Infinite Horizon Problems on Stratifiable State Constraints sets

Cristopher Hermosilla, Hasnaa Zidani

Infinite Horizon Problems on Stratifiable State Constraints sets*

C. Hermosilla† H. Zidani‡

Abstract

This paper deals with a state-constrained control problem. It is well known that, unless some compatibility condition between constraints and dynamics holds, the value function has not enough regularity, or can fail to be the unique constrained viscosity solution of a Hamilton-Jacobi-Bellman (HJB) equation. Here, we consider the case of a set of constraints having a stratified structure. Under this circumstance, the interior of this set may be empty or disconnected, and the admissible trajectories may have the only option to stay on the boundary without possible approximation in the interior of the constraints. In such situations, the classical pointing qualification hypothesis are not relevant. The discontinuous value function is then characterized by means of a system of HJB equations on each stratum that composes the state constraints. This result is obtained under a local controllability assumption which is required only on the strata where some chattering phenomena could occur.

Keywords. State constrained, infinite horizon problem, stratified systems, HJB equations, optimal control.

1 introduction.

We are concerned with the optimal control problem of infinite horizon for trajectories lying in a closed set $\mathcal{K} \subseteq \mathbb{R}^N$. The main issue is to characterize the value function of this problem as the unique solution to a HJB equation. More precisely, given a dynamic $f(\cdot,\cdot)$, a nonempty compact set $\mathcal{A} \subseteq \mathbb{R}^m$, a Borel measurable function $u : [0,\infty) \to \mathcal{A}$ and a point $x \in \mathcal{K}$ we consider trajectories $y_{x,u}(\cdot)$ solutions to the differential equation

$$(D_{u,x}) \qquad \begin{cases} \dot{y}(t) = f(y(t), u(t)) & \text{a.e. } t > 0 \\ y(0) = x, \end{cases}$$

which are feasible on the set $\mathcal{K}$, that is,

$$y_{x,u}(t) \in \mathcal{K}, \quad \forall t \geq 0. \tag{1}$$

We denote the set of admissible control as

$$\mathbb{A}(x) = \{u : [0,+\infty) \to \mathcal{A} \mid y_{x,u}(\cdot) \text{ satisfies } (1)\}. \tag{2}$$

*This work was supported by the European Union under the 7th Framework Programme **FP7-PEOPLE-2010-ITN** Grant agreement number 264735-SADCO

†Commands team at INRIA Saclay & ENSTA ParisTech, 828, Boulevard des Maréchaux, 91762 Palaiseau Cedex. Email: cristopher.hermosilla@inria.fr

‡Mathematics Department (UMA), ENSTA ParisTech & Commands team at INRIA Saclay, 828, Boulevard des Maréchaux, 91762 Palaiseau Cedex. Email: hasnaa.zidani@ensta-paristech.fr

Then, for some discount factor $\lambda > 0$ and some running cost function $\ell(\cdot,\cdot)$, the value function related to this problem is

$$v(x) := \inf\left\{ \int_0^\infty e^{-\lambda t}\ell(y_{x,u}(t),u(t))dt \;\middle|\; u \in \mathbb{A}(x)\right\}. \tag{3}$$

In the case when $\mathcal{K} = \mathbb{R}^N$ and under standard hypothesis on the data, it is well known that $v(\cdot)$ is a uniformly continuous function which can be characterized as the unique viscosity solution to a HJB equation in that class of functions; see for instance [6].

When the control problem is in the presence of state constraints ($\mathcal{K} \neq \mathbb{R}^N$) a state-space constrained HJB equation can be associated with the value function as done in [31, 32]. In our setting, the HJB equation takes the form

$$\lambda v + H(x, \nabla v) = 0 \quad \text{in } \mathcal{K}, \tag{4}$$

where $H(x,p) := \max\{-\langle f(x,u),p\rangle - \ell(x,u) \mid u \in \mathcal{A}\}$. It is straightforward to check that the value function satisfies (4) in the constrained viscosity sense, that means that v is a subsolution in $\text{int}(\mathcal{K})$ and a supersolution on $\mathcal{K}$. However, it is complicated to prove the uniqueness of the solution to (4). The main difficulty comes from the fact that the state-space HJB equation may admit several solutions (in the constrained viscosity sense) if the behavior of the solution on the boundary is not taken into account; see for instance the discussion in [12, 22].

One possible way to overcome the problem exposed above is to consider some compatibility assumptions between the dynamics and the state-constraints. The most classical of these is called the *inward pointing condition* (IPC). It was first introduced by Soner in [31, 32] for the case that $\text{int}(\mathcal{K}) \neq \emptyset$ with smooth boundary and it has been object of subsequence generalization to more general cases; see [22, 33, 16, 28, 19] among others. This condition basically says that at each point of the boundary of $\mathcal{K}$ there exists a vector field of the system pointing into $\mathcal{K}$. Under this assumption the value function is Lipschitz continuous and then uniqueness can be established. From the point of view of the dynamical system, the IPC ensures the existence of the so-called *neighboring feasible trajectories* which make possible to approximate any trajectory hitting the boundary by a sequence of arcs which remain in the interior of $\mathcal{K}$; see for instance [21, 9, 10]. We refer to [13, 26, 27] for weaker inward pointing assumptions, and to [24, 25] for more properties and the numerical approximation of continuous constrained viscosity solutions.

Another compatibility assumption of similar nature, called the *outward pointing condition* (OPC), has been considered by [11] in the context of exit time problems. This assumption states that each point of the boundary of $\mathcal{K}$ can be reached by a trajectory coming from the interior of the set and it implies a certain monotonicity of the solution to the HJB equation which allows to treat the case when the value function is discontinuous. Under this assumption, it is possible to characterize the value function as the unique lower semi-continuous solution of a HJB equation; see for instance [21, 20, 19].

Let us stress that all the works mentioned above assume the same compatibility assumption on the whole boundary, that is, no mix type of pointing condition has been so far considered and, with exception of [20], they all require that $\mathcal{K}$ has non-empty interior.

However, there are many control problems in which these compatibility assumptions are never satisfied. For example, if we consider a mechanical system governed by a second order equation:

$$\ddot{y} = \varphi(y,\dot{y},u) \qquad y \in [a,b], \qquad \dot{y} \in [c,d] \tag{5}$$

using the transformation $y_1 = y$ and $y_2 = \dot{y}$ we can rewrite system (5) as:

$$\begin{pmatrix}\dot{y}_1\\ \dot{y}_2\end{pmatrix} = \begin{pmatrix} y_2 \\ \varphi(y_1,y_2,u)\end{pmatrix} \qquad (y_1,y_2) \in \mathcal{K}_0 = [a,b]\times[c,d]. \tag{6}$$

It is clear that in this case, the direction of the vector field on $\{a\} \times (c, d)$ and on $\{b\} \times (c, d)$ does not depend on the control nor in the initial dynamic φ but only on the sign of y_2, and so, for some values of y_2 the vector field will point into $\mathcal{K}_0$ and for others will point into $\mathbb{R}^2 \setminus \mathcal{K}_0$.

On the other hand, as mentioned before, the *Pointing Conditions* are sufficient hypothesis whose main target is to approximate any trajectory of the control system by a sequence of trajectories (of the same controlled dynamic) living completely in the interior of the constraints. This fact may suggest that the right requirement may be the *interior approximation of trajectories.* Nevertheless, it will automatically rule out cases where junctions are present, for example

$$\mathcal{K}_1 = \{x \in \mathbb{R}^N : x_N^2 \geq x_1^2 + \ldots + x_{N-1}^2\}.$$

The previous example also illustrates the importance of considering the information passing through the boundary in cases when the interior of the state-constraint is disconnected, in that context the approach of stratification we propose here seems to suit well.

Furthermore, optimal control problems on networks as studied in [1, 2] (see also the references therein) can also be seen as state-constrained problems, where the interior of the state constraints set is always empty and where the junction plays an essential role.

In the general case where $\mathcal{K}$ is assumed to be any closed set of $\mathbb{R}^N$, and under some convexity assumptions on the dynamics, the value function is lower semi-continuous and it can be characterized as the smallest supersolution to (4); see [14] for more details. In [3], it has been shown that the epigraph of the value function v can always be described by an auxiliary optimal control problem without state constraints for which the value function is Lipschitz continuous and characterized, without any further assumptions, as the unique viscosity solution to a HJB equation. This approach leads to a constructive way for determining the epigraph of v and to its numerical approximation. It can also be extended to more general situations of time-dependent state constraints [18].

In this paper, we follow the same line of investigation as in [22, 12]. We aim at characterizing the value function by a completed system of HJB equations. The proof used here is based on nonsmooth analysis as in [33] where the notion of HJB equation is understood in the proximal sense by means of the theory of weak and strong invariance. We investigate the characterization of the value function for a class of control problems where the set of constraints enjoys a regular stratification property (i.e, $\mathcal{K}$ is a collection of strata of different dimensions; see section 2.2 for a precise definition). Moreover, the discontinuous value function is characterized by means of a system of HJB equations on each stratum of $\mathcal{K}$. This result is obtained under a local controllability assumption which is required only on the stratum where some chattering phenomena could occur.

The paper is organised as follows: Section 2 presents the setting of the control problem and the main results. Properties of the value function are given in Section 3. Some useful results on invariance principles are given in Section 4. And finally the main results are proved in Section 5.

2 Setting of the problem and main result.

2.1 Notation.

Throughout this paper, $\mathbb{R}$ denotes the sets of real numbers, $|\cdot|$ the Euclidean norm on $\mathbb{R}^N$, $\mathbb{B}$ the unit open ball $\{x \in \mathbb{R}^N : \ |x| < 1\}$ and $\mathbb{B}(x, r) = x + r\mathbb{B}$. For a set $S \subseteq \mathbb{R}^N$, $\text{int}(S), \overline{S}, \text{bdry}(S)$ and $\text{co}(S)$ denote its interior, closure, boundary and convex hull, respectively. Also for S convex we denote $\text{ri}(S)$ and $\text{rbd}(S)$ its relative interior and boundary, respectively. The indicator function is given by $\mathbb{1}_S(x) = 1$ if $x \in S$ and $\mathbb{1}_S(x) = 0$ if it is not. The distance function to $\mathcal{S}$ is $\text{dist}_{\mathcal{S}}(x) = \inf\{|x - y| : \ y \in \mathcal{S}\}$ and in the case the infimum is attained we call the set of solution the projections of x over $\mathcal{S}$ and we denote it by $\text{proj}_{\mathcal{S}}(x)$. Let S_1 and S_2 be two compact set, then the Hausdorff distance is given by

$$d_H(S_1, S_2) = \max\left\{\sup_{x \in S_2} \text{dist}_{S_1}(x), \sup_{x \in S_1} \text{dist}_{S_2}(x)\right\}.$$

We adopt the convention that $d_H(\emptyset,\emptyset)=0$ and $d_H(\emptyset,S)=+\infty$ if $S\neq\emptyset$. For a given function $\varphi:\mathbb{R}^N\to\mathbb{R}\cup\{+\infty\}$, the epigraph of this function is the set

$$\operatorname{epi}\varphi=\{(x,r)\in\mathbb{R}^N\times\mathbb{R}\mid r\geq\varphi(x)\}.$$

Let $\mathcal{M}$ be an embedded manifold of $\mathbb{R}^N$, then $\mathcal{T}_{\mathcal{M}}(x)$ stands for the tangent space to $\mathcal{M}$ at $x\in\mathcal{M}$.

2.2 Stratified Systems and Hypothesis.

Consider the optimal control problem of infinite horizon:

$$\text{(P)}\qquad \upsilon(x)=\begin{cases}\inf & \displaystyle\int_0^\infty e^{-\lambda t}\ell(y_{x,u}(t),u(t))dt\\ & y_{x,u}(\cdot)\text{ solution to }(D_{u,x}),\\ & y_{x,u}(t)\in\mathcal{K},\quad\forall t\geq 0,\\ & u(t)\in\mathcal{A}\text{ measurable.}\end{cases}\qquad \forall x\in\mathcal{K},$$

where $\mathcal{A}$ is a compact set of $\mathbb{R}^m$ and $\lambda>0$ is a given discount factor $\lambda>0$. We assume standard hypothesis on the dynamic f and the running cost ℓ. Namely, the dynamic $f:\mathbb{R}^N\times\mathcal{A}\to\mathbb{R}^N$ satisfies:

$$(H_f)\qquad\begin{cases}(i) & f(\cdot,\cdot)\text{ is continuous on }\mathbb{R}^N\times\mathcal{A}.\\ (ii) & \forall u\in\mathcal{A},\ f(\cdot,u)\text{ is locally Lipschitz continuous on }\mathcal{K}.\\ (iii) & \exists c_f>0\text{ such that }\forall x\in\mathcal{K}:\\ & \quad\max\{|f(x,u)|:\ u\in\mathcal{A}\}\leq c_f(1+|x|).\end{cases}$$

Let $x\in\mathcal{K}$ and $u:[0,+\infty)\to\mathcal{A}$ be a measurable control. By a solution to $(D_{u,x})$ we mean an absolutely continuous function $y(\cdot)$ that satisfies

$$y(t)=x+\int_0^t f(y(s),u(s))ds\quad\text{for all }t\geq 0.$$

By (H_f), the solution is uniquely determined by x and u. Furthermore, the maximal (forward and backward) solution can always be defined for positive times.

