Problem (HP) in the next Section 3.3.

(ii) Consider the framework of Proposition 3.1. Given an admissible triplet $(y,v,\mathbb{T})$ for Problem (CP*), one can easily invert the augmentation procedure and obtain a pair (x,u) which satisfies all the constraints of Problem (HP), except one. Precisely, even if (x,u) follows the same sequence $(f_k^*)_{k=1,\ldots,N}$ of dynamics than the pair (x^*,u^*), it does not necessarily follow the same sequence of regions $(E_k^*)_{k=1,\ldots,N}$ (and thus it is not necessarily admissible for Problem (HP)). This is the major difficulty of the proof of Proposition 3.1 and, as we will see with a counterexample in Section 3.4, the notion of L^1–local solution (which consists in considering the triplet $(y,v,\mathbb{T})$ in a standard neighborhood of $(y^*,v^*,\mathbb{T}^*)$) fails to guarantee this property. This is because, even if transverse conditions are satisfied by the pair (x^*,u^*), allowing L^1-perturbations of u^* (with possibly far values in U from the ones of u^*) in the neighborhoods of the crossing times τ_k^* may lead to a perturbed pair (x,u) that does not satisfy the transverse conditions, and thus to a perturbed trajectory x that may visit a different sequence of regions than x^*. On the contrary, the new notion of $\mathrm{L}^1_{[\varepsilon,1-\varepsilon]}$–local solution, for $0 < \varepsilon < 1/2$, addresses this issue by allowing L^1-perturbations of u^* only outside neighborhoods of the crossing times τ_k^*.

3.3. HMP and comments. The Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ associated with Problem (HP) is defined by $H(x,u,p) := \langle p, h(x,u) \rangle$ for all $(x,u,p) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$. We are now in a position to state the main result of this paper.

THEOREM 3.1 (HMP). *If* $(x^*, u^*) \in \mathrm{AC}([0,T],\mathbb{R}^n) \times \mathrm{L}^\infty([0,T],\mathbb{R}^m)$ *is a global solution to Problem* (HP)*, that is moreover a regular solution to* (HS)*, associated with a partition* $\mathbb{T}^* = \{\tau_k^*\}_{k=0,\dots,N}$*, such that* g *is submersive at* $(x^*(0), x^*(T))$*, then there exists a nontrivial pair* $(p, p^0) \in \mathrm{PAC}_{\mathbb{T}^*}([0,T],\mathbb{R}^n) \times \mathbb{R}_+$ *satisfying:*

(i) the Hamiltonian system $\dot{x}^*(t) = \nabla_p H(x^*(t), u^*(t), p(t))$ *and* $-\dot{p}(t) = \nabla_x H(x^*(t), u^*(t), p(t))$ *for almost every* $t \in [0,T]$*;*

(ii) the endpoint transversality condition

$$\begin{pmatrix} p(0) \\ -p(T) \end{pmatrix} = p^0 \nabla\phi(x^*(0), x^*(T)) + \nabla g(x^*(0), x^*(T))\xi,$$

for some $\xi \in \mathrm{N_S}[g(x^*(0), x^*(T))]$*;*

(iii) the discontinuity condition $p^+(\tau_k^*) - p^-(\tau_k^*) = \sigma_k \nabla F_k^*(x^*(\tau_k^*))$ *for some* $\sigma_k \in \mathbb{R}$*, for all* $k \in \{1,\dots,N-1\}$*;*

(iv) the Hamiltonian maximization condition $u^*(t) \in \arg\max_{\omega \in \mathrm{U}} H(x^*(t), \omega, p(t))$ *for almost every* $t \in [0,T]$*;*

(v) the Hamiltonian constancy condition $H(x^*(t), u^*(t), p(t)) = c$ *for almost every* $t \in [0,T]$*, for some* $c \in \mathbb{R}$*.*

Proof. The proof of Theorem 3.1 is postponed to Appendix B. It is based on Proposition 3.1 and on the application of Theorem 2.1 to the triplet $(y^*, v^*, \mathbb{T}^*)$. ◻

Remark 3.3. (i) In the classical PMP (that is, when the dynamics is not heterogeneous), the costate p is absolutely continuous over the entire interval $[0,T]$ and satisfies Items (i), (ii), (iv) and (v) of Theorem 3.1 (see, *e.g.*, [27]). In the present setting of heterogeneous dynamics, the costate p is (only) piecewise absolutely continuous over $[0,T]$, admitting at each crossing time τ_k^* a discontinuity jump satisfying Item (iii) of Theorem 3.1. Under the (slightly) stronger transverse conditions (TC'), the Hamiltonian constancy condition allows to obtain

$$\sigma_k = -\frac{\left\langle p^-(\tau_k^*), f_{k+1}^*(x^*(\tau_k^*), (u^*)^+(\tau_k^*)) - f_k^*(x^*(\tau_k^*), (u^*)^-(\tau_k^*)) \right\rangle_{\mathbb{R}^n}}{\left\langle \nabla F_k^*(x^*(\tau_k^*)), f_{k+1}^*(x^*(\tau_k^*), (u^*)^+(\tau_k^*)) \right\rangle_{\mathbb{R}^n}}$$

$$= -\frac{\left\langle p^+(\tau_k^*), f_{k+1}^*(x^*(\tau_k^*), (u^*)^+(\tau_k^*)) - f_k^*(x^*(\tau_k^*), (u^*)^-(\tau_k^*)) \right\rangle_{\mathbb{R}^n}}{\left\langle \nabla F_k^*(x^*(\tau_k^*)), f_k^*(x^*(\tau_k^*), (u^*)^-(\tau_k^*)) \right\rangle_{\mathbb{R}^n}},$$

for all $k \in \{1,\dots,N-1\}$, and thus the discontinuity conditions can be expressed as forward (or backard) discontinuity jumps. Such discontinuity jumps are very standard in the literature on hybrid optimal control problems (see, *e.g.*, [8, 26]) and the discontinuity conditions have even been announced in our setting of spatially heterogeneous dynamics in the papers [1, 23]. However, as explained in Introduction, we recall that the proofs in [1, 23] are not satisfactory for several and different reasons.

(ii) Similarly to Item (iii) of Remark 2.3, and as explained in [6, 7], the submersiveness hypothesis made in Theorem 3.1 can be removed but, in that case, all items of Theorem 3.1 remain valid, except Item (ii).

