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

$$\begin{cases} \dot{y}(t) = f(y(t), u(t)) & \text{a.e. } t > 0 \\ y(0) = x, \end{cases} \tag{1}$$

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

$$y_{x,u}(t) \in \mathcal{K}, \quad \forall t \geq 0. \tag{2}$$

We denote the set of admissible controls as

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

*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

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

In the case when $\mathcal{K} = \mathbb{R}^N$ and under standard hypothesis on the data, it is well known that $\vartheta(\cdot)$ 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, Chapter 3].

When the control problem is in the presence of state constraints ($\mathcal{K} \neq \mathbb{R}^N$), a constrained HJB equation can be associated with the value function as done in [39]. In our setting, the HJB equation takes the form

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

where $H(x,p) := \max\{-\langle f(x,u),p\rangle - \ell(x,u) \mid \ u \in \mathcal{A}\}$. It is well-known that the value function satisfies (4) in the constrained viscosity sense, that means that ϑ is a subsolution on $\text{int}(\mathcal{K})$ 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 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 [15, 27].

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 [39] for the case when $\text{int}(\mathcal{K}) \neq \emptyset$ with smooth boundary and it has been object of subsequence generalization to various cases; see [27, 41, 19, 34, 22] and the references therein. This condition basically says that at each point of the boundary of $\mathcal{K}$ there exists a controlled vector field pointing into $\mathcal{K}$. Under this assumption the value function is Lipschitz continuous and then uniqueness can be established. Furthermore, 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 [24, 11, 13]. We refer to [16, 32, 33] for weaker inward pointing assumptions, and to [30, 31] 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 [14] 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 [24, 23, 22].

Let us stress that all the works mentioned above assume the same compatibility assumption on the whole boundary, that is, no mixed type of pointing condition has been so far considered and, with exception of [23], 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:

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

using the transformation $y_1 = y$ and $y_2 = \dot{y}$ the systems can be rewrite as:

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

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 hypotheses whose main purpose is to approximate any trajectory of the control system by a sequence of trajectories (of the same controlled dynamic) lying completely in the interior of the constraints. This fact may suggest that a more appropriated requirement would be the *interior approximation of trajectories*. Nevertheless, it will automatically rule out cases where junctions are present, for instance

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

The previous example also illustrates the importance of considering the information passing through the boundary in the case when the interior of the set of feasible states is disconnected, the approach proposed below seems to be well adapted to these situations. Furthermore, optimal control problems on networks as studied in [1] (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 [17] for more details. In [3], it has been shown that the epigraph of the value function ϑ 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 assumption, as the unique viscosity solution to a HJB equation. This approach leads to a constructive way for determining the epigraph of ϑ and to its numerical approximation. It can also be extended to more general situations of time-dependent state constraints [21].

In this paper, we follow the same line of investigation as in [27, 15]. 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 [41] 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 2.1 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.

1.1 Notation.

Throughout this paper, $\mathbb{R}$ denotes the sets of real numbers, $|\cdot|$ is the Euclidean norm and $\langle \cdot, \cdot \rangle$ is the Euclidean inner product 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 $S \subseteq \mathbb{R}^N$, $\mathrm{int}(S), \overline{S}, \mathrm{bdry}(S)$ and $\mathrm{co}(S)$ denote its interior, closure, boundary and convex hull, respectively. Also for S convex we denote by $\mathrm{ri}(S)$ and $\mathrm{rbd}(S)$ 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 $\mathrm{dist}_{\mathcal{S}}(x) = \inf\{|x - y| : \ y \in \mathcal{S}\}$ 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 $\mathrm{proj}_S(x)$. Let S_1 and S_2 be two compact set, then the Hausdorff distance is given by

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

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

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

The effective domain of φ is given by $\operatorname{dom}\varphi = \{x \in \mathbb{R}^N \mid \varphi(x) \in \mathbb{R}\}$. If Γ is a set-valued map, then $\operatorname{dom}\Gamma$ is the set points for which $\Gamma(x) \neq \emptyset$.

For an embedded manifold of $\mathbb{R}^N$, the tangent space to $\mathcal{M}$ at x is $\mathcal{T}_{\mathcal{M}}(x)$.

2 Setting of the problem and main result.

Throughout the paper the abbreviation l.s.c (respectively, u.s.c) stands for "lower semicontinuous" (respectively, upper semicontinuous).

2.1 Stratified systems and hypothesis.

We consider the optimal control problem of infinite horizon given by (3) where $\mathcal{A}$ is a compact set of $\mathbb{R}^m$ and $\lambda > 0$ is a given discount factor. Henceforward, the notation ϑ is reserved to denote the value function of the problem.

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:

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

And the running cost $\ell : \mathbb{R}^N \times \mathcal{A} \to \mathbb{R}$ satisfies:

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

Let $x \in \mathcal{K}$ and $u : [0,+\infty) \to \mathcal{A}$ be a measurable control. By a solution to (1) 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, and thus it is denoted by $y_{x,u}$. Furthermore, the maximal solution is defined for all times.

Remark 2.1. *By the Gronwall Lemma and (H_f), each solution to (1) satisfies:*

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

Furthermore, since $\lambda_\ell \geq 1$

$$\ell(y(t),u(t)) \leq c_\ell(1+|x|)^{\lambda_\ell} e^{\lambda_\ell c_f t}.$$

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

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

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

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

It is not difficult to see that by (H_ℓ) this set-valued map has compact and nonempty images on a neighborhood of $[0, +\infty) \times \mathcal{K}$. Moreover, throughout the paper, we will also suppose that

(H_0) $\qquad G(\cdot)$ has convex images on a neighborhood of $[0, +\infty) \times \mathcal{K}$.

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

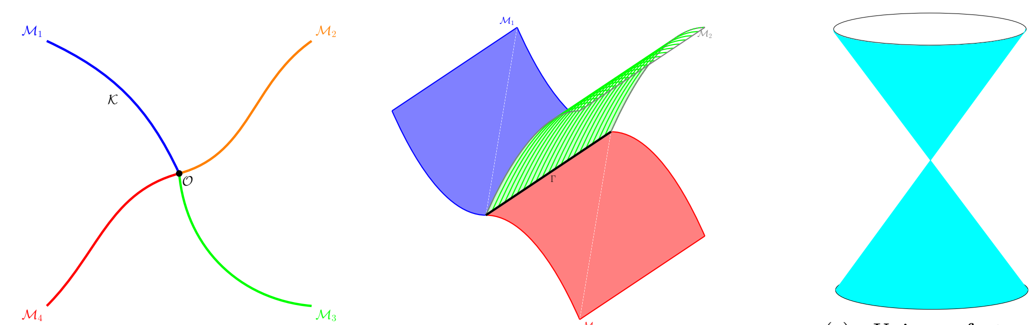
that is, there exists a locally finite collection $\{\mathcal{M}_i\}_{i\in\mathcal{I}}$ of embedded manifolds of $\mathbb{R}^N$ such that:

- $\mathcal{K} = \bigcup_{i\in\mathcal{I}} \mathcal{M}_i$ 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 $\mathcal{M}_i \subseteq \overline{\mathcal{M}_j}$ and $\dim(\mathcal{M}_i) < \dim(\mathcal{M}_j)$.

In particular, if $\operatorname{int} \mathcal{K} \neq$ and $\partial\mathcal{K}$ is smooth, as in [39, 35], then (H_1) with only two strata, namely, $\mathcal{M}_0 = \operatorname{int} \mathcal{K}$ and $\mathcal{M}_1 = \partial\mathcal{K}$.

Remark 2.2. *A network as in [1] can be considered as a particular case of stratified sets. Indeed, in case $\mathcal{K}$ is a network, the stratification consists only of edges and junctions. Figure 1a shows an example of a network consisting of edges $\mathcal{M}_1, \ldots, \mathcal{M}_4$ and a junction $\mathcal{M}_0 := \{\mathcal{O}\}$ (note that here $\mathcal{M}_0$ is a manifold of dimension 0).*

More general networks can also be considered as in Figure 1b where the set $\mathcal{K}$ is a network embeded in the space $\mathbb{R}^3$. The stratification consists of three branches that are smooth surfaces $\mathcal{M}_1, \ldots, \mathcal{M}_3$, and a junction $\mathcal{M}_0$ that corresponds to the curve Γ



(a) A network in $\mathbb{R}^2$.

(b) A generalized network in $\mathbb{R}^3$

(c) Union of two cones

Remark 2.3. *An important class of sets that admits 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:*

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

Furthermore, the class of stratifiable sets is quite broad, it includes sub-analytic and semi-algebraic sets and also definable sets of an o-minimal structure; see for instance [42, 28].

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

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

This map in general is only upper semicontinuous with possibly empty images. However, for the purposes of this paper we require more regularity. We assume the following, for every $i \in \mathcal{I}$:

(H_2) $\qquad$ $\mathcal{A}_i$ is locally Lipschitz on $\mathcal{M}_i$ with respect to the Hausdorff metric.