Remark 2.1. *Note that, by Gronwall Lemma and hypothesis (H_f), each solution to $(D_{u,x})$ satisfies the following bounds:*

$$\begin{aligned}&1+|y(t)|\leq(1+|x|)e^{c_f t}\qquad t\geq 0;\\ &|y(t)-x|\leq(1+|x|)(e^{c_f t}-1)\qquad t\geq 0;\\ &|\dot{y}(t)|\leq c_f(1+|x|)e^{c_f t}\qquad a.e.\ t>0;\end{aligned}$$

The running cost $\ell:\mathbb{R}^N\times\mathcal{A}\to\mathbb{R}$ is also continuous and satisfies:

$$(H_\ell)\qquad\begin{cases}(i) & \ell(\cdot,\cdot)\text{ is continuous on }\mathbb{R}^N\times\mathcal{A}.\\ (ii) & \forall u\in\mathcal{A},\ \ell(\cdot,u)\text{ is locally Lipschitz continuous on }\mathcal{K}.\\ (iii) & \exists c_\ell>0,\ \lambda_\ell\geq 1\text{ such that }\forall(x,u)\in\mathcal{K}\times\mathcal{A}:\\ & \quad 0\leq\ell(x,u)\leq c_\ell(1+|x|^{\lambda_\ell}).\end{cases}$$

When dealing with integral cost, it is usual to introduce an augmented dynamics. For this end, we define

$$\beta(\tau,x,u):=e^{-\lambda\tau}[c_\ell(1+|x|^{\lambda_\ell})-\ell(x,u)]\qquad\forall(\tau,x,u)\in\mathbb{R}\times\mathcal{K}\times\mathcal{A}.$$

Then, we consider the augmented dynamic $G : \mathbb{R} \times \mathcal{K} \rightrightarrows \mathbb{R}^N \times \mathbb{R}$ defined by

$$G(\tau, x) = \left\{ \begin{pmatrix} f(x,u) \\ e^{-\lambda\tau}\ell(x,u) + r \end{pmatrix} \;\middle|\; \begin{array}{c} u \in \mathcal{A}, \\ 0 \le r \le \beta(\tau, x, u) \end{array} \right\}, \qquad \forall (\tau, x) \in \mathbb{R} \times \mathcal{K}.$$

It is not difficult to see that by (H_ℓ) this set-valued map has compact and nonempty images. Also, in all our analysis we will always suppose

$$(H_0) \qquad\qquad G(\tau, x) \text{ is convex for any } (\tau, x) \in \mathbb{R} \times \mathcal{K}.$$

The class of control problems we are considering in this paper do not necessarily satisfy any qualification hypothesis. Here, we require two principal assumptions. The first one is that the state-constraints set admits a sufficiently regular partition into smooth manifolds or *strata*. More precisely,

$$(H_1) \qquad\qquad \mathcal{K} \text{ is a closed and stratifiable subset of } \mathbb{R}^N,$$

that is, there exists a locally finite collection $\{\mathcal{M}_i\}_{i\in\mathcal{I}}$ of embedded manifolds of $\mathbb{R}^N$ such that:

- $\mathcal{K} = \bigcup_{i\in\mathcal{I}} \mathcal{M}_i$ and $\mathcal{M}_i \cap \mathcal{M}_j = \emptyset$ when $i \neq j$.
- If $\mathcal{M}_i \cap \overline{\mathcal{M}_j} \neq \emptyset$, necessarily $\mathcal{M}_i \subseteq \overline{\mathcal{M}}_j$.

Remark 2.2. *A network set as defined in [1, 2] can be considered as a particular case of stratified sets. Indeed, in this case the stratification is very simple and consists in edges and junctions.*

Remark 2.3. *An important class of sets that admit a stratification as described above is the class of polytopes in $\mathbb{R}^N$. In fact, these sets can be decomposed into a finite number of open convex polytopes of the form:*

$$P = \left\{ x \in \mathbb{R}^N \;\middle|\; \begin{array}{ll} \langle \eta_k, x\rangle = \alpha_k, & k = 1, \dots, n, \\ \langle \eta_k, x\rangle < \alpha_k, & k = n+1, \dots, m \end{array} \right\}$$

where $\eta_1, \dots, \eta_m \in \mathbb{R}^N$.

Furthermore, the class of sets that admit a stratification is quite broad, it includes sub-analytic and semi-algebraic sets and also definable sets of an o-minimal structure; see for instance [34, 23].

The second hypothesis is related to the dynamics obtained as the intersection between the original dynamic and the tangent space to each stratum $\mathcal{M}_i$. For each index $i \in \mathcal{I}$, let us consider the compact-valued multifunction $\mathcal{A}_i : \mathcal{M}_i \rightrightarrows \mathcal{A}$ given by

$$\mathcal{A}_i(x) := \{u \in \mathcal{A} \;\mid\; f(x,u) \in \mathcal{T}_{\mathcal{M}_i}(x)\}. \tag{7}$$

Thus, we assume that the following also holds

$$(H_2) \qquad\qquad \text{Each } \mathcal{A}_i \text{ is locally Lipschitz on } \mathcal{M}_i \text{ w.r.t. the Hausdorff metric.}$$

Using an appropriated convention for the Hausdorff distance, this hypothesis allows to consider the case when no tangential dynamic is defined, that is, when $\mathcal{A}_i$ has empty image. In that situation, (H_2) states that $\mathcal{A}_i$ has empty images all along $\mathcal{M}_i$.

Example 2.1. *Consider the following dynamic:*

$$\begin{pmatrix} \dot y_1 \\ \dot y_2 \end{pmatrix} = \begin{pmatrix} y_2 \\ u \end{pmatrix}, \qquad\qquad u \in \mathcal{A} := [-1, 1], \quad y_1(t), y_2(t) \in [-r, r],$$

where the stratification of the state-constraints set is given by Figure 1. Here, the strata $\mathcal{M}_0$ represents the interior of the square, and the rest of the stratification consists of edges and nodes $\mathcal{M}_i$, for $i = 1, \cdots, 12$. It is straightforward to check that (H_2) holds because $\mathcal{A}_0 = \mathcal{A}$, $\mathcal{A}_i = \{0\}$ for $i = 1, \dots, 4$ and $\mathcal{A}_j = \emptyset$ for $i = 5, \dots, 12$.

It is clear in this example that neither the IPC not the OPC condition is satisfied. In figure 1, the green zone corresponds to the viable set, that it, the set of points for which $\mathbb{A}(x) \neq \emptyset$.

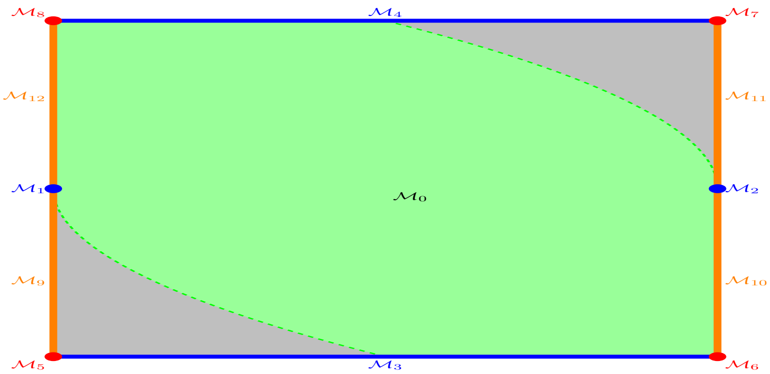

Figure 1: Stratification of Example 2.1.

Remark 2.4. *Under assumption* (H_2)*, the set-valued map* $\mathcal{A}_i$ *can be extended by continuity up to the boundary in such a way that* (H_2) *holds up to* $\overline{\mathcal{M}}_i$*. For sake of simplicity, this extension is denoted by* $\mathcal{A}_i$ *as well and so,* $\mathcal{A}_i(x)$ *for* $x \in \overline{\mathcal{M}} \setminus \mathcal{M}$ *is well defined.*

On the other hand, it is also convenient to introduce a backward augmented dynamic defined for any $(\tau, x) \in \mathbb{R} \times \overline{\mathcal{M}}_i$ as follows:

$$G_i(\tau, x) = \left\{ -\begin{pmatrix} f(x,u) \\ e^{-\lambda\tau}\ell(x,u) + r \end{pmatrix} \;\middle|\; \begin{matrix} u \in \mathcal{A}_i(x), \\ 0 \le r \le \beta(\tau, x, u) \end{matrix} \right\}. \tag{8}$$

It is not difficult to see that thanks to (H_0) and the definition of $\mathcal{A}_i(\cdot)$, the augmented multifunction G_i has convex compact images. These images would be nonempty if and only if A_i has nonempty images.

Finally, for technical reasons, an extra hypothesis of controllability on certain strata will be required in order to complete the proof of the main theorem. For this purpose, we introduce the set-valued map of tangents control to closure of each stratum:

$$\mathcal{A}_i^\sharp(x) := \left\{ u \in \mathcal{A} \;\mid\; f(x,u) \in \mathcal{T}^B_{\overline{\mathcal{M}}_i}(x) \right\}, \tag{9}$$

where $\mathcal{T}^B_{\overline{\mathcal{M}}_i}(x)$ stands for the Bouligand tangent cone to $\overline{\mathcal{M}}_i$ at a point $x \in \overline{\mathcal{M}}_i$ (see Section 4.1 for more details) and $\mathcal{A}_i(\cdot)$ is the extended set-valued map defined previously in Remark 2.4. Then the controllability assumption can be stated as follows:

$$\begin{cases} \forall j \in \mathcal{I}, \text{ one of the following holds:} \\ i) \;\; \forall i \in \mathcal{I} \text{ with } \mathcal{M}_j \subseteq \overline{\mathcal{M}}_i,\ \mathcal{A}_i(x) = \mathcal{A}_i^\sharp(x)\ \forall x \in \mathcal{M}_j. \\ ii) \; \exists r_j > 0 \text{ such that } \mathcal{T}_{\mathcal{M}_j}(x) \cap \mathbb{B}(0, r_j) \subseteq f(x, \mathcal{A}_j(x))\ \forall x \in \mathcal{M}_j. \end{cases} \tag{H_3}$$

Remark 2.5. *The assumptions* (H_2) *and* (H_3) *amount saying that for any* $j \in \mathcal{I}$*. whether the set of viable or admissible trajectories is locally Lipschitz up to the boundary of any stratum close to* $\mathcal{M}_j$ *and of bigger dimension, or the controlled system is small time locally controllable on the stratum* $\mathcal{M}_j$*.*

2.3 Main result.

As stated in the introduction, the main aim of this paper is to characterize the value function of the infinite horizon problem in term of a bilateral Hamilton-Jacobi equation. This equation consists principally in two part, one corresponding to the classical notion of supersolution and another corresponding to a subsolution in a stratified sense.