Remark 3.4. (i) Consider the framework of Proposition 3.1. From Item (i) of Remark 3.2, we know that $(y^*, v^*, \mathbb{T}^*)$ is not a L^1–local solution to Problem (CP*) in general. Nevertheless, according to the ideas presented in Item (ii) of Remark 3.2, it may be possible to avoid the use of the notion of $\mathrm{L}^1_\square$–local solution introduced in the present paper. However, to our best knowledge, this would not be possible without obtaining a weaker result and/or without restricting the framework. Let us develop two options in that direction:

- First, under the (slightly) stronger transverse conditions (TC'), it can be proved that $(y^*, v^*, \mathbb{T}^*)$ is a L^∞–local solution to Problem (CP*), in the sense that there exists $\eta > 0$ such that $\phi^*(y^*(0), y^*(1), \mathbb{T}^*) \leq \phi^*(y(0), y(1), \mathbb{T})$ for all admissible triplets $(y, v, \mathbb{T})$ satisfying $\|y - y^*\|_{\mathrm{C}} + \|v - v^*\|_{\mathrm{L}^\infty} + \|\mathbb{T} - \mathbb{T}^*\|_{\mathbb{R}^{N+1}} \leq \eta$. This idea is in-line with the approach developed in [5]. In that context, assuming for simplicity that U is closed and convex and applying a *weak version* of the classical PMP (that is, a version adapted to L^∞–local solutions, see [13] and discussion therein), one can derive a weaker version of Theorem 3.1, that is, without the Hamiltonian constancy condition and, above all, where the Hamiltonian maximization condition is replaced by the weaker *Hamiltonian gradient condition* $\nabla_u H(x^*(t), u^*(t), p(t)) \in \mathrm{N}_{\mathrm{U}}[u^*(t)]$ for a.e. $t \in [0, T]$.
- Second, under the (very) stronger transverse conditions given by

$$\text{(TC")} \qquad \forall \omega \in \mathrm{U}, \quad \begin{cases} \langle \nabla F_k^*(x^*(\tau_k^*)), f_k^*(x^*(\tau_k^*), \omega) \rangle_{\mathbb{R}^n} \geq \beta_k, \\ \langle \nabla F_k^*(x^*(\tau_k^*)), f_{k+1}^*(x^*(\tau_k^*), \omega) \rangle_{\mathbb{R}^n} \geq \beta_k, \end{cases}$$

  for some $\beta_k > 0$ at each crossing time τ_k^*, it can be proved that $(y^*, v^*, \mathbb{T}^*)$ is a L^1–local solution to Problem (CP*). In that context one can derive Theorem 3.1 from the application of the classical PMP. However the strong transverse conditions (TC") are quite restrictive and are not satisfied in practice (see the counterexample presented in the next Section 3.4).

From a general point of view, it can be observed that the choice of the transverse conditions (more or less strong) influences the *local quality* (L^1, L^∞ or $\mathrm{L}^1_\square$) of the solution $(y^*, v^*, \mathbb{T}^*)$ to Problem (CP*) and thus the version of the PMP that can be applied to it, and finally the version of the HMP obtained on the original pair (x^*, u^*).

(ii) For simplicity, Definition 2.1 allows trajectories x such that $x(0)$ and $x(T)$ belong to regions only (and not to their boundaries). This restriction may limit the scope of our results. To overcome this restriction, some adjustments have to be performed. For instance, consider the framework of Theorem 3.1 with $x^*(0) \in E_1$ and $x^*(T) \in \partial E_N$ (other cases can be handled similarly). To deal with this situation, one has to add in Definition 3.2 the existence of a local C^1 description F_N of ∂E_N in a neighborhood of $x^*(T)$ and an adapted transverse condition of the form

$$\langle \nabla F_N^*(x^*(T)), f_N^*(x^*(T), u^*(t)) \rangle_{\mathbb{R}^n} \geq \beta_N, \quad \text{a.e. } t \in [T - \alpha_N, T),$$

with $\alpha_N > 0$ and $\beta_N > 0$. Then the augmented problem (CP*) must be adjusted carefully by adding the inequality constraint $F_N^*(y_N(1)) \leq 0$ to keep the validity of Proposition 3.1. Finally, adapting the submersiveness hypothesis (involving both g and F_N^*), applying Theorem 2.1 and inverting the augmentation procedure, the conclusion of Theorem 3.1 remains valid, but with an additional term of the form $\zeta \nabla F_N^*(x^*(T))$ with $\zeta \geq 0$ in the expression of $-p(T)$.

(iii) In addition to the comments made in the previous Item (ii), we would like to emphasize that certain cases where $x^*(0)$ and $x^*(T)$ belong to boundaries of the regions can be treated without the adjusted procedure discussed above. For instance, if the initial condition is fixed on a boundary, then no information is expected for $p(0)$ and, furthermore, with the approach developed in this paper, only perturbations of the control over $[\varepsilon, T]$ for small $\varepsilon > 0$ are considered. Hence the corresponding perturbed trajectories coincide with the nominal trajectory over $[0, \varepsilon]$ and thus satisfy the initial condition. Another example is provided with minimum time problems where the target belongs to a boundary of a region. In that context, a simple dynamical programming argument can eliminate the need of a transverse condition at T (see [4]).

(iv) Here we focus on possible extensions and perspectives of our work.

- First, one may consider a setting where the subdynamics $h_j : \mathbb{R}^n \times \mathbb{R}^{m_j} \to \mathbb{R}^n$ have possibly different control dimensions $m_j \in \mathbb{N}^*$ and with possibly different control constraint sets $\mathrm{U}_j \subset \mathbb{R}^{m_j}$. This generalized context is interesting to impose specific values for the control in certain regions (for example, by taking $\mathrm{U}_j = \{0_{\mathbb{R}^{m_j}}\}$ for some $j \in \mathcal{J}$). We believe that our methodology can be adapted to this framework without any major difficulty.
- Second, one may consider an extended setting that includes a *regionally switching parameter* (see [2]), meaning that the control system depends on a parameter that remains constant in each region but can change its value when the state crosses a boundary. This framework enables us to handle, as a specific case, control systems with *loss control regions* (see [3, 4]). This extension is the subject of a work in progress by the authors.

3.4. A counterexample. Consider the framework of Proposition 3.1. This section is dedicated to an explicit counterexample showing that the triplet $(y^*, v^*, \mathbb{T}^*)$ is not a L^1–local solution to Problem (CP*) in general. To this aim consider the two-dimensional case $n = 2$, the state space partition $\mathbb{R}^2 = \overline{X_1} \cup \overline{X_2}$ where $X_1 := (-\infty, 1) \times \mathbb{R}$ and $X_2 := (1, +\infty) \times \mathbb{R}$, and the hybrid optimal control problem given by

$$(\mathrm{HP}_{\mathrm{ex}}) \qquad \begin{array}{ll} \text{minimize} & -(x_1(2) - 2)^3 - \rho x_2(2), \\ \text{subject to} & (x, u) \in \mathrm{AC}([0, 2], \mathbb{R}^2) \times \mathrm{L}^\infty([0, 2], \mathbb{R}), \\ & \dot{x}(t) = h(x(t), u(t)), \quad \text{a.e. } t \in [0, 2], \\ & x(0) = 0_{\mathbb{R}^2}, \\ & u(t) \in [-1, 1], \quad \text{a.e. } t \in [0, 2], \end{array}$$

where the spatially heterogeneous dynamics $h : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2$ is defined by

$$h(x, u) := \begin{cases} \Big(1, ((1 - x_1)^+)^2\Big), & \text{if } x \in X_1, \\ \Big(u, ((1 - x_1)^+)^2\Big), & \text{if } x \in X_2, \end{cases}$$

for all $x = (x_1, x_2) \in X_1 \cup X_2$ and all $u \in \mathbb{R}$, and where $\rho > 96$.