Remark 2.4. *Note that $\mathcal{A}_i$ can be extended up to $\overline{\mathcal{M}}_i$ by density. Furthermore, this extension turns out to be locally Lipschitz as well. So without loss of generality we assume that $\mathcal{A}_i$ is defined up to $\overline{\mathcal{M}}_i$ in a locally Lipchitz way.*

Remark 2.5. *By the convention adopted in Section 1.1 for the Hausdorff distance, (H_2) allows to consider the case when no tangential dynamic is defined, that is, when $\mathcal{A}_i$ has empty images. In that situation, (H_2) states that if $\mathcal{A}_i(x) = \emptyset$ for some $x \in \mathcal{M}_i$, then it has empty images all along $\mathcal{M}_i$.*

Example 2.1. *Consider the following dynamic:*

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} y_2 \\ u \end{pmatrix}, \qquad\qquad u \in \mathcal{A} := [-1,1], \quad y_1(t), y_2(t) \in [-r,r].$$

Many stratifications are possible for the set of state-constraints. We represent one particular stratification in Figure 1. In this case, $\mathcal{M}_0$ is the interior of the square, $\mathcal{M}_3$, $\mathcal{M}_4$, $\mathcal{M}_9$, $\mathcal{M}_{10}$, $\mathcal{M}_{11}$ and $\mathcal{M}_{12}$ are segments, and $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_5$, $\mathcal{M}_6$, $\mathcal{M}_7$ and $\mathcal{M}_8$ are single points. We can check easily (H_2), indeed, $\mathcal{A}_0 = \mathcal{A}$, $\mathcal{A}_i = \{0\}$ for $i = 1, \ldots, 4$ and $\mathcal{A}_j = \emptyset$ for $i = 5, \ldots, 12$.

It is clear in this example that neither the IPC nor the OPC condition is satisfied. In figure 1, the green zone corresponds to the viable set, that is, the set of points for which $\mathbb{A}(x) \neq \emptyset$. Note that in this case, the viable set can also be decomposed into a regular stratification.

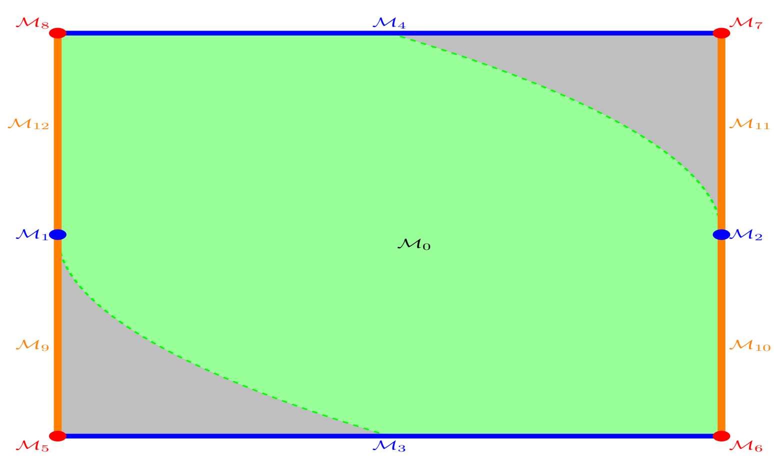


Figure 1: Stratification of Example 2.1.

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, for any $x \in \mathcal{K}$ and

$t \geq 0$, we denote by $\mathcal{R}(x;t) = \cup_{u\in\mathbb{A}(x)}\{y_{x,u}(t)\}$ the reachable set at time t, that is the set of all possible positions that can be reached by an admissible trajectory solution of (1) associated to a feasible control $u \in \mathbb{A}(x)$ (and then lying in the set $\mathcal{K}$ on $[0,t]$). On the other hand, we define for any $x \in \mathcal{M}_i, t \geq 0$, the reachable set in the stratum $\mathcal{M}_i$ that is the set of all possible positions that can be reached, at time t, by an admissible trajectory solution of (1) associated to a feasible control $u \in \mathbb{A}(x)$ and lying in $\mathcal{M}_i$ on the time interval $[0,t]$:

$$\mathcal{R}_i(x;t) = \bigcup_{u\in\mathbb{A}(x)} \{y_{x,u}(t) \mid \ y_{x,u}(s) \in \mathcal{M}_i \ \forall s \in [0,t]\}.$$

Hence the controllability hypothesis can be stated as follows, for every $i \in \mathcal{I}$:

$$(H_3) \qquad \begin{cases} \text{If } \mathcal{A}_i \neq \emptyset, \quad \text{then } \exists \varepsilon_i, \Delta_i > 0 \text{ such that} \\ \mathcal{R}(x;t) \cap \overline{\mathcal{M}}_i \subseteq \displaystyle\bigcup_{s\in[0,\Delta_i t]} \mathcal{R}_i(x;t) \quad \forall x \in \mathcal{M}_i, \ \forall t \in [0,\varepsilon_i]. \end{cases}$$

This assumption is made in order to approximate trajectories that may switch between two or more strata infinitely many times on a short interval (this could happen if the set $\mathcal{A}_i$ is nonempty). Finally let us notice that the same assumption has been already used in [35, Hypothesis (H_4)] for similar purposes in the context of optimal syntheses analysis.

Remark 2.6. *Note that (H_3) is trivial if $\mathcal{M}_i$ is an open set or more generally if $\mathcal{M}_i$ is of maximal dimension among the strata that forms $\mathcal{K}$ (indeed, in this case there exists $\epsilon > 0$ such that for any $t \in [0,\epsilon]$, we have: $\mathcal{R}(x;t) \cap \overline{\mathcal{M}}_i = \mathcal{R}_i(x;t)$). The same remark holds whenever $\mathcal{A}_i = \mathcal{A}$. Moreover, (H_3) is straightforward if $\mathcal{M}_i$ is a single point (since in this case, if $\mathcal{A}_i \neq \emptyset$ then $\mathcal{R}(x;t) \cap \overline{\mathcal{M}}_i = \overline{\mathcal{M}}_i = \mathcal{R}_i(x;t)$).*

Let us also point out the fact that (H_3) can be satisfied under an easy criterion of full *controllability condition on manifolds*. The most classical assumption of controllability is the following: $\forall i \in \mathcal{I}$ with $\mathcal{A}_i \neq \emptyset$, it holds:

$$\exists r_i > 0 \text{ such that } \mathcal{T}_{\mathcal{M}_i}(x) \cap \mathbb{B}(0,r_i) \subseteq f(x,\mathcal{A}_i(x)) \ \forall x \in \mathcal{M}_i. \tag{5}$$

Indeed, this corresponds to the Petrov condition on manifolds. Hence, by adapting the classical arguments to this setting, we can see that (5) implies the Lipschitz regularity of the minimum time function of the controlled dynamics restricted to the manifold $\mathcal{M}_i$, and so (H_3) follows; see for instance [6, Chapter 4.1]. However, let us emphasis on that (5) is only a *sufficient* condition to satisfy assumption (H_3). Indeed, (H_3) is still satisfied in some cases where Petrov condition does not hold (for instance, the double-integrator system in Example 2.1 fulfills the requirement (H_3) and clearly does not satisfy the Petrov condition (5)).

2.2 Main results

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 HJB equation. The notion of solution that will be introduced here is based on the classical notion of supersolution and on a new subsolution notion in a stratified sense.

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

$$H(x,\zeta) = \max_{u\in\mathcal{A}} \{-\langle \zeta, f(x,u)\rangle - \ell(x,u)\}, \quad \forall x, \zeta \in \mathbb{R}^N.$$

For each index $i \in \mathcal{I}$ we define $H_i : \mathcal{M}_i \times \mathbb{R}^N \rightrightarrows \mathbb{R}$, the tangential Hamiltonian on $\mathcal{M}_i$, by

$$H_i(x,\zeta) = \max_{u\in\mathcal{A}_i(x)} \{-\langle \zeta, f(x,u)\rangle - \ell(x,u)\}, \quad \forall x \in \mathcal{M}_i, \ \forall \zeta \in \mathbb{R}^N.$$

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*

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

Definition 2.2. *Let $\varphi : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be l.s.c.. A vector $\zeta \in \mathbb{R}^N$ is called a viscosity subgradient of φ at $x \in \operatorname{dom} \varphi$ if and only if there exists a continuously differentiable function g so that*

$$\nabla g(x) = \zeta \text{ and } \varphi - g \text{ attains a local minimum at } x.$$

Furthermore, ζ is called a proximal subgradient of φ at x if for some $\sigma > 0$,

$$g(y) := \langle \zeta, y - x \rangle - \sigma |y - x|^2.$$

The set of all proximal subgradients at x is denoted by $\partial_P \varphi(x)$.

