RELATIONSHIP BETWEEN THE MAXIMUM PRINCIPLE AND DYNAMIC PROGRAMMING FOR MINIMAX PROBLEMS

CRISTOPHER HERMOSILLA AND HASNAA ZIDANI

Abstract. This paper is concerned with the relationship between the maximum principle and dynamic programming for a large class of optimal control problems with maximum running cost. Inspired by a technique introduced by Vinter in the 1980s, we are able to obtain jointly a global and a partial sensitivity relation that link the coextremal with the value function of the problem at hand. One of the main contributions of this work is that these relations are derived by using a single perturbed problem, and therefore, both sensitivity relations hold, at the same time, for the same coextremal.

As a by-product, and thanks to the level-set approach, we obtain a new set of sensitivity relations for Mayer problems with state constraints. One important feature of this last result is that it holds under mild assumptions, without the need of imposing strong compatibility assumptions between the dynamics and the state constraints set. Minimax optimal control problems and Maximum principle and Dynamic programming and Sensitivity analysis and State constraints

1. Introduction

This work is devoted to study the relationship between the maximum principle and dynamic programming for optimal control problems governed by a control system of ordinary differential equation on $\mathbb{R}^N$. In particular in this paper, we are concerned with two types of optimization problems: (i) the Minimax problems, where the goal is to minimize a maximum running cost of the form

$$\mathbf{y} \mapsto \max_{t\in[0,T]} \Phi(t, \mathbf{y}(t)) \bigvee \Psi(\mathbf{y}(T)),$$

and (ii) Mayer problems with pathwise state inequality constraints

$$\Phi(t, \mathbf{y}(t)) \leq 0, \quad \forall t \in [0, T].$$

Our task is to provide some insights on the connexions between the coextremal that appears in the corresponding maximum principle and the value function associated with perturbations in the initial data.

This type of connexions, called by some authors the *sensitivity relations*, have been intensively studied in the literature for problems without state constraints.

This work was supported by ECOS-ANID (Grant number ECOS200064). C. Hermosilla has received research support from FONDECYT grant number 11190456 and H. Zidani received research support from Région Normandie and EU through ERDF program (Chaire COPTI)..

For example, given $s \in [0,T]$ and $y \in \mathbb{R}^N$, let us consider the following Mayer problem:

$$\begin{cases} \text{Minimize} & \Psi(\mathbf{y}(T)) \\ \text{subject to} & \dot{\mathbf{y}}(t) = \mathbf{F}(t,\mathbf{y}(t),\mathbf{u}(t)), \quad \text{a.e. on } [0,T], \\ & \mathbf{u}(t) \in U(t) \subseteq \mathbb{R}^m \quad \text{a.e. on } [0,T]. \\ & \mathbf{y}(s) = y. \end{cases}$$

Assume in addition that the data of the problem is regular enough. Then, classical results (see for instance [11, 24]) lead to the so-called Pontryagin's maximum principle (PMP), which allows to associate any optimal pair $(\mathbf{y},\mathbf{u})$ with a costate function $\mathbf{p}$ satisfying a *transversality condition* involving the cost Ψ and the following equations:

$$\begin{aligned} \dot{\mathbf{y}}(t) &= \nabla_p H(t,\mathbf{y}(t),\mathbf{u}(t),\mathbf{p}(t)) \\ \dot{\mathbf{p}}(t) &= -\nabla_y H(t,\mathbf{y}(t),\mathbf{u}(t),\mathbf{p}(t)) \\ H(t,\mathbf{y}(t),\mathbf{u}(t),\mathbf{p}(t)) &= \max_{u\in U(t)} H(t,\mathbf{y}(t),u,\mathbf{p}(t)), \end{aligned}$$

where $H(t,y,u,p) := p \cdot \mathbf{F}(t,y,u)$, for every $(t,y,u,p) \in [0,T]\times\mathbb{R}^N\times\mathbb{R}^m\times\mathbb{R}^N$. The PMP provides a set of necessary conditions for any (locally) optimal process, whereas the dynamic programming approach provides insightful information on the value function V, that is, the mapping that associates the initial time and position of the problem (s,y) with the optimal value $V(s,y)$ of the corresponding control problem. One can guarantee, under rather mild assumptions, the Lipschitz continuity, in a local sense, of the value function (see for instance [4, 24]) as well as the fact that it satisfies a dynamic programming principle; The latter claiming that optimal trajectories remain optimal on each subinterval. For Mayer problems, the relationship between the PMP and value function has been known since the seminal works of Bellman in the 1950s, and can be described as follows:

$$-\mathbf{p}(t) = \nabla_y V(t,\mathbf{y}(t)), \quad \mathcal{H}(t,\mathbf{y}(t),\mathbf{p}(t)) = \partial_t V(t,\mathbf{y}(t)), \tag{1}$$

where $\mathcal{H}$ is the maximized Hamiltonian. These relations highlight the fact that the costate function somewhat measures the sensitivity of the optimal cost of the problem with respect small variations in the optimal trajectory.

In the particular case that the value function is continuously differentiable as well as all the functions involved in the problem setting, the proof of the relations (1) is rather simple. Moreover, if the value function is merely locally Lipschitz continuous, and the costate function is unique, the relations (1) can be extended by using variational techniques (super and sub-differentials) as in [26] (we refer also to [4] where different arguments are provided).

In [12, 23] the authors studied the sensitivity relations for Mayer problems with locally Lipschitz continuous cost; the value function turns out to be Lipschitz continuous as well. Since the final cost is not necessarily smooth, the costate function may not be uniquely determined. In this setting, the sensitivity relations claim that at least one of the coextremals $\mathbf{p}$ given by the PMP, satisfies the endpoints conditions

$$-\mathbf{p}(0) \in \partial_y V(0,\mathbf{y}(0)), \qquad -\mathbf{p}(T) \in \partial_y V(T,\mathbf{y}(T)),$$

and either a *partial sensitivity relation*

$$-\mathbf{p}(t) \in \partial_y V(t,\mathbf{y}(t)), \quad \text{a.e. on } (0,T), \tag{2}$$

or a *global sensitivity relation*

$$(\mathcal{H}(t, \mathbf{x}(t), \mathbf{p}(t)), -\mathbf{p}(t)) \in \partial V(t, \mathbf{x}(t)), \quad \text{for all } t \in [0, T]. \tag{3}$$

These relations improve (1) by using the generalized gradient of the value function (which is well defined in this context). It is worth pointing out that some examples are presented in [12, 24] to demonstrate that not all coextremals satisfy the sensitivity relations, even though they are associated with the same optimal control problem and same minimizer.

Another worth mentioning nonsmooth case is when the dynamics is governed by a linear system and the cost is a convex function. These are the so-called fully convex optimal control problems. It's well known that in this setting, the PMP takes an equivalent form in terms of a Hamiltonian system; see for instance [19]. Moreover, in this context, the value function is a convex function on the state variable (see [21]) and the partial sensitivity relations (2) hold true *for any coextremal*; see for instance [15, Lemma 3.1].

The proofs for (2) (in [12]) and the one for (3) (in [23]) are based on the construction of a new optimal control problem, which is a perturbation of the initial one, however, involving additional control variables, and for which the original minimizer, is as well a solution. This new problem has, by construction, a richer class of variations. Necessary conditions for optimality are then evoked, which makes new information to come out, leading to the sensitivity relations for the optimal control problem we have considered at the beginning.

In the present work, we first derive the sensitivity relations for the control problem with a maximum running cost. The general idea of the proof follows similar arguments developed in [23, 9]. However, some specific difficulties arise in the maximum running cost case. First, the choice of the perturbed problem should be adapted to take into account the nature of the running cost and also to take into account the nature of the sensitivity of the value function with respect to disturbances around the optimal trajectory. This sensitivity is more complex than in the case of Mayer problem. Let us also highlight that in the present work, the partial and global sensitivity relations can be jointly obtained from a single perturbed problem (whereas in the original proofs in [9], a different perturbed problem was introduced for each sensitivity relation).

Motivated by the well-known connexion between minimax problems and state constrained problems (e.g. [1]), the sensitivity relations for the minimax problem can be used as a pivot for getting a relationship between the maximum principle and the dynamic programming for problems with state constraints. As a long list of works demonstrates, the presence of state constraints in optimal control problems poses many challenging difficulties. Indeed, the value function is no longer finite nor locally Lipschitz continuous everywhere, and unless some strong controllability assumptions are enforced, it may even be discontinuous on the interior of the state constraints set. Several works have been devoted to analyze of the regularity of the value function for state-constrained optimal control problems; we refer for instance to [22, 14, 13, 18, 17, 16] and the references therein. See also [3, 5, 7] for similar results for the case of differential games.

Connections between the PMP and the dynamic programming have been reported for state-constrained optimal control problems in [8] under the so-called *Inward Pointing Condition.* This condition forces the state constraint set to be a viability domain, while guaranteeing that the value function is Lipschitz continuous on the interior of that set. In this work, we consider a general control problem with pathwise state inequality constraints. We follow some ideas introduced in [1, 10] and reformulate the original control problem as a control problem with a maximum running cost whose value function is locally Lipschitz continuous everywhere. Then, by using the results that link the original control problem with the auxiliary problem, we are able to derive a set of sensitivity relations without requiring any type of compatibility assumptions such as the mentioned above.

The paper is organized as follows. Section 2 presents some preliminary results in Nonsmooth Analysis. Section 3 introduces the minimax control problem and presents the sensitivity relations for this problem. The proof of the sensitivity relations for the minimax problem is given in Section 4. Section 5 is devoted to the relation between the PMP and the dynamic programming principle for state constrained optimal control problems.

2. Preliminaries

In this paper standard notation is used. For example, $\mathbb{R}$ denotes the set of real numbers, $x \cdot y$ the Euclidean inner product of $x, y \in \mathbb{R}^N$, $\|x\|$ is the norm of $x \in \mathbb{R}^N$ (for any $N \geq 2$), $\mathbb{B}_N$ is the unit closed ball centered at the origin of $\mathbb{R}^N$ (also denoted $\mathbb{B}$ if there is no ambiguity) and $\mathbb{B}(x; r) = x + r\mathbb{B}$.

For any set $S \subseteq \mathbb{R}^N$, $\text{co}(S)$ denotes its convex envelope (the smallest convex set that contains S). For any $a, b \in \mathbb{R}$, we define

$$a \bigvee b := \max(a, b).$$

We write $L^1\left([0,T];\mathbb{R}^N\right)$, $L^\infty\left([0,T];\mathbb{R}^N\right)$ and $W^{1,1}\left([0,T];\mathbb{R}^N\right)$ for the usual Lebesgue spaces and Sobolev space of functions $\mathbf{x} : [0,T] \to \mathbb{R}^N$, respectively. The abbreviation "w.r.t." means "with respect to", and "a.e." stands for "almost everywhere". Also, $\mathcal{L}$ denotes the Lebesgue subsets of $[0,T]$, $\mathcal{B}^m$ stands for the Borel subsets of $\mathbb{R}^m$ and $\mathcal{L} \times \mathcal{B}^m$ is the corresponding product σ-algebra. In this paper, the δ-tube around a curve $\mathbf{x} : [0,T] \to \mathbb{R}^N$ is the open set defined via the formula

$$\left\{x \in \mathbb{R}^N \mid \exists t \in [0,T],\ \|x - \mathbf{x}(t)\| < \delta\right\}, \qquad \text{for some } \delta > 0.$$

A set-valued map $U : [0,T] \rightrightarrows \mathbb{R}^m$ is said to be *locally selectionable* ([2, Definition 1.10.1]) if for any $(t,u) \in \text{Gr}(U)$ there exists a selection v of U, which is continuous on a neighborhood of t and satisfies $v(t) = u$. Here $\text{Gr}(U)$ stands for the graph of the set-valued map U.

2.1. **Nonmooth analysis tools.**

Before going any further, we evoke some definitions and properties from sub-differential calculus that we will need in the rest of the paper. For more details on nonsmooth analysis, we refer to [11, Chapter 10] or [24, Chapter 5].