3.4.1. A global solution (x^*, u^*) to Problem (HP$_{\mathrm{ex}}$). In view of the definition of h in the region X_1 and following Definition 2.1, any admissible pair (x, u) for Problem (HP$_{\mathrm{ex}}$) has exactly one crossing time $\tau_1 = 1$, and satisfies $x_1(t) = t$ for all $t \in [0, 1]$ and $x_1(t) > 1$ for all $t \in (1, 2]$. Moreover an easy computation shows that

$$x_2(t) = \begin{cases} \frac{1}{3}((t-1)^3 + 1) & \text{if } t \in [0, 1], \\ \frac{1}{3} & \text{if } t \in [1, 2], \end{cases}$$

for all $t \in [0, 2]$. Since the value $x_2(2)$ is fixed to $\frac{1}{3}$ for any admissible pair, Problem (HP$_{\mathrm{ex}}$) simply amounts to maximize the value of $x_1(2)$. In view of the definition of h in the region X_2, one can easily deduce that a global solution (x^*, u^*) to Problem (HP$_{\mathrm{ex}}$) is given by

$$x_1^*(t) := t, \quad x_2^*(t) := \begin{cases} \frac{1}{3}((t-1)^3 + 1) & \text{if } t \in [0, 1], \\ \frac{1}{3} & \text{if } t \in [1, 2], \end{cases} \qquad u^*(t) := 1,$$

for all $t \in [0,2]$, and the corresponding optimal cost is given by $\mathcal{C}^* := -\frac{\rho}{3}$. Furthermore one can observe that the pair (x^*, u^*) is a regular solution to the corresponding hybrid control system (Definition 3.2) with exactly one crossing time $\tau_1^* = 1$.

3.4.2. The corresponding triplet $(y^*, v^*, \mathbb{T}^*)$. Now consider the framework of Proposition 3.1. The corresponding triplet $(y^*, v^*, \mathbb{T}^*)$ is given by

$$(y_1^1)^*(s) = s, \quad (y_2^1)^*(s) = s+1, \quad (y_1^2)^*(s) = \frac{1}{3}((s-1)^3+1), \quad (y_2^2)^*(s) = \frac{1}{3},$$

and $v_1^*(s) = v_2^*(s) = 1$ for all $s \in [0,1]$, and $\mathbb{T}^* = \{0,1,2\}$. As expected the triplet $(y^*, v^*, \mathbb{T}^*)$ is admissible for the classical optimal control problem with parameter

$$(\mathrm{CP}^*_{\mathrm{ex}}) \qquad \begin{array}{ll} \text{minimize} & -(y_2^1(1)-2)^3 - \rho y_2^2(1), \\ \text{subject to} & (y,v,\mathbb{T}) \in \mathrm{AC}([0,1],\mathbb{R}^4) \times \mathrm{L}^\infty([0,1],\mathbb{R}^2) \times \mathbb{R}^3, \\ & \dot{y}_1^1(s) = \tau_1, \quad \text{a.e. } s \in [0,1], \\ & \dot{y}_1^2(s) = \tau_1((1-y_1^1(s))^+)^2, \quad \text{a.e. } s \in [0,1], \\ & \dot{y}_2^1(s) = (2-\tau_1)v_2(s), \quad \text{a.e. } s \in [0,1], \\ & \dot{y}_2^2(s) = (2-\tau_1)((1-y_2^1(s))^+)^2, \quad \text{a.e. } s \in [0,1], \\ & y_1^1(0) = 0, \quad y_1^2(0) = 0, \quad y_1^1(1) - 1 = 0, \\ & y_2^1(0) - y_1^1(1) = 0, \quad y_2^2(0) - y_1^2(1) = 0, \\ & \tau_0 = 0, \quad \tau_1 \in [0,2], \quad \tau_2 = 2, \\ & v_1(s),\, v_2(s) \in [-1,1], \quad \text{a.e. } s \in [0,1], \end{array}$$

with the cost $\mathcal{C}^* = -\frac{\rho}{3}$.

3.4.3. The triplet $(y^*, v^*, \mathbb{T}^*)$ is not a L^1–local solution to Problem $(\mathrm{CP}^*_{\mathrm{ex}})$. For any $\varepsilon > 0$ small enough, we introduce the triplet $(y^\varepsilon, v^\varepsilon, \mathbb{T}^\varepsilon)$ defined by $(y_1^1)^\varepsilon := (y_1^1)^*$, $(y_1^2)^\varepsilon := (y_1^2)^*$, $v_1^\varepsilon = v_1^*$, $\mathbb{T}^\varepsilon = \mathbb{T}^*$, and by

$$(y_2^1)^\varepsilon(s) := \begin{cases} s+1, & \text{if } s \in [0,\varepsilon], \\ 2\varepsilon - s + 1, & \text{if } s \in [\varepsilon, 3\varepsilon], \\ s - 4\varepsilon + 1, & \text{if } s \in [3\varepsilon, 1], \end{cases} \qquad v_2^\varepsilon(s) := \begin{cases} 1, & \text{if } s \in [0,\varepsilon], \\ -1, & \text{if } s \in [\varepsilon, 3\varepsilon], \\ 1, & \text{if } s \in [3\varepsilon, 1], \end{cases}$$

and

$$(y_2^2)^\varepsilon(s) := \begin{cases} \frac{1}{3}, & \text{if } s \in [0, 2\varepsilon], \\ \frac{1}{3}((s-2\varepsilon)^3+1), & \text{if } s \in [2\varepsilon, 3\varepsilon], \\ \frac{1}{3}((s-4\varepsilon)^3 + 2\varepsilon^3 + 1), & \text{if } s \in [3\varepsilon, 4\varepsilon], \\ \frac{1}{3}(2\varepsilon^3+1), & \text{if } s \in [4\varepsilon, 1], \end{cases}$$

for all $s \in [0,1]$. One can easily conclude that the triplet (y^*, v^*, τ^*) is not a L^1–local solution to Problem $(\mathrm{CP}^*_{\mathrm{ex}})$ since:

– The triplet $(y^\varepsilon, v^\varepsilon, \tau^\varepsilon)$ is admissible for Problem $(\mathrm{CP}^*_{\mathrm{ex}})$ for any $\varepsilon > 0$.