Let us consider the (maximized) Hamiltonian $H : \mathcal{K} \times \mathbb{R}^N \rightrightarrows \mathbb{R}$ which is given by

$$H(x,\zeta) = \max\{-\langle \zeta, f(x,u)\rangle - \ell(x,u) \mid u \in \mathcal{A}\}, \tag{10}$$

and for each index $i \in \mathcal{I}$ we define $H_i : \mathcal{M}_i \times \mathbb{R}^N \rightrightarrows \mathbb{R}$, the restricted Hamiltonian to $\mathcal{M}_i$, by

$$H_i(x,\zeta) = \max\{-\langle \zeta, f(x,u)\rangle - \ell(x,u) \mid u \in \mathcal{A}_i(x)\}. \tag{11}$$

Definition 2.1. *Let $\psi : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be a given function. We say that ψ has σ-superlinear growth on its domain if there exists $c_\psi > 0$ such that*

$$|\psi(x)| \leq c_\psi(1+|x|)^\sigma \qquad \forall x \in \operatorname{dom}\psi.$$

Whence, the main result of the paper can be stated as follows:

Theorem 2.1. *Suppose that (H_0), (H_1), (H_2) and (H_3) hold in addition to (H_f) and (H_ℓ). Assume also that $\lambda > \lambda_\ell c_f$ (where $\lambda_\ell > 0$ and $c_f > 0$ are the constants given by (H_ℓ) and (H_f), respectively). Then the value function $\upsilon(\cdot)$ of the problem (P) is the only l.s.c. function with λ_ℓ-superlinear growth which is $+\infty$ on $\mathbb{R}^N \setminus \mathcal{K}$ and that satisfies:*

$$\lambda\upsilon(x) + H(x,\zeta) \geq 0 \qquad \forall x \in \mathcal{K},\ \forall \zeta \in \partial_P \upsilon(x), \tag{12}$$

$$\lambda\upsilon(x) + H_i(x,\zeta) \leq 0 \qquad \forall x \in \mathcal{M}_i,\ \forall \zeta \in \partial_P \upsilon_i(x),\ \forall i \in \mathcal{I}, \tag{13}$$

where $\upsilon_i(x) = \begin{cases} \upsilon(x) & \text{if } x \in \overline{\mathcal{M}}_i \\ +\infty & \text{otherwise.} \end{cases}$ *and $\partial_P\psi(\cdot)$ denotes the proximal subdiffererential of a l.s.c function ψ (see Section 4.1 for more details).*

Remark 2.6. *Recall that when* $\operatorname{int}(\mathcal{K})$ *is a nonempty set, it is a smooth manifold of $\mathbb{R}^N$ and therefore, there is no loss of generality in assuming that it is one of the stratum, say $\mathcal{M}_0$, of the stratification of $\mathcal{K}$. In that case, $H_0 = H$, and so, the constrained Hamiton-Jacobi equation proposed by Soner in [31] is included in the set of equations proposed in Theorem 2.1.*

Comments The idea of characterizing the value function by a system of HJB equations on whole the domain $\mathcal{K}$, including its boundary, appears already in the work of Ishii-Koike [22]. However, in that paper the set $\mathcal{A}(x)$ is assumed nonempty everywhere on $\mathcal{K}$, requiring in particular that the viable set is whole the set $\mathcal{K}$. Moreover, the result in [22] assume some restrictive hypothesis on $\mathcal{K}$ and on the set-valued map $x \longmapsto \mathcal{A}(x)$.

Let us also mention the work in [12] where it is shown that the HJB equation should be completed by an additional *information* on the increasing property of the solution along trajectories lying on the boundary. In the present work, we explicitly express the additional information in terms of HJB equations on each strata. The regularity assumptions on the set $\mathcal{K}$ are quite general and allow several situations that are not covered by the known literature. However, Theorem 2.1 requires a new controllability assumption that is needed only on the strata where some chattering behaviour may occur.

Let us point out that the main motivation for introducing (H_3) relies in the property obtained in Lemma 5.1. Indeed, this assumption is only a sufficient condition to ensure the conclusion of that lemma. Clearly, there are many other cases where (H_3) does not hold whereas the conclusion of lemma 5.1 still holds as for instance in Example 2.1. Therefore, it should be noticed that the result of Theorem 2.1 still holds if we replace (H_3) by any other assumption that ensures the conclusion of Lemma 5.1.

Finally, let us stress on that the notion of viscosity solution of HJB equation as stated in Theorem 2.1 uses nonsmooth analysis tools. A reformulation of the main result of this paper could be stated by using a l.s.c viscosity notion based on test functions as in [17, 8]. More precisely, Theorem 2.1 is equivalent to say that υ is the unique l.s.c. viscosity solution with λ_ℓ-superlinear growth of the HJB equation:

$$\lambda \upsilon(x) + H(x, D_x \upsilon(x)) \geq 0 \qquad \forall x \in \mathcal{K}, \tag{14a}$$

$$\lambda \upsilon(x) + H_i(x, D_x \upsilon(x)) \leq 0 \qquad \forall x \in \mathcal{M}_i,\ \forall i \in \mathcal{I}. \tag{14b}$$

Here the l.s.c viscosity solution of (14) has to be understood in the sense given in the next definition.

Definition 2.2. *Let u be a l.s.c function defined on $\mathbb{R}^N$ with Dom $u \subset \mathcal{K}$.*

The function u is said a supersolution of (14a) if for every $\varphi \in C^1(\mathcal{K})$ and for every $x_0 \in \mathcal{K}$ such that $u - \varphi$ achieved a local minimum at x_0 on $\mathcal{K}$, we have:

$$\lambda u(x_0) + H(x_0, \nabla\varphi(x_0)) \geq 0.$$

Moreover, u is said a subsolution of (14b) if for every $\varphi \in C(\mathcal{K})$, with $\varphi \in C^1(\mathcal{M}_i)$ for $i \in \mathcal{I}$, and for every $x_0 \in \mathcal{M}_i$ such that $u - \varphi$ achieved a local minimum at x_0 on $\mathcal{M}_i$, we have:

$$\lambda u(x_0) + H_i(x_0, \nabla\varphi(x_0)) \leq 0.$$

Finally, u is said a l.s.c viscosity solution of (14) if it is supersolution of (14a) and subsolution of (14b).

3 Properties of the Value function.

In this section we study the principal properties of the value function associated with the optimization problem (P). In particular, we will show where the assumptions on the dynamics, the running cost and the discount factor are used for.

Remark 3.1. *Suppose that hypothesis (H_ℓ) holds, then by Remark 2.1 the value function (3) satisfies the following inequality on its domain:*

$$|\upsilon(x)| \leq \int_0^\infty e^{-\lambda t} c_\ell (1 + (1 + |x|)^{\lambda_\ell} e^{\lambda_\ell c_f t}) dt \qquad \forall x \in \operatorname{dom} \upsilon.$$

Therefore, if $\lambda > \lambda_\ell c_f$ we obtain that the value function has λ_ℓ-superlinear growth on its domain, that is, there exists a constant $c_\upsilon > 0$ such that

$$|\upsilon(x)| \leq c_\upsilon (1 + |x|)^{\lambda_\ell} \quad \forall x \in \operatorname{dom} \upsilon.$$

3.1 Existence of optimal controls.

The next proposition is a classical type of results in optimal control and states that if the optimal value of (P) is finite, then the control problem admits an optimal control.

Proposition 3.1. *Suppose that (H_f), (H_ℓ) and (H_0) hold, and that $\lambda > \lambda_\ell c_f$. If $x \in \mathcal{K}$ and $\upsilon(x) \in \mathbb{R}$ then there exists $u \in \mathbb{A}(x)$ an optimal solution to (P).*

Proof. Let $x \in \mathcal{K}$ such that $\upsilon(x) \in \mathbb{R}$. This means that for every $n \geq 0$, there exists a control law $u_n \in \mathbb{A}(x)$ and the associated trajectory y_n solution of ($D_{u,x}$) (with the initial condition $y_n(0) = x$) such that $y_n(t)$ belongs to $\mathcal{K}$ for every $t \in [0, \infty]$ and such that:

$$\lim_{n\to+\infty} \int_0^\infty e^{-\lambda t} \ell(y_n(t), u_n(t))\, dt = \upsilon(x). \tag{15}$$

Set $z_n(t) = \ell(y_n(t), u_n(t))$, then from Remarks 2.1 and 3.1 we obtain the bounds:

$$
\begin{cases} |y_n(t)| \le (1+|x|)e^{c_f t} & t \ge 0; \\ |\dot{y}_n(t)| \le c_f(1+|x|)e^{c_f t} & \text{a.e. } t > 0; \\ |z_n(t)| \le c_\ell(1+(1+|x|)^{\lambda_\ell})e^{\lambda_\ell c_f t} & \text{a.e. } t > 0; \end{cases} \tag{16}
$$

Recall that $\lambda > \lambda_\ell c_f$ and consider the measure $d\mu = e^{-\lambda t}dt$. Let $W^{1,1}([0,+\infty); d\mu)$ be the Soboloev space of functions $z \in L^1([0,+\infty); d\mu)$ integrables for the measure $d\mu$ such that its weak derivative $\dot{z} \in L^1([0,+\infty); d\mu))$. Note that by (16) the sequences $\{y_n\}$ and $\{z_n\}$ are bounded in $W^{1,1}([0,+\infty); d\mu)$ and $L^1([0,+\infty); d\mu)$, respectively. Whence, by Arzela-Ascoli and Alaoglu Theorems, we derive that there exists a function $y \in W^{1,1}([0,+\infty); d\mu)$ and a subsequence, still denoted by y_n, such that

$$
\begin{aligned} &y_n \text{ converges uniformly to } y \text{ on compact subsets of } [0,+\infty), \\ &\dot{y}_n \text{ converges weakly to } \dot{y} \text{ in } L^1([0,+\infty); d\mu). \end{aligned}
$$

See for instance [4, Theorem 0.3.4]. Note that $y(t) \in \mathcal{K}$ for every $t \ge 0$

On the other hand, it is not difficult to see that $\{z_n\}$ is equi-integrable with respect to the measure $d\mu$, then by the Dunford-Pettis Theorem there exists a function $z \in L^1([0,+\infty); d\mu)$ and a subsequence, still denoted by z_n, such that z_n converges weakly to z in $L^1([0,+\infty); d\mu)$.

Let $G_0(x) = G(0,x)$ for every $x \in \mathcal{K}$, by (H_f), (H_ℓ) and (H_0), we have that G_0 is locally Lipschitz and has closed convex images. Then by the Convergence Theorem [4, Theorem 1.4.1], we conclude that $(\dot{y}, z) \in G_0(y)$. Thus, by the Filippov selection Theorem (see Lemma 4.1), there exists a measurable control $u : [0,+\infty) \to$ such that satisfies

$$
\begin{aligned} &\dot{y}(t) = f(y(t), u(t)) \quad \text{a.e. } t > 0, \qquad y(0) = x. \\ &z(t) = \ell(y(t), u(t)) \quad \text{a.e. } t > 0. \end{aligned}
$$

Finally, since $y(t) \in \mathcal{K}$ for every $t \ge 0$, $u \in \mathbb{A}(x)$ and since $\phi \equiv 1 \in L^\infty([0,+\infty); d\mu)$, we have

$$
\int_0^\infty e^{-\lambda t} z(t)dt = \lim_{n\to+\infty} \int_0^\infty e^{-\lambda t} z_n(t)dt = v(x).
$$

So, u is a solution to the problem. □

3.2 Lower semicontinuity

In contrast with unconstrained optimal control problems, the value function under presence of state constraints is not in general continuous. However, thanks to the compactness of trajectories of the augmented dynamic G it is l.s.c.

Proposition 3.2. *Suppose that* (H_f), (H_ℓ) *and* (H_0) *hold, and that* $\lambda > \lambda_\ell c_f$. *Then the value function* $x \mapsto v(x)$ *is l.s.c. on* $\operatorname{dom} \mathbb{A}$.

Proof. Let $x \in \operatorname{dom} \mathbb{A}$, then $v(x) \in \mathbb{R}$. Let $\{x_n\} \subseteq \mathcal{K} \cap \mathbb{B}(x,1)$ be a sequence such that $x_n \to x$, we need to prove that

$$
\liminf_{n\to+\infty} v(x_n) \ge v(x).
$$

Suppose that there exists a subsequence, denoted equally, so that $\{x_n\} \subseteq \operatorname{dom} \mathbb{A}$ otherwise the inequality holds immediately. Then, by Proposition 3.1 there exists an optimal control $u_n \in \mathbb{A}(x_n)$. Note that (16) holds with a possible larger constant. Whence, using the same technique as in Proposition 3.1 we can prove that there exists $u \in \mathbb{A}(x)$ such that

$$
\int_0^\infty e^{-\lambda t} \ell(y_{x,u}(t), u(t))(t)dt \le \liminf_{n\to+\infty} \int_0^\infty e^{-\lambda t} \ell(y_n(t), u_n(t))dt = \liminf_{n\to+\infty} v(x_n).
$$

Finally, using the definition of the value function we conclude the proof. □

3.3 Dynamic Programming Principle and consequences

It is well known that the value function satisfies the classical dynamic programming principle. In this case, it states as follows: for any $T > 0$

$$\text{(DPP)} \qquad \upsilon(x) = \inf\left\{ \int_0^T e^{-\lambda t}\ell(y_{x,u}(t), u(t))dt + e^{-\lambda T}\upsilon(y_{x,u}(T)) : \ u \in \mathbb{A}(x) \right\}.$$

This principle include two different increasing properties along admissible trajectories of the control system. Indeed, the two elementary inequalities that define (DPP) can be interpreted as a *weakly decreasing* and a *strongly decreasing* principle, respectively. More precisely,

Definition 3.1. *Let $\varphi : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function, we say that φ is:*

i) weakly decreasing for the control system if for all $x \in \operatorname{dom}\varphi$, there exists a control $u \in \mathbb{A}(x)$ such that

$$\text{(17)} \qquad e^{-\lambda t}\varphi(y_{x,u}(t)) + \int_0^t e^{-\lambda s}\ell(y_{x,u}(s), u(s))ds \leq \varphi(x) \quad \forall t \geq 0.$$

ii) strongly increasing for the control system if $\operatorname{dom}\mathbb{A} \subseteq \operatorname{dom}\varphi$ and for any $x \in \mathcal{K}$ and $u : [0, +\infty) \to \mathcal{A}$ measurable we have

$$\text{(18)} \qquad e^{-\lambda t}\varphi(y_{x,u}(t)) + \int_0^t e^{-\lambda s}\ell(y_{x,u}(s), u(s))ds \geq \varphi(x) \quad \forall t \in [0, T),$$

where $T := \inf\{t \geq 0 : \ y_{x,u}(t) \notin \mathcal{K}\}$.