Let $\phi : \mathbb{R}^N \to \mathbb{R}$ be a Lipschitz continuous function in a neighborhood of a point $x \in \mathbb{R}^N$. The *generalized directional derivative* of ϕ at x, for any direction $v \in \mathbb{R}^N$, is given by:

$$D^\circ\phi(x; v) := \limsup_{y \to x, h \downarrow 0} \frac{\phi(y + hv) - \phi(y)}{h}.$$

Consequently, the Clarke's sub-differential of ϕ at x is given by:

$$\partial\phi(x) = \left\{\xi \in \mathbb{R}^N \mid D^\circ\phi(x; v) \geq \xi \cdot v,\ \ \forall v \in \mathbb{R}^N\right\}.$$

The generalized directional derivative and the Clarke's sub-differential are related to each other in the following way (see for instance [24, Proposition 4.7.4]):

$$D^\circ\phi(x;v) := \max_{\xi\in\partial\phi(x)} \xi\cdot v, \qquad \forall v\in\mathbb{R}^N. \tag{4}$$

The gradient formula for the Clarke's sub-differential ([24, Theorem 4.7.7]) reads as follows: for any $\Omega\subseteq\mathbb{R}^N$ that has (Lebesgue) measure zero it follows that:

$$\partial\phi(x) = \mathrm{co}\left\{\xi\in\mathbb{R}^N \mid \exists x_i\to x,\ x_i\notin\Omega,\ \nabla\phi(x_i) \text{ exists and } \nabla\phi(x_i)\to\xi\right\}. \tag{5}$$

The Clarke's sub-differential of ϕ at x is a compact, convex and nonempty set of $\mathbb{R}^N$, and if $L>0$ is the Lipschitz constant of ϕ in a neighborhood of x then $\partial\phi(x)\subseteq\mathbb{B}_N(0;L)$. This sub-differential satisfies some properties reminiscent of the classical differential calculus as the *sum rule* $\partial(\phi+\psi)(x)\subseteq\partial\phi(x)+\partial\psi(x)$ for ϕ and ψ Lipschitz continuous functions in a neighborhood of x. Moreover, when ψ is continuously differentiable (in a neighborhood of x) then

$$\partial(\phi+\psi)(x) = \partial\phi(x)+\{\nabla\psi(x)\}.$$

Notice too that for any $\alpha\in\mathbb{R}$, we have $\partial(\alpha\phi)(x)=\alpha\partial\phi(x)$.

If ϕ is a bivariate function with $x=(x_1,x_2)$, then for a given $(\bar{x}_1,\bar{x}_2)\in\mathbb{R}^N$, we write $\partial_{x_1}\phi(\bar{x}_1,\bar{x}_2)$ and $\partial_{x_2}\phi(\bar{x}_1,\bar{x}_2)$ for the sub-differentials of the functions $x_1\mapsto\phi(x_1,\bar{x}_2)$ at $x_1=\bar{x}_1$ and $x_2\mapsto\phi(\bar{x}_1,x_2)$ at $x_2=\bar{x}_2$, respectively. These are the partial sub-differentials of the function ϕ at $\bar{x}=(\bar{x}_1,\bar{x}_2)$. It is noteworthy to mention that in general

$$\partial\phi(\bar{x}_1,\bar{x}_2)\neq\partial_{x_1}\phi(\bar{x}_1,\bar{x}_2)\times\partial_{x_2}\phi(\bar{x}_1,\bar{x}_2). \tag{6}$$

The lower *Dini directional derivative* at x of a function $\psi:\mathbb{R}^N\to\mathbb{R}$ in the direction $v\in\mathbb{R}^N$ is defined as:

$$D^\downarrow\psi(x;v) := \liminf_{w\to v,h\downarrow 0}\frac{\psi(x+hw)-\psi(x)}{h}.$$

If ψ is Lipschitz in a neighborhood of x then we have the following expression:

$$D^\downarrow\psi(x;v) = \liminf_{h\downarrow 0}\frac{\psi(x+hv)-\psi(x)}{h}.$$

Following [10, Lemma 2.1], if $\psi:\mathbb{R}^N\to\mathbb{R}$ is Lipschitz continuous function on a neighborhood of x, then

$$D^\downarrow\psi(x;v_1+v_2)\geq D^\downarrow\psi(x;v_1)-D^\circ\psi(x;-v_2), \qquad \forall v_1,v_2\in\mathbb{R}^N.$$

Therefore, the following inequality holds for any $v_1,\dots,v_k\in\mathbb{R}^N$:

$$D^\downarrow\psi\left(x;\sum_{i=0}^k v_i\right)\geq D^\downarrow\psi(x;v_0)-\sum_{i=1}^k D^\circ\psi(x;-v_i). \tag{7}$$

2.2. **Optimality conditions for minimax problems.** Let $T>0$ be a finite time horizon and consider now the following minimax optimal control:

$$(\mathsf{P}_{\mathrm{m}})\qquad \begin{cases}\text{Minimize} & \left(\max_{t\in[0,T]}\varphi(t,\mathsf{x}(t))\bigvee\psi_T(\mathsf{x}(T))\right)+\psi_0(\mathsf{x}(0))\\ \text{over all} & \mathsf{x}\in W^{1,1}([0,T];\mathbb{R}^{\mathsf{n}}) \text{ and } \mathsf{u}:[0,T]\to\mathbb{R}^{\mathsf{m}} \text{ measurable}\\ \text{such that} & \dot{\mathsf{x}}(t)=\mathcal{F}(t,\mathsf{x}(t),\mathsf{u}(t)),\quad \text{a.e. on } [0,T],\\ & \mathsf{u}(t)\in\mathsf{U}(t),\qquad \text{for any } t\in[0,T].\end{cases}$$

Let us recall that a feasible process $(\bar{\mathsf{x}},\bar{\mathsf{u}})$ for problem $(\mathsf{P}_{\rm m})$ is said to be a *strong local minimizer* if there exists $\varrho>0$ such that

$$\left(\max_{t\in[0,T]}\varphi(t,\bar{\mathsf{x}}(t))\bigvee\psi_T(\bar{\mathsf{x}}(T))\right)+\psi_0(\bar{\mathsf{x}}(0))\leq\left(\max_{t\in[0,T]}\varphi(t,\mathsf{x}(t))\bigvee\psi_T(\mathsf{x}(T))\right)+\psi_0(\mathsf{x}(0))$$

for any feasible process (x,u) for problem $(\mathsf{P}_{\rm m})$ such that $\|\mathsf{x}-\bar{\mathsf{x}}\|_{L^\infty}\leq\varrho$.

Optimality conditions for minimax optimal control problems such as $(\mathsf{P}_{\rm m})$ in form of a maximum principle are well-known nowadays; see for instance [6]. The following is a suitable version for the case we are considering.

Lemma 1. *Let $(\bar{\mathsf{x}},\bar{\mathsf{u}})$ be a strong local minimizer for the problem $(\mathsf{P}_{\rm m})$. Assume that for some $\varrho>0$ the following conditions hold:*

(A1) *the mapping $(t,\mathsf{u})\mapsto\mathcal{F}(t,\mathsf{x},\mathsf{u})$ is $\mathcal{L}\times\mathcal{B}^{\mathsf{m}}$ measurable on $[0,T]\times\mathbb{R}^{\mathsf{m}}$ for $\mathsf{x}\in\mathbb{R}^{\mathsf{n}}$ fixed, and there are $\kappa_{\mathcal{F}},c_{\mathcal{F}}\in L^1([0,T];\mathbb{R})$ such that*

$$\|\mathcal{F}(t,\mathsf{x},\mathsf{u})-\mathcal{F}(t,\mathsf{y},\mathsf{u})\|\leq\kappa_{\mathcal{F}}(t)\|\mathsf{x}-\mathsf{y}\|\quad\text{and}\quad\|\mathcal{F}(t,\mathsf{x},\mathsf{u})\|\leq c_{\mathcal{F}}(t)$$

for all $\mathsf{x},\mathsf{y}\in\mathbb{B}_{\mathsf{n}}(\bar{\mathsf{x}}(t),\varrho)$, $\mathsf{u}\in\mathsf{U}(t)$ and for a.e. $t\in[0,T]$.

(A2) *ψ_0 and ψ_T are Lipschitz continuous on $\mathbb{B}(\bar{\mathsf{x}}(0),\varrho)$ and $\mathbb{B}(\bar{\mathsf{x}}(T),\varrho)$, respectively. Also, φ is Lipschitz continuous on the ϱ-tube around $\bar{\mathsf{x}}$.*

(A3) $\mathrm{Gr}(\mathsf{U})\subseteq\mathbb{R}^{\mathsf{m}+1}$ *is a nonempty $\mathcal{L}\times\mathcal{B}^{\mathsf{m}}$ measurable set.*

(A4) *For a.e. $t\in[0,T]$ fixed, $\mathsf{x}\mapsto\mathcal{F}(t,\mathsf{x},\mathsf{u})$ is continuously differentiable on $\mathbb{B}(\bar{\mathsf{x}}(t),\varrho)$ for all $\mathsf{u}\in\mathsf{U}(t)$ fixed and $\mathcal{F}(t,\mathsf{x},\mathsf{U}(t))$ is a compact set for all $\mathsf{x}\in\mathbb{B}(\bar{\mathsf{x}}(t),\varrho)$.*

Then there exist $\mathsf{p}\in W^{1,1}([0,T];\mathbb{R}^{\mathsf{n}})$, $\lambda\in[0,1]$, a finite regular (nonnegative) Borel measure μ on $[0,T]$ and a Borel measurable function $\gamma:[0,T]\to\mathbb{R}^{\mathsf{n}}$ such that the following conditions hold:

(i) Non-triviality: $\lambda+\mu([0,T])=1$ *with*

$$\begin{cases}\lambda=1 & \text{if } \psi_T(\bar{\mathsf{x}}(T))>\max\limits_{t\in[0,T]}\varphi(t,\bar{\mathsf{x}}(t)),\\ \lambda=0 & \text{if } \psi_T(\bar{\mathsf{x}}(T))<\max\limits_{t\in[0,T]}\varphi(t,\bar{\mathsf{x}}(t));\end{cases}$$

(ii) $\gamma(t)\in\partial_x\varphi(t,\bar{\mathsf{x}}(t))$ *for μ-a.e. $t\in[0,T]$;*

(iii) $\mathrm{supp}(\mu)\subseteq\left\{t\in[0,T]\mid\varphi(t,\bar{\mathsf{x}}(t))=\max\limits_{s\in[0,T]}\varphi(s,\bar{\mathsf{x}}(s))\right\}$.

(iv) Costate equation: $-\dot{\mathsf{p}}(t)=\mathsf{q}(t)D_{\mathsf{x}}\mathcal{F}(t,\bar{\mathsf{x}}(t),\bar{\mathsf{u}}(t))$ *for a.e. $t\in[0,T]$;*

(v) Transversality condition: $\mathsf{p}(0)\in\partial\psi_0(\bar{\mathsf{x}}(0))$ *and* $-\mathsf{q}(T)\in\lambda\partial\psi_T(\bar{\mathsf{x}}(T))$*;*

(vi) Maximum principle: $\mathsf{q}(t)\cdot\mathcal{F}(t,\bar{\mathsf{x}}(t),\bar{\mathsf{u}}(t))=\max\limits_{\mathsf{u}\in\mathsf{U}(t)}\mathsf{q}(t)\cdot\mathcal{F}(t,\bar{\mathsf{x}}(t),\mathsf{u})$ *for a.e. $t\in[0,T]$;*

Here

$$\mathsf{q}(t)=\begin{cases}\mathsf{p}(t)+\displaystyle\int_{[0,t[}\gamma(s)\mu(ds) & \text{if } t\in[0,T[\\ \mathsf{p}(T)+\displaystyle\int_{[0,T]}\gamma(s)\mu(ds) & \text{if } t=T\end{cases}$$

Proof. This is a straightforward consecuence of [24, Proposition 9.5.4]. Indeed, it is enough to take the data $\tilde{g}(x,y,z)=h(y,z)+\psi_0(x)$ with $h(y,z):=z\vee\psi_T(y)$ in [24, Proposition 9.5.4]. Then, by the sum rule ([11, Theorem 10.13]) we get

$$\partial\tilde{g}(x,y,z)\subseteq\partial\tilde{g}_1(x,y,z)+\partial\tilde{g}_2(x,y,z),$$

where $\tilde{g}_1(x,y,z) := h \circ \pi_{(y,z)}(x,y,z)$ and $\tilde{g}_2(x,y,z) := \psi_0 \circ \pi_x(x,y,z)$; here $\pi_{(y,z)}(x,y,z) = (y,z)$ and $\pi_x(x,y,z) = x$. Therefore, by the chain rule ([11, Theorem 10.19]), it follows that

$$\partial \tilde{g}(x,y,z) \subseteq \partial \psi_0(x) \times \partial h(y,z).$$

Notice too that by the max rule ([24, Theorem 5.5.2]), we also have

$$\partial h(y,z) \subseteq \left\{ \lambda \partial \psi_T(y) \times \{0\} + (0, 1-\lambda) \mid \begin{matrix} \lambda = 0 \text{ if } \psi_T(y) < z \\ \lambda = 1 \text{ if } \psi_T(y) > z \end{matrix} \right\}.$$

Using this and the transversality condition in [24, Proposition 9.5.4], we get the non-triviality condition in the lemma. □ □

The preceding result will be our vehicle for establishing a relationship between the maximum principle and the dynamic programming principle.