– It holds that $\lim_{\varepsilon \to 0}(\|y^\varepsilon - y^*\|_{\mathrm{C}} + \|v^\varepsilon - v^*\|_{\mathrm{L}^1} + \|\mathbb{T}^\varepsilon - \mathbb{T}^*\|_{\mathbb{R}^3}) = 0$.

– For any $\varepsilon > 0$, the cost $\mathcal{C}^\varepsilon$ associated with the triplet $(y^\varepsilon, v^\varepsilon, \mathbb{T}^\varepsilon)$ is given by

$$\mathcal{C}^\varepsilon = -\frac{\rho}{3} - \left(\frac{2\rho}{3} - 64\right)\varepsilon^3 < -\frac{\rho}{3} = \mathcal{C}^*.$$

Appendix A. Proof of Proposition 3.1. Consider the framework of Proposition 3.1 and let us prove that the triplet $(y^*, v^*, \mathbb{T}^*)$ is a $\mathrm{L}^1_{[\varepsilon,1-\varepsilon]}$–local solution to

Problem (CP*) for any $0<\varepsilon<\frac{1}{2}$. Therefore let $0<\varepsilon<\frac{1}{2}$ and $R\geq\|v^*\|_{\mathrm{L}^\infty}$. Our aim is to prove that there exists $\eta>0$ such that $\phi^*(y^*(0),y^*(1),\mathbb{T}^*)\leq\phi^*(y(0),y(1),\mathbb{T})$ for any triplet $(y,v,\mathbb{T})$ that is admissible for Problem (CP*) and satisfying

$$\text{(A.1)}\qquad \begin{cases} \|y-y^*\|_{\mathrm{C}}+\|v-v^*\|_{\mathrm{L}^1}+\|\mathbb{T}-\mathbb{T}^*\|_{\mathbb{R}^{N+1}}\leq\eta, \\ \|v\|_{\mathrm{L}^\infty}\leq R, \\ v(s)=v^*(s) \text{ a.e. } s\in[0,\varepsilon]\cup[1-\varepsilon,1]. \end{cases}$$

To this aim we need to introduce several technical positive parameters:

($\mathfrak{P}_1$) Let $\underline{\theta}:=\min_{k\in\{1,\dots,N\}}|\tau^*_k-\tau^*_{k-1}|>0$ and $\overline{\theta}:=\max_{k\in\{1,\dots,N\}}|\tau^*_k-\tau^*_{k-1}|>0$.

($\mathfrak{P}_2$) From the transverse conditions (see Definition 3.2) and the (uniform) continuities of the functions ∇F^*_k and f^*_k on compact sets, there exist $0<\nu\leq\min_{k\in\{1,\dots,N-1\}}\nu^*_k$ and $0<\alpha\leq\min\{\frac{\underline{\theta}}{3},\min_{k\in\{1,\dots,N-1\}}\alpha^*_k\}$ such that

$$\begin{cases} \langle\nabla F^*_k(z),f^*_k(z,u^*(t))\rangle_{\mathbb{R}^n}>0, & \text{a.e. } t\in[\tau^*_k-\alpha,\tau^*_k), \\ \langle\nabla F^*_k(z),f^*_{k+1}(z,u^*(t))\rangle_{\mathbb{R}^n}>0, & \text{a.e. } t\in(\tau^*_k,\tau^*_k+\alpha], \end{cases}$$

for all $z\in\overline{\mathrm{B}}_{\mathbb{R}^n}(x^*(\tau^*_k),\nu)$ and all $k\in\{1,\dots,N-1\}$.

($\mathfrak{P}_3$) From continuity of y^* over $[0,1]$, there exists $0<\chi<\frac{1}{2}$ such that $\|y^*_k(s)-y^*_k(0)\|_{\mathbb{R}^n}\leq\frac{\nu}{2}$ for all $s\in[0,\chi]$ and $\|y^*_k(s)-y^*_k(1)\|_{\mathbb{R}^n}\leq\frac{\nu}{2}$ for all $s\in[1-\chi,1]$, for all $k\in\{1,\dots,N\}$.

($\mathfrak{P}_4$) Define $\gamma:=\frac{\underline{\theta}}{3}\min\{\varepsilon,\chi,\frac{\alpha}{\overline{\theta}}\}>0$ and $r:=\frac{\gamma}{\overline{\theta}+\underline{\theta}}>0$. Note that $0<\gamma\leq\alpha\leq\frac{\underline{\theta}}{3}$ and $0<r<\frac{1}{2}$.

($\mathfrak{P}_5$) From continuity of y^*, from (3.3) and the openness of the regions E^*_k, there exists $\delta>0$ such that

$$\begin{cases} \overline{\mathrm{B}}_{\mathbb{R}^n}(y^*_1(s),\delta)\subset E^*_1, & \forall s\in[0,1-r], \\ \overline{\mathrm{B}}_{\mathbb{R}^n}(y^*_k(s),\delta)\subset E^*_k, & \forall s\in[r,1-r], \quad \forall k\in\{2,\dots,N-1\}, \\ \overline{\mathrm{B}}_{\mathbb{R}^n}(y^*_N(s),\delta)\subset E^*_N, & \forall s\in[r,1]. \end{cases}$$

We are now in a position to continue the proof. To this aim let $\eta:=\min\{\frac{\underline{\theta}}{3},\frac{\nu}{2},\delta\}>0$ and $(y,v,\mathbb{T})$ be an admissible triplet for Problem (CP*) satisfying (A.1). Our aim is to prove that $\phi^*(y^*(0),y^*(1),\mathbb{T}^*)\leq\phi^*(y(0),y(1),\mathbb{T})$.

Step 1. Since $0=\tau^*_0<\tau^*_1<\cdots<\tau^*_{N-1}<\tau^*_N=T$ and $\mathbb{T}\in\Delta$ with $\|\mathbb{T}-\mathbb{T}^*\|_{\mathbb{R}^{N+1}}\leq\eta\leq\frac{\underline{\theta}}{3}$, one can easily deduce that $0=\tau_0<\tau_1<\cdots<\tau_{N-1}<\tau_N=T$. Therefore we are in a position to define $(x,u)\in\mathrm{AC}([0,T],\mathbb{R}^n)\times\mathrm{L}^\infty([0,T],\mathbb{R}^m)$ by

$$\text{(A.2)}\qquad x(t):=y_k\left(\frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}}\right)\quad\text{and}\quad u(t):=v_k\left(\frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}}\right),$$

for all $t\in[\tau_{k-1},\tau_k]$ and all $k\in\{1,\dots,N\}$. Note that x is well defined since $y_{k+1}(0)=y_k(1)$ for all $k\in\{2,\dots,N\}$ (from admissibility of the triplet $(y,v,\mathbb{T})$). Observe that $(y_1(0),y_N(1))=(x(0),x(T))$ and recall that $(y^*_1(0),y^*_N(1))=(x^*(0),x^*(T))$. Therefore, from the definition of ϕ^* (see Section 3.2) and since (x^*,u^*) is a global solution to Problem (HP), to obtain that $\phi^*(y^*(0),y^*(1),\mathbb{T}^*)\leq\phi^*(y(0),y(1),\mathbb{T})$, we only need to prove that the pair (x,u) is admissible for Problem (HP).