The following definition and lemma are required to single the value function out among other l.s.c. function.

Lemma 3.1. *Suppose that (H_f), (H_ℓ) and (H_0) hold, and that $\lambda > \lambda_\ell c_f$. Let $\varphi : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function with λ_ℓ-superlinear growth. If φ is:*

1. *weakly decreasing for the control system, then $\upsilon(x) \leq \varphi(x)$ for all $x \in \mathcal{K}$.*
2. *strongly increasing for the control system then $\upsilon(x) \geq \varphi(x)$ for all $x \in \mathcal{K}$.*

Proof. First of all, note that if $\lambda > \lambda_\ell c_f$, then for any function φ with λ_ℓ-superlinear growth and for any trajectory $y(\cdot)$ of ($D_{u,x}$) such that $y(t) \in \operatorname{dom}\varphi$, we have that

$$\text{(19)} \qquad \lim_{t \to +\infty} e^{-\lambda t}\varphi(y(t)) = 0.$$

Case 1. Suppose φ is weakly decreasing for the control system. Let $x \in \mathcal{K}$, if $x \notin \operatorname{dom}\varphi$ then the inequality is trivial. Let x be in $\operatorname{dom}\varphi$, there exists a control $u \in \mathbb{A}(x)$ such that for all $n \in \mathbb{N}$

$$e^{-\lambda n}\varphi(y(n)) + \int_0^\infty e^{-\lambda s}\ell(y_{x,u}(s), u(s))\mathbb{1}_{[0,n]}ds \leq \varphi(x).$$

Therefore, by using the monotone convergence theorem, and by taking into account equation (19) and the definition of the value function we obtain the desired inequality $\upsilon(x) \leq \varphi(x)$.

Case 2. Suppose φ is strongly increasing for the control system and let $x \in \mathcal{K}$. Assume that $\upsilon(x) \in \mathbb{R}$, otherwise the result is direct. Let $\bar{u} \in \mathbb{A}(x)$ be the optimal control associated with (P) and denote by $y = y_{x,\bar{u}}$ the optimal trajectory associated with $\bar{u}$. Then we have

$$e^{-\lambda t}\varphi(y(t)) + \int_0^t e^{-\lambda s}\ell(y(s), \bar{u}(s))ds \geq \varphi(x) \quad \forall t \geq 0.$$

Then using again equation (19) and taking the limit of $t \to +\infty$ we conclude the proof. □

In view of the previous comparison lemma we can state an intermediate characterization of the value function of the problem in terms of the increasing principles of definition 3.1.

Proposition 3.3. *The value function $v(\cdot)$ is the only l.s.c. function with λ_ℓ-superlinear growth that is weakly decreasing and strongly increasing for the control system at the same time.*

Proof. Recall that the value function satisfies (DPP). So, it is easy to see that $v(\cdot)$ is weakly decreasing and strongly decreasing for the control system. The uniqueness is a direct consequence of Lemma 3.1. □

4 Invariance and smooth manifolds.

Before entering into the details of the proof of the main theorem, we need to state and prove some some fundamental results on proximal analysis which will be useful in the proof of our main theorem.

4.1 Background in Nonsmooth Analysis and Smooth Manifolds.

We recall some basic notions in proximal analysis and cones. For a further discussion about this topic we refer the reader to [4, 5, 15].

Let $\mathcal{S} \subseteq \mathbb{R}^k$ be a locally closed set and $x \in \mathcal{S}$. A vector $\eta \in \mathbb{R}^k$ is called proximal normal to $\mathcal{S}$ at x if there exists $\sigma = \sigma(x, \eta) > 0$ so that

$$\frac{|\eta|}{2\sigma}|x-y|^2 \geq \langle \eta, y - x\rangle \quad \forall y \in \mathcal{S}.$$

The set of all such vectors η is known as the *Proximal normal cone* to $\mathcal{S}$ at x and is denoted by $\mathcal{N}^P_{\mathcal{S}}(x)$. This cone is a convex set and possibly contains only the zero vector of $\mathbb{R}^k$.

For a lower semicontinuous function $\varphi : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in \operatorname{dom} \varphi$, the proximal subdifferential is the set of vector $\zeta \in \mathbb{R}^k$ for which $(\zeta, -1) \in \mathcal{N}^P_{\operatorname{epi} \varphi}(x, \varphi(x))$. In such case we write that $\zeta \in \partial_P \varphi(x)$.

The *proximal subgradient inequality* is the most common way to describe the proximal subdifferential and states that $\zeta \in \partial_P \varphi(x)$ if and only if $\exists \sigma, \delta > 0$ such that

$$\varphi(y) \geq \varphi(x) + \langle \zeta, y - x\rangle - \sigma|y - x|^2 \quad \forall y \in \mathbb{B}(x, \delta) \cap \operatorname{dom} \varphi.$$

The *Bouligand tangent cone* to $\mathcal{S}$ at $x \in \mathcal{S}$, known also as the *Contingent cone*, is defined as follows:

$$\mathcal{T}^B_{\mathcal{S}}(x) = \left\{ v \in \mathbb{R}^k : \liminf_{t \to 0^+} \frac{\operatorname{dist}_{\mathcal{S}}(x + tv)}{t} \leq 0 \right\}.$$

Let $\mathcal{M}$ be an embedded manifold of $\mathbb{R}^N$ we denote by $\mathcal{T}_{\mathcal{M}}(x)$ the tangent space to $\mathcal{M}$ at $x \in \mathcal{M}$. In particular we have that $\mathcal{T}^B_{\mathcal{M}}(x)$ coincides with $\mathcal{T}_{\mathcal{M}}(x)$ at any $x \in \mathcal{M}$.

4.2 Selections Theorems and trajectories of the controlled system.

The relation between open-loop control systems and differential inclusions is well understood throughout the Filippov Selection Theorem. For sake of completeness we recall its statement for parameterized set-valued maps, the proof can be found in [4] or [30].

Lemma 4.1. *Assume $U \subseteq \mathbb{R}^m$ is a compact set and $-\infty \leq a < b \leq +\infty$ are given. Let $\phi : [a, b] \times U \to \mathbb{R}^k$ be a continuous function, and let $\gamma : [a, b] \to \mathbb{R}^k$ be a measurable function. Suppose that $\gamma(t) \in \phi(t, U)$ for almost all $t \in [a, b]$. Then there exists a measurable function $u : [a, b] \to U$ satisfying*

$$\gamma(t) = \phi(t, u(t)) \qquad a.e.\ t \in [a, b].$$

In the lemma, the selected control is only measurable, however a stronger version of this theorem, known as the Michael Selection Theorem, says that the control can in fact be continuous. For an exact statement of Michael's theorem, we refer to [5, Theorem 9.1.2] for instance.

Lemma 4.2. *Assume $U \subseteq \mathbb{R}^m$ is a compact set and $-\infty \leq a < b \leq +\infty$ are given. Let $\phi : [a,b] \times U \to \mathbb{R}^k$ be a continuous function with $\phi(t,U)$ is a convex set for any $t \in [a,b]$, and let $\gamma : [a,b] \to \mathbb{R}^k$ be a measurable function. Suppose that $\gamma(t) \in \phi(t,U)$ for almost all $t \in [a,b]$. Then there exists a continuous function $u : [a,b] \to U$ satisfying*

$$\gamma(t) = \phi(t, u(t)) \qquad a.e.\ t \in [a,b].$$

For future purpose, we require the existence of smooth functions which are locally backward trajectories of the augmented dynamics G_i given by (8).

Proposition 4.1. *Suppose that (H_f), (H_ℓ), (H_0), (H_1) and (H_2) hold. Let $x \in \mathcal{K}$ and $u_x \in \mathcal{A}_i(x)$. Then there exist $\varepsilon > 0$, a measurable control map $u : (-\varepsilon, 0] \to \mathcal{A}$ and a smooth arc $y : (-\varepsilon, 0] \to \mathcal{M}_i$ with $y(0) = x$, $\dot{y}(0) = f(x, u_x)$ and satisfying*

$$\dot{y}(t) = f(y(t), u(t)) \quad \textit{and} \quad \lim_{t \to 0^-} \frac{1}{t} \int_t^0 e^{-\lambda s} \ell(y(s), u(s)) ds = \ell(x, u_x).$$

Proof. Note that $(f(x,u_x), \ell(x,u_x)) \in G_i(0,x)$, therefore, by the (H_2) and (H_0) and the Michael Selection Theorem (see Lemma 4.2), there exists a continuous selection of G_i denoted by g_i such that $g_i(0,x) = (f(x,u_x), \ell(x,u_x))$ and the differential equation

$$(\dot{y}, \dot{z}) = g_i(t,y), \quad y(0) = x, \ z(0) = 0$$

admits a solution which is smooth and such that $(\dot{y}(0), \dot{z}(0)) = g_i(0,x)$. By Lemma 4.1, there exists a measurable control $u : (-\varepsilon, 0] \to \mathcal{A}$ such that $u(t) \in \mathcal{A}_i(y(t))$ a.e. $t \in (-\varepsilon, 0]$,

$$\dot{y}(t) = f(y(t), u(t)) \quad \text{and} \quad z(t) = \int_t^0 e^{-\lambda s} \ell(y(s), u(s)) ds, \qquad \forall t \in (-\varepsilon, 0].$$

Since $\dot{z}(0) = \ell(x, u_x)$ the conclusion follows. ∎

4.3 Invariance.

From the theoretical point of view, the proximal Invariance is a powerful tool for optimal control theory because, in the unconstrained case, it allows to characterize the value function as a the unique solution to an appropriated Hamilton-Jacobi equation by means of invariance concepts. This theory has been studied by many authors; see for instance [15, Chapter 4] among others for a more concise explanation.

Here we do not pretend to give a complete introduction to this subject but to present the fundamental results we need in the forthcoming sections. We present a sort of local versions of weak invariance and a suitable adaptation of strong invariance to embedded manifolds. The first part is taken from [35] and the second is similar in spirit to the work exposed in [7].

Definition 4.1. *Let $\mathcal{S} \subseteq \mathbb{R}^k$ nonempty, $U \subseteq \mathbb{R}^k$ open and $\Gamma : \mathbb{R}^k \rightrightarrows \mathbb{R}^k$ given. The system $(\mathcal{S}, \Gamma)$ is called weakly invariant in U if for all $x \in \mathcal{S} \cap U$, there exists a Γ-trajectory $\gamma(\cdot)$ which remains in U on a maximal interval $[0,T)$ and that satisfies*

$$\gamma(0) = x \quad \textit{and} \quad \gamma(t) \in \mathcal{S} \quad \forall t \in [0,T).$$

A very useful characterization of weakly invariance can be stated in term of minimized Hamiltonians and proximal normals.