3. Statement of the problem

3.1. **Control systems.** Consider the following dynamical system:

$$\dot{\mathbf{y}}(t) = \mathbf{F}(t, \mathbf{y}(t), \mathbf{u}(t)), \qquad \text{for a.e. } t \in [0,T], \tag{8}$$

where the control input $\mathbf{u} : [0,T] \longrightarrow \mathbb{R}^m$ is a measurable function such that $\mathbf{u}(t) \in U(t)$ for a.e. $t \in [0,T]$ and $U \colon [0,T] \rightrightarrows \mathbb{R}^m$ is a given set-valued map with nonempty images. Throughout this paper, we assume that:

($\mathbf{H}_1$) (a) the mapping $(t,u) \mapsto \mathbf{F}(t,y,u)$ is $\mathcal{L} \times \mathcal{B}^m$ measurable on $[0,T] \times \mathbb{R}^m$ for every $y \in \mathbb{R}^N$ fixed.

(b) there exists $k_{\mathbf{F}} > 0$ such that

$$\|\mathbf{F}(t,x,u) - \mathbf{F}(t,y,u)\| \leq k_{\mathbf{F}} \|x - y\|, \qquad \forall x, y \in \mathbb{R}^N, \ (t,u) \in \mathrm{Gr}(U).$$

(c) there exists $c_{\mathbf{F}} > 0$ such that

$$\|\mathbf{F}(t,y,u)\| \ \leq \ c_{\mathbf{F}}(1 + \|y\|), \qquad \forall y \in \mathbb{R}^N, \ (t,u) \in \mathrm{Gr}(U).$$

($\mathbf{H}_2$) $\mathrm{Gr}(U) \subseteq \mathbb{R}^{m+1}$ is a nonempty $\mathcal{L} \times \mathcal{B}^m$ measurable set.

We denote by $\mathcal{U}$ the set of all admissible controls, that is, all the possible (Lebesgue) measurable selections of the set-valued mapping $t \mapsto U(t)$; $\mathcal{U}$ is nonempty thanks to Aumann's Measurable Selection Theorem ([24, Theorem 2.3.12]). Notice that assumption **($\mathbf{H}_1$)** guarantees that for every $\mathbf{u} \in \mathcal{U}$, the mapping $(t,y) \mapsto \mathbf{F}(t,y,u(t))$ is a Carathéodory function ([24, Corollary 2.3.3]), and so, by the Generalized Filippov Existence Theorem ([24, Theorem 2.4.3]), there exists an absolutely continuous curve $\mathbf{y} : [0,T] \to \mathbb{R}^N$ satisfying the dynamic equation (8), which can be uniquely determined if some intermediate data is fixed. Given $(t,x) \in [0,T] \times \mathbb{R}^N$, let us denote by $\mathbb{X}^{\mathbf{F}}(t,x)$ the set of all state-control pairs $(\mathbf{y}, \mathbf{u}) \in W^{1,1}\left([0,T]; \mathbb{R}^N\right) \times \mathcal{U}$ that satisfy (8) and such that $\mathbf{y}(t) = x$.

By **($\mathbf{H}_1$)** and the Gronwall's inequality (see [24, Lemma 2.4.4]), it follows that for every $(\mathbf{y}, \mathbf{u}) \in \mathbb{X}^{\mathbf{F}}(t,x)$ we have

$$1 + \|\mathbf{y}(s)\| \leq (1 + \|x\|) e^{c_{\mathbf{F}}(s-t)}, \quad \forall s \in [t,T]. \tag{9}$$

$$\|\dot{\mathbf{y}}(s)\| \leq c_{\mathbf{F}}(1 + \|x\|) e^{c_{\mathbf{F}}(s-t)}, \quad \text{for a.e. } s \in [t,T]. \tag{10}$$

3.2. **Minimax optimal control problems.** Let us consider now the following optimal control problem with maximum running cost:

$$(\mathrm{P}) \qquad \begin{cases} \text{Minimize} & \max\limits_{t\in[0,T]} \Phi(t,\mathbf{y}(t)) \bigvee \Psi(\mathbf{y}(T)) \\ \text{over all} & \mathbf{y}\in W^{1,1}\left([0,T];\mathbb{R}^N\right) \text{ and } \mathbf{u}\in\mathcal{U} \\ \text{such that} & \dot{\mathbf{y}}(t)=\mathbf{F}(t,\mathbf{y}(t),\mathbf{u}(t)), \quad \text{a.e. on } [0,T], \\ & \mathbf{y}(0)=y_0. \end{cases}$$

In addition to the assumptions on the dynamical system, we consider as well the following conditions on the costs:

($\mathbf{H}_3$) Φ and Ψ are locally Lipschitz continuous.

Notice that in problem (P) the initial condition is fixed, while in (P_m), it is free but it has an associated cost. It is important to note that [24, Proposition 9.5.4] can also be used for deriving a maximum principle for (P) with the appropriate modifications; namely, the initial condition for the coextremal $\mathsf{p}(0)$ being free. In particular, under an additional smoothness assumption that we specified later, we can deduce the following maximum principle: if $(\bar{\mathbf{y}},\bar{\mathbf{u}})$ is a minimizer of problem (P), then there exist $\mathbf{p}\in W^{1,1}\left([0,T];\mathbb{R}^N\right)$, $\lambda\in[0,1]$, a finite regular (nonnegative) Borel measure μ on $[0,T]$ and a Borel measurable function $\gamma:[0,T]\to\mathbb{R}^N$ such that

(MP_i) Non-triviality: $\lambda+\mu([0,T])=1$ with

$$\begin{cases} \lambda=1 & \text{if } \Psi(\bar{\mathbf{y}}(T)) > \max\limits_{t\in[0,T]} \Phi(t,\bar{\mathbf{y}}(t)) \\ \lambda=0 & \text{if } \Psi(\bar{\mathbf{y}}(T)) < \max\limits_{t\in[0,T]} \Phi(t,\bar{\mathbf{y}}(t)); \end{cases}$$

(MP_{ii}) $\gamma(t)\in\partial_y\Phi(t,\bar{\mathbf{y}}(t))$ for μ-a.e. $t\in[0,T]$;

(MP_{iii}) $\mathrm{supp}(\mu)\subseteq\left\{t\in[0,T] \mid \Phi(t,\bar{\mathbf{y}}(t)) = \max\limits_{s\in[0,T]}\Phi(s,\bar{\mathbf{y}}(s))\right\}$;

(MP_{iv}) Costate equation and transversality condition:

$$\begin{cases} -\dot{\mathbf{p}}(t)=\mathbf{q}(t)D_y\mathbf{F}(t,\bar{\mathbf{y}}(t),\bar{\mathbf{u}}(t)), & \text{for a.e. } t\in[0,T], \\ -\mathbf{q}(T)\in\lambda\partial\Psi(\bar{\mathbf{y}}(T)), \end{cases}$$

with

$$\mathbf{q}(t)=\begin{cases} \mathbf{p}(t)+\displaystyle\int_{[0,t[}\gamma(s)\mu(ds) & \text{if } t\in[0,T[\\ \mathbf{p}(T)+\displaystyle\int_{[0,T]}\gamma(s)\mu(ds) & \text{if } t=T \end{cases}$$

(MP_v) Maximality condition:

$$H^{\mathbf{F}}(t,\bar{\mathbf{y}}(t),\bar{\mathbf{u}}(t),\mathbf{q}(t)) = \max_{u\in U(t)} H^{\mathbf{F}}(t,\bar{\mathbf{y}}(t),u,\mathbf{q}(t)) \quad \text{for a.e. } t\in[0,T].$$

where the Hamiltonian $H^{\mathbf{F}}:[0,T]\times\mathbb{R}^N\times\mathbb{R}^m\times\mathbb{R}^N\to\mathbb{R}$ is given by

$$H^{\mathbf{F}}(t,y,u,q):=q\cdot\mathbf{F}(t,y,u) \quad \forall (t,y,u,q)\in[0,T]\times\mathbb{R}^N\times\mathbb{R}^m\times\mathbb{R}^N.$$

The aim of this work is to study the relationship between the dynamic programming principle and the maximum principle for this type of optimal control problems. To do so, let us consider the value function corresponding to problem (P), which is given by

$$\mathcal{V}(t,x) = \inf_{(\mathbf{y},\mathbf{u})\in\mathbb{X}^{\mathbf{F}}(t,x)} \left\{ \max_{s\in[t,T]} \Phi(s,\mathbf{y}(s)) \bigvee \Psi(\mathbf{y}(T)) \right\}.$$

It is not difficult to see that $\mathcal{V}$ satisfies the dynamic programming principle, which in this case reads as follows: for any $t \in [0,T]$ and $h \geq 0$ such that $t+h \leq T$, and for all $x \in \mathbb{R}^N$ we have

$$\mathcal{V}(t,x) = \inf_{(\mathbf{y},\mathbf{u})\in\mathbb{X}^{\mathbf{F}}(t,x)} \left\{ \mathcal{V}(t+h,\mathbf{y}(t+h)) \bigvee \max_{s\in[t,t+h]} \Phi(s,\mathbf{y}(s)) \right\}. \tag{11}$$

Similarly as done in [1], it can be proven that this value function is locally Lipschitz continuous. For convenience of the reader, we provide the details of the proof of the following lemma in the appendix.

Lemma 2. *Assume that the hypotheses* $(\mathbf{H}_1)$, $(\mathbf{H}_2)$ *and* $(\mathbf{H}_3)$ *hold, then* $\mathcal{V}$ *is locally Lipschitz continuous on* $[0,T]\times\mathbb{R}^N$.

Consequently, $\partial\mathcal{V}(t,x)$ and $\partial_x\mathcal{V}(t,x)$, the Clarke's sub-differentials of the functions $(t,x)\mapsto\mathcal{V}(t,x)$ and $x\mapsto\mathcal{V}(t,x)$ for $t\in[0,T]$ fixed, respectively, are well-defined compact convex and non-empty sets. Taking this into account, our main theorem reads as follows.

Theorem 3.1. *Assume that the hypotheses* $(\mathbf{H}_1)$, $(\mathbf{H}_2)$ *and* $(\mathbf{H}_3)$ *hold. Let* $(\bar{\mathbf{y}},\bar{\mathbf{u}})$ *be a minimizer of problem* (P) *and suppose in addition that:*

- *(A_1) for every* $(t,u)\in\mathrm{Gr}(U)$, $y\mapsto\mathbf{F}(t,y,u)$ *is continuously differentiable;*
- *(A_2) there is a finite set* $D\subseteq[0,T]$ *such that the map* $(t,u)\mapsto\mathbf{F}(t,y,u)$ *is continuous on* $\{(t,u)\in\mathrm{Gr}(U)\mid t\notin D\}$ *for every* $y\in\mathbb{R}^N$ *fixed;*[1]
- *(A_3)* U *is a locally selectionable set-valued map.*

Then there exist $\mathbf{p}\in W^{1,1}\left([0,T];\mathbb{R}^N\right)$, $\lambda\in[0,1]$, *a finite regular (nonnegative) Borel measure* μ *on* $[0,T]$ *and a Borel measurable function* $\gamma:[0,T]\to\mathbb{R}^N$ *such that (MP_i), (MP_{ii}), (MP_{iii}), (MP_{iv}) and (MP_v) hold together with the following sensitivity relations:*

$$-\mathbf{p}(0)\in\partial_x\mathcal{V}(0,\bar{\mathbf{y}}(0))$$

and for a.e. $t\in[0,T]$

$$-\mathbf{q}(t)\in\nu(t)\partial_x\mathcal{V}(t,\bar{\mathbf{y}}(t)) \quad \textit{and} \quad (H^{\mathbf{F}}(t,\bar{\mathbf{y}}(t),\bar{\mathbf{u}}(t),\mathbf{q}(t)),-\mathbf{q}(t))\in\nu(t)\partial\mathcal{V}(t,\bar{\mathbf{y}}(t)).$$

where $\nu(s):=1-\mu([0,s[)$ *for any* $s\in[0,T]$.