From admissibility of the triplet $(y,v,\mathbb{T})$, it is clear that $g(x(0),x(T))\in\mathrm{S}$ and $u(t)\in\mathrm{U}$ for almost every $t\in[0,T]$. Therefore it only remains to prove that (x,u) is a solution to the hybrid control system (HS) (see Definition 3.1). From (A.2) and the admissibility of the triplet $(y,v,\mathbb{T})$, one can easily obtain that

$$\text{(A.3)}\qquad \dot{x}(t)=f^*_k(x(t),u(t)),\quad\text{a.e. } t\in(\tau_{k-1},\tau_k),$$

for all $k \in \{1,\dots,N\}$. Therefore, to conclude the proof, we only need to prove that

$$\begin{cases} x(t) \in E_1^*, & \forall t \in [\tau_0,\tau_1), \\ x(t) \in E_k^*, & \forall t \in (\tau_{k-1},\tau_k), \quad \forall k \in \{2,\dots,N-1\}, \\ x(t) \in E_N^*, & \forall t \in (\tau_{N-1},\tau_N]. \end{cases}$$

This is exactly our goal in the next two steps.

Step 2. Since $\|\mathbb{T}-\mathbb{T}^*\|_{\mathbb{R}^{N+1}} \leq \eta \leq \frac{\theta}{3}$, note that $\tau_k - \tau_{k-1} \leq \overline{\theta} + 2\eta \leq \overline{\theta} + \underline{\theta}$ for all $k \in \{1,\dots,N\}$. Hence, since moreover $r := \frac{\gamma}{\overline{\theta}+\underline{\theta}}$, observe that

$$\begin{cases} \frac{t-\tau_0}{\tau_1-\tau_0} \in [0,1-r], & \forall t \in [\tau_0,\tau_1-\gamma], \\ \frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}} \in [r,1-r], & \forall t \in [\tau_{k-1}+\gamma,\tau_k-\gamma], \quad \forall k \in \{2,\dots,N-1\}, \\ \frac{t-\tau_{N-1}}{\tau_N-\tau_{N-1}} \in [r,1], & \forall t \in [\tau_{N-1}+\gamma,\tau_N]. \end{cases}$$

As a consequence, from (A.2) and ($\mathfrak{P}_5$), and since $\|y_k - y_k^*\|_{\mathrm{C}} \leq \|y-y^*\|_{\mathrm{C}} \leq \eta \leq \delta$, one can easily obtain that

$$\begin{cases} x(t) \in E_1^*, & \forall t \in [\tau_0,\tau_1-\gamma], \\ x(t) \in E_k^*, & \forall t \in [\tau_{k-1}+\gamma,\tau_k-\gamma], \quad \forall k \in \{2,\dots,N-1\}, \\ x(t) \in E_N^*, & \forall t \in [\tau_{N-1}+\gamma,\tau_N]. \end{cases}$$

Therefore, to conclude the proof, it only remains to prove that $x(t) \in E_k^*$ for all $t \in [\tau_k-\gamma,\tau_k)$ and $x(t) \in E_{k+1}^*$ for all $t \in (\tau_k,\tau_k+\gamma]$, for all $k \in \{1,\dots,N-1\}$. This is the objective of the following last step.

Step 3. Let us start with two observations. First, since $\|\mathbb{T}-\mathbb{T}^*\|_{\mathbb{R}^{N+1}} \leq \eta \leq \frac{\theta}{3}$, it holds that $|\tau_k - \tau_{k-1}| \geq \frac{\theta}{3}$ for all $k \in \{1,\dots,N\}$. Second, since $\gamma := \frac{\theta}{3}\min\{\varepsilon,\chi,\frac{\alpha}{\overline{\theta}}\}$, one can get that

$$\frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}} \in [1-\varepsilon,1], \quad \frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}} \in [1-\chi,1], \quad \tau_{k-1}^* + (\tau_k^*-\tau_{k-1}^*)\frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}} \in [\tau_k^*-\alpha,\tau_k^*],$$

for all $t \in [\tau_k-\gamma,\tau_k]$ and all $k \in \{1,\dots,N-1\}$. We deduce the following results:

(i) Since $v_k(s) = v_k^*(s)$ for almost every $s \in [1-\varepsilon,1]$, one can easily obtain from (A.2) and (3.1) that $u(t) = u^*(\tau_{k-1}^* + (\tau_k^*-\tau_{k-1}^*)\frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}})$, with $\tau_{k-1}^* + (\tau_k^*-\tau_{k-1}^*)\frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}} \in [\tau_k^*-\alpha,\tau_k^*]$, for almost every $t \in [\tau_k-\gamma,\tau_k)$ and all $k \in \{1,\dots,N-1\}$.

(ii) Since $\|y_k - y_k^*\|_{\mathrm{C}} \leq \|y-y^*\|_{\mathrm{C}} \leq \eta \leq \frac{\nu}{2}$, one can easily obtain from (A.2), from the equality $x^*(\tau_k^*) = y_k^*(1)$ and from ($\mathfrak{P}_3$) that $x(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x^*(\tau_k^*),\nu)$ for all $t \in [\tau_k-\gamma,\tau_k]$ and all $k \in \{1,\dots,N-1\}$.

(iii) We obtain from (A.3), from the previous two items and from ($\mathfrak{P}_2$) that the derivative of $F_k^* \circ x$ satisfies

$$\left\langle \nabla F_k^*(x(t)), f_k^*\left(x(t), u^*\left(\tau_{k-1}^* + (\tau_k^*-\tau_{k-1}^*)\frac{t-\tau_{k-1}}{\tau_k-\tau_{k-1}}\right)\right)\right\rangle_{\mathbb{R}^n} > 0,$$

for almost every $t \in [\tau_k-\gamma,\tau_k)$ and all $k \in \{1,\dots,N-1\}$. From admissibility of the triplet $(y,v,\mathbb{T})$ and (A.2), we know that $F_k^*(x(\tau_k)) = F_k^*(y_k(1)) = 0$ for all $k \in \{1,\dots,N-1\}$. As a consequence we obtain that $F_k^*(x(t)) < 0$ for all $t \in [\tau_k-\gamma,\tau_k)$ which implies from Definition 3.2, since $x(t) \in \overline{\mathrm{B}}_{\mathbb{R}^n}(x^*(\tau_k^*),\nu)$ and $\nu \leq \nu_k^*$, that $x(t) \in E_k^*$ for all $t \in [\tau_k-\gamma,\tau_k)$ and all $k \in \{1,\dots,N-1\}$.