Hence, 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 $\vartheta(\cdot)$ of the problem (3) 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:*

$$\lambda \vartheta(x) + H(x, \zeta) \geq 0 \qquad \forall x \in \mathcal{K}, \; \forall \zeta \in \partial_P \vartheta(x), \tag{6}$$

$$\lambda \vartheta(x) + H_i(x, \zeta) \leq 0 \qquad \forall x \in \mathcal{M}_i, \; \forall \zeta \in \partial_P \vartheta_i(x), \; \forall i \in \mathcal{I}, \tag{7}$$

where $\vartheta_i(x) = \vartheta(x)$ if $x \in \overline{\mathcal{M}_i}$ and $\vartheta_i(x) = +\infty$ otherwise.

In the above theorem, the definition of v_i is not simply to make sense of relation (7), but also a compatibility condition on the (v_i)'s. This point is crucial to the comparison principle analysis.

Remark 2.7. *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 HJB equation proposed by Soner in [39] is included in the set of equations proposed in Theorem 2.1.*

Remark 2.8. *If for some $i \in \mathcal{I}$, $\mathcal{M}_i = \{\bar{x}\}$ and $\mathcal{A}_i(\bar{x}) \neq \emptyset$ (this is the case when for instance $\mathcal{K}$ is a network with $\bar{x}$ being one of the junctions), then $f(\bar{x}, u) = 0$ for any $u \in \mathcal{A}_i(\bar{x})$ and so $H_i(x, \zeta) = -\min\{\ell(\bar{x}, u) \mid \; u \in \mathcal{A}_i(\bar{x})\}$ for any $\zeta \in \mathbb{R}^N$. Hence, (7) for this index corresponds the following inequality:*

$$\lambda \vartheta(\bar{x}) \leq \min_{u \in \mathcal{A}_i(\bar{x})} \ell(\bar{x}, u).$$

A particular situation of interest is when $\mathcal{K}$ is a network as in Remark 2.2. Let $\mathcal{O}$ be the junction and $\mathcal{M}_1, \ldots, \mathcal{M}_p$ be its edges. For sake of simplicity assume that for each $i \in \{1, \ldots, p\}$, $\exists \mathcal{A}_i \subseteq \mathcal{A}$ s.t.

$$f(x, \mathcal{A}) \cap \mathcal{T}_{\mathcal{M}_i}(x) = f(x, \mathcal{A}_i), \quad \forall x \in \mathcal{M}_i. \tag{$H_3^\sharp$}$$

Corollary 2.1. *Suppose that (H_f), (H_ℓ) and (H_0) hold with $\mathcal{K}$ a network as before and $\lambda > \lambda_0$. Assume that ($H_3^\sharp$) holds and let*

$$\mathcal{A}_0 = \{u \in \mathcal{A} \mid \; f(\mathcal{O}, u) = 0\}.$$

Then the value function $\vartheta(\cdot)$ of the problem (3) 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:

$$\lambda \vartheta(x) + \max_{u \in \mathcal{A}_i} \{-\langle \zeta, f(x, u) \rangle - \ell(x, u)\} = 0 \quad \forall x \in \mathcal{M}_i, \; \forall \zeta \in \partial_P \vartheta(x),$$

$$\lambda \vartheta(\mathcal{O}) + H(\mathcal{O}, \zeta) \geq 0 \quad \forall \zeta \in \partial_P \vartheta(\mathcal{O}),$$

$$\lambda \vartheta(\mathcal{O}) - \min_{u \in \mathcal{A}_0} \ell(\mathcal{O}, u) \leq 0.$$

Proof. Let $\mathcal{M}_0 = \{\mathcal{O}\}$, then $\mathcal{M}_0, \ldots, \mathcal{M}_p$ is an admissible stratification for $\mathcal{K}$. Furthermore, since $\mathcal{A}_i(x) = \mathcal{A}_i$ whenever $x \in \mathcal{M}_i$ and $\mathcal{A}_0(\mathcal{O}) = \mathcal{A}_0$, ($H_2$) holds. By Remark 2.3, (H_3) also holds. So, the corollary follows from Theorem 2.1 and Remark 2.8. □

2.3 Discussion and comments

The main contribution of Theorem 2.1 relies on the characterization of the value function in situations where the set $\mathcal{K}$ is not necessarily the closure of its interior, and the value function is not necessarily Lipschitz continuous. As already mentioned in the introduction, several contributions have been devoted to the case where Inward Pointing conditions (IPC) are satisfied and the interior of $\mathcal{K}$ is not empty; see the pioneering works [39, 40], the more recent works [12, 13] and the references therein.

When the IPC is not satisfied, 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 [27]. However, in that paper the set $\mathbb{A}(x)$ is assumed nonempty everywhere on $\mathcal{K}$, requiring in particular that the viable set is whole the set $\mathcal{K}$. Moreover, the result in [27] assume some restrictive hypothesis on $\mathcal{K}$ and on the set-valued map $x \longmapsto \mathbb{A}(x)$.

Let us also mention the work in [15] 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 behavior may occur.

More recently, there is an increasingly interest in control problems in stratified domains, see [36, 7, 8, 37]. In those papers, the control problem is formulated in the whole space $\mathbb{R}^N$ with a given stratification, and under a strong controllability assumption that guarantees the continuity of the value function and provides an appropriate framework for analysing the transmission conditions. In the present manuscript, the stratification is used in a completely different way for characterizing the l.s.c value function of state-constrained control problems.

On another hand, several papers have been devoted to the case of control problems on networks [25, 2, 26], where the framework is also different from the one considered in Corrollary 2.1. Indeed, in the above cited papers, the dynamics is defined only on each branch and is not Lipschitz continuous in the whole network. Here, we prefer to focus on the state-constrained setting and the general result we obtain indicates that in the particular case of optimal control problems in networks, it is possible to avoid the controllability assumption usually considered in the literature at the junction points. The characterization of the value function could be then considered in the bilateral viscosity sense. Moreover, we believe that the same arguments developed in this paper can be extended to more general control problems on networks (with discontinuous dynamics and also in higher dimension than 1). This study will be developed further in an ongoing work.

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 [20, 10]. More precisely, Theorem 2.1 is equivalent to say that ϑ is the unique l.s.c. viscosity solution with λ_ℓ-superlinear growth of the HJB equation:

$$\lambda\vartheta(x) + H(x, D_x\vartheta(x)) \geq 0 \qquad \forall x \in \mathcal{K}, \tag{8a}$$

$$\lambda\vartheta(x) + H_i(x, D_x\vartheta(x)) \leq 0 \qquad \forall x \in \mathcal{M}_i, \ \forall i \in \mathcal{I}. \tag{8b}$$

Here the l.s.c viscosity solution of (8) has to be understood in the Bilateral sense given in the next definition.

Definition 2.3. *Let V be a l.s.c function defined on $\mathbb{R}^N$ with Dom $V \subset \mathcal{K}$.*

The function V is said a supersolution of (8a) if for every $\varphi \in C^1(\mathbb{R}^N)$ and for every $x_0 \in \mathcal{K}$ such that $V - \varphi$ achieved a local minimum at x_0 on $\mathcal{K}$, we have:

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

Moreover, V is said a subsolution of (8b) 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 $V - \varphi$ achieved a local minimum at x_0 on $\mathcal{M}_i$, we have:

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

Finally, V is said a l.s.c viscosity solution of (8) if it is supersolution of (8a) and subsolution of (8b).

3 Properties of the Value function.

In this section we study the principal properties of the value function of problem (3). In particular, we present an intermediate characterization of ϑ.

Proposition 3.1. *Suppose that (H_f) and (H_ℓ) hold and assume that $\lambda > \lambda_\ell c_f$. Then, the value function has λ_ℓ-superlinear growth on its domain.*

Proof. By Remark 2.1 the value function (3) satisfies the following inequality:

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

Therefore, since $\lambda > \lambda_\ell c_f$ we obtain the desired result. □

3.1 Existence of optimal controls.

The next proposition is a classical type of result in optimal control and states the existence of minimizer for the problem (3).

Proposition 3.2. *Suppose that (H_f), (H_ℓ) and (H_0) hold, and that $\lambda > \lambda_\ell c_f$. If $\vartheta(x) \in \mathbb{R}$ for some $x \in \mathcal{K}$ then there exists $u \in \mathbb{A}(x)$ a minimizer of (3).*

Proof. Let $x \in \mathcal{K}$ such that $\vartheta(x) \in \mathbb{R}$. This means that for every $n \geq 0$, there exists a control law $u_n \in \mathbb{A}(x)$ such that:

$$\lim_{n\to+\infty} \int_0^\infty e^{-\lambda t} \ell(y_n(t), u_n(t))\, dt = \vartheta(x), \tag{9}$$

where y_n is the solution to (1) with the initial condition $y_n(0) = x$.