Proposition 4.2. *[35, Theorem 3.1(a)] Suppose $\mathcal{S} \subseteq \mathbb{R}^k$ is nonempty closed, $U \subseteq \mathbb{R}^k$ is open and $\Gamma : \mathbb{R}^k \rightrightarrows \mathbb{R}^k$ is a multifunction satisfying:*

i) Γ has convex and nonempty images on $\mathbb{R}^k$ and $\mathcal{S} \cap U$, respectively.

ii) Γ has graph closed $\mathbb{R}^k \times \mathbb{R}^k$.

iii) For each compact set $K \subseteq \mathbb{R}^k$, there exists a constant $c_\Gamma > 0$ so that

$$\sup\{|v| : \ x \in K, \ v \in \Gamma(x)\} \leq c_\Gamma.$$

Then $(\mathcal{S}, \Gamma)$ is weakly invariant in U if and only if

$$\min_{v \in \Gamma(x)} \langle \eta, v \rangle \leq 0 \qquad \forall x \in \mathcal{S} \cap U, \ \forall \eta \in \mathcal{N}^P_{\mathcal{S}}(x). \tag{20}$$

We now present a useful criterion for strong invariance adapted to smooth manifolds. This proposition is similar to Theorem 4.1 in [7].

Remark 4.1. *The statement of the next proposition is very close to the one of Theorem 4.1 in [7]. For the convenience of the reader, we give here a detailed and complete proof.*

Proposition 4.3. *Suppose $\mathfrak{M}$ is an embedded manifold of $\mathbb{R}^k$, $\mathcal{S} \subseteq \mathbb{R}^k$ is a nonempty closed set such that $\mathcal{S} \cap \overline{\mathfrak{M}} \neq \emptyset$ and $\Gamma : \mathfrak{M} \rightrightarrows \mathbb{R}^k$ satisfies:*

$$\begin{cases} \forall K \subseteq \overline{\mathfrak{M}} \text{ compact } \exists L_\Gamma, c_\Gamma > 0 \text{ so that:} \\ i) \ \Gamma(x) \subseteq \Gamma(y) + L_\Gamma |x - y| \mathbb{B}, \quad \forall x, y \in \mathfrak{M} \cap K. \\ ii) \sup\{|v| : \ x \in K, \ v \in \Gamma(x)\} \leq c_\Gamma. \end{cases} \tag{21}$$

Suppose that for all $R > 0$ there exists $\kappa > 0$ such that

$$\sup_{v \in \Gamma(x)} \langle x - s, v \rangle \leq \kappa \, dist_{\mathcal{S} \cap \overline{\mathfrak{M}}}(x)^2 \qquad \forall x \in \mathfrak{M} \cap \mathbb{B}(0, R), \ \forall s \in proj_{\mathcal{S} \cap \overline{\mathfrak{M}}}(x). \tag{22}$$

Then for any $R > 0$ there exists $\kappa > 0$ such that for any Γ-trajectory lying on $\mathbb{B}(0, R)$ denoted by $\gamma : [0, T] \to \overline{\mathfrak{M}}$ with $\gamma(t) \in \mathfrak{M}$ on $(0, T)$ we have

$$dist_{\mathcal{S} \cap \overline{\mathfrak{M}}}(\gamma(t)) \leq e^{\kappa t} dist_{\mathcal{S} \cap \overline{\mathfrak{M}}}(\gamma(0)) \qquad \forall t \in [0, T].$$

Proof. Let $R > 0$ and γ as in the statement so that $\gamma([0, T]) \subseteq \mathbb{B}(0, R)$. We denote by c_Γ be the corresponding bound for the velocities, κ the constant of (22) associated with R, let L_Γ the Lipschitz constant of Γ on $\mathbb{B}(0, R)$ and $C_1 > 0$ such that

$$\max_{t \in [0, T]} \text{dist}_{\mathcal{S} \cap \overline{\mathfrak{M}}}(\gamma(t)) \leq C_1$$

Let $\varepsilon > 0$ and set $t_0 = 0$, we construct inductively a partition of $[0, T]$ in the following way: Given $t_i \in [0, T)$ take $t_{i+1} \in (t_i, T]$ as satisfying

$$t_{i+1} \leq t_i + \varepsilon \quad \text{and} \quad |\gamma((1 - s)t_i + st_{i+1}) - \gamma(t_i)| \leq \frac{1}{L_\Gamma} \varepsilon, \ \forall s \in [0, 1].$$

Note that $|\gamma((1 - s)t_i + st) - \gamma(t_i)| \leq c_\Gamma(t - t_i)$ for any $s \in [0, 1]$, so the choice of such t_{i+1} is possible. Moreover, we can do this in such a way it produces a finite partition of $[0, T]$ which we denote $\pi_\varepsilon = \{0 = t_0 < t_1 < \ldots < t_n < t_{n+1} = T\}$. Note that $\|\pi_\varepsilon\| = \max_{i=0,\ldots,n}(t_{i+1} - t_i) \leq \varepsilon$. For any $i \in \{0, \ldots, n + 1\}$, we set $\gamma_i = \gamma(t_i)$ and choose $s_i \in \text{proj}_{\mathcal{S} \cap \overline{\mathfrak{M}}}(\gamma_i)$ arbitrary. Suppose first that $\gamma(0) \in \mathfrak{M}$. We will show the inequality only for $t = T$. For $t \in (0, T)$ it suffices to restrict the partition to $[0, t]$.

Let $s \mapsto \psi_i(s) := \gamma((1-s)t_i + st_{i+1})$ defined on $[0,1]$. Whence, ψ is an absolutely continuous function with $\dot{\psi}(s) = \dot{\gamma}((1-s)t_i + st_{i+1})(t_{i_1} - t_i)$ a.e. $s \in [0,1]$. Thus

$$\gamma_{i+1} - \gamma_i = (t_{i+1} - t_i) \int_0^1 \dot{\gamma}((1-s)t_i + st_{i+1})ds$$

On the other hand, since Γ is locally Lipschitz, we have that

$$\Gamma(\gamma((1-s)t_i + st_{i+1})) \subseteq \Gamma(\gamma_i) + L_\Gamma|\gamma((1-s)t_i + st_{i+1}) - \gamma(t_i)|\mathbb{B}, \quad \forall s \in [0,1].$$

By construction $L_\Gamma|\gamma((1-s)t_i+st_{i+1})-\gamma(t_i)| \leq \varepsilon$. Therefore, there exist two measurable functions $v_i : [0,1] \to \Gamma(\gamma_i)$ and $b_i : [0,1] \to \mathbb{B}$ such that

$$\dot{\gamma}((1-s)t_i + st_{i+1}) = v_i(s) + \varepsilon b_i(s), \quad \text{a.e. } s \in [0,1].$$

Whence

$$\begin{aligned}
\text{dist}_{\mathcal{S}\cap\overline{\mathfrak{M}}}(\gamma_{i+1})^2 &\leq |\gamma_{i+1} - s_i|^2 \\
&= |\gamma_i - s_i|^2 + 2(t_{i+1} - t_i) \int_0^1 \langle \gamma_i - s_i, v_i(s) + \varepsilon b_i(s) \rangle ds + |\gamma_{i+1} - \gamma_i|^2 \\
&\leq (1 + 2(t_{i+1} - t_i)\kappa)\text{dist}_{\mathcal{S}\cap\overline{\mathfrak{M}}}(\gamma_i)^2 + \varepsilon(t_{i+1} - t_i)[2C_1 + c_\Gamma^2],
\end{aligned}$$

where this last comes from (22), the definition of b_i and the choice of t_i.

Let us denote $\sigma_i = \text{dist}_{\mathcal{S}\cap\overline{\mathfrak{M}}}(\gamma_i)$ and $\delta_i = t_{i+1} - t_i$. Then, using an inductive argument it is not difficult to show that

$$\begin{aligned}
\sigma_{n+1}^2 &\leq \prod_{i=0}^{n}(1 + 2\delta_i\kappa)\sigma_0^2 + \varepsilon[2C_1 + c_\Gamma^2] \sum_{j=0}^{n} \prod_{i=j+1}^{n} (1 + 2\delta_i\kappa)\delta_j. \\
&\leq \left(\prod_{i=0}^{n}(1 + 2\delta_i\kappa) \right) \left(\sigma_0^2 + \varepsilon[2C_1 + c_\Gamma^2] \sum_{j=0}^{n} \delta_j \right).
\end{aligned}$$

Note that

$$\sum_{j=0}^{n} \delta_j = T \quad \text{and} \quad \prod_{i=0}^{n}(1 + 2\delta_i\kappa) \leq e^{2\kappa T},$$

so we obtain

$$\sigma_{n+1}^2 \leq e^{2\kappa T}(\sigma_0^2 + \varepsilon[2C_1 + c_\Gamma^2]T).$$

Since $\sigma_{n+1} = \text{dist}_{\mathcal{S}\cap\overline{\mathfrak{M}}}(\gamma(T))$ and $\sigma_0 = \text{dist}_{\mathcal{S}\cap\overline{\mathfrak{M}}}(\gamma(0))$, letting $\varepsilon \to 0$ we obtain the desired result.

Suppose now that $\gamma(0) \notin \mathfrak{M}$. Then it is clear that for any $\delta > 0$ small enough the trajectory $\tilde{\gamma} = \gamma|_{[\delta,T]}$ satisfies the previous assumptions, so the inequality is valid on the interval $[\delta, T]$ for any $\delta > 0$. Finally, since the distance function is continuous, we can extend the inequality up to $t = 0$ by taking limits. □

5 Proof of the main result.

In this section we present a characterization of the value function as the unique solution to a bilateral Hamilton-Jabobi among a class of lower semi-continuous functions. This analysis is done in several steps, first we show that a function is weakly decreasing for the control system if and only if it is a supersolution to a Hamilton-Jabobi equation on $\mathcal{K}$. Secondly, we show that if a function is strongly increasing for the control system then it is a subsolution to a Hamilton-Jabobi equation on each stratum $\mathcal{M}_i$. The final step and most technical consist in characterizing the strongly increasing principle in terms of Hamilton-Jacobi inequalities on each strata that is component of $\mathcal{K}$.

In particular, by gathering Proposition 5.1, 5.2 and 5.3 the proof of Theorem 2.1 follows easily.

5.1 Weakly decreasing principle.

This principle states, as we mentioned above, that satisfying inequality (17) is equivalent to be a *supersolution* to a Hamilton-Jacobi equation. The idea of the proof is quite classical and do not require nothing but standing assumptions of control theory; see for instance [15, Chapter 4]. Nevertheless, we provide the proof for sake of completeness.

Proposition 5.1. *Suppose that* (H_f), (H_ℓ) *and* (H_0) *hold. Consider a given l.s.c. function with real-extended values* $\varphi : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$. *Then* φ *is is weakly decreasing for the control system if and only if*

$$\lambda\varphi(x) + H(x, \zeta) \geq 0 \qquad \forall x \in \mathcal{K},\ \forall \zeta \in \partial_P \varphi(x) \tag{23}$$

Proof. Let us first prove the implication ($\Rightarrow$).

Suppose φ is is weakly decreasing for the control system. Let $x \in \mathcal{K}$, if $\partial_P\varphi(x) = \emptyset$ then (23) holds by vacuity. If on the contrary, let $\zeta \in \partial_P\varphi(x)$, then $x \in \operatorname{dom}\varphi$ and there exists $u \in \mathbb{A}(x)$ such that (17) holds. Let us denote $y(\cdot)$ the trajectory associated with the control u starting from x at time t=0. By the proximal subgradient inequality we have that $\exists \sigma, \delta > 0$ such that

$$\varphi(y(t)) \geq \varphi(x) + \langle \zeta, y(t) - x\rangle - \sigma|y(t) - x|^2 \quad \forall t \in [0, \delta).$$

Using that $y(\cdot)$ is a trajectory and (17) we get for any t small enough

$$(1 - e^{\lambda t})\varphi(x) + \int_0^t [\langle \zeta, f(y(s), u(s))\rangle + \ell(y(s), u(s))]\, ds \leq \sigma|y(t) - x|^2$$

By (H_f) and (H_ℓ) we get

$$\frac{(1 - e^{\lambda t})}{t}\varphi(x) + \frac{1}{t}\int_0^t [\langle \zeta, f(x, u(s))\rangle + \ell(x, u(s))]\, ds \leq h(t)$$

where $h(t)$ is such that $\lim_{t\to 0^+} h(t) = 0$. Therefore taking infimum over all $u \in \mathcal{A}$ inside the integral and letting $t \to 0^+$ we obtain (23).