Following the same strategy one can obtain that $x(t) \in E^*_{k+1}$ for all $t \in (\tau_k, \tau_k + \gamma]$ and all $k \in \{1, \dots, N-1\}$. The proof of Proposition 3.1 is complete.

Appendix B. Proof of Theorem 3.1. Consider the framework of Theorem 3.1. From Proposition 3.1, the corresponding triplet $(y^*, v^*, \mathbb{T}^*)$ constructed in Section 3.2 is a $\mathrm{L}^1_\square$–local solution to Problem (CP*). Before applying Theorem 2.1, we need to verify that g^* is submersive at $(y^*(0), y^*(1), \mathbb{T}^*)$. From the definition of the function g^* (see Section 3.2), note that the matrix $\nabla g^*(y^*(0), y^*(1), \mathbb{T}^*) \in \mathbb{R}^{(nN+nN+(N+1))\times \ell^*}$ is

$$\left[\begin{array}{c|c|c|c}
\nabla_1 g(y_1^*(0), y_N^*(1)) & 0_{\mathbb{R}^{n\times n(N-1)}} & 0_{\mathbb{R}^{n\times(N-1)}} & 0_{\mathbb{R}^{n\times(N+1)}} \\ \hline
0_{\mathbb{R}^{n(N-1)\times \ell}} & \mathrm{Id}_{\mathbb{R}^{n(N-1)\times n(N-1)}} & 0_{\mathbb{R}^{n(N-1)\times(N-1)}} & 0_{\mathbb{R}^{n(N-1)\times(N+1)}} \\ \hline
0_{\mathbb{R}^{n(N-1)\times \ell}} & -\mathrm{Id}_{\mathbb{R}^{n(N-1)\times n(N-1)}} & \begin{matrix} \nabla F_1^*(y_1^*(1)) & & \\ & \ddots & \\ & & \nabla F_{N-1}^*(y_{N-1}^*(1)) \end{matrix} & 0_{\mathbb{R}^{n(N-1)\times(N+1)}} \\ \hline
\nabla_2 g(y_1^*(0), y_N^*(1)) & 0_{\mathbb{R}^{n\times n(N-1)}} & 0_{\mathbb{R}^{n\times(N-1)}} & 0_{\mathbb{R}^{n\times(N+1)}} \\ \hline
0_{\mathbb{R}^{(N+1)\times \ell}} & 0_{\mathbb{R}^{(N+1)\times n(N-1)}} & 0_{\mathbb{R}^{(N+1)\times(N-1)}} & \mathrm{Id}_{\mathbb{R}^{(N+1)\times(N+1)}}
\end{array}\right]$$

From Definition 3.2, it holds that $\nabla F^*_k(y^*_k(1)) = \nabla F^*_k(x^*(\tau^*_k)) \neq 0_{\mathbb{R}^n}$ for all $k \in \{1, \dots, N-1\}$. Since g is submersive at $(x^*(0), x^*(T)) = (y^*_1(0), y^*_N(1))$, one can easily conclude that g^* is submersive at $(y^*(0), y^*(1), \mathbb{T}^*)$.

B.1. Application of Theorem 2.1. Let us introduce the Hamiltonian $\mathcal{H} : \mathbb{R}^{nN} \times \mathbb{R}^{mN} \times \mathbb{R}^{N+1} \times \mathbb{R}^{nN} \to \mathbb{R}$ associated with Problem (CP*) given by

$$\mathcal{H}(y, v, \mathbb{T}, q) := \langle q, f^*(y, v, \mathbb{T})\rangle_{\mathbb{R}^{nN}} = \sum_{k=1}^{N} \langle q_k, (\tau_k - \tau_{k-1}) f^*_k(y_k, v_k)\rangle_{\mathbb{R}^n},$$

for all $y = (y_1, \dots, y_N) \in \mathbb{R}^{nN}$, $v = (v_1, \dots, v_N) \in \mathbb{R}^{mN}$, $\mathbb{T} = \{\tau_0, \dots, \tau_N\} \in \mathbb{R}^{N+1}$ and $q = (q_1, \dots, q_N) \in \mathbb{R}^{nN}$. From Theorem 2.1, there exists a nontrivial pair $(q, q^0) \in \mathrm{AC}([0,1], \mathbb{R}^{nN}) \times \mathbb{R}_+$ satisfying:

(i) the Hamiltonian system $\dot{y}^*(s) = \nabla_q \mathcal{H}(y^*(s), v^*(s), \mathbb{T}^*, q(s))$ and $-\dot{q}(s) = \nabla_y \mathcal{H}(y^*(s), v^*(s), \mathbb{T}^*, q(s))$ for almost every $s \in [0,1]$;

(ii) the endpoint transversality condition

$$\begin{pmatrix} q(0) \\ -q(1) \\ \displaystyle\int_0^1 \nabla_{\mathbb{T}} \mathcal{H}(y^*(s), v^*(s), \mathbb{T}^*, q(s))\, ds \end{pmatrix} = q^0 \nabla \phi^*(y^*(0), y^*(1), \mathbb{T}^*) + \nabla g^*(y^*(0), y^*(1), \mathbb{T}^*)\tilde{\xi},$$

for some $\tilde{\xi} \in \mathrm{N}_{\mathrm{S}^*}[g^*(y^*(0), y^*(1), \mathbb{T}^*)]$;

(iii) the Hamiltonian maximization condition $v^*(s) \in \arg\max_{\widetilde{\omega} \in \mathrm{U}^N} \mathcal{H}(y^*(s), \widetilde{\omega}, \mathbb{T}^*, q(s))$ for almost every $s \in [0,1]$.

B.2. Introduction of the nontrivial pair (p, p^0). Since the pair (q, q^0) is not trivial, it is clear that the pair $(p, p^0) \in \mathrm{PAC}_{\mathbb{T}^*}([0,T], \mathbb{R}^n) \times \mathbb{R}_+$ defined by $p^0 := q^0$

and

$$p(t) := \begin{cases} q_1\left(\frac{t-\tau_0^*}{\tau_1^*-\tau_0^*}\right), & \forall t \in [\tau_0^*, \tau_1^*), \\ q_k\left(\frac{t-\tau_{k-1}^*}{\tau_k^*-\tau_{k-1}^*}\right), & \forall t \in (\tau_{k-1}^*, \tau_k^*), \quad \forall k \in \{2,\dots,N-1\}, \\ q_N\left(\frac{t-\tau_{N-1}^*}{\tau_N^*-\tau_{N-1}^*}\right), & \forall t \in (\tau_{N-1}^*, \tau_N^*], \end{cases}$$

is not trivial.