Set $z_n(t) = \ell(y_n(t), u_n(t))$, then from Remark 2.1 we have:

$$|y_n(t)| \leq (1+|x|)e^{c_f t} \qquad \text{for any } t \geq 0, \tag{10}$$

$$|\dot{y}_n(t)| \leq c_f(1+|x|)e^{c_f t} \qquad \text{a.e. } t > 0, \tag{11}$$

$$|z_n(t)| \leq c_\ell(1+|x|)^{\lambda_\ell} e^{\lambda_\ell c_f t} \qquad \text{a.e. } t > 0. \tag{12}$$

Consider the measure $d\mu = e^{-\lambda t}dt$ and let $L^1 := L^1([0, +\infty); d\mu)$ be the Banach space of integrable functions on $[0, +\infty)$ for the measure $d\mu$. Consequently, we denote by $W^{1,1}$ the Sobolev space of functions L^1 which have their weak derivative also in L^1.

Let $\omega : [0, +\infty) \to \mathbb{R}$ be given by $\omega(t) := c_f(1+|x|)e^{c_f t}$ for any $t \geq 0$. By (H_ℓ), $\lambda > c_f$ because $\lambda_\ell \geq 1$. So, (11) implies that $\omega(\cdot)$ is a positive function in L^1 which dominates $|\dot{y}_n|$. Moreover, by (10) the sequence $\{y_n(t)\}$ is relatively compact for any $t \geq 0$, hence the hypothesis of theorem [4, Theorem 0.3.4] are satisfied and so, there exist a function $y \in W^{1,1}$ and a subsequence, still denoted by $\{y_n\}$, such that

$$y_n \text{ converges uniformly to } y \text{ on compact subsets of } [0,+\infty),$$
$$\dot{y}_n \text{ converges weakly to } \dot{y} \text{ in } L^1.$$

On the other hand, given that $\lambda > \lambda_\ell c_f$ and (12) holds, it is not difficult to see that $\{z_n\}$ is equi-integrable with respect to $d\mu$, then by the Dunford-Pettis Theorem there exist a function $z \in L^1$ and a subsequence, still denoted by z_n, such that z_n converges weakly to z in L^1.

Let $\Gamma(x) = G(0,x) \subseteq \mathbb{R}^N \times \mathbb{R}$ for every $x \in \mathcal{K}$. Hence, by (H_f) and (H_ℓ), Γ is locally Lipschitz with closed images and by (H_0) it has convex images. Then the Convergence Theorem [4, Theorem 1.4.1] implies that $(\dot{y}, z) \in \Gamma(y)$ for almost every $t \geq 0$. Thus, by the Measurable Selection Theorem (see [4, Theorem 1.14.1]), there exist two measurable functions $u : [0,+\infty) \to \mathcal{A}$ and $r : [0,+\infty) \to [0,+\infty)$ such that satisfies

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

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

$$\int_0^\infty e^{-\lambda t}\ell(y(t), u(t))dt \leq \int_0^\infty e^{-\lambda t} z(t)dt = \lim_{n\to+\infty} \int_0^\infty e^{-\lambda t} z_n(t)dt = \vartheta(x).$$

Therefore, u is a minimizer of 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 3.3. *Suppose that (H_f), (H_ℓ) and (H_0) hold, and that $\lambda > \lambda_\ell c_f$. Then the value function $x \mapsto \vartheta(x)$ is l.s.c. on $\mathcal{K}$. In particular, the viable set* $\operatorname{dom} \mathbb{A}$ *is closed.*

Proof. Let $\{x_n\} \subseteq \mathcal{K}$ be a sequence such that $x_n \to x$. Without lost of generality we assume that $|x_n| \leq |x| + 1$. We need to prove that

$$\liminf_{n\to+\infty} \vartheta(x_n) \geq \vartheta(x).$$

Suppose that there exists a subsequence, denoted equally, so that $\{x_n\} \subseteq \operatorname{dom} \vartheta$. Otherwise the inequality holds immediately. Then, by Proposition 3.2, for any $n \in \mathbb{N}$ there exists an optimal control $u_n \in \mathbb{A}(x_n)$. Let y_n the optimal trajectory associated with u_n and x_n. Notice that (10), (11) and (12) hold with x_n instead of x. Hence, since $|x_n|$ is uniformly bounded ($|x_n| \leq |x| + 1$) we can use the same technique as in Proposition 3.2 to find that there exists $u \in \mathbb{A}(x)$ such that

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

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

3.3 Dynamic Programming Principle (DPP)

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

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

The (DPP) includes two different increasing properties along admissible controlled arcs. Indeed, the two elementary inequalities that define it can be interpreted as a *weakly decreasing* and a *strongly increasing* principle, respectively.

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

$$(13) \qquad e^{-\lambda t} \varphi(y_{x,u}(t)) + \int_0^t e^{-\lambda s} \ell(y_{x,u}(s), u(s))ds \leq \varphi(x) \quad \forall t \geq 0.$$

ii) strongly increasing for the control system if $\operatorname{dom} \mathbb{A} \subseteq \operatorname{dom} \varphi$ and for any $x \in \mathcal{K}$ and $u \in \mathbb{A}(x)$ measurable we have

$$(14) \qquad e^{-\lambda t} \varphi(y_{x,u}(t)) + \int_0^t e^{-\lambda s} \ell(y_{x,u}(s), u(s))ds \geq \varphi(x) \quad \forall t \geq 0.$$

The following lemma is required to single out the value function among other l.s.c. functions.

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 $\vartheta(x) \leq \varphi(x)$ for all $x \in \mathcal{K}$.*
2. *strongly increasing for the control system then $\vartheta(x) \geq \varphi(x)$ 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 (1) such that $y(t) \in \operatorname{dom} \varphi$,

$$(15) \qquad \lim_{t \to +\infty} e^{-\lambda t} \varphi(y(t)) = 0.$$

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 the Monotone Convergence Theorem, (15) and the definition of the value function we obtain the desired inequality $\vartheta(x) \leq \varphi(x)$.

Case 2. Suppose φ is strongly increasing for the control system and let $x \in \mathcal{K}$. Assume that $\vartheta(x) \in \mathbb{R}$, otherwise the result is direct. Let $\bar{u} \in \mathbb{A}(x)$ be the optimal control associated with (3) and let $\bar{y}$ be the optimal trajectory associated with $\bar{u}$ and x. Then

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

Then by (15), letting $t \to +\infty$ we conclude the proof. $\square$

In view of the previous comparison lemma we can state an intermediate characterization of the value function in terms of the Definition 3.1.

Proposition 3.4. *The value function $\vartheta(\cdot)$ 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 $\vartheta(\cdot)$ satisfies (DPP). So, it is weakly decreasing and strongly decreasing for the control system. The uniqueness and the growth condition are consequences of Lemma 3.1 and Proposition 3.1. $\square$

4 Trajectories and Invariance on smooth manifolds.

Before entering into the details of the proof of the main theorem, we need to prove some fundamental results. The first proposition states the existence of smooth trajectories for a given initial data, namely, initial point and initial velocity.

Proposition 4.1. *Suppose that (H_f), (H_ℓ), (H_0), (H_1) and (H_2) hold. Then, for any $i \in \mathcal{I}$ such that $\mathcal{A}_i$ has nonempty images, for every $x \in \mathcal{M}_i$ and any $u_x \in \mathcal{A}_i(x)$ there exist $\varepsilon > 0$, a measurable control map $u : (-\varepsilon, \varepsilon) \to \mathcal{A}$, a measurable function $r : (-\varepsilon, \varepsilon) \to [0, +\infty)$ and a continuously differentiable arc $y : (-\varepsilon, \varepsilon) \to \mathcal{M}_i$ with $y(0) = x$ and $\dot{y}(0) = f(x, u_x)$, such that*

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

Proof. Let $R > 0$ and set $\mathcal{M}_i^R = \mathcal{M}_i \cap \mathbb{B}(x, R)$. Consider the set valued map $\Gamma_i : \mathcal{M}_i^R \times (-1, 1) \to \mathbb{R}^N \rightrightarrows \mathbb{R}$ given by

$$\Gamma_i(y, t) = \left\{ \begin{pmatrix} f(y, u) \\ e^{-\lambda t}\ell(y, u) + r \end{pmatrix} \;\middle|\; \begin{array}{c} u \in \mathcal{A}_i(y), \\ 0 \le r \le \beta(y, u) \end{array} \right\}, \quad \forall (y, t) \in \mathcal{M}_i^R \times (-1, 1).$$

Note that by the definition of $\mathcal{A}_i$ and thanks to (H_f) and (H_ℓ), Γ_i has closed images and since $\mathcal{A}_i$ has nonempty images, Γ_i has nonempty images as well. The definition of $\mathcal{A}_i$ and (H_0) imply that it also has convex images.