Remark 1. *The sensitivity relations for the minimax problem involve an additional multiplier* $\nu(t):=1-\mu([0,t[)\in[0,1]$. *Taking into account the non-triviality condition (MP_i), some direct consequences can be drawn.*

(i) Let us consider the case when $\Psi(\bar{\mathbf{y}}(T))>\max_{s\in[0,T]}\Phi(s,\bar{\mathbf{y}}(s))$. *In this situation, the optimal solution is also optimal for the problem with only the final cost* Ψ *(a Mayer type problem). By (MP_i), we get that* $\lambda=1$ *and the measure* $\mu\equiv0$. *Therefore,* $\nu\equiv-1$ *and* $\mathbf{p}=\mathbf{q}$ *on* $[0,T]$. *The sensitivity relations reduce to the classical relations known for Mayer problems:*

$$-\mathbf{p}(t)\in\partial_x\mathcal{V}(t,\bar{\mathbf{y}}(t)) \quad \textit{and} \quad (H^{\mathbf{F}}(t,\bar{\mathbf{y}}(t),\bar{\mathbf{u}}(t),\mathbf{p}(t)),-\mathbf{p}(t))\in\partial\mathcal{V}(t,\bar{\mathbf{y}}(t)).$$

[1] This holds for instance if $\mathbf{F}$ that is piecewise continuous with respect to t in the sense that there are $t_1,\dots t_n\in[0,T]$ such that $\mathbf{F}$ is continuous on $[0,T]\setminus\{t_1,\dots t_n\}\times\mathbb{R}^N\times\mathbb{R}^m$.

(ii) Now, assume that $\Psi(\bar{\mathbf{y}}(T)) < \max_{s\in[0,T]} \Phi(s,\bar{\mathbf{y}}(s)) = \Phi(t_0,\bar{\mathbf{y}}(t_0))$*, and set*

$$\bar{t} := \max\left\{ t\in[0,T] \mid \Phi(t,\bar{\mathbf{y}}(t)) = \max_{s\in[0,T]} \Phi(s,\bar{\mathbf{y}}(s))\right\}.$$

Here again, the non-triviality condition gives more information on λ *and* $\nu(\cdot)$*. Indeed, (MP*$_i$*) implies, in this case, that* $\lambda = 0$ *and* $\mu([0,T]) = 1$*. This property combined with (MP*$_{iii}$*) imply that in fact* $\mu([0,t[) = 1$ *for every* $t > \bar{t}$*. Therefore* $\nu(t) = 0$ *on* $]\bar{t},T]$*, and the sensitivity relations hold along the optimal trajectory until the time* $\bar{t}$*, which corresponds to the last time when the maximal cost is reached. Moreover, we have* $\mathbf{q}(t) = 0$ *on* $]\bar{t},T]$*.*

(iii) Consider the specific case when

$$\{\bar{t}\} = \left\{ t\in[0,T] \mid \Phi(t,\bar{\mathbf{y}}(t)) = \max_{s\in[0,T]} \Phi(s,\bar{\mathbf{y}}(s))\right\},$$

and assume that $\Psi(\bar{\mathbf{y}}(T)) < \Phi(\bar{t},\bar{\mathbf{y}}(\bar{t}))$*. Notice that in this case* $\bar{t} < T$*, the measure is atomic (i.e.,* $\mu = \delta_{\bar{t}}$*) and we have* $\nu(t) = \begin{cases} -1 & \text{for } t\leq \bar{t},\\ 0 & \text{for } t > \bar{t}.\end{cases}$

4. Proof of Theorem 3.1

The proof of Theorem 3.1 includes several technical arguments, so for this reason we have subdivided it into several stages.

In what follows, let $(\bar{\mathbf{y}},\bar{\mathbf{u}}) \in \mathbb{X}^{\mathbf{F}}(t,x)$ be a minimizer of problem (P). We also assume that the assumptions of Theorem 3.1 are in force.

4.1. Basic definitions and perturbed problem.

For a given $\varepsilon > 0$ small enough and $t\in[0,T]$, we set:

$$G_0^\varepsilon(t) := \{(\alpha,\beta)\in \partial\mathcal{V}(t,x) \mid x\in \mathbb{B}_N((\bar{\mathbf{y}}(t);\varepsilon)\}$$

and

$$G_1^\varepsilon(t) := \{\beta\in \partial_y\mathcal{V}(t,x) \mid x\in \mathbb{B}_N(\bar{\mathbf{y}}(t);\varepsilon)\}.$$

Since $\mathcal{V}$ is locally Lipschitz continuous (Lemma 2), for any bounded set S of $\mathbb{R}^N$, $\partial\mathcal{V}$ is bounded on $[0,T]\times S$, hence $G_0^\varepsilon(t)$ and $G_1^\varepsilon(t)$ are bounded sets, uniformly w.r.t. the t variable.

Now, let us introduce the *support* functions $\sigma_0^\varepsilon : [0,T]\times\mathbb{R}\times\mathbb{R}^N \to \mathbb{R}$ and $\sigma_1^\varepsilon : [0,T]\times\mathbb{R}^N\to\mathbb{R}$ defined as:

$$\sigma_0^\varepsilon(t,\omega,\theta) := \sup_{(\alpha,\beta)\in G_0^\varepsilon(t)} (\alpha,\beta)\cdot(\omega,-\theta), \qquad \forall t\in[0,T],\ \omega\in\mathbb{R},\ \theta\in\mathbb{R}^N$$

and

$$\sigma_1^\varepsilon(t,b) := \sup_{\beta\in G_1^\varepsilon(t)} \beta\cdot(-b), \qquad \forall t\in[0,T],\ b\in\mathbb{R}^N.$$

The function σ_1^ε will play a fundamental role when it comes to obtaining the *partial sensitivity relations*, that is, the sensitivity relations associated only with the state variable.

By similar arguments as the ones pointed out in [23, Lemma 3.1], since the subdifferential of a locally Lipschitz continuous function has a closed graph when it is seen as a set-valued map ([11, Proposition 10.10]), it also follows that G_0^ε and G_1^ε have closed graphs. This fact implies in turn that the functions σ_0^ε and σ_1^ε are both upper semi-continuous. Moreover, for any $t\in[0,T]$, the mapping $(\omega,\theta)\mapsto \sigma_0^\varepsilon(t,\omega,\theta)$ is continuous on $\mathbb{R}\times\mathbb{R}^N$. On the other hand, the function σ_0^ε is bounded on any bounded set of $[0,T]\times\mathbb{R}\times\mathbb{R}^N$. In particular, for any essentially bounded and measurable functions $(\boldsymbol{\omega},\boldsymbol{\theta}):[0,T]\to\mathbb{R}\times\mathbb{R}^N$, the function $t\mapsto\sigma_0^\varepsilon(t,\boldsymbol{\omega}(t),\boldsymbol{\theta}(t))$ is integrable. Similarly, for any measurable mapping $\mathbf{b}:[0,T]\to\mathbb{R}^N$, the function $t\mapsto\sigma_1^\varepsilon(t,\mathbf{b}(t))$ is measurable and bounded on $[0,T]$.

Consider now the following perturbed optimal control problem:

$$(\mathcal{P}_\varepsilon)\qquad\begin{cases}\text{Minimize} & J(\mathbf{y},\boldsymbol{\zeta})\\ \text{over all} & (\mathbf{y},\boldsymbol{\zeta})\in W^{1,1}\left([0,T];\mathbb{R}^{N+1}\right),\ \mathbf{u}\in\mathcal{U}\text{ and measurable}\\ & \ \text{functions }(\boldsymbol{\omega},\boldsymbol{\theta},\mathbf{b}):[0,T]\to\mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^N\\ \text{such that} & \dot{\mathbf{y}}(t)=(1+\boldsymbol{\omega}(t))\mathbf{F}(t,\mathbf{y}(t),\mathbf{u}(t))+\boldsymbol{\theta}(t)+\mathbf{b}(t),\\ & \dot{\boldsymbol{\zeta}}(t)=\sigma_0^\varepsilon(t,\boldsymbol{\omega}(t),\boldsymbol{\theta}(t))+\sigma_1^\varepsilon(t,\mathbf{b}(t)),\\ & |\boldsymbol{\omega}(t)|\le\varepsilon,\ \|\boldsymbol{\theta}(t)\|\le\varepsilon,\ \|\mathbf{b}(t)\|\le\varepsilon,\\ & \text{for a.e. } t\in[0,T].\end{cases}$$

where the perturbed cost function is defined via the formula:

$$J(\mathbf{y},\boldsymbol{\zeta}):=\left[\max_{t\in[0,T]}(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(t))\bigvee(\Psi(\mathbf{y}(T))+\boldsymbol{\zeta}(T))\right]-\mathcal{V}(0,\mathbf{y}(0))-\boldsymbol{\zeta}(0).$$

Notice that in problem $(\mathcal{P}_\varepsilon)$ the initial position for the state variable is free. Besides, the *perturbations* $(\mathbf{u},\boldsymbol{\omega},\boldsymbol{\theta},\mathbf{b})$ are considered as control inputs that enrich the class of variations for this new optimal control problem.

4.2. **Persistence of optimality.** In what follows let us consider $\bar{\boldsymbol{\zeta}}\equiv 0$, $\bar{\boldsymbol{\omega}}\equiv 0$, $\bar{\boldsymbol{\theta}}\equiv 0$ and $\bar{\mathbf{b}}\equiv 0$. For any given $\varepsilon\in]0,1]$, let us show that $((\bar{\mathbf{y}},\bar{\boldsymbol{\zeta}}),(\bar{\mathbf{u}},\bar{\boldsymbol{\omega}},\bar{\boldsymbol{\theta}},\bar{\mathbf{b}}))$ is a strong local minimizer of the problem $(\mathcal{P}_\varepsilon)$. For this, we first remark that $(\bar{\mathbf{y}},\bar{\boldsymbol{\zeta}})$ is a feasible trajectory for problem $(\mathcal{P}_\varepsilon)$ associated with the control law $(\bar{\mathbf{u}},\bar{\boldsymbol{\omega}},\bar{\boldsymbol{\theta}},\bar{\mathbf{b}})$ and also

$$J(\bar{\mathbf{y}},\bar{\boldsymbol{\zeta}})=\left[\max_{t\in[0,T]}\Phi(t,\bar{\mathbf{y}}(t))\bigvee\Psi(\bar{\mathbf{y}}(T))\right]-\mathcal{V}(0,\bar{\mathbf{y}}(0))=0.$$

Now, consider any feasible process $(\mathbf{y},\boldsymbol{\zeta},\mathbf{u},\boldsymbol{\omega},\boldsymbol{\theta},\mathbf{b})$ for problem $(\mathcal{P}_\varepsilon)$, such that $\mathbf{y}$ lies in the ε-tube around $\bar{\mathbf{y}}$, i.e.,

$$\|\mathbf{y}(t)-\bar{\mathbf{y}}(t)\|<\varepsilon,\quad\forall t\in[0,T].$$

We will show that

$$J(\mathbf{y},\boldsymbol{\zeta})\ge 0,\tag{12}$$

which in turn will lead directly to the desired conclusion that $((\bar{\mathbf{y}},\bar{\boldsymbol{\zeta}}),(\bar{\mathbf{u}},\bar{\boldsymbol{\omega}},\bar{\boldsymbol{\theta}},\bar{\mathbf{b}}))$ is a strong local minimizer of the problem $(\mathcal{P}_\varepsilon)$.

As $t\mapsto\mathcal{V}(t,\mathbf{y}(t))$ is an absolutely continuous function on $[0,T]$ (thanks to Lemma 2), it is then almost everywhere differentiable on $[0,T]$. Furthermore, in the same spirit as [9, Proposition 5.4], we have the following key result.