B.3. Hamiltonian system and Hamiltonian maximization condition of Theorem 3.1. From the above Items (i) and (iii) and from (3.2), the Hamiltonian system and the Hamiltonian maximization condition of Theorem 3.1 are satisfied.

B.4. Endpoint transversality condition of Theorem 3.1. From the definitions of g^* and S^* (see Section 3.2) and since $\tilde{\xi} \in \mathrm{N}_{\mathrm{S}^*}[g^*(y^*(0), y^*(1), \mathbb{T}^*)]$, we can write $\tilde{\xi} := (\xi, \xi^2, \xi^3, \xi^4) \in \mathbb{R}^\ell \times \mathbb{R}^{n(N-1)} \times \mathbb{R}^{N-1} \times \mathbb{R}^{N+1}$ with

$$\xi \in \mathrm{N}_{\mathrm{S}}[g(y_1^*(0), y_N^*(1))] \quad \text{and} \quad \xi^4 \in \mathrm{N}_\Delta[\mathbb{T}^*].$$

Since $(y_1^*(0), y_N^*(1)) = (x^*(0), x^*(T))$, note that $\xi \in \mathrm{N}_{\mathrm{S}}[g(x^*(0), x^*(T))]$. Furthermore, from the first two components of the above Item (ii), from the expression of $\nabla g^*(y^*(0), y^*(1), \mathbb{T}^*)$ given at the beginning of Appendix B and from the expression of $\nabla \phi^*(y^*(0), y^*(1), \mathbb{T}^*)$ (see Section 3.2 for the definition of ϕ^*), we obtain that

$$\begin{aligned} p(0) = q_1(0) &= q^0 \nabla_1 \phi(y_1^*(0), y_N^*(1)) + \nabla_1 g(y_1^*(0), y_N^*(1))\xi \\ &= p^0 \nabla_1 \phi(x^*(0), x^*(T)) + \nabla_1 g(x^*(0), x^*(T))\xi, \end{aligned}$$

and

$$\begin{aligned} -p(T) = -q_N(1) &= q^0 \nabla_2 \phi(y_1^*(0), y_N^*(1)) + \nabla_2 g(y_1^*(0), y_N^*(1))\xi \\ &= p^0 \nabla_2 \phi(x^*(0), x^*(T)) + \nabla_2 g(x^*(0), x^*(T))\xi. \end{aligned}$$

Therefore the endpoint transversality condition of Theorem 3.1 is proved.

B.5. Discontinuity condition of Theorem 3.1. From the first two components of the above Item (ii), from the expression of $\nabla g^*(y^*(0), y^*(1), \mathbb{T}^*)$ given at the beginning of Appendix B and from the expression of $\nabla \phi^*(y^*(0), y^*(1), \mathbb{T}^*)$ (see Section 3.2 for the definition of ϕ^*), we obtain that

$$\forall k \in \{2,\dots,N\},\ q_k(0) = \xi_{k-1}^2 \quad \text{and} \quad \forall k \in \{1,\dots,N-1\},\ -q_k(1) = -\xi_k^2 + \xi_k^3 \nabla F_k^*(y_k^*(1)).$$

We deduce that

$$p^+(\tau_k^*) - p^-(\tau_k^*) = q_{k+1}(0) - q_k(1) = \xi_k^3 \nabla F_k^*(y_k^*(1)) = \xi_k^3 \nabla F_k^*(x^*(\tau_k^*)),$$

for all $k \in \{1,\dots,N-1\}$. Therefore the discontinuity condition of Theorem 3.1 is satisfied with $\sigma_k := \xi_k^3$ for all $k \in \{1,\dots,N-1\}$.

B.6. Hamiltonian constancy condition of Theorem 3.1. From the Hamiltonian system and the maximization condition and applying [21, Theorem 2.6.1] on each interval $[\tau_{k-1}^*, \tau_k^*]$, we obtain that, for all $k \in \{1,\dots,N\}$, there exists a constant $c_k \in \mathbb{R}$ such that $\langle p(t), f_k^*(x^*(t), u^*(t)) \rangle_{\mathbb{R}^n} = c_k$ for almost every $t \in [\tau_{k-1}^*, \tau_k^*]$.

Furthermore, from the definition of Δ (see Section 3.2) and since $0 = \tau_0^* < \tau_1^* < \ldots < \tau_{N-1}^* < \tau_N^* = T$, we deduce from $\xi^4 \in \mathrm{N}_\Delta[\mathbb{T}^*]$ that all components of ξ^4 are zero, except possibly the first and last components. Thus, from the third component of the above Item (ii), from the expression of $\nabla g^*(y^*(0), y^*(1), \mathbb{T}^*)$ given at the beginning of Appendix B and from the expression of $\nabla \phi^*(y^*(0), y^*(1), \mathbb{T}^*)$ (see Section 3.2 for the definition of ϕ^*), we obtain that

$$\int_0^1 \langle q_{k+1}(s), f_{k+1}^*(y_{k+1}^*(s), v_{k+1}^*(s))\rangle_{\mathbb{R}^n} \, ds = \int_0^1 \langle q_k(s), f_k^*(y_k^*(s), v_k^*(s))\rangle_{\mathbb{R}^n} \, ds,$$

for all $k \in \{1, \ldots, N-1\}$. From affine changes of time variable, we obtain that

$$\frac{1}{\tau_{k+1}^* - \tau_k^*} \int_{\tau_k^*}^{\tau_{k+1}^*} \langle p(t), f_{k+1}^*(x^*(t), u^*(t))\rangle_{\mathbb{R}^n} \, dt = \frac{1}{\tau_k^* - \tau_{k-1}^*} \int_{\tau_{k-1}^*}^{\tau_k^*} \langle p(t), f_k^*(x^*(t), u^*(t))\rangle_{\mathbb{R}^n} \, dt,$$

for all $k \in \{1, \ldots, N-1\}$. From constancy of the above two integrands, we deduce that $c_{k+1} = c_k$ for all $k \in \{1, \ldots, N-1\}$. Therefore the Hamiltonian constancy condition is satisfied and the proof of Theorem 3.1 is complete.