Besides, by (H_2), Γ_i is Lipschitz on $\mathcal{M}_i^R \times (-1, 1)$, so it admits a Lipschitz selection, $g_i : \mathcal{M}_i^R \times (-1, 1) \to \mathbb{R}^N \times \mathbb{R}$ such that $g_i(x, 0) = (f(x, u_x), \ell(x, u_x))$; see [5, Theorem 9.4.3] and the subsequent remark. Notice also that

$$g(y, t) \in f(y, \mathcal{A}_i(y)) \times \mathbb{R} \subseteq \mathcal{T}_{\mathcal{M}_i}(y) \times \mathbb{R}, \quad \forall (y, t) \in \mathcal{M}_i^R \times (-1, 1).$$

Hence, by the Nagumo theorem (see for instance [4, Theorem 4.2.2]) and the Lipschitz continuity of g_i, there exists $\varepsilon > 0$ such that the differential equation

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

admits a unique solution which is continuously differentiable on $(-\varepsilon, \varepsilon)$ such that $y(t) \in \mathcal{M}_i$ for every $t \in (-\varepsilon, \varepsilon)$, $\dot{y}(0) = f(x, u_x)$ and $\dot{z}(0) = \ell(x, u_x)$.

On the other hand, since $\Gamma_i(y, t) \subseteq G(t, y)$ for any $(t, y) \in (-1, 1) \times \mathcal{M}_i^R$, by the Measurable Selection Theorem (see [4, Theorem 1.14.1]), there exist a measurable control $u : (-\varepsilon, \varepsilon) \to \mathcal{A}$ and a measurable function $r : (-\varepsilon, \varepsilon) \to [0, +\infty)$ such that

$$(\dot{y}, \dot{z}) = (f(y, u), e^{-\lambda t}\ell(y, u) + r), \quad \text{a.e. on } (-\varepsilon, \varepsilon).$$

Finally, since $y(t) \in \mathcal{M}_i$, we have that $u \in \mathcal{A}_i(y)$ a.e. on $(-\varepsilon, \varepsilon)$, and so the conclusion follows, because

$$z(t) = \int_0^t \left(e^{-\lambda s}\ell(y(s), u(s)) + r(s)\right) ds, \qquad \forall t \in (-\varepsilon, \varepsilon).$$

$\square$

We now present a useful criterion for strong invariance adapted to smooth manifolds. This proposition is similar in spirit to Theorem 4.1 in [9].

Proposition 4.2. *Suppose $\mathfrak{M} \subseteq \mathbb{R}^k$ is locally closed, $\mathcal{S} \subseteq \mathbb{R}^k$ is closed with $\mathcal{S} \cap \overline{\mathfrak{M}} \neq \emptyset$ and $\Gamma : \overline{\mathfrak{M}} \rightrightarrows \mathbb{R}^k$ is locally Lipschitz and locally bounded.*

Let $R > 0$ and set $\mathfrak{M}^R = \mathfrak{M} \cap \mathbb{B}(0, R)$. Assume that there exists $\kappa = \kappa(R) > 0$ such that

$$\sup_{v \in \Gamma(x)} \langle x - s, v \rangle \leq \kappa \, dist_{\mathcal{S} \cap \overline{\mathfrak{M}}}(x)^2 \qquad \forall x \in \mathfrak{M}^R, \ \forall s \in proj_{\mathcal{S} \cap \overline{\mathfrak{M}}}(x). \tag{16}$$

Then for any absolutely continuous arc $\gamma : [0, T] \to \overline{\mathfrak{M}}$ that satisfies

$$\dot{\gamma} \in \Gamma(\gamma) \quad \text{a.e. on } [0, T] \qquad \text{and} \qquad \gamma(t) \in \mathfrak{M}^R \qquad \forall t \in (0, T),$$

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 $\tilde{R} > 0$ so that $\gamma([0, T]) \subseteq \mathbb{B}(0, \tilde{R})$. We denote by c_Γ and L_Γ the corresponding bound for the velocities of Γ and the Lipschitz constant of Γ on $\overline{\mathfrak{M}} \cap \mathbb{B}(0, \tilde{R})$. We take $C_1 > 0$ such that

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

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 $t_{i+1} \in (t_i, T]$ 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 $s_i \in \mathrm{proj}_{\mathcal{S} \cap \overline{\mathfrak{M}}}(\gamma_i)$ arbitrary. Suppose first that $\gamma(0) \in \mathfrak{M}$. We will show the inequality only for $t = T$. For $t \in (0, T)$ the proof is similar.

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

$$\omega(1) - \omega(0) = \gamma_{i+1} - \gamma_i = (t_{i+1} - t_i) \int_0^1 \dot{\gamma}((1 - s)t_i + st_{i+1}) ds$$

On the other hand, since Γ is locally Lipschitz

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

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

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

Hence

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

where this last comes from (16), the definition of b_i and the choice of t_i.

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

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

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 $\sigma_{n+1} = \mathrm{dist}_{\mathcal{S}\cap\overline{\mathfrak{M}}}(\gamma(T))$ and $\sigma_0 = \mathrm{dist}_{\mathcal{S}\cap\overline{\mathfrak{M}}}(\gamma(0))$, 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. $\square$

5 Proof of the main result.

In this section we present a characterization of the value function as the unique solution to a bilateral HJB 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 on $\mathcal{K}$. Secondly, we show that if a function is strongly increasing for the control system then it is a subsolution on each stratum $\mathcal{M}_i$. The final and most technical step consists in characterizing the strongly increasing principle in terms of HJB inequalities on each stratum of $\mathcal{K}$.

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

5.1 Normal Cones and Proximal subgradients.

For sake of the exposition, we recall the definition of the Proximal normal cone and its relation with the proximal subgradient of Definition 2.2. For a further discussion about this topic we refer the reader to [18].

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 $\mathcal{N}_{\mathcal{S}}^P(x)$. If $\mathcal{S} = \operatorname{epi}\varphi$ where $\varphi : \mathbb{R}^k \to \mathbb{R}\cup\{+\infty\}$ is a l.s.c. function, then for every $x \in \operatorname{dom}\varphi$, the following relation is valid:

$$\partial_P\varphi(x) \times \{-1\} \subseteq \mathcal{N}^P_{\operatorname{epi}\varphi}(x, \varphi(x)), \quad \forall x \in \operatorname{dom}\varphi.$$

By definition of the proximal subdifferential, $\zeta \in \partial_P\varphi(x)$ if and only if there exist $\sigma, \delta > 0$ such that

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

This inequality is called the *proximal subgradient inequality.*

5.2 Weakly decreasing principle.

This section aims at proving that the weakly decreasing principle (13) is equivalent to the *super-solution* of an appropriate HJB equation. The idea of the proof uses very classical arguments and requires only standing assumptions of control theory; see [18, Chapter 4]. Nevertheless, since the statement in the same framework as considered in this paper seems not to be available in the literature, we provide here an outline of 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*

$$\lambda\varphi(x) + H(x, \zeta) \geq 0 \qquad \forall x \in \mathcal{K},\ \forall \zeta \in \partial_P \varphi(x) \tag{17}$$

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

Suppose φ is weakly decreasing for the control system. Let $x \in \mathcal{K}$, if $\partial_P\varphi(x) = \emptyset$ then (17) holds by vacuity. If on the contrary, there exists $\zeta \in \partial_P\varphi(x)$, then $x \in \operatorname{dom}\varphi$ and there exists $u \in \mathbb{A}(x)$ such that (13) holds. Let us denote $y(\cdot)$ the trajectory of (1) associated with the control u and x. 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 (13) we get for any t small enough

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

Since f and ℓ are locally Lipschitz we get

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

where $h(t)$ is such that $\lim_{t\to 0^+} h(t) = 0$. Therefore taking infimum over $u \in \mathcal{A}$ inside the integral and letting $t \to 0^+$ we get (17) after some algebraic steps.

Now, we turn to the second part of the proof ($\Leftarrow$). Let $\mathcal{O} \subseteq \mathbb{R}^{N+1}$ be the neighborhood of $[0, +\infty) \times \mathcal{K}$ given by (H_0) which we assume is open. Consider $\psi : [0, +\infty) \times \mathcal{K} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ defined as

$$\psi(\tau, x, z) = \begin{cases} e^{-\lambda\tau}\varphi(x) + z & \text{if } x \in \mathcal{K}, \\ +\infty & \text{otherwise,} \end{cases} \qquad \forall (\tau, x, z) \in [0, +\infty) \times \mathcal{K} \times \mathbb{R},$$

and $\Gamma : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R} \times \mathbb{R} \rightrightarrows \mathbb{R} \times \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}$ given by

$$\Gamma(\tau, x, z, w) = \{1\} \times G(\tau, x) \times \{0\}, \quad \forall (\tau, x, z, w) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}.$$

To prove that φ is weakly decreasing for the control system let us first show that for any $\gamma_0 \in \operatorname{epi}\psi$, there exists an absolutely continuous arc $\gamma : [0, T) \to \mathcal{O} \times \mathbb{R}^2$ that satisfies

$$\dot\gamma \in \Gamma(\gamma) \quad \text{a.e. on } [0, T) \quad \text{and} \quad \gamma(0) = \gamma_0 \tag{18}$$

such that $\gamma(t) \in \operatorname{epi}\psi$ for every $t \in [0, T)$, or in term of [43, Definition 3.1], $(\Gamma, \operatorname{epi}\psi)$ is weakly invariant on $\mathcal{O} \times \mathbb{R}^2$. We seek to apply [43, Theorem 3.1(a)].