Lemma 3. *Under the assumptions of Theorem 3.1,, there is a measurable set $\Theta\subseteq[0,T]$ of full measure (i.e., $[0,T]\backslash\Theta$ is a null set) such that for every $(t,x)\in\Theta\times\mathbb{R}^N$ and $u\in U(t)$, we have*

$$\min\left(D^\circ\mathcal{V}\big((t,x);-(1,\mathbf{F}(t,x,u))\big),\mathcal{V}(t,x)-\Phi(t,y)\right)\le 0,\tag{13}$$

and therefore, for almost every $t \in [0,T]$, we have

$$\left(\frac{d}{dt}\mathcal{V}(t,\mathbf{y}(t)) + \sigma_0^\varepsilon(t,\boldsymbol{\omega}(t),\boldsymbol{\theta}(t)) + \sigma_1^\varepsilon(t,\mathbf{b}(t))\right) \bigvee \left(\Phi(t,\mathbf{y}(t)) - \mathcal{V}(t,\mathbf{y}(t))\right) \geq 0. \tag{14}$$

Proof. Let $(t,x) \in [0,T]\times\mathbb{R}^N$. By definition, we know that $\mathcal{V}(t,x) \geq \Phi(t,x)$. Thus, to prove (13), we only need to consider the case $\mathcal{V}(t,x) > \Phi(t,x)$. In this case, since Φ and $\mathcal{V}$ are locally Lipschitz continuous ($(\mathbf{H}_3)$ and Lemma 2), there is $\delta_0 > 0$ such that

$$\mathcal{V}(s,y) > \Phi(s,y), \qquad \forall (s,y) \in [0,T]\times\mathbb{R}^N \bigcap \mathbb{B}_{N+1}((t,x);\delta_0).$$

Let $D \subseteq [0,T]$ be a discrete set such that $(s,u) \mapsto \mathbf{F}(s,y,u)$ is continuous on $\{(s,u) \in \mathrm{Gr}(U) \mid s \notin D\}$ for every $y \in \mathbb{R}^N$ fixed. Let $\Theta = [0,T]\setminus D$, and take $u \in U(t)$ and $(s,y) \in]0,T]\times\mathbb{R}^N \bigcap \mathbb{B}_{N+1}((t,x);\delta_0)$. Notice that it is possible to take $\delta_0 > 0$ such that $s \in \Theta$ and such that there is a continuous selection v of U on $[0,T]\cap(t-\delta_0,t+\delta_0)$ that satisfies $v(t) = u$; the existence of such selection is thanks to the fact that U is locally selectionable. Let $\mathbf{y} \in W^{1,1}\left([0,T];\mathbb{R}^N\right)$ be the unique solution to (8) that satisfies the condition $\mathbf{y}(s) = y$ associated with the control

$$\mathbf{u}(\tau) := \begin{cases} v(\tau) & \text{if } \tau \in (t-\delta_0, t+\delta_0), \\ \bar{\mathbf{u}}(\tau) & \text{otherwise,} \end{cases} \qquad \forall \tau \in [0,T].$$

It follows then that $\mathbf{y}$ is continuously differentiable at $\tau = s$ with $\dot{\mathbf{y}}(s) = \mathbf{F}(s,y,v(s))$. This last assertion comes from the fact that $(\tau,z) \mapsto \mathbf{F}(\tau,z,v(\tau))$ is continuous at $(\tau,z) = (s,y)$ because the following inequality holds

$$\|\mathbf{F}(\tau,z,v(\tau)) - \mathbf{F}(s,y,v(s))\| \leq \|\mathbf{F}(\tau,y,v(\tau)) - \mathbf{F}(s,y,v(s))\| + k_{\mathbf{F}}\|z-y\|$$

Since $\mathcal{V}$ satisfies the dynamic programming principle (11), for any $h > 0$ small enough we have

$$\mathcal{V}(s,y) \geq \mathcal{V}(s,y) \bigvee \max_{\tau\in[s-h,s]} \Phi(\tau,\mathbf{y}(\tau)) \geq \mathcal{V}(s-h,\mathbf{y}(s-h)). \tag{15}$$

Assume that $\mathcal{V}$ is differentiable at (s,y), then from (15) we get

$$-\vartheta'(s) = \lim_{h\to 0} \frac{\mathcal{V}(s-h,\mathbf{y}(s-h)) - \mathcal{V}(s,\mathbf{y}(s))}{h} \leq 0.$$

where $\vartheta'(s)$ is the derivative of the function $\tau \mapsto \vartheta(\tau) := \mathcal{V}(\tau,\mathbf{y}(\tau))$ at $\tau = s$. It follows then that

$$\nabla\mathcal{V}(s,y)\cdot(-1,-\mathbf{F}(s,y,u)) = \nabla\mathcal{V}(s,\mathbf{y}(s))\cdot(-1,-\dot{\mathbf{y}}(s)) = -\vartheta'(s) \leq 0.$$

Therefore, by the continuity properties of F, given $(t,x) \in \Theta\times\mathbb{R}^N$ and $u \in U(t)$, we also have that $\xi\cdot(-1,-\mathbf{F}(t,x,u)) \leq 0$ for any ξ belonging to the set

$$\mathrm{co}\left\{\xi \in \mathbb{R}^{N+1} \mid \exists (s_i,y_i) \to (t,x),\ s_i \in \Theta,\ \nabla\mathcal{V}(s_i,y_i) \text{ exists and } \nabla\mathcal{V}(s_i,y_i) \to \xi\right\}.$$

Thus the proof of the first claim of the lemma follows from (4) combined with (5).

Notice that for a.e. $t \in [0,T]$ it holds:

$$\frac{d}{dt}\mathcal{V}(t,\mathbf{y}(t)) = D^{\downarrow}\mathcal{V}((t,\mathbf{y}(t));(1,\dot{\mathbf{y}}(t)))$$

because, since $\dot{\mathbf{y}}$ is essentially bounded (by Gronwall's inequality (9)) and the value function $\mathcal{V}$ is locally Lipschitz continuous, we have

$$\lim_{h\downarrow 0}\frac{1}{h}\Big(\mathcal{V}(t+h,\mathbf{y}(t+h))-\mathcal{V}(t,\mathbf{y}(t))\Big)=\lim_{h\downarrow 0}\frac{1}{h}\Big(\mathcal{V}(t+h,\mathbf{y}(t)+h\dot{\mathbf{y}}(t))-\mathcal{V}(t,\mathbf{y}(t))\Big).$$

Furthermore (with the notation $\mathfrak{F}=\mathbf{F}(t,\mathbf{y}(t),\mathbf{u}(t))$) we have

$$\begin{pmatrix}1\\ \dot{\mathbf{y}}\end{pmatrix}=\begin{pmatrix}1\\ (1+\boldsymbol{\omega})\mathfrak{F}+\boldsymbol{\theta}+\mathbf{b}\end{pmatrix}=\begin{pmatrix}0\\ \mathbf{b}\end{pmatrix}+(1+\boldsymbol{\omega})\begin{pmatrix}1\\ \mathfrak{F}\end{pmatrix}+\begin{pmatrix}-\boldsymbol{\omega}\\ \boldsymbol{\theta}\end{pmatrix}.$$

By using (7) and the relation $D^{\circ}\phi(x;\lambda v)=\lambda D^{\circ}\phi(x;v)$ (for $\lambda\geq 0$), we get:

$$\begin{aligned}\frac{d}{dt}\mathcal{V}(t,\mathbf{y}(t)) &= D^{\downarrow}\mathcal{V}((t,\mathbf{y}(t));(1,\dot{\mathbf{y}}(t)))\\ &\geq\ D^{\downarrow}\mathcal{V}((t,\mathbf{y}(t));(0,\mathbf{b}(t)))\\ &-\ D^{\circ}\mathcal{V}((t,\mathbf{y}(t));(\boldsymbol{\omega}(t),-\boldsymbol{\theta}(t)))\\ &-(1+\boldsymbol{\omega}(t))D^{\circ}\mathcal{V}((t,\mathbf{y}(t));-(1,\mathbf{F}(t,\mathbf{y}(t),\mathbf{u}(t)))\end{aligned}\tag{16}$$

Notice that for any given values of ω, θ and b, it holds

$$D^{\circ}\mathcal{V}((t,\mathbf{y}(t));(\omega,-\theta))=\max_{(\alpha,\beta)\in\partial\mathcal{V}(t,\mathbf{y}(t))}(\alpha,\beta)\cdot(\omega,-\theta)\leq\sigma_0^{\varepsilon}(t,\omega,\theta).\tag{17}$$

and

(18)

$$\begin{aligned}D^{\downarrow}\mathcal{V}((t,\mathbf{y}(t));(0,\mathbf{b}))=D_x^{\downarrow}\mathcal{V}(t,\cdot)(\mathbf{y}(t);\mathbf{b})&=-\limsup_{h\downarrow 0}\frac{\mathcal{V}(t,\mathbf{y}(t))-\mathcal{V}(t,\mathbf{y}(t)+h\mathbf{b})}{h}\\ &=-\limsup_{h\downarrow 0}\frac{\mathcal{V}(t,\mathbf{y}(t)+h\mathbf{b}+h(-\mathbf{b}))-\mathcal{V}(t,\mathbf{y}(t)+h\mathbf{b})}{h}\\ &\geq -D_x^{\circ}\mathcal{V}(t,\cdot)(\mathbf{y}(t);-\mathbf{b})=-\max_{\beta\in\partial_x\mathcal{V}(t,\mathbf{y}(t))}\beta\cdot(-\mathbf{b})\ \geq\ -\sigma_1^{\varepsilon}(t,\mathbf{b}).\end{aligned}$$

Finally, by combining inequalities (13),(16), (17), (18) and by using the fact that $1+\boldsymbol{\omega}(t)\geq 0$, we conclude that (14) is satisfied. □ □

Let $\mathcal{J}:=\{t\in]0,T[|\ \ \mathcal{V}(t,\mathbf{y}(t))>\Phi(t,\mathbf{y}(t))\}$. Since $\mathcal{V}$, Φ and $\mathbf{y}$ are continuous functions, $\mathcal{J}$ is an open subset of $]0,T[$. If $\mathcal{J}=\emptyset$, then for every $s\in[0,T]$, $\mathcal{V}(t,\mathbf{y}(t))=\Phi(t,\mathbf{y}(t))$. In this situation we have

$$\mathcal{V}(0,\mathbf{y}(0))+\boldsymbol{\zeta}(0)=\Phi(0,\mathbf{y}(0))+\boldsymbol{\zeta}(0)\leq\max_{s\in[0,T]}\left(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(s)\right)\bigvee\left(\Psi(\mathbf{y}(T))+\boldsymbol{\zeta}(T)\right),$$

which implies that $J(\mathbf{y},\boldsymbol{\zeta},\mathbf{u},\boldsymbol{\omega},\boldsymbol{\theta},\mathbf{b})\geq 0$, and then claim (12) is proved.

Now, assume that $\mathcal{J}\neq\emptyset$. Introduce

$$\bar{a}:=\inf\{t\mid t\in\mathcal{J}\},\qquad \bar{b}:=\sup\{t\mid t\in\mathcal{J}\}.$$

First, we can show that either $\bar{a}=0$ or $\mathcal{V}(\bar{a},\mathbf{y}(\bar{a}))=\Phi(\bar{a},\mathbf{y}(\bar{a}))$. Indeed, if $\bar{a}>0$ and $\mathcal{V}(\bar{a},\mathbf{y}(\bar{a}))>\Phi(\bar{a},\mathbf{y}(\bar{a}))$, then by continuity of $t\mapsto\mathcal{V}(t,\mathbf{y}(t))-\Phi(t,\mathbf{y}(t))$, there exists $\delta>0$ small enough such that $0\leq\bar{a}-\delta$ and $\mathcal{V}(t,\mathbf{y}(t))>\Phi(t,\mathbf{y}(t))$ on $]\bar{a}-\delta,\bar{a}[$, which contradicts the definition of $\bar{a}$. By similar arguments, we can also show that $\bar{b}=T$ or $\mathcal{V}(\bar{b},\mathbf{y}(\bar{b}))=\Phi(\bar{b},\mathbf{y}(\bar{b}))$. Since $\mathcal{J}$ is an open nonempty set, there exist two non-decreasing sequences $(\alpha_i)_{i\in\mathbb{N}}$ and $(\beta_i)_{i\in\mathbb{N}}$ in $[0,T]$ such that $\bigcup_i]\alpha_i,\beta_i[\subseteq\mathcal{J}$ with $\alpha_i<\beta_i$, satisfying

$$\mathcal{V}(\alpha_i,\mathbf{y}(\alpha_i))=\Phi(\alpha_i,\mathbf{y}(\alpha_i))\quad\text{and}\quad\mathcal{V}(\beta_i,\mathbf{y}(\beta_i))=\Phi(\beta_i,\mathbf{y}(\beta_i)),\qquad\forall i\in\mathbb{N}.$$