Appendix C. Nonadmissibility of needle-like perturbations. Here we prove that needle-like perturbations of the control are not admissible (in a sense to precise) in our setting of spatially heterogeneous dynamics. This is a major difference with respect to the classical optimal control theory. Consider the one-dimensional case $n = 1$, the state space partition $\mathbb{R} = \overline{X_1} \cup \overline{X_2}$, where $X_1 := (-\infty, 1)$ and $X_2 := (1, +\infty)$, and the hybrid control system given by

$$(\mathrm{HS}_{\mathrm{ex}}) \qquad \dot{x}(t) = h(x(t), u(t)), \quad \text{for a.e. } t \in [0, 2],$$

where the spatially heterogeneous dynamics $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is defined by $h(x, u) := u$ if $x \in X_1$ and by $h(x, u) := -u$ if $x \in X_2$. Now consider the trajectory x given by $x(t) = t$ for all $t \in [0, 2]$ and the corresponding control u given by $u(t) = 1$ over $[0, 1]$ and $u(t) = -1$ over $(1, 2]$. Note that all conditions from Definitions 3.1 and 3.2 are satisfied, with $\tau_1 = 1$ as unique crossing time. For any small $\alpha > 0$, denote by x_α the solution to the hybrid control system $(\mathrm{HS}_{\mathrm{ex}})$ associated with the initial condition $x_\alpha(0) = x(0) = 0$ and the needle-like perturbation $u_\alpha : [0, 2] \to \mathbb{R}$ of u defined by $u_\alpha(t) = -1$ over $(\frac{1}{2} - \alpha, \frac{1}{2}]$ and by $u_\alpha(t) = u(t)$ elsewhere. Then the perturbed trajectory x_α satisfies $x_\alpha(t) \in X_1$ over the whole interval $[0, 2]$ and thus x_α does not uniformly converge to x over $[0, 2]$ when $\alpha \to 0$.

REFERENCES

[1] 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.

[2] T. Bayen, A. Bouali, and L. Bourdin, *Hybrid maximum principle with regionally switching parameter*, submitted, hal-03638701, (2022).

[3] T. Bayen, A. Bouali, and L. Bourdin, *Optimal control problems with non-control regions: necessary optimality conditions*, IFAC-PapersOnLine, 55 (2022), pp. 68–73.

[4] T. Bayen, A. Bouali, and L. Bourdin, *Minimum time problem for the double integrator with a loss control region*, submitted, hal-03928967v2, (2023).

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

[6] 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).

[7] P. Bettiol and L. Bourdin, *Pontryagin maximum principle for state constrained optimal sampled-data control problems on time scales*, ESAIM Control Optim. Calc. Var., 27 (2021).
[8] V. Boltyanski, *The maximum principle for variable structure systems*, Internat. J. Control, 77 (2004), pp. 1445–1451.
[9] J. F. Bonnans, *Course on optimal control*, OROC Ensta Paris-Tech and optimization master, U. Paris-Saclay, (2017).
[10] J. F. Bonnans and C. S. F. de la Vega, *Optimal control of state constrained integral equations*, Set-Valued Var. Anal., 18 (2010), pp. 307–326.
[11] J. F. Bonnans, X. Dupuis, and L. Pfeiffer, *Second-order necessary conditions in Pontryagin form for optimal control problems*, SIAM J. Control Optim., 52 (2014), pp. 3887–3916.
[12] L. Bourdin, *Note on Pontryagin maximum principle with running state constraints and smooth dynamics–proof based on the Ekeland variational principle*, Research notes, arXiv:1604.04051, (2016).
[13] L. Bourdin and E. Trélat, *Pontryagin maximum principle for finite dimensional nonlinear optimal control problems on time scales*, SIAM Journal on Control and Optimization, 51 (2013), pp. 3781–3813.
[14] A. Bressan and Y. Hong, *Optimal control problems on stratified domains*, Netw. Heterog. Media, 2 (2007), pp. 313–331.
[15] A. Bressan and B. Piccoli, *Introduction to the mathematical theory of control*, vol. 2 of AIMS Ser. Appl. Math., Springfield, MO, 2007.
[16] L. Cesari, *Lagrange and Bolza Problems of optimal control and other problems*, Springer, New York, 1983, p. 196–205.
[17] F. Clarke, *Functional analysis, calculus of variations and optimal control*, vol. 264 of Grad. Texts in Math., Springer, London, 2013.
[18] F. H. Clarke and R. B. Vinter, *Optimal multiprocesses*, SIAM J. Control Optim., 27 (1989), pp. 1072–1091.
[19] A. Dmitruk and A. Kaganovich, *The hybrid maximum principle is a consequence of Pontryagin maximum principle*, Systems Control Lett., 57 (2008), pp. 964–970.
[20] I. Ekeland, *Nonconvex minimization problems*, Bull. Amer. Math. Soc. (N.S.), 1 (1979), pp. 443–474.
[21] H. O. Fattorini, *Infinite dimensional optimization and control theory*, Cambridge University Press, 1999.
[22] M. Garavello and B. Piccoli, *Hybrid necessary principle*, SIAM J. Control Optim., 43 (2005), pp. 1867–1887.
[23] 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.
[24] X. J. Li and J. M. Yong, *Optimal control theory for infinite-dimensional systems*, Systems Control Found. Appl., Birkhäuser Boston, Inc., Boston, MA, 1995.
[25] A. A. Milyutin and N. P. Osmolovskii, *Calculus of variations and optimal control*, vol. 180 of Transl. Math. Monogr., 1998.
[26] A. Pakniyat and P. E. Caines, *On the hybrid minimum principle: The hamiltonian and adjoint boundary conditions*, IEEE Trans. Automat. Control, 66 (2020), pp. 1246–1253.
[27] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, A Pergamon Press Book. The Macmillan Co., New York, 1964.
[28] P. Riedinger, C. Lung, and F. Kratz, *An optimal control approach for hybrid systems*, Eur. J. Control, 9 (2003), pp. 449–458.
[29] M. S. Shaikh and P. E. Caines, *On the optimal control of hybrid systems: optimization of trajectories, switching times, and location schedules*, Springer Berlin Heidelberg, 2003.
[30] M. S. Shaikh and P. E. Caines, *On the hybrid optimal control problem: theory and algorithms*, IEEE Trans. Automat. Control, 52 (2007), pp. 1587–1603.
[31] H. J. Sussmann, *A maximum principle for hybrid optimal control problems*, in Proceedings of the 38th IEEE Conf. Decis. Control, vol. 1, 1999, pp. 425–430.
[32] A. J. Van Der Schaft and J. M. Schumacher, *An introduction to hybrid dynamical systems*, vol. 251, Springer London, 2000.
[33] R. Vinter, *Optimal control*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 2000.
[34] M. I. Zelikin and V. F. Borisov, *Theory of chattering control*, Systems & Control: Foundations & Applications. With applications to astronautics, robotics, economics, and engineering, Birkhäuser Boston, Inc., Boston, MA, 1994.