Note that $\operatorname{epi}\psi$ is closed because φ is l.s.c. and Γ has nonempty convex compact images on $\mathcal{O} \times \mathbb{R}^2$ because of (H_0). Moreover, by (H_f) and (H_ℓ), Γ has closed graph 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, to prove the weak invariance of $(\Gamma, \operatorname{epi}\psi)$ we only need to show that, for $\mathcal{S} = \operatorname{epi}\psi$, (17) implies

$$\min_{v\in\Gamma(\chi)} \langle \eta, v\rangle \le 0 \qquad \forall \chi \in \mathcal{S}\cap U,\ \forall \eta \in \mathcal{N}^P_{\mathcal{S}}(\chi). \tag{19}$$

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 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

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

Hence, by (17) we get $\min\{\langle \eta, v\rangle \mid\ v \in \Gamma(\tau, x, z, w)\} \le 0$.

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 [38]), there exist some sequences $\{(\tau_n, x_n, z_n)\} \subseteq \operatorname{dom}\psi$, $\{(\xi_n)\} \subseteq \mathbb{R}^{N+2}$ and $\{p_n\} \subseteq (0,\infty)$ such that

$$\begin{aligned}(\tau_n, x_n, z_n) \to (\tau, x, z), &\qquad \psi(\tau_n, x_n, z_n) \to \psi(\tau, x, z),\\ (\xi_n, p_n) \to (\xi, 0), &\qquad \frac{1}{p_n}\xi_n \in \partial_P\psi(\tau_n, x_n, z_n).\end{aligned}$$

Thus, using the same argument as above we can show

$$\min\{\langle(\xi_n, -p_n), v\rangle \mid v \in \Gamma(\tau_n, x_n, z_n, \psi(\tau_n, x_n, z_n))\} \le 0.$$

Hence, 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 (19).

So, by [43, Theorem 3.1(a)], for every $\gamma_0 = (\tau_0, x_0, z_0, w_0) \in \mathcal{S}\cap\mathcal{O}\times\mathbb{R}^2$ there exists an absolutely continuous arc $\gamma(t) = (\tau(t), y(t), z(t), w(t))$ which lies in $\mathcal{O}\times\mathbb{R}^2$ for a maximal period of time $[0, T)$ so that (18) holds and

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

By the Measurable Selection Theorem (see [4, Theorem 1.14.1]), $y(\cdot)$ is a solution of (1) for some $u : [0, T) \to \mathcal{A}$. Also, $y(t) \in \operatorname{dom}\varphi \subseteq \mathcal{K}$, $\forall t \in [0, T)$.

Moreover, since $w(t) = w_0$ and $\tau(t) = \tau_0 + t$

$$z(t) = \int_0^t [e^{-\lambda(\tau_0+s)}\ell(y(s), u(s)) + r(s)]ds, \qquad \text{with } r(s) \ge 0 \text{ a.e.}$$

Notice that $\gamma_0 = (0, x, 0, \varphi(x)) \in \operatorname{epi}\psi$ for any $x \in \operatorname{dom}\varphi$, so to conclude the proof we just need to show that $T = +\infty$. By contradiction, suppose $T < +\infty$, then $(\tau(t), y(t)) \to \operatorname{bdry}\mathcal{O}$ as $t \to T^-$. Nevertheless, since $\mathcal{O}$ is a neighborhood of $[0, +\infty)\times\mathcal{K}$ and $\tau(t) = t$ and $y(t) \in \mathcal{K}$ for any $t \in [0, T)$ this is not possible. Therefore, the conclusion follows.

□

5.3 Strongly increasing principle.

Now we show that satisfying inequality (14) is equivalent to be a *subsolution* of the HJB equation on each stratum. We first prove the necessary part.

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. Suppose that* φ *is strongly increasing for the control system, then*

$$\lambda\varphi(x) + H_i(x, \zeta) \leq 0 \qquad \forall x \in \mathcal{M}_i,\ \forall \zeta \in \partial_P \varphi_i(x), \tag{20}$$

where $\varphi_i(x) = \varphi(x)$ *if* $x \in \overline{\mathcal{M}}_i$ *and* $\varphi_i(x) = +\infty$ *otherwise.*

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 (20) for any $(i, x) \in \mathcal{I} \times \mathcal{K}$ 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

$$-\lambda\varphi(x) + \langle \zeta, f(x, u_x)\rangle + \ell(x, u_x) \geq 0, \qquad \forall \zeta \in \partial_P \varphi_i(x). \tag{21}$$

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

$$u_\tau(t) := u(t - \tau)\mathbb{1}_{[0,\tau]}(t) + \bar{u}(t - \tau)\mathbb{1}_{(\tau,+\infty)}(t) \quad \text{ for a.e. } t \in [0, +\infty).$$

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 $t \in [0, \tau]$.

Moreover, $u_\tau \in \mathbb{A}(y(-\tau))$, so since φ is strongly increasing

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

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

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

Hence,

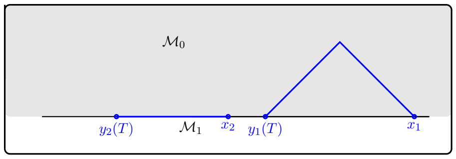
$$\frac{e^{-\lambda\tau} - 1}{\tau}\varphi(x) + \frac{e^{-\lambda\tau}}{\tau}\int_{-\tau}^0 \left(e^{-\lambda s}\ell(y(s), u(s)) + r(s)\right) S ds + \left\langle \zeta, \frac{x - y(-\tau)}{\tau}\right\rangle \geq h(\tau),$$

with $\lim_{\tau \to 0^+} h(\tau) = 0$. Therefore, by Proposition 4.1, passing to the limit in the last inequality we obtain (21) and so (20) follows. ∎

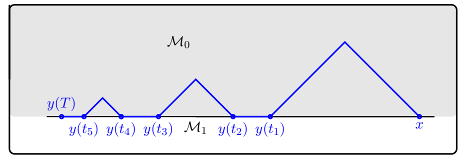
5.4 Characterization of Strongly increasing principle

In this section we prove the converse of Proposition 5.2 under the controllability assumption (H_3). The proof consists in analyze three different types of trajectories defined on a finite interval of time $[0, T]$. The first case corresponds to trajectories that dwell on a single manifold but whose extremal points may not do so, as for instance in Figure 2a. This case is treated independently in Lemma 5.1. The second type is studied in Step 1 of the proof of Proposition 5.3, these trajectories have the characteristic that can be decomposed into a finite number of first type trajectories; see an example in Figure 2b.

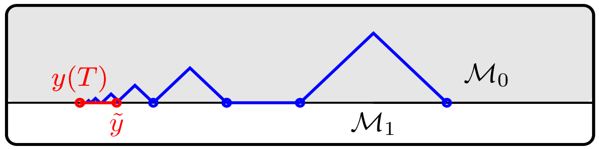
The third and more delicate type of trajectories to treat are those one that switch from one stratum to another infinitely many times in a finite interval as in Figure 2c. The hypothesis (H_3) is made to handle these trajectories. It allows to construct an approximate trajectory of type 2, as in Figure 2c, whose the corresponding cost is almost the same.



(a) Extremal switching times.



(b) Finite switching times.



(c) Chattering trajectory and its approximation.

Figure 2: Situation to be considered.

Lemma 5.1. *Suppose that* (H_0), (H_1) *and* (H_2) *hold in addition of* (H_f) *and* (H_ℓ). *Let* $\varphi : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ *be a l.s.c. function. Assume that* (20) *holds. Then for any* $x \in \mathcal{K}$, $u \in \mathbb{A}(x)$ *and any* $0 \leq a < b < +\infty$, *if* $y(t) := y_{x,u}(t) \in \mathcal{M}_i$ *for every* $t \in (a,b)$ *with* $i \in \mathcal{I}$, *we have*

$$\varphi(y(a)) \leq e^{-\lambda(b-a)}\varphi(y(b)) + e^{\lambda a}\int_a^b e^{-\lambda s}\ell(y,u)ds. \tag{22}$$

Proof. First of all we consider a backward augmented dynamic defined for any $(\tau, x) \in \mathbb{R} \times \mathcal{M}_i$ as follows:

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

Thanks to (H_0) and the definition of $\mathcal{A}_i(\cdot)$, the mapping G_i has convex compact images and by the statement of the proposition, G_i has nonempty images as well. Additionally, G_i is locally Lipschitz by (H_2).