Now, we turn to the second par of the proof ($\Leftarrow$). To prove that φ is is weakly decreasing for the control system let us first show that $(\operatorname{epi}\psi, \Gamma)$ is weakly invariant in $U := (-\rho, +\infty) \times \operatorname{int}\mathcal{K}_\rho \times \mathbb{R}^2$, where $\rho > 0$ is given, $\mathcal{K}_\rho$ stands for $\mathcal{K} + \mathbb{B}(0, \rho)$,

$$\psi(\tau, x, z) = \begin{cases} e^{-\lambda\tau}\varphi(x) + z & \text{if } \tau \in \mathbb{R},\ x \in \mathcal{K},\ z \in \mathbb{R} \\ +\infty & \text{otherwise }, \end{cases}$$

and $\Gamma : \overline{U} \rightrightarrows \mathbb{R}^{N+3}$ is the augmented dynamic given by

$$\Gamma(\tau, x, z, w) = \{1\} \times G(\tau, x) \times \{0\}.$$

Note that $\operatorname{epi}\psi$ is closed because φ is l.s.c. and Γ has nonempty convex compact images because of (H_0). Moreover, by (H_f) and (H_ℓ), Γ is locally Lipschitz and satisfies the following growth condition:

$$\exists c_\Gamma > 0 \text{ so that } \sup\{|v| \mid v \in \Gamma(\tau, x, z, w)\} \leq c_\Gamma(1 + |x| + e^{-\lambda\tau}|x|^{\lambda_\ell}).$$

Therefore, Γ satisfies all the assumption of Proposition 4.2 and this can be applied (extending Γ to $\mathbb{R}^{N+3}$ as the empty set outside $\overline{U}$). So, let us show that (23) implies (20), where $\mathcal{S} = \operatorname{epi}\psi$.

Let $(\tau, x, z, w) \in \mathcal{S} \cap U$, then $x \in \operatorname{dom}\varphi$. Consider $\eta \in \mathcal{N}^P_{\mathcal{S}}(\tau, x, z, w)$, since this is the normal cone to an epigraph, we can write $\eta = (\xi, -p)$ with p is nonnegative. Suppose $p > 0$ then $w = \psi(\tau, x, z)$ and

$$\frac{1}{p}\xi \in \partial_P\psi(\tau, x, z) \subseteq \{-\lambda e^{-\lambda\tau}\varphi(x)\} \times e^{-\lambda\tau}\partial_P\varphi(x) \times \{1\}.$$

Therefore, for some $\zeta \in \partial_P \varphi(x)$ we have

$$\min_{v\in\Gamma(\tau,y,z,w)} \langle \eta, v\rangle \le \min_{\substack{u\in\mathcal{A},\\ 0\le r\le \beta(\tau,x,u)}} p\left\{e^{-\lambda\tau}(-\lambda\varphi(x) + \langle \eta, f(x,u)\rangle + \ell(x,u)) + r\right\}$$
$$\le e^{-\lambda\tau} p \min_{u\in\mathcal{A}} \{-\lambda\varphi(x) + \langle \zeta, f(x,u)\rangle + \ell(x,u)\}.$$

Whence, by (23)

$$\min_{v\in\Gamma(\tau,y,z,w)} \langle \eta, v\rangle \le e^{-\lambda\tau} p\,(-\lambda\varphi(x) - H(x,\zeta)) \le 0.$$

Suppose now that $p = 0$, then $(\xi, 0) \in \mathcal{N}^P_{\mathcal{S}}(\tau, x, z, \psi(\tau, x, z))$ and by Rockafellar's horizontality theorem (see for instance [29]), there exist some sequences $\{(\tau_i, x_i, z_i)\} \subseteq \operatorname{dom}\psi$, $\{(\xi_i)\} \subseteq \mathbb{R}^{N+2}$ and $\{p_i\} \subseteq (0,\infty)$ such that

$$\begin{aligned}
&(\tau^i, x^i, z^i) \to (\tau, x, z),\\
&\psi(\tau^i, x^i, z^i) \to \psi(\tau, x, z),\\
&(\xi_i, p_i) \to (\xi, 0),\\
&\frac{1}{p_i}\xi_i \in \partial_P \psi(\tau^i, x^i, z^i).
\end{aligned}$$

Thus, using the same argument as above we can show

$$\min\{\langle(\xi_i, -p_i), v\rangle \mid v \in \Gamma(\tau_i, x_i, z_i, \psi(\tau^i, x^i, z^i))\} \le 0.$$

Whence, since Γ is locally Lipschitz, we can take the liminf in the last inequality and since $\Gamma(\tau, x, z, \psi(\tau, y, z)) = \Gamma(\tau, x, z, w)$, we obtain (20).

So, by Proposition 4.2, for every $\gamma_0 = (\tau_0, x_0, z_0, w_0) \in \mathcal{S} \cap U$ there exists a Γ-trajectory $\gamma(t) = (\tau(t), y(t), z(t), w(t))$ which lies in U for a maximal period of time $[0, T)$, with $\gamma(0) = \gamma_0$ so that

$$e^{-\lambda\tau(t)}\varphi(y(t)) + z(t) \le w(t) \quad \forall t \in [0, T).$$

By the Filippov selection theorem, $y(\cdot)$ is a solution of ($D_{u,x}$) for some $u : [0, T) \to \mathcal{A}$. Also, $y(t) \in \operatorname{dom}\varphi \subseteq \mathcal{K}$, $\forall t \in [0, T)$, so

$$f(y(s), u(s)) \in \mathcal{T}^B_{\mathcal{K}}(y(s)), \qquad \text{a.e. } s \in [0, T).$$

Moreover,

$$z(t) = \int_0^t [e^{-\lambda t}\ell(y(s), u(s)) + r(s)]ds, \qquad \text{with } r(s) \ge 0 \text{ a.e.}$$

Take $\gamma_0 = (0, x, 0, \varphi(x))$ for any $x \in \operatorname{dom}\varphi$, to conclude the proof we just need to show that $T = +\infty$. By contradiction, suppose $T < +\infty$, note that $\tau(t) = t$, so by the structure of U we have that $y(t)$ approaches to $\operatorname{bdry}\mathcal{K}_\rho$ as $t \to T^-$ and $y(t) \in \mathcal{K}$, $\forall t \in [0, T)$, which is not possible since $\operatorname{dist}_{\operatorname{bdry}\mathcal{K}_\rho}(\bar{x}) > \rho$ for any $\bar{x} \in \mathcal{K}$. So the conclusion follows.

□

5.2 Strongly increasing principle.

This principle has the same nature than that of previous section in the sense that states that satisfying inequality (18) is equivalent to be a *subsolution* to a Hamilton-Jacobi equation.

Proposition 5.2. *Suppose that* (H_f), (H_ℓ), (H_0), (H_1) *and* (H_2) *hold. Let* $\varphi : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ *be a l.s.c. function. Consider a given l.s.c. function with real-extended values* $\varphi : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ *which is in addition strongly increasing for the control system. Then*

$$\lambda\varphi(x) + H_i(x, \zeta) \leq 0 \qquad \forall x \in \mathcal{M}_i,\ \forall \zeta \in \partial_P \varphi_i(x), \tag{24}$$

where $\varphi_i(x) = \begin{cases} \varphi(x) & \text{if } x \in \overline{\mathcal{M}}_i \\ +\infty & \text{otherwise.} \end{cases}$

Proof. First of all note that $\zeta \in \partial_P \varphi_i(x)$ if and only if $\exists \sigma, \delta > 0$ such that

$$\varphi(y) \geq \varphi(x) + \langle \zeta, y - x\rangle - \sigma|y - x|^2 \quad \forall y \in \mathbb{B}(x, \delta) \cap \overline{\mathcal{M}}_i.$$

We only show (24) for any $(i, x) \in I \times \mathcal{K}$ such that $x \in \operatorname{dom} \partial_P \varphi_i \cap \mathcal{M}_i \cap \operatorname{dom} \mathcal{A}_i$. Otherwise, the conclusion is direct.

Let $(i, x) \in I \times \mathcal{K}$ as before and take $u_x \in \mathcal{A}_i(x)$, it suffices to prove

$$-\lambda\varphi(x) + \langle \zeta, f(x, u_x)\rangle + \ell(x, u_x) \geq 0, \qquad \forall \zeta \in \partial_P \varphi_i(x). \tag{25}$$

Let $u : (-\varepsilon, 0] \to \mathcal{A}$ and $y : (-\varepsilon, 0] \to \mathcal{M}_i$ be the measurable control and smooth arc given by Proposition 4.1, respectively, where $\varepsilon > 0$ is also given by this proposition. Let $\bar{u} : [0, +\infty) \to \mathcal{A}$ be a measurable control map, then for all $\tau \in (0, \varepsilon)$ we define the control map $u_\tau : [0, +\infty) \to \mathcal{A}$ as follows:

$$u_\tau(t) := \begin{cases} u(t - \tau) & \text{if } t \in [0, \tau] \\ \bar{u}(t - \tau) & \text{a.e. } t \in (\tau, +\infty). \end{cases}$$

Let $y_\tau(\cdot)$ be the trajectory associated with u_τ starting from $y_\tau(0) = y(-\tau)$. Clearly, $y_\tau(t) = y(t - \tau)$ for any $t \in [0, \tau]$.

Note that $\inf\{t \geq 0 : \ y_\tau(t) \notin \mathcal{K}\} \geq \tau$ and since φ is strongly increasing for the control system, we obtain

$$e^{-\lambda\tau}\varphi(x) + \int_0^\tau e^{-\lambda s}\ell(y(s - \tau), u(s - \tau))ds \geq \varphi(y(-\tau)).$$

Take $\zeta \in \partial_P \varphi_i(x)$ and τ small enough, so that the proximal subgradient inequality is valid. Then

$$\varphi(y(-\tau)) \geq \varphi(x) + \langle \zeta, y(-\tau) - x\rangle - \sigma|y(-\tau) - x|^2.$$

Whence,

$$\left(\frac{e^{-\lambda\tau} - 1}{\tau}\right)\varphi(x) + \frac{e^{-\lambda\tau}}{\tau}\int_{-\tau}^0 e^{-\lambda s}\ell(y(s), u(s))ds + \left\langle \eta, \frac{x - y(-\tau)}{\tau}\right\rangle \geq h(\tau),$$

with $\lim_{\tau\to 0^+} h(\tau) = 0$. Therefore, by Proposition 4.1, passing to the limit in the last inequality we obtain (25) and so (24). □

5.3 Characterization of Strongly increasing principle

In this section we prove the converse of Proposition 5.2 under an extra controllability assumption (H_3). As mentioned in section 2.1 (paragraph of comments), the most important consequence of this hypothesis is the result stated in the lemma 5.1.

Let $x \in \mathcal{K}$, let $u : [0, +\infty) \to \mathcal{A}$ be a measurable control, and let y be the trajectory associated with u. Assume that $y(t)$ belongs to $\mathcal{K}$ for any positive time. In the context of stratified systems it is useful to define the notation

$$J_i(y) = \{t \geq 0 : \ y(t) \in \mathcal{M}_i\} \tag{26}$$

for denoting the set of times for which the arc $y(\cdot)$ belongs to the stratum $\mathcal{M}_i$.

Lemma 5.1. *Suppose that* (H_f) *and assertion* (*ii*) *of* (H_3) *hold for some* $i \in \mathcal{I}$. *Let* $T > 0$ *be a finite time and* $x \in \mathcal{K}$. *Let* $u : (0, +\infty) \to \mathcal{A}$ *be a given admissible control with its corresponding trajectory* $y(\cdot)$ *solution of* ($D_{u,x}$). *Assume that* y *stays in* $\mathcal{K}$ *on* $[0, T]$ *and consider the set* $J_i(y)$ *defined in* (26).

Then, there exists $\varepsilon_0 > 0$ *such that for any* $b < a < b + \varepsilon_0$ *with* $y(b)$ *and* $y(a)$ *in* $\mathcal{M}_i$, *there exist* $c > b$, *a map* $\tau : [b, a] \to [b, c]$, *and an admissible control* $\tilde{u} : [b, c] \to \mathcal{A}$ *such that if* $\tilde{y}(\cdot)$ *is the trajectory of* ($D_{u,x}$) *associated with* $\tilde{u}$ *starting from* $y(b)$, *then*

1. $\tilde{y}(s) \in \mathcal{M}_i$ *for any* $s \in [b, c]$.
2. $\tilde{y}(b) = y(b)$ *and* $\tilde{y}(c) = y(a)$.
3. τ *is piecewise smooth on* $[b, a]$ *with*
   (a) $\tau'(s) = 1$ *for a.e.* $s \in [b, a] \cap J_i(y)$,
   (b) $\exists \delta_l \in (0, 1)$ *such that* $\tau'(s) \geq \delta_l$ *for a.e.* $s \in [b, a] \setminus J_i(y)$.
   (c) $\exists \tau_0 = \tau_0(T, x) > 0$ *such that* $\sup_{s \in [b,a]} \|\tau'(s)\| \leq \tau_0$.
4. $\tilde{u}(\tau(s)) = u(s)$ *for a.e.* $s \in [b, a] \cap J_i(y)$.