The sequence $(\beta_i)_{i\in\mathbb{N}}$ can be taken such that either $\beta_i = \bar{b}$ for any $i \in \mathbb{N}$ large enough or $(\beta_i)_{i\in\mathbb{N}}$ is strictly increasing with $\beta_i \nearrow \bar{b}$ if $i \to +\infty$. Notice that this cover the case when $\mathcal{J}$ is the union of a finite number of open and disjoint intervals ($\alpha_0 = \bar{a}$ and $\beta_i = \bar{b}$ for any $i \in \mathbb{N}$ large enough). From (14), we deduce that on every sub-interval $]\alpha_i, \beta_i[$

$$\frac{d}{dt}\mathcal{V}(t,\mathbf{y}(t)) + \sigma_0^\varepsilon(t,\boldsymbol{\omega}(t),\boldsymbol{\theta}(t)) + \sigma_1^\varepsilon(t,\mathbf{b}(t)) \geq 0.$$

By integrating this inequality over $[\alpha_i, \beta_i]$, we get

$$\mathcal{V}(\alpha_i,\mathbf{y}(\alpha_i)) + \boldsymbol{\zeta}(\alpha_i) \leq \mathcal{V}(\beta_i,\mathbf{y}(\beta_i)) + \boldsymbol{\zeta}(\beta_i),$$

which implies

$$\mathcal{V}(\alpha_i,\mathbf{y}(\alpha_i)) + \boldsymbol{\zeta}(\alpha_i) \leq (\mathcal{V}(\beta_i,\mathbf{y}(\beta_i)) + \boldsymbol{\zeta}(\beta_i)) \bigvee \max_{t\in[\alpha_i,\beta_i]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)).$$

Besides, on every sub-interval $[\beta_i, \alpha_{i+1}]$ such that $\beta_i < \bar{b}$, we have:

$$\begin{aligned}
\mathcal{V}(\beta_i,\mathbf{y}(\beta_i)) + \boldsymbol{\zeta}(\beta_i) = \Phi(\beta_i,\mathbf{y}(\beta_i)) + \boldsymbol{\zeta}(b_i) &\leq \max_{t\in[\beta_i,\alpha_{i+1}]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)) \\
&\leq (\mathcal{V}(\alpha_{i+1},\mathbf{y}(\alpha_{i+1})) + \boldsymbol{\zeta}(\alpha_{i+1})) \bigvee \max_{t\in[\beta_i,\alpha_{i+1}]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)) \\
&\leq (\mathcal{V}(\beta_{i+1},\mathbf{y}(\beta_{i+1})) + \boldsymbol{\zeta}(\beta_{i+1})) \bigvee \max_{t\in[\beta_i,\beta_{i+1}]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)).
\end{aligned}$$

By induction, for any $i \in \mathbb{N}$ fixed we get then for any $k \in \mathbb{N}$

$$\mathcal{V}(\beta_i,\mathbf{y}(\beta_i)) + \boldsymbol{\zeta}(\beta_i) \leq (\mathcal{V}(\beta_{i+k},\mathbf{y}(\beta_{i+k})) + \boldsymbol{\zeta}(\beta_{i+k})) \bigvee \max_{t\in[\beta_i,\beta_{i+k}]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)).$$

Therefore, since $[\alpha_i, \beta_{i+k}] \subseteq [\bar{a}, \bar{b}]$ for any $i \in \mathbb{N}$ and $k \in \mathbb{N}$, we get

$$\mathcal{V}(\alpha_i,\mathbf{y}(\alpha_i)) + \boldsymbol{\zeta}(\alpha_i) \leq \big(\mathcal{V}(\beta_{i+k},\mathbf{y}(\beta_{i+k})) + \boldsymbol{\zeta}(\beta_{i+k})\big) \bigvee \max_{t\in[\bar{a},\bar{b}]}\big(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)\big).$$

Thus, for any $i \in \mathbb{N}$ fixed, letting $k \to +\infty$ we obtain

$$\mathcal{V}(\alpha_i,\mathbf{y}(\alpha_i)) + \boldsymbol{\zeta}(\alpha_i) \leq (\mathcal{V}(\bar{b},\mathbf{y}(\bar{b})) + \boldsymbol{\zeta}(\bar{b})) \bigvee \max_{t\in[\bar{a},\bar{b}]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)), \qquad \forall i \in \mathbb{N}.$$

If $\alpha_0 = \bar{a}$, then it follows that

$$\mathcal{V}(\bar{a},\mathbf{y}(\bar{a})) + \boldsymbol{\zeta}(\bar{a}) \leq (\mathcal{V}(\bar{b},\mathbf{y}(\bar{b})) + \boldsymbol{\zeta}(\bar{b})) \bigvee \max_{t\in[\bar{a},\bar{b}]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)). \tag{19}$$

This inequality can also be deduced if $\bar{a} < \alpha_0$. Indeed, it is enough to consider the case in which $\mathcal{J}$ is not the union of a finite number of open and disjoint intervals; recall that this case can be covered by the preceding analysis. In particular, we can find two strictly decreasing sequences $(\mathfrak{a}_j)_{j\in\mathbb{N}}$ and $(\mathfrak{b}_j)_{j\in\mathbb{N}}$ in $[0,T]$ such that $\bigcup_j]\mathfrak{a}_j, \mathfrak{b}_j[\subseteq \mathcal{J}$ with $\mathfrak{a}_0 = \alpha_0$, $\mathfrak{b}_0 = \beta_0$, $\mathfrak{a}_j < \mathfrak{b}_j$ and $\mathfrak{a}_j \searrow \bar{a}$ if $j \to +\infty$, satisfying as well

$$\mathcal{V}(\mathfrak{a}_j,\mathbf{y}(\mathfrak{a}_j)) = \Phi(\mathfrak{a}_j,\mathbf{y}(\mathfrak{a}_j)) \quad \text{and} \quad \mathcal{V}(\mathfrak{b}_j,\mathbf{y}(\mathfrak{b}_j)) = \Phi(\mathfrak{b}_j,\mathbf{y}(\mathfrak{b}_j)), \qquad \forall j \in \mathbb{N}.$$

Repeating the arguments exposed above, but using $\mathfrak{a}_{i+j}$ and $\mathfrak{b}_{i+j}$ instead of α_i and β_i, respectively, it follows that

$$\mathcal{V}(\mathfrak{a}_j,\mathbf{y}(\mathfrak{a}_j)) + \boldsymbol{\zeta}(\mathfrak{a}_j) \leq (\mathcal{V}(\bar{b},\mathbf{y}(\bar{b})) + \boldsymbol{\zeta}(\bar{b})) \bigvee \max_{t\in[\bar{a},\bar{b}]}(\Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t)), \qquad \forall j \in \mathbb{N}.$$

Therefore, letting $j \to +\infty$ we get (19) as claimed above.

Now, we have several cases to consider:

(1) Suppose that $\bar{a}=0$ and $\bar{b}=T$. Since $\mathcal{V}(T,x)=\Phi(T,x)\vee\Psi(x)$, then

$$\mathcal{V}(0,\mathbf{y}(0))+\boldsymbol{\zeta}(0)\leq(\Psi(\mathbf{y}(T))+\boldsymbol{\zeta}(T))\bigvee\max_{t\in[0,T]}(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(t)). \tag{20}$$

(2) If $\bar{a}=0$ and $\bar{b}<T$. In this case, we know that $\mathcal{V}(t,\mathbf{y}(t))=\Phi(t,\mathbf{y}(t))$ on $[\bar{b},T]$. Therefore, from (19), we get

$$\begin{aligned}\mathcal{V}(0,\mathbf{y}(0))+\boldsymbol{\zeta}(0) &\leq (\mathcal{V}(\bar{b},\mathbf{y}(\bar{b}))+\boldsymbol{\zeta}(\bar{b}))\bigvee\max_{t\in[0,\bar{b}]}(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(t))\\ &= \max_{t\in[0,\bar{b}]}(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(t))\leq\max_{t\in[0,T]}(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(t)).\end{aligned}$$

From this inequalities, we deduce that (20) holds also in this case.

(3) If $\bar{a}>0$, in this case we have $\mathcal{V}(t,\mathbf{y}(t))=\Phi(t,\mathbf{y}(t))$ on $[0,\bar{a}]$. Hence,

$$\begin{aligned}\mathcal{V}(0,\mathbf{y}(0))+\boldsymbol{\zeta}(0) &= (\Phi(0,\mathbf{y}(0))+\boldsymbol{\zeta}(0))\leq\max_{t\in[0,T]}(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(t))\\ &\leq \max_{t\in[0,T]}(\Phi(t,\mathbf{y}(t))+\boldsymbol{\zeta}(t))\bigvee(\Psi(\mathbf{y}(T))+\boldsymbol{\zeta}(T)).\end{aligned}$$

In all cases, we conclude that $J(\mathbf{y},\boldsymbol{\zeta},\mathbf{u},\boldsymbol{\omega},\boldsymbol{\theta},\mathbf{b})\geq 0$ which is the desired claim to prove in this step.

4.3. **Optimality conditions for the perturbed problem.** Since $(\bar{\mathbf{y}},\bar{\boldsymbol{\zeta}})$ is a strong local minimizer of the optimal control problem $(\mathcal{P}_\varepsilon)$, which is a min-max problem, such as $(\mathsf{P}_{\mathrm{m}})$, with the data

$$\bar{\mathsf{x}}=(\bar{\mathbf{y}},\bar{\boldsymbol{\zeta}}),\quad \bar{\mathsf{u}}=(\bar{\mathbf{u}},\bar{\boldsymbol{\omega}},\bar{\boldsymbol{\theta}},\bar{\mathbf{b}}),\quad \mathsf{U}(t)=U(t)\times[-\varepsilon,\varepsilon]\times\mathbb{B}_N(0;\varepsilon)\times\mathbb{B}_N(0;\varepsilon),$$

the dynamics

$$\mathcal{F}(t,(y,\zeta),(u,\omega,\theta,b))=((1+\omega)\mathbf{F}(t,y,u)+\theta+b,\sigma_0^\varepsilon(t,\omega,\theta)+\sigma_1^\varepsilon(t,b))$$

and the costs

$$\psi_0(y,\zeta)=-\mathcal{V}(0,y)-\zeta,\quad \psi_T(y,\zeta)=\Psi(y)+\zeta,\quad \varphi(t,(y,\zeta))=\Phi(t,y)+\zeta.$$

It is not difficult to see that the assumptions of Lemma 1 hold for these data, and so we can apply it to obtain a set of optimality conditions that depends on the penalization parameter ε. It follows then that, by Lemma 1 there exist $\mathbf{p}_y^\varepsilon\in W^{1,1}\left([0,T];\mathbb{R}^N\right)$, $\mathbf{p}_\zeta^\varepsilon\in W^{1,1}([0,T];\mathbb{R})$, $\lambda^\varepsilon\in[0,1]$, a finite regular (nonnegative) Borel measure μ^ε on $[0,T]$ and some Borel measurable functions $\gamma_y^\varepsilon:[0,T]\to\mathbb{R}^N$ and $\gamma_\zeta^\varepsilon:[0,T]\to\mathbb{R}^N$ such that:

(i-ε) $\lambda^\varepsilon+\mu^\varepsilon([0,T])=1$ with

$$\begin{cases}\lambda^\varepsilon=1 & \text{if } \Psi(\bar{\mathbf{y}}(T))>\max\limits_{t\in[0,T]}\Phi(t,\bar{\mathbf{y}}(t))\\ \lambda^\varepsilon=0 & \text{if } \Psi(\bar{\mathbf{y}}(T))<\max\limits_{t\in[0,T]}\Phi(t,\bar{\mathbf{y}}(t))\end{cases}$$

(ii-ε) $\gamma_y^\varepsilon(t)\in\partial_y\Phi(t,\bar{\mathbf{y}}(t))$ and $\gamma_\zeta^\varepsilon(t)=1$ for μ^ε-a.e. $t\in[0,T]$;

(iii-ε) $\operatorname{supp}(\mu^\varepsilon)\subseteq\left\{t\in[0,T]\mid\Phi(t,\bar{\mathbf{y}}(t))=\max\limits_{s\in[0,T]}\Phi(s,\bar{\mathbf{y}}(s))\right\}$.