Since $y = y_{x,u} \in \mathcal{M}_i$ on (a,b), then $\mathcal{A}_i$ has nonempty images we set $\mathfrak{M}_i = \mathbb{R} \times \mathcal{M}_i \times \mathbb{R}^2$ and define $\Gamma_i : \mathfrak{M}_i \rightrightarrows \mathbb{R}^{N+3}$ as

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

Note that $\mathfrak{M}_i$ is an embedded manifold of $\mathbb{R}^{N+3}$ and Γ_i satisfies the same assumptions than G_i with nonempty images. Consider the closed set $\mathcal{S}_i = \text{epi}(\psi_i)$ where

$$\psi_i(\tau, x, z) = \begin{cases} e^{-\lambda\tau}\varphi_i(x) + z & \text{if } x \in \overline{\mathcal{M}_i}, \\ +\infty & \text{otherwise}, \end{cases} \quad \forall (\tau, x, z) \in [0, +\infty) \times \overline{\mathcal{M}_i} \times \mathbb{R}.$$

Then if (20) holds, the following also holds

$$\sup_{v \in \Gamma_i(\tau,x,z,w)} \langle \eta, v \rangle \leq 0 \qquad \forall (\tau, x, z, w) \in \mathcal{S}_i, \; \forall \eta \in \mathcal{N}^P_{\mathcal{S}_i}(\tau, x, z, w). \tag{23}$$

Indeed, if $\mathcal{S}_i = \emptyset$ it holds by vacuity. Otherwise, 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. Recall that $\Gamma_i(\tau, x, z, w) \neq \emptyset$ because $\mathcal{A}_i(x) \neq \emptyset$. Consider $p > 0$, then, by the same arguments used in 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 $\zeta \in \partial_P \varphi_i(x)$

$$\begin{aligned} \langle (\xi, -p), v \rangle &= pe^{-\lambda\tau}(\lambda\varphi_i(x) - \langle \zeta, f(x,u) \rangle - \ell(x,u) - r) \\ &\leq pe^{-\lambda\tau}(\lambda\varphi_i(x) - \langle \zeta, f(x,u) \rangle - \ell(x,u)) \\ &\leq pe^{-\lambda\tau}(\lambda\varphi_i(x) + H_i(x, \zeta)). \end{aligned}$$

Since $\varphi_i(x) = \varphi(x)$, (20) holds and $v \in \Gamma_i(\tau, x, z, w)$ is arbitrary, we can take supremum over v to obtain the desired inequality (23). Similarly as done for Proposition 5.1, if $p = 0$ we use the Rockafellar Horizontal Theorem and the continuity of Γ_i to obtain (23) for any η.

Let $R > \tilde{R} > 0$ large enough so that $y_{x,u}([a,b]) \subseteq \mathbb{B}(0, \tilde{R})$ and

$$\sup_{X \in \mathfrak{M} \cap \mathbb{B}(0,\tilde{R})} |\mathrm{proj}_{\overline{\mathfrak{M}}_i \cap \mathcal{S}_i}(X)| < R.$$

Let L_i be the Lipschitz constant for Γ_i on $\mathfrak{M}_i \cap \mathbb{B}(0, R)$, so (23) implies (16) with $\kappa = L_i$. In particular, by Proposition 4.2 we have that for any absolutely continuous arc $\gamma : [a,b] \to \overline{\mathfrak{M}}_i$ which satisfies (18) (with Γ_i instead of Γ) and $\gamma(t) \in \mathfrak{M}_i$ for any $t \in (a,b)$,

$$\mathrm{dist}_{\mathcal{S} \cap \overline{\mathfrak{M}}_i}(\gamma(t)) \leq e^{L_i t} \mathrm{dist}_{\mathcal{S} \cap \overline{\mathfrak{M}}_i}(\gamma(a)) \qquad \forall t \in [a,b]. \tag{24}$$

Finally, consider the absolutely continuous arc defined on $[a,b]$ by

$$\gamma(t) = \left(a - t, y(a + b - t), -\int_a^t e^{\lambda(s-a)} \ell(y(a+b-s), u_l(a+b-s))ds, \varphi(b) \right).$$

Since $\dot{\gamma} \in \Gamma_i(\gamma)$ a.e. on $[a,b]$, $\gamma(t) \in \mathfrak{M}_i$ for any $t \in (a,b)$ and $\gamma(a) \in \mathcal{S}_i$ we get that $\gamma(b) \in \mathcal{S}_i$ which implies (22) after some algebraic steps. □

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 with $\mathrm{dom}\, \mathbb{A} \subseteq \mathrm{dom}\, \varphi$. If (20) holds, then φ is strongly increasing for the controlled system.*

Proof. Let $x \in \mathrm{dom}\, \varphi$ and $u \in \mathbb{A}(x)$. We want to show that inequality (14) holds for $y = y_{x,u}$. For this purpose we fix $T > 0$ and we set $\mathcal{I}_T(y) = \{i \in \mathcal{I} : \exists t \in [0,T],\ y(t) \in \mathcal{M}_i\}$. Note that $\mathcal{I}_T(y)$ is finite because the stratification is locally finite and so

$$[0,T] = \bigcup_{i \in \mathcal{I}_T(y)} J_i(y), \quad \text{with } J_i(y) := \{t \in [0,T] \mid\ y(t) \in \mathcal{M}_i\}.$$

We split 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 intervals, 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 $i_l \in \mathcal{I}_T(y)$ satisfying $(t_l, t_{l+1}) \subseteq J_{i_l}(y)$. Therefore, for any $l \in \{0, \ldots, n\}$ such that $t_l < t_{l+1}$ by Lemma 5.1 we have

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

Hence, using inductively the previous estimation and noticing that $t_0 = 0$ and $t_{n+1} = T$ we get exactly (14), 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 $y(t) \in \mathcal{M}$ for every $t \in (a,b)$, where $0 \leq a < b \leq T$ then (22) holds.*

By Lemma 5.1, 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 (which may be disconnected) 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: Suppose that $y([a, b]) \subseteq \mathcal{M} \cup \mathcal{M}_i$. Without loss of generality we can assume that $y(a), y(b) \in \mathcal{M}_i$. Therefore, $J := [a, b] \setminus J_i(y)$ is open and so, for any $\varepsilon > 0$ there exists a partition of $[a, b]$

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

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

$$\varphi(y(a_l)) \leq e^{-\lambda(b_l - a_l)}\varphi(y(b_l)) + e^{\lambda a_l}\int_{a_l}^{b_l} e^{-\lambda s}\ell(y, u)ds. \tag{25}$$

Notice also that

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

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

We now prove that there exists $L > 0$ so that

$$\varphi(y(b_l)) \leq e^{\lambda\varepsilon_l}\left(e^{-\lambda(a_{l+1} - b_l)}\varphi(y(a_{l+1})) + e^{\lambda b_l}\int_{b_l}^{a_{l+1}} e^{-\lambda s}\ell(y, u)ds\right) + L\varepsilon_l. \tag{26}$$

On the other hand, there exists a countable family of intervals $(\alpha_p, \beta_p) \subseteq [b_l, a_{l+1}]$ (not necessarily pairwise different) such that $\varepsilon_l = \sum_{p\in\mathbb{N}}(\beta_p - \alpha_p)$, $y(t) \in \mathcal{M}$ for any $t \in (\alpha_p, \beta_p)$ and $y(\alpha_p), y(\beta_p) \in \mathcal{M}_i$. If the number of intervals turns out to be finite, then (26) follows by the same arguments as in Step 1. So we assume that $\{(\alpha_p, \beta_p)\}_{p\in\mathbb{N}}$ is an infinite family of pairwise disjoint intervals.

Since ε is arbitrary, we can assume that it is small enough such that $\varepsilon_l < \varepsilon_i$ where ε_i is given by (H_3). So, for any $p \in \mathbb{N}$, there exists $u_p : [0, +\infty) \to \mathcal{A}$ measurable and $\delta_p > \alpha_p - \beta_p$ such that

$$y_p(t) \in \mathcal{M}_i,\ \forall t \in [\alpha_p, \beta_p + \delta_p], \quad y_p(\alpha_p) = y(\alpha_p), \quad \text{and} \quad y_p(\beta_p + \delta_p) = y(\beta_p)$$

where y_p is the solution to (1) associated with u_p. Furthermore, there exists $\Delta_i > 0$ such that $\delta_p < (1 - \Delta_i)(\beta_p - \alpha_p)$.

Let $J_i^l := [b_l, a_{l+1}] \cap J_i(y)$ and the measurable function $\omega : [b_l, a_{l+1}] \to \mathbb{R}$

$$\omega(t) = \mathbb{1}_{J_i^l}(t) + \sum_{p\in\mathbb{N}} \frac{\beta_p - \alpha_p + \delta_p}{\beta_p - \alpha_p}\mathbb{1}_{(\alpha_p, \beta_p)}(t) > 0, \quad \forall t \in [b_l, a_{l+1}].$$

Define $\nu(t) = b_l + \int_{b_l}^{t} \omega(s)ds$ for every $t \in [b_l, a_{l+1}]$. Note that it is absolutely continuous, strictly increasing and bounded from above by $c_{l+1} := \nu(a_{l+1})$ on $[b_l, a_{l+1}]$, so it is an homeomorphism from $[b_l, a_{l+1}]$ into $[b_l, c_{l+1}]$.