Proof. Let $y(\cdot)$ is the admissible trajectory fixed in the statement of the Lemma.

By assertion (*ii*) of (H_3), there exist $\rho_0 > 0$ and $\varepsilon_0 > 0$ such that for any $\theta \in J_i(y) \cap [b, a]$ with $b < a < b + \varepsilon_0$ and any $z \in \mathcal{M}_i \cap \overline{\mathbb{B}}(y(\theta), \rho_0)$, with $z \neq y(\theta)$, there exists $\theta_z > \theta$, an admissible control $\tilde{u}_z$ and its admissible trajectory $y^{\tilde{u}_z}_{y(\theta)}$ starting from $y(\theta)$ such that $y^{\tilde{u}_z}_{y(\theta)}(s) \in \mathcal{M}_i$ for every $s \in [\theta, \theta_z]$ and $y^{\tilde{u}_z}_{y(\theta)}(\theta_z) = z$.

Let $\theta_1 = \sup\{s \in [b, a] : \ y(t) \in \mathcal{M}_i \ \forall t \in [b, s]\}$ be the first switching time of $\mathcal{M}_i$ on $[b, a]$. If $\theta_1 = a$ the proof is direct. Suppose that $\theta_1 < a$ and take

$$\theta_2 = \sup\{s \in J_i \cap (\tau_1, a] : \ y(s) \in \overline{\mathbb{B}}(y(\theta_1), \rho_0)\}.$$

By the previous argument, there exists an admissible trajectory $\tilde{y} : [\theta_1, \tilde{\theta}_2] \to \mathcal{M}_i$ such that $\tilde{y}(\tilde{\theta}_2) = y(\theta_2)$. If $\theta_2 < a$, we can repeat the process and define θ_3 and so on. Now, by Remark 2.1, $\theta_{k+1} - \theta_k$ is bounded from below by a constant which depends un x, T and ρ_0. Therefore, since the interval $[b, a]$ is compact, the process will eventually finish in a finite number of steps, so the claim is proved. □

Now we are in position to state a result on the converse of Proposition 5.2.

Proposition 5.3. *Suppose that* (H_0), (H_1), (H_2) *and* (H_3) *hold in addition of* (H_f) *and* (H_ℓ). *Let* $\varphi : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ *be a l.s.c. function. If* (24) *holds, then* φ *is strongly increasing for the control system*

Proof. Let us set $\mathfrak{M}_i = \mathbb{R} \times \mathcal{M}_i \times \mathbb{R}^2$ and let $\Gamma_i : \mathfrak{M}_i \rightrightarrows \mathbb{R}^{N+3}$ given by

$$\Gamma_i(\tau, x, z, w) = \{-1\} \times G_i(\tau, x) \times \{0\}, \quad \forall (\tau, x, z, w) \in \mathfrak{M}_i.$$

Note that $\mathfrak{M}_i$ is an embedded manifold of $\mathbb{R}^{N+3}$ and for any $i \in \mathcal{I}$ such that $\mathcal{A}_i$ has nonempty images, Γ_i satisfies the same assumptions as G_i.

Let $\mathcal{S}_i = \overline{\mathfrak{M}}_i \cap \mathcal{S} = \mathrm{epi}(\varphi_i)$ which is closed. Then if (24) holds we have

$$\sup_{v \in \Gamma_i(\tau,x,z,w)} \langle \eta, v \rangle \leq 0 \qquad \forall (\tau, x, z, w) \in \mathcal{S}_i, \ \forall \eta \in \mathcal{N}^P_{\mathcal{S}_i}(\tau, x, z, w). \tag{27}$$

Indeed, take $(\tau, x, z, w) \in \mathcal{S}_i$ and $(\xi, -p) \in \mathcal{N}^P_{\mathcal{S}_i}(\tau, x, z, w)$. Therefore, we have $p \geq 0$ because $\mathcal{S}_i$ is the epigraph of a function. Suppose $\Gamma_i(\tau, x, z, w) \neq \emptyset$, otherwise the inequality is obtained directly.

Consider $p > 0$, then, by Proposition 5.1, for any $v \in \Gamma_i(\tau, x, z, w)$ we have, for some $u \in \mathcal{A}_i(x)$, $r \geq 0$ and $\zeta \in \partial_P \varphi(x)$

$$\begin{aligned}\langle (\xi, -p), v \rangle &= p \left\{ e^{-\lambda \tau}(\lambda \varphi(x) - \langle \zeta, f(x,u) \rangle - \ell(x,u)) - r \right\} \\ &\leq p e^{-\lambda \tau}(\lambda \varphi(x) - \langle \zeta, f(x,u) \rangle - \ell(x,u)) \\ &\leq p e^{-\lambda \tau}(\lambda \varphi(x) + H_i(x, \zeta)) \\ &\leq 0\end{aligned}$$

Since $v \in \Gamma_i(\tau, x, z, w)$ is arbitrary we can take supremum and we obtain the desired inequality. For the case $p = 0$ we can use the Rockafellar Horizontal Theorem and the continuity of Γ_i to obtain (27).

Recall that Γ_i is locally Lipschitz, then (22) holds and so, in particular, by Proposition 4.3 we have that for any index $i \in \mathcal{I}$ such that $\operatorname{dom} \mathcal{A}_i \neq \emptyset$, for any Γ_i-trajectory $\gamma : [a,b] \to \overline{\mathfrak{M}}_i$ with $\gamma(t) \in \mathfrak{M}_i$ for any $t \in (a,b)$, there exists $L > 0$ such that

$$\operatorname{dist}_{\mathcal{S} \cap \overline{\mathfrak{M}}_i}(\gamma(t)) \leq e^{Lt} \operatorname{dist}_{\mathcal{S} \cap \overline{\mathfrak{M}}_i}(\gamma(a)) \qquad \forall t \in [a,b]. \tag{28}$$

On the other hand, let $x \in \operatorname{dom} \varphi$ and $u : [0, +\infty) \to \mathcal{A}$ measurable. We want to show that inequality (18) holds for $y = y_{x,u}(\cdot)$. For this purpose we set

$$T := \sup\{t \geq 0 : \; y(s) \in \mathcal{K}, \forall s \in [0,t]\}.$$

If $T = 0$ there is nothing to prove, let us focus on the case $T > 0$.

Let $\gamma : [a,b] \to \mathcal{K}$ be a given arc, then we set $\mathcal{I}_{[a,b]}(\gamma) = \{i \in \mathcal{I} : \; \exists s \in [a,b], \; \gamma(s) \in \mathcal{M}_i\}$. Note that for every $t \in [0,T]$, the set $\mathcal{I}_{[0,t]}(y)$ is finite and

$$[0,t] = \bigcup_{i \in \mathcal{I}_{[0,t]}(y)} J_i(y).$$

We split the rest of the proof into two parts:

Step 1. Suppose first that each $J_i(y)$ can be written as the union of a finite number of interval, this means that there exists a partition of $[0,t]$

$$\pi = \{0 = t_0 \leq t_1 \leq \ldots \leq t_n \leq t_{n+1} = t\}$$

so that if $t_l < t_{l+1}$ for some $l \in \{0, \ldots, n\}$, then there exists $i_l \in \mathcal{I}_{[0,t]}(y)$ satisfying $(t_l, t_{l+1}) \subseteq J_{i_l}(y)$.

Thus, let $l \in \{0, \ldots, n\}$ be an arbitrary index such that $t_l < t_{l+1}$ and i_l as above. Let $x_l = y_{x,u}(t_{l+1})$, $y_l(s) = y_{x,u}(t_{l+1} + t_l - s)$ and $u_l(s) = u(t_{l+1} + t_l - s)$ for any $s \in [t_l, t_{l+1}]$.

Consider the arc

$$\gamma(s) = \left(t_l - s, y_l(s), -\int_{t_l}^{s} e^{\lambda r} \ell(y_l(r), u_l(r)) dr, \varphi(x_l) \right) \quad \forall s \in [t_l, t_{l+1}].$$

Then, $\gamma(\cdot)$ is a trajectory of Γ_{i_l} on $[t_l, t_{l+1}]$ which starts from $\gamma_l = (0, x_l, 0, \varphi(x_l))$ and that belongs to $\mathcal{S}_{i_l}$. So by (28), $\gamma(t) \in \mathcal{S}_{i_l}$, and we have

$$\varphi(y(t_l)) \leq e^{-\lambda(t_{l+1} - t_l)} \varphi(y(t_{l+1})) + \int_{t_l}^{t_{l+1}} e^{-\lambda s} \ell(y,u) ds.$$

Since the previous estimation is for any l arbitrary and taking into account the fact that $e^{-\lambda(t_{l+1} - t_l)} \leq 1$, for any $0 \leq l \leq n$, we obtain

$$\varphi(y(t_0)) \leq e^{-\lambda(t_{n+1} - t_0)} \varphi(y(t_{n+1})) + \int_{t_0}^{t_{n+1}} e^{-\lambda s} \ell(y,u) ds,$$

which is exactly (18), so the result follows.

Step 2. In general, the admissible trajectories may cross a stratum, infinitely many times in arbitrary small periods of times. In order to deal with this general situation, we will use an inductive argument in the number of strata where the trajectory can pass, let us denote this number by κ. The induction hypothesis $(\mathcal{P}_\kappa)$ is:

Suppose $\mathcal{M}$ is the union of κ strata and $y(s) \in \mathcal{M}$ for every $s \in (a, b)$, where $0 \leq a < b \leq t$ then

$$\varphi(y(a)) \leq e^{-\lambda(b-a)}\varphi(y(b)) + \int_a^b e^{-\lambda s}\ell(y,u)ds.$$

As shown in Step 1, the induction property holds true for the case when $\kappa = 1$ because the arc remains in only one stratum. So, let us assume that the induction hypothesis holds for some $\kappa \geq 1$. Let us prove it also holds for $\kappa + 1$.

Suppose that for some $0 \leq a < b \leq t$, the arc y is contained in the union of $\kappa + 1$ strata on the interval (a, b). By the stratified structure of $\mathcal{K}$, we can always assume that there exists a unique stratum of minimal dimension where the trajectory passes, we denote it by $\mathcal{M}_i$ and by $\mathcal{M}$ the union of the remaining κ strata. Note that, $\mathcal{M}_i \subseteq \overline{\mathcal{M}}$ and $\mathcal{M}$ is relatively open with respect to $\overline{\mathcal{M}}$. Two cases have to be considered:

Case 1: If statement (i) of (H_3) holds, then the map $x \mapsto f(x, \mathcal{A}) \cap \mathcal{T}^B_{\overline{\mathcal{M}}}(x)$ is locally Lipschitz around $\mathcal{M}_i \cap \mathcal{M}$ because of (H_2). So, by Proposition 4.3 the induction hypothesis holds.

Case 2: If statement (ii) of (H_3) holds. Without loss of generality we can assume that $y(a), y(b) \in \mathcal{M}_i$. Therefore, $J = [0, t] \setminus J_i(y)$ is open and so, for any $\varepsilon > 0$ there exists a partition of $[0, t]$

$$b_0 := a \leq a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_n < b_n \leq b =: a_{n+1}$$

such that

$$\text{meas}\left(J \setminus \bigcup_{l=1}^{n}(a_l, b_l)\right) \leq \varepsilon.$$

with $y(a_l), y(b_l) \in J_i$ and $(a_l, b_l) \subseteq J$ for any $l = 1, \ldots, n$. In particular, by the induction hypothesis we have

$$\varphi(y(a_l)) \leq e^{-\lambda(b_l-a_l)}\varphi(y(b_l)) + \int_{a_l}^{b_l} e^{-\lambda s}\ell(y,u)ds. \tag{29}$$

Note also that

$$\bigcup_{l=0}^{n}[b_l, a_{l+1}] \setminus J_i(y) \subseteq J \setminus \bigcup_{l=1}^{n}(a_l, b_l).$$

Whence, if we set $\varepsilon_l = \text{meas}([b_l, a_{l+1}] \setminus J_i(y))$, we have $\sum_{l=0}^{n} \varepsilon_l \leq \varepsilon$.