(iv-ε) $-\dot{\mathbf{p}}_y^\varepsilon(t)=\mathbf{q}_y^\varepsilon(t)D_y\mathbf{F}(t,\bar{\mathbf{y}}(t),\bar{\mathbf{u}}(t))$ and $\dot{\mathbf{p}}_\zeta^\varepsilon(t)=0$ for a.e. $t\in[0,T]$;

(v-ε) $-\mathbf{p}_y^\varepsilon(0)\in\partial_y\mathcal{V}(0,\bar{\mathbf{y}}(0))$, $\mathbf{p}_\zeta^\varepsilon(0)=-1$, $-\mathbf{q}_y^\varepsilon(T)\in\lambda^\varepsilon\partial\Psi(\bar{\mathbf{y}}(T))$ and $-\mathbf{q}_\zeta^\varepsilon(T)=\lambda^\varepsilon$;

(vi-ε) $\mathbf{q}_y^\varepsilon(t) \cdot \mathbf{F}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t)) = \max_{(u,\omega,\theta,b)\in \mathsf{U}(t)} H_\varepsilon(t, \bar{\mathbf{y}}(t), \bar{\boldsymbol{\zeta}}(t), u, \theta, \omega, b, \mathbf{q}_y^\varepsilon(t), \mathbf{q}_\zeta^\varepsilon(t))$, a.e. on $[0,T]$, where

$$H_\varepsilon(t, y, \zeta, u, \theta, \omega, b, q_y, q_\zeta) := q_y \cdot ((1+\omega)\mathbf{F}(t, y, u) + \theta + b) + q_\zeta \left(\sigma_0^\varepsilon(t, \theta, \omega) + \sigma_1^\varepsilon(t, b)\right).$$

Here,

$$\mathbf{q}_y^\varepsilon(t) = \begin{cases} \mathbf{p}_y^\varepsilon(t) + \displaystyle\int_{[0,t[} \gamma_y^\varepsilon(s))\mu^\varepsilon(ds) & \text{if } t \in [0,T[, \\ \mathbf{p}_y^\varepsilon(T) + \displaystyle\int_{[0,T]} \gamma_y^\varepsilon(s))\mu^\varepsilon(ds) & \text{if } t = T, \end{cases}$$

and

$$\mathbf{q}_\zeta^\varepsilon(t) = \begin{cases} \mathbf{p}_\zeta^\varepsilon(t) + \displaystyle\int_{[0,t[} \gamma_\zeta^\varepsilon(s))\mu^\varepsilon(ds) & \text{if } t \in [0,T[, \\ \mathbf{p}_\zeta^\varepsilon(T) + \displaystyle\int_{[0,T]} \gamma_\zeta^\varepsilon(s))\mu^\varepsilon(ds) & \text{if } t = T. \end{cases}$$

Notice that from (iv-ε) and (v-ε) we get $\mathbf{p}_\zeta^\varepsilon \equiv -1$, and so from (ii-$\varepsilon$) we get

$$\mathbf{q}_\zeta^\varepsilon(t) = -1 + \mu^\varepsilon([0,t[) \quad \text{for any } t \in [0,T[. \tag{21}$$

Assertion (vi-ε) means that for every $(\omega, \theta, b) \in \mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^N$ with $|\omega| \le \varepsilon$, $\|\theta\| \le \varepsilon$, $\|b\| \le \varepsilon$, for a.e. $t \in [0,T]$ and every $u \in U(t)$, we have

$$(1+\omega)\mathbf{q}_y^\varepsilon(t) \cdot \mathbf{F}(t, \bar{\mathbf{y}}(t), u) + \mathbf{q}_y^\varepsilon(t) \cdot (\theta + b) + \mathbf{q}_\zeta^\varepsilon(t)(\sigma_0^\varepsilon(t, \theta, \omega) + \sigma_1^\varepsilon(t, b)) \le \mathbf{q}_y^\varepsilon(t) \cdot \mathbf{F}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t)). \tag{22}$$

In particular, taking $\theta = 0$, $\omega = 0$ and $b = 0$ in (22), it follows that for a.e. $t \in [0,T]$ and every $u \in U(t)$

$$\mathbf{q}_y^\varepsilon(t) \cdot \mathbf{F}(t, \bar{\mathbf{y}}(t), u) \le \mathbf{q}_y^\varepsilon(t) \cdot \mathbf{F}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t)),$$

that is, we get the maximality condition:

$$H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{q}_y^\varepsilon(t)) = \max_{u\in U(t)} H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), u, \mathbf{q}_y^\varepsilon(t)) \quad \text{a.e. on } [0,T]. \tag{23}$$

On the other hand, with $u = \bar{\mathbf{u}}(t)$ in (22), for every $(\omega, \theta) \in \mathbb{R}^{N+1}$ such that $|\omega| \le \varepsilon$, $\|\theta\| \le \varepsilon$ and $\|b\| \le \varepsilon$, we obtain for a.e. $t \in [0,T]$

$$\omega\mathbf{q}_y^\varepsilon(t) \cdot \mathbf{F}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t)) + \mathbf{q}_y^\varepsilon(t) \cdot (\theta + b) + \mathbf{q}_\zeta^\varepsilon(t)(\sigma_0^\varepsilon(t, \theta, \omega) + \sigma_1^\varepsilon(t, b)) \le 0. \tag{24}$$

Now, by setting $b = 0$, the inequality (24) implies that for every $(\omega, \theta) \in \mathbb{R}^{N+1}$ such that $|\omega| \le \varepsilon$, $\|\theta\| \le \varepsilon$,

$$\omega\mathbf{q}_y^\varepsilon(t) \cdot \mathbf{F}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t)) + \mathbf{q}_y^\varepsilon(t) \cdot \theta + \mathbf{q}_\zeta^\varepsilon(t)\sigma_0^\varepsilon(t, \theta, \omega) \le 0.$$

Notice that since $0 \le \mu^\varepsilon([0,T]) \le 1$, by (21), we deduce that $-\mathbf{q}_\zeta^\varepsilon \ge 0$. Thus, since the support function σ_0^ε is positively homogeneous on the variables (θ, ω) for $t \in [0,T]$ fixed, it follows that for every $(\omega, \theta) \in \mathbb{R}^{N+1}$ we have

$$(H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{q}_y^\varepsilon(t)), -\mathbf{q}_y^\varepsilon(t)) \cdot (\omega, -\theta) \le \sup_{(\alpha,\beta)\in -\mathbf{q}_\zeta^\varepsilon(t)G_0^\varepsilon(t)} (\alpha, \beta) \cdot (\omega, -\theta).$$

By [20, Theorem 13.1], we conclude that

$$\left(H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{q}_y^\varepsilon(t)), -\mathbf{q}_y^\varepsilon(t)\right) \in -\mathbf{q}_\zeta^\varepsilon(t)\overline{\operatorname{co}\,(G_0^\varepsilon(t))}, \quad a.e.\ t \in [0,T]. \tag{25}$$

On the other hand, if we set $\theta = 0$ and $\omega = 0$ in (24), and use the fact that the support function σ_1^ε is positively homogeneous on the variable b for $t \in [0,T]$ fixed, we obtain

$$-\mathbf{q}_y^\varepsilon(t) \cdot b \leq -\mathbf{q}_\zeta^\varepsilon(t)\sigma_1^\varepsilon(t,-b) = \sup_{\beta \in -\mathbf{q}_\zeta^\varepsilon(t) G_1^\varepsilon(t)} \beta \cdot b, \qquad \forall b \in \mathbb{R}^N.$$

Similarly as done before, by [20, Theorem 13.1] we conclude that

$$-\mathbf{q}_y^\varepsilon(t) \in -\mathbf{q}_\zeta^\varepsilon(t)\overline{\mathrm{co}\,(G_1^\varepsilon(t))}, \quad a.e.\ t \in [0,T]. \tag{26}$$

4.4. **Passage to the limit and conclusion.** All the previous properties were established for a given $\varepsilon \in]0,1]$. Now let us consider a monotonically decreasing sequence $\varepsilon_k \searrow 0$ and let us study the convergence when k goes to $+\infty$. In what follows we will use the same arguments as in [24, Chapter 9]. First notice that, since $\mathbf{p}_y^{\varepsilon_k}(0) \in \partial_x \mathcal{V}(0, \bar{\mathbf{y}}(0))$ for any $k \in \mathbb{N}$ and $y \mapsto \mathcal{V}(0,y)$ is locally Lipschitz continuous, the sequence $\{\mathbf{p}_y^{\varepsilon_k}(0)\}_{k\in\mathbb{N}}$ is bounded. Moreover, from (iv-ε), the definition of $\mathbf{q}_y^{\varepsilon_k}$ and $(\mathbf{H}_1)$, it follows that

$$\|\dot{\mathbf{p}}_y^{\varepsilon_k}(t)\| \leq k_{\mathbf{F}} \left(\|\mathbf{p}_y^{\varepsilon_k}(t)\| + \int_{[0,t[} \|\gamma_y^{\varepsilon_k}(s)\| \mu^{\varepsilon_k}(ds) \right), \qquad \text{for a.e. } t \in [0,T].$$

By the non-triviality condition and the fact that Φ is Locally Lipschitz continuous, we get that the integral terms in the above inequality is uniformly bounded w.r.t. k. Therefore, by Gronwall's inequality and [2, Theorem 0.3.4], it follows that, along a sub-sequence which we do not relabel, $\mathbf{p}_y^{\varepsilon_k}$ converges uniformly to some $\mathbf{p} \in W^{1,1}\left([0,T];\mathbb{R}^N\right)$ and $\dot{\mathbf{p}}_y^{\varepsilon_k}$ converges weakly in $L^1\left([0,T];\mathbb{R}^N\right)$ to $\dot{\mathbf{p}}$, when $k \to +\infty$.

Besides, since $\{\lambda^{\varepsilon_k}\}_{k\in\mathbb{N}}$ is bounded and $0 \leq \mu^{\varepsilon_k}([0,T]) \leq 1$ for any $k \in \mathbb{N}$, we deduce that for a sub-sequence (which again we do not relabel), there exist $\lambda \in [0,1]$ and a finite regular (nonnegative) Borel measure μ such that

$$\lambda^{\varepsilon_k} \to \lambda, \quad \text{and} \quad \mu^{\varepsilon_k} \rightharpoonup^* \mu \text{ for the weak}^* \text{ topology on } C\left([0,T];\mathbb{R}\right)^*.$$

In particular, it follows that $\mu^{\varepsilon_k}([0,T]) \to \mu([0,T])$, and so we recover the non-triviality condition (MP$_i$). It also follows that for any relatively open set $B \subseteq [0,T]$, we have

$$\mu(B) \leq \liminf_{k\to+\infty} \mu^{\varepsilon_k}(B).$$

Therefore, by (iii-ε) we obtain (MP$_{iii}$). Furthermore, since $\partial_y \Phi(t, \bar{\mathbf{y}}(t))$ is a compact non-empty set and $(\gamma_y^{\varepsilon_k})_n$ is a sequence of Borel measurable functions satisfying $\gamma_y^{\varepsilon_k}(t) \in \partial_y \Phi(t, \bar{\mathbf{y}}(t))$, by [24, Proposition 9.2.1], we can arrange another sub-sequence extraction (which we do not relabel) such that

$$\mu^{\varepsilon_k}(ds) \rightharpoonup^* \mu(ds) \quad \text{and} \quad \gamma_y^{\varepsilon_k}(s)\mu^{\varepsilon_k}(ds) \rightharpoonup^* \gamma(s)\mu(ds),$$

for the weak* topology on $C\left([0,T];\mathbb{R}\right)^*$, where $\gamma : [0,T] \longrightarrow \mathbb{R}^N$ is a Borel measurable function that satisfies

$$\gamma(t) \in \partial_y \Phi(t, \bar{\mathbf{y}}(t)) \text{ for } \mu\text{-a.e. } t \in [0,T].$$

Notice that for a.e. $t \in [0,T[$

$$\int_{[0,t[} \gamma^{\varepsilon_k}(s)\mu^{\varepsilon_k}(ds) \to \int_{[0,t[} \gamma(s)\mu(ds) \quad \text{and} \quad \mu^{\varepsilon_k}([0,t[) \to \mu([0,t[).$$

Thus, by defining the function $\mathbf{q}$ via the formula

$$\mathbf{q}(t) = \begin{cases} \mathbf{p}(t) + \displaystyle\int_{[0,t[} \gamma(s)\mu(ds) & \text{if } t \in [0,T[, \\ \mathbf{p}(T) + \displaystyle\int_{[0,T]} \gamma(s)\mu(ds) & \text{if } t = T, \end{cases}$$

we deduce with the help of [2, Theorem 1.4.1], that $(\mathbf{p}, \mathbf{q})$ verifies

$$\begin{cases} -\dot{\mathbf{p}}(t) = \mathbf{q}(t) D_y \mathbf{F}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t)) & \text{for a.e. } t \in [0,T], \\ \mathbf{q}(T) \in -\lambda \partial \Psi(\bar{\mathbf{y}}(T)). \end{cases}$$

Now, we can pass to the limit in (25) and in (26) using similar arguments as in [23]. For this, we set

$$\nu(t) := \begin{cases} 1 - \mu([0,t[) & \text{if } t \in [0,T[, \\ 1 - \mu([0,T]) & \text{if } t = T. \end{cases}$$

Let $S \subseteq [0,T]$ of full measure such that (25) and (26) hold for each $\varepsilon = \varepsilon_k$ and such that

$$\mathbf{q}_y^{\varepsilon_k}(t) \to \mathbf{q}(t) \quad \text{and} \quad \mu^{\varepsilon_k}([0,t[) \to \mu([0,t[), \qquad \forall t \in S.$$

In particular, we may also assume from (21) that $-\mathbf{q}_\zeta^{\varepsilon_k}(t) \to \nu(t)$ and

$$\left(H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{q}_y^\varepsilon(t)), -\mathbf{q}_y^\varepsilon(t)\right) \to \left(H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{q}(t)), -\mathbf{q}(t)\right), \qquad \forall t \in S.$$

Notice that ν is a non-negative function. So, let us consider first the case in which $\nu(t) > 0$ for all $t \in S$. We may assume as well, from (21), that $-\mathbf{q}_\zeta^{\varepsilon_k}(t) > 0$ for any $t \in S$.