Let $\tilde{u} : [b_l, c_{l+1}] \to \mathcal{A}$ measurable defined as

$$\tilde{u} = u(\nu^{-1})\mathbb{1}_{J_i^l}(\nu^{-1}) + \sum_{p\in\mathbb{N}} u_p \mathbb{1}_{(\alpha_p,\beta_p)}(\nu^{-1}), \quad \text{a.e. on } [b_l, c_{l+1}],$$

and let $\tilde{y}$ be trajectory of (1) associated with u_p such that $\tilde{y}(b_l) = y(b_l)$. Note that by construction $\tilde{y}(\nu(t)) = y(t)$ for any $t \in J_i^l$ and $\tilde{y}(t) \in \mathcal{M}_i$ for any $t \in [b_l, c_{l+1}]$. Hence by Lemma 5.1

$$\varphi(y(b_l)) \le e^{-\lambda(c_{l+1}-b_l)}\varphi(y(a_{l+1})) + e^{\lambda b_l}\int_{b_l}^{c_{l+1}} e^{-\lambda s}\ell(\tilde{y}(s), \tilde{u}(s))ds. \tag{27}$$

By the Change of Variable Theorem for absolutely continuous function (see for instance [29, Theorem 3.54]) we get

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

Furthermore, $\ell(\tilde{y}(\nu), \tilde{u}(\nu))\nu' = \ell(y, u)$ a.e. on J_i^l and by Remark 2.1

$$\ell(\tilde{y}(\tau), \tilde{u}(\tau))\nu' \le L := \max\{1, \Delta\}c_\ell(1 + |x|)^{\lambda_\ell} e^{\lambda_\ell c_f(T+\Delta\varepsilon_l)} \quad \text{on } [b_l, a_{l+1}].$$

On the other hand, since $\ell \ge 0$ we get

$$\int_{b_l}^{c_{l+1}} e^{-\lambda s}\ell(\tilde{y}(s), \tilde{u}(s))ds \le \int_{b_l}^{a_{l+1}} e^{-\lambda\nu(s)}\ell(y, u)ds + L\varepsilon_l, \tag{28}$$

and we finally get (26) from (27) and (28) since

$$\nu(t) \ge b_l + \text{meas}(J_i^l \cap [b_l, t]) = t - \text{meas}([b_l, t] \cap J_l) \ge t - \varepsilon_l, \quad \forall t \in [b_l, a_{l+1}].$$

By (25) and (26) the following also holds

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

Therefore, by using an inductive argument we can prove that

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

and using (27) on the interval $[b_n, a_{n+1}]$ we get

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

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

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

Thus, letting $\varepsilon \to 0$ we obtain the induction hypothesis for $\kappa + 1$.

Case 2: We consider the case $y(a) \notin \mathcal{M} \cup \mathcal{M}_i$ or $y(b) \notin \mathcal{M} \cup \mathcal{M}_i$.

Suppose first that $y(a) \notin \overline{\mathcal{M}}_i \setminus \mathcal{M}_i$ and $y(b) \notin \overline{\mathcal{M}}_i \setminus \mathcal{M}_i$, then there exists $\delta > 0$ such that $y(t) \in \mathcal{M} \cup \mathcal{M}_i$ for every $t \in [a+\delta, b-\delta]$ and $\mathrm{dist}_{\overline{\mathcal{M}}_i \setminus \mathcal{M}_i}(y(t)) > 0$ for every $t \in [a, a+\delta] \cup [b-\delta, b]$. So, we can partitionate $[0, T]$ into three parts $[a, a+\delta]$, $[a+\delta, b-\delta]$ and $[b-\delta, b]$. In view of Case 2 and the inductive hypothesis, (22) holds in each of the previous intervals. So, gathering the three inequalities we get the induction hypothesis for $\kappa + 1$.

Secondly, suppose that only $y(a) \notin \mathcal{M} \cup \mathcal{M}_i$, then there exists a sequence $\{a_n\} \subseteq (a, b)$ such that $a_n \to a$ and $y([a_n, b]) \subseteq \mathcal{M} \setminus \mathcal{M}_i$. So, by Case 1,

$$\varphi(y(a_n)) \leq e^{-\lambda(b-a_n)}\varphi(y(b)) + e^{\lambda a_n} \int_{a_n}^{b} e^{-\lambda s} \ell(y_{x,u}, u) ds.$$

Furthermore, since φ is l.s.c. and $y(\cdot)$ is continuous we can pass to the limit to get (22), so the result also holds in this situation.

Finally, it only remains the situation $y(b) \in \overline{\mathcal{M}}_i \setminus \mathcal{M}_i$. Similarly as above, there exists a sequence $\{b_n\} \subseteq (a, b)$ such that $b_n \to b$ and $y([a, b_n]) \subseteq \mathcal{M} \setminus \mathcal{M}_i$ such that

$$\varphi(y(a)) \leq e^{-\lambda(b_n-a)}\varphi(y(b_n)) + e^{\lambda a} \int_{a}^{b_n} e^{-\lambda s} \ell(y_{x,u}, u) ds.$$

By (H_3), for $n \in \mathbb{N}$ large enough, there exists a control $u_n : (b_n, b+\delta_n) \to \mathcal{A}$ and a trajectory $y_n : [b_n, b+\delta_n] \to \overline{\mathcal{M}}_i$ with $y_n(b_n) = y(b_n)$, $y_n(b+\delta_n) = y(b)$ and $y_n(t) \in \mathcal{M}_i$ for any $t \in [b_n, b+\delta_n)$. By Lemma 5.1

$$\varphi(y(b_n)) \leq e^{-\lambda(b-b_n)}\varphi(y(b)) + \varepsilon_n,$$

with $\varepsilon_n \to 0$ as $n \to +\infty$, then gathering both inequalities and letting $n \to +\infty$ we get the induction hypothesis and so the proof is complete.

□

References

[1] Y. Achdou, F. Camilli, A. Cutrì and N. Tchou, Hamilton-Jacobi equations on networks. *Nonlinear Differential Equations and Applications*, 20(3):413–445, 2013.

[2] Y. Achdou, S. Oudet and N. Tchou, Hamilton-Jacobi equations for optimal control on junctions and networks. Preprint submitted, 2014.

[3] A. Altarovici, O. Bokanowski, and H. Zidani. A general Hamilton-Jacobi framework for nonlinear state-constrained control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(2): 337–357, 2013.

[4] J.-P. Aubin and A. Cellina, *Differential inclusions*, vol. 264 of Comprehensive studies in mathematics, Springer, Berlin, Heidelberg, New York, Tokyo, 1984.

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

[7] G. Barles, A. Briani and E. Chasseigne, A Bellman approach for two-domains optimal control problems in $\mathbb{R}^N$, *ESAIM: Control, Optimisation and Calculus of Variations*, 19(03):710–739, 2013.

[8] G. Barles, A. Briani and E. Chasseigne, A Bellman approach for regional optimal control problems in $\mathbb{R}^N$, with A. Briani and G. Barles, *SIAM Journal on Control and Optimization* , 2014.

[9] R. C. Barnard and P. R. Wolenski. Flow invariance on stratified domains. *Set-Valued and Variational Analysis*, 21:377–403, 2013.

[10] 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.

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

[12] P. Bettiol, H. Frankowska, et al. Lipschitz regularity of solution map of control systems with multiple state constraints. *Discrete and Continuous Dynamical Systems-Series A*, 32(1):1–26, 2012.

[13] 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.

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

[15] 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.

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

[17] P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre. Optimal times for constrained nonlinear control problems without local controllability. *Applied Mathematics & Optimization*, 36:21–42, 1997.

[18] F. Clarke, Y. Ledyaev, R. Stern, and P. Wolenski. *Nonsmooth Analysis and Control Theory*, volume 178. Springer, 1998.

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

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

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

[22] 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.

[23] 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.

[24] 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.

[25] C. Imbert, R. Monneau and H. Zidani, A Hamilton-Jacobi approach to junction problems and application to traffic flows, *ESAIM: Control, Optimisation and Calculus of Variations*, 19(01): 129–166, 2013.

[26] C. Imbert and R. Monneau, Flux-limited solutions for quasi-convex Hamilton-Jacobi equations on networks, *preprint.*

[27] 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.

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

[29] G. Leoni. *A first course in Sobolev spaces*, volume 105. American Mathematical Soc., 2009.

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

[31] 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.

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

[33] 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.

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

[35] B. Piccoli. Optimal syntheses for state constrained problems with application to optimization of cancer therapies. *Mathematical Control and Related Fields*, 2(4):383–398, 2012.

[36] Z. Rao and H. Zidani, Hamilton-Jacobi-Bellman Equations on Multi-Domains, In Control and Optimization with PDE Constraints, International Series of Numerical Mathematics, vol. 164, Birkhäuser Basel, pp. 93-116, 2013.

[37] Z. Rao, A. Siconolfi and H. Zidani, Transmission conditions on interfaces for Hamilton–Jacobi–Bellman equations, submitted, 2013.

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

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

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

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

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

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