For sake of simplicity, assume that τ given by Lemma 5.1 is smooth in the whole time interval (otherwise the same analysis can be repeated in each interval where τ is smooth). By Lemma 5.1 we have that for any $l \in \{0, \ldots, n\}$ such that $b_l < a_{l+1}$, there exist $c_{l+1} > b_l$, a map $\tau : [b_l, a_{l+1}] \to [b_l, c_{l+1}]$, and an admissible control $\tilde{u} : [b_l, c_{l+1}] \to \mathcal{A}$ such that if $\tilde{y}$ is the trajectory associated with $\tilde{u}$ starting from $y(b_l)$, then

1. $\tilde{y}(s) \in \mathcal{M}_i$ for any $s \in [b_l, c_{l+1}]$.
2. $\tilde{y}(b_l) = y(b_l)$ and $\tilde{y}(c_{l+1}) = y(a_{l+1})$.
3. τ is piecewise smooth on $[b_l, a_{l+1}]$ with
   (a) $\tau'(s) = 1$ if $s \in [b_l, a_{l+1}] \cap J_i(y)$,
   (b) $\exists \delta_l \in (0, 1)$ such that $\tau'(s) \geq \delta_l$ whenever $s \in [b_l, a_{l+1}] \setminus J_i(y)$.

(c) $\exists \tau_0 = \tau_0(t,x) > 0$ such that $\sup_{s\in[b_l,a_{l+1}]} \|\tau'(s)\| \le \tau_0$.

4. $\tilde{u}(\tau(s)) = u(s)$ for any $s \in [b_l, a_{l+1}] \cap J_i(y)$.

By Remark 2.1, we have $1 + |y(s)| \le (1+|x|)e^{c_f t}$ for every $s \in [0,t]$, so since $\tilde{y}$ is indeed a trajectory of the control system starting from $y(b_l)$, we have $|\tilde{y}(s)| \le (1+|x|)e^{2c_f t}$ for any $s \in [b_l, a_{l+1}]$. In particular, setting $r = (1+|x|)e^{2c_f t}$, we have

$$y(s), \tilde{y}(\tau(s)) \in \mathbb{B}(0,r), \quad \forall s \in [b_l, a_{l+1}].$$

Let us denote $J_i^l := [b_l, a_{l+1}] \cap J_i(y)$ and $J^l := [b_l, a_{l+1}] \cap J$. Recall that $[b_l, a_{l+1}] = J_i^l \cup J^l$. Therefore, by assertions 3.(a) and 3.(b) we get

$$\varphi(y(b_l)) \le e^{-\lambda(c_{l+1}-b_l)}\varphi(y(a_{l+1})) + \int_{b_l}^{c_{l+1}} e^{-\lambda s}\ell(\tilde{y}(s),\tilde{u}(s))ds. \tag{30}$$

Note that by 3.(c) we have that $s - \varepsilon_l \le \tau(s) \le s + \varepsilon_l \tau_0$ and

$$\int_{b_l}^{c_{l+1}} e^{-\lambda s}\ell(\tilde{y}(s),\tilde{u}(s))ds = \int_{b_l}^{a_{l+1}} e^{-\lambda\tau(s)}\ell(\tilde{y}(\tau(s)),\tilde{u}(\tau(s)))\tau'(s)ds.$$

On the other hand, note that by (d), $|\tilde{y}(\tau) - y| = 0$ on J_i^l. Therefore, since y and $\tilde{y}(\tau)$ are bounded on $[b_l, a_{l+1}]$, and ℓ is locally Lipchitz, there exists $L_\ell = L_\ell(t,x) > 0$ such that

$$\left| \int_{J_i^l} e^{-\lambda\tau}\left[\ell(\tilde{y}(\tau),\tilde{u}(\tau))\tau' - \ell(y,u)\right]ds \right| \le \int_{J_i^l} L_\ell |\tilde{y}(\tau) - y| = 0.$$

and by (H_ℓ), it comes that:

$$\left| \int_{J^l} e^{-\lambda\tau}\left[\ell(\tilde{y}(\tau),\tilde{u}(\tau))\tau' - \ell(y,u)\right]ds \right| \le 2c_\ell(1+r^{\lambda_\ell})\tau_0\varepsilon_l.$$

Thus, setting $L := 2c_\ell\tau_0(1+r^{\lambda_\ell})$ we get

$$\int_{b_l}^{c_{l+1}} e^{-\lambda s}\ell(\tilde{y}(s),\tilde{u}(s))ds \le e^{\lambda\varepsilon_l}\int_{J_i^l} e^{-\lambda s}\ell(y,u)ds + L\varepsilon_l.$$

Therefore, since $\ell \ge 0$ and by (29) and (30) we get

$$\begin{aligned}
\varphi(y(b_l)) &\le e^{\lambda\varepsilon_l}\left(e^{-\lambda(a_{l+1}-b_l)}\varphi(y(a_{l+1})) + \int_{b_l}^{a_{l+1}} e^{-\lambda s}\ell(y,u)ds\right) + L\varepsilon_l, \\
&\le e^{\lambda\varepsilon_l}\left(e^{-\lambda(b_{l+1}-b_l)}\varphi(y(b_{l+1})) + \int_{b_l}^{b_{l+1}} e^{-\lambda s}\ell(y,u)ds\right) + L\varepsilon_l.
\end{aligned}$$

In the case $b_l = a_{l+1}$ this inequality trivially holds by (29). So, using an inductive argument we can prove that

$$\begin{aligned}
\varphi(y(b_0)) \le\; & e^{\lambda\sum_{l=0}^{n-1}\varepsilon_l}\left(e^{-\lambda(b_n-b_0)}\varphi(y(b_n)) + \int_{b_0}^{b_n} e^{-\lambda s}\ell(y,u)ds\right) \\
&+ L\left(\sum_{l=0}^{n-1}\varepsilon_l e^{\lambda\sum_{k=0}^{l-1}\varepsilon_k}\right),
\end{aligned}$$

and using (30) on the interval $[b_n, a_{n+1}]$ we get

$$\varphi(y(b_0)) \leq e^{\lambda \sum_{l=0}^{n} \varepsilon_l} \left(e^{-\lambda(a_{n+1}-b_0)} \varphi(y(a_{n+1})) + \int_{b_0}^{a_{n+1}} e^{-\lambda s} \ell(y,u) ds \right) + L \left(\sum_{l=0}^{n} \varepsilon_l e^{\lambda \sum_{k=0}^{l-1} \varepsilon_k} \right).$$

Finally, by the definition of b_0 and a_{n+1} we finally get

$$\varphi(y(a)) \leq e^{\lambda\varepsilon} \left(e^{-\lambda(b-a)} \varphi(y(b)) + \int_a^b e^{-\lambda s} \ell(y,u) ds \right) + L e^{\lambda\varepsilon} \varepsilon.$$

So, letting $\varepsilon \to 0$ we obtain the induction hypothesis for $\kappa + 1$. So (18) follows easily from the induction hypothesis and the proof is completed.

□

References

[1] Y. Achdou, F. Camilli, A. Cutrì, and N. Tchou. Hamilton–jacobi equations constrained on networks. *Nonlinear Differential Equations and Applications NoDEA*, pages 1–33, 2012.

[2] Y. Achdou, S. Oudet, and N. Tchou. Hamilton-jacobi equations for optimal control on junctions and networks. *hal.archives-ouvertes.fr:hal-00847210*, 2013.

[3] A. Altarovici, O. Bokanowski, and H. Zidani. A general hamilton-jacobi framework for non-linear state-constrained control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 2012.

[4] J.-P. Aubin and A. Cellina. *Differential inclusions: set-valued maps and viability theory.* Springer-Verlag New York, Inc., 1984.

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

[6] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[7] R. C. Barnard and P. R. Wolenski. Flow invariance on stratified domains. *Set-Valued and Variational Analysis*, pages 1–27, 2012.

[8] E. N. Barron and R. Jensen. Semicontinuous viscosity solutions for Hamilton-Jacobi equations with convex Hamiltonians. *Comm. Partial Differential Equations*, 15(12):1713–1742, 1990.

[9] P. Bettiol and H. Frankowska. Regularity of solution maps of differential inclusions under state constraints. *Set-Valued Analysis*, 15(1):21–45, 2007.

[10] P. Bettiol, H. Frankowska, and R. B. Vinter. l^∞ estimates on trajectories confined to a closed subset. *Journal of Differential Equations*, 252(2):1912–1933, 2012.

[11] A. Blanc. Deterministic exit time problems with discontinuous exit cost. *SIAM J. Control Optim.*, 35:399–434, 1997.

[12] O. Bokanowski, N. Forcadel, and H. Zidani. Deterministic state-constrained optimal control problems without controllability assumptions. *ESAIM: Control, Optimisation and Calculus of Variations*, 17(04):995–1015, 2011.

[13] I. Capuzzo-Dolcetta and P.-L. Lions. Hamilton-Jacobi equations with state constraints. *Trans. Amer. Math. Soc.*, 318(2):643–683, 1990.

[14] P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre. Optimal times for constrained nonlinear control problems without local controllability. *Appl. Math. Optim*, 36:21–42, 1997.

[15] F. Clarke, Y. Ledyaev, R. Stern, and P. Wolenski. *Nonsmooth Analysis and Control Theory (Graduate Texts in Mathematics)*. Springer, 1 edition, 1997.

[16] F. Clarke and R. Stern. Hamilton-jacobi characterization of the state constrained value. *Nonlinear Analysis: Theory, Methods & Applications*, 61(5):725–734, 2005.

[17] M. Crandall and P.-L. Lions. Viscosity solutions of Hamilton Jacobi equations. *Bull. American Mathematical Society*, 277:1–42, 1983.

[18] N. Forcadel, Z. Rao, and H. Zidani. State-constrained optimal control problems of impulsive differential equations. *Applied Mathematics & Optimization,*, 68:1–19, 2013.

[19] H. Frankowska and M. Mazzola. Discontinuous solutions of hamilton–jacobi–bellman equation under state constraints. *Calculus of Variations and Partial Differential Equations*, 46(3-4):725–747, 2013.

[20] H. Frankowska and S. Plaskacz. Semicontinuous solutions of Hamilton-Jacobi-Bellman equations with degenerate state constraints. *J. Math. Anal. Appl.*, 251(2):818–838, 2000.

[21] H. Frankowska and R. B. Vinter. Existence of neighboring feasible trajectories: Applications to dynamic programming for state-constrained optimal control problems. *Journal of Optimization Theory and Applications*, 104(1):20–40, 2000.

[22] H. Ishii and S. Koike. A new formulation of state constraint problems for first-order pdes. *SIAM Journal on Control and Optimization*, 34(2):554–571, 1996.

[23] V. Y. Kaloshin. A geometric proof of the existence of whitney stratifications. *Mosc. Math. J*, 5(1):125–133, 2005.

[24] P. Loreti. Some properties of constrained viscosity solutions of Hamilton-Jacobi-Bellman equations. *SIAM J. Control Optim.*, 25:1244–1252, 1987.

[25] P. Loreti and E. Tessitore. Approximation and regularity results on constrained viscosity solutions of Hamilton-Jacobi-Bellman equations. *J. Math. Systems, Estimation Control*, 4:467–483, 1994.

[26] M. Motta. On nonlinear optimal control problems with state constraints. *SIAM J. Control Optim.*, 33:1411–1424, 1995.

[27] M. Motta and F. Rampazzo. Multivalued dynamics on a closed domain with absorbing boundary. applications to optimal control problems with integral constraints. *Nonlinear Analysis*, 41:631–647, 2000.

[28] C. Nour and R. Stern. The state constrained bilateral minimal time function. *Nonlinear Analysis: Theory, Methods & Applications*, 69(10):3549–3558, 2008.

[29] R. T. Rockafellar. Proximal subgradients, marginal values, and augmented lagrangians in nonconvex optimization. *Mathematics of Operations Research*, 6(3):424–436, 1981.

[30] G. Smirnov. *Introduction to the theory of differential inclusions.* American Mathematical Society, 2002.

[31] H. Soner. Optimal control with state-space constraint I. *SIAM Journal on Control and Optimization*, 24(3):552–561, 1986.

[32] H. Soner. Optimal control with state-space constraint II. *SIAM Journal on Control and Optimization*, 24(6):1110–1122, 1986.

[33] R. Stern. Characterization of the state constrained minimal time function. *SIAM journal on control and optimization*, 43(2):697–707, 2004.

[34] L. Van den Dries and C. Miller. Geometric categories and o-minimal structures. *Duke Math. J*, 84(2):497–540, 1996.

[35] P. Wolenski and Y. Zhuang. Proximal analysis and the minimal time function. *SIAM J. Control Optim.*, 36(3):1048–1072 (electronic), 1998.