Take $t \in S$ fixed. From (25) and (26) we get

$$\frac{1}{-\mathbf{q}_\zeta^{\varepsilon_k}(t)} \left(H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{q}_y^{\varepsilon_k}(t)), -\mathbf{q}_y^{\varepsilon_k}(t)\right) \in \overline{\mathrm{co}\,(G_0^{\varepsilon_k}(t))}, \qquad \forall k \in \mathbb{N} \tag{27}$$

and

$$-\frac{1}{-\mathbf{q}_\zeta^{\varepsilon_k}(t)} \mathbf{q}_y^{\varepsilon_k}(t) \in \overline{\mathrm{co}\,(G_1^{\varepsilon_k}(t))}, \qquad \forall k \in \mathbb{N}. \tag{28}$$

Notice that by definition we have

$$G_0^{\varepsilon_{k+1}}(t) \subseteq G_0^{\varepsilon_k}(t) \quad \text{and} \quad G_1^{\varepsilon_{k+1}}(t) \subseteq G_1^{\varepsilon_k}(t), \qquad \forall k \in \mathbb{N}.$$

Therefore, the pointwise convergence of the multipliers combined with (27) and (28), lead to

$$(\mathbf{H}(t), -\mathbf{q}(t)) \in \nu(t) \bigcap_{k \in \mathbb{N}} \overline{\mathrm{co}\,(G_0^{\varepsilon_k}(t))} \quad \text{and} \quad -\mathbf{q}(t) \in \nu(t) \bigcap_{k \in \mathbb{N}} \overline{\mathrm{co}\,(G_1^{\varepsilon_k}(t))}. \tag{29}$$

where $\mathbf{H}(t) = H^{\mathbf{F}}(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{q}(t))$ for any $t \in S$. Assume now that

$$(\mathbf{H}(t), -\mathbf{q}(t)) \notin \nu(t) \partial \mathcal{V}(t, \bar{\mathbf{y}}(t)).$$

Since $\nu(t) \partial \mathcal{V}(t, \bar{\mathbf{y}}(t))$ is a compact convex and nonempty subset of $\mathbb{R} \times \mathbb{R}^N$ and (4) holds, there are $\omega \in \mathbb{R}$, $\theta \in \mathbb{R}^N$ and $\rho > 0$ such that

$$\mathbf{H}(t)\omega + \mathbf{q}(t) \cdot \theta - \rho > \nu(t) D^\circ \mathcal{V}((t, \bar{\mathbf{y}}(t)); (\omega, -\theta)).$$

Moreover, given that the generalized directional derivative is upper semicontinuous ([11, 10.2 Proposition]), it follows that there is $\bar{k} \in \mathbb{N}$ so that

$$\frac{1}{\nu(t)}\left(\mathbf{H}(t)\omega + \mathbf{q}(t)\cdot\theta - \rho\right) > D^{\circ}\mathcal{V}((t,x);(\omega,-\theta)), \qquad \forall x \in \mathbb{B}_N(\bar{\mathbf{y}}(t);\varepsilon_{\bar{k}}).$$

Therefore, from (4) and the preceding inequality we get

$$\mathbf{H}(t)\omega + \mathbf{q}(t)\cdot\theta > \rho + \nu(t) \sup_{(\alpha,\beta)\in G_0^{\varepsilon_{\bar{k}}}(t)} (\alpha,\beta)\cdot(\omega,-\theta).$$

The latter inequality implies that

$$(\mathbf{H}(t), -\mathbf{q}(t)) \notin \nu(t)\overline{\mathrm{co}\left(G_0^{\varepsilon_{\bar{k}}}(t)\right)},$$

which, in the light of (29), is absurd. Therefore, the global sensibility relation holds true for any $t \in S$ such that $\nu(t) > 0$.

On the other hand, if for some $\tau \in S$ we have $\nu(\tau) = 0$, then since $G_0^{\varepsilon_k}(\tau)$ is uniformly bounded w.r.t. $k \in \mathbb{N}$, from (25) it follows that

$$(\mathbf{H}(\tau), -\mathbf{q}(\tau)) = 0 \in \nu(\tau)\partial\mathcal{V}(t, \bar{\mathbf{y}}(t)).$$

Thus, the global sensibility relation holds as well in this case.

A similar argument can be used to prove the partial sensitivity relation

$$-\mathbf{q}(t) \in \nu(t)\partial_x\mathcal{V}(t,\bar{\mathbf{y}}(t)), \qquad \text{a.e. on } [0,T].$$

This concludes therefore the proof of Theorem 3.1.

5. Application to problems with state constraints

Consider the following control problem with state constraints:

$$(\mathrm{P_{sc}}) \qquad \begin{cases} \text{Minimize } \ \psi(\mathbf{x}(T)) + \displaystyle\int_0^T \ell(t,\mathbf{x}(t),\mathbf{u}(t))\,dt, \\ \qquad \dot{\mathbf{x}}(t) = f(t,\mathbf{x}(t),\mathbf{u}(t)), \text{ a.e. on } [0,T] \\ \qquad \varphi(t,\mathbf{x}(t)) \le 0 \quad \text{for every } s \in [0,T], \\ \qquad \mathbf{x}(0) = x_0 \\ \qquad \mathbf{u} \in \mathcal{U}, \end{cases}$$

Our goal in this section is to derive sensitivity relations for this type of problems. The main difficulty one may encounter in doing so is that, in general, the value function associated with problem ($\mathrm{P_{sc}}$) is not continuous, or even finite on the state constraint set $\{\varphi \le 0\}$. It is well-known that to obtain the Lipschitz character of the value function one may be forced to adopt rather strong controllability assumptions. To avoid these issues we consider an auxiliary optimal control problem of minimax type, whose main feature is that its value function is Lipschitz continuous under mild assumptions.

Following ideas introduced in [1], we consider the auxiliary minimax optimal control problem

$$(\mathrm{P_{aux}}) \qquad \begin{cases} \text{Minimize } \ \displaystyle\max_{t\in[0,T]} \varphi(t,\mathbf{x}(t)) \bigvee (\psi(\mathbf{x}(T)) - \mathbf{z}(T)) \\ \qquad \dot{\mathbf{x}}(t) = f(t,\mathbf{x}(t),\mathbf{u}(t)), \text{ a.e. on } [0,T] \\ \qquad \dot{\mathbf{z}}(t) = -\ell(t,\mathbf{x}(t),\mathbf{u}(t)), \text{ a.e. on } [0,T] \\ \qquad \mathbf{x}(0) = x_0, \ , \mathbf{z}(0) = z_0 \\ \qquad \mathbf{u} \in \mathcal{U}, \end{cases}$$

Notice that the connection between (P_{sc}) and (P_{aux}) is the following: Let $\mathbf{u} \in \mathcal{U}$ be a given control and $\mathbf{x}$ be the corresponding solution of the dynamical equation that defines (P_{sc}). Let $\vartheta_0(x_0)$ and $\omega_0(x_0, z_0)$ be the value of the optimization problem (P_{sc}) and (P_{aux}), respectively. It follows then that if $\mathbf{u}$ is an optimal control for (P_{sc}) then it is also optimal for (P_{aux}) provided $z_0 = \vartheta_0(x_0)$. In this case, we also have $\omega_0(x_0, \vartheta_0(x_0)) = 0$. Notice that the condition $\vartheta_0(x_0) < +\infty$ is equivalent to saying that the (P_{sc}) is feasible.

The sensitivity relations for problem (P_{sc}) can be then written in terms of

$$\mathcal{W}(t,x,z) := \inf_{\mathbf{u}\in\mathcal{U}} \left\{ \max_{s\in[t,T]} \varphi(s,\mathbf{x}(s)) \bigvee (\psi(\mathbf{x}(T)) - \mathbf{z}(T)) \mid \begin{array}{l} \text{for a.e. } s \in [t,T]: \\ \dot{\mathbf{x}}(s) = f(s,\mathbf{x}(s),\mathbf{u}(s)), \\ \dot{\mathbf{z}}(s) = -\ell(s,\mathbf{x}(s),\mathbf{u}(s)), \\ \text{and } \mathbf{x}(t) = x,\ \mathbf{z}(t) = z \end{array} \right\}$$

the value function of the problem (P_{aux}) and the *Hamiltonian* function $H_\lambda : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ defined for any $\lambda \in \mathbb{R}$ via the formula

$$H_\lambda(t,x,u,q) := q \cdot f(t,x,u) - \lambda \ell(t,x,u), \qquad \forall (t,x,u,q) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n.$$

Theorem 5.1. *Assume that* $(\mathbf{H}_2)$ *holds in addition to*

($\mathbf{C}_1$) *(a) For every $x \in \mathbb{R}^n$, $(t,u) \mapsto (f(t,x,u), \ell(t,x,u))$ is $\mathcal{L} \times \mathcal{B}^m$ measurable on $[0,T] \times \mathbb{R}^m$;*

(b) there exist $k_f, k_\ell > 0$ such that

$$\|f(t,x,u) - f(t,y,u)\| \le k_f \|x - y\| \qquad \forall x,y \in \mathbb{R}^n,\ (t,u) \in \mathrm{Gr}(U),$$
$$|\ell(t,x,u) - \ell(t,y,u)| \le k_\ell \|x - y\| \qquad \forall x,y \in \mathbb{R}^n,\ (t,u) \in \mathrm{Gr}(U);$$

(c) there exist $c_f, c_\ell > 0$ and $\alpha \in \mathbb{N}$ such that

$$\|f(t,x,u)\| \le \ c_f(1 + \|x\|), \quad \forall x \in \mathbb{R}^n,\ (t,u) \in \mathrm{Gr}(U).$$
$$|\ell(t,x,u)| \le \ c_\ell(1 + \|x\|^\alpha), \quad \forall x \in \mathbb{R}^n,\ (t,u) \in \mathrm{Gr}(U).$$

($\mathbf{C}_2$) *φ and ψ are locally Lipschitz continuous.*

Let $x_0 \in \mathbb{R}^n$ be such that $\vartheta_0(x_0) < +\infty$, and assume that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{X}_{[0,T]}(x_0)$ is a solution to the control problem (P_{sc}). *Assume as well that*

(1) $x \mapsto (f(t,x,u), \ell(t,x,u))$ is continuously differentiable for any $(t,u) \in \mathrm{Gr}(U)$ fixed;
(2) there is a finite set $D \subseteq [0,T]$ such that the map $(t,u) \mapsto (f(t,x,u), \ell(t,x,u))$ is continuous on $\{(t,u) \in \mathrm{Gr}(U) \mid t \notin D\}$ for every $x \in \mathbb{R}^n$ fixed;
(3) U is a locally selectionable set-valued map.

Let $\bar{\mathbf{z}}(t) = \vartheta(t, \bar{\mathbf{x}}(t))$ for every $t \in [0,T]$. Then there exist $\mathbf{p} \in W^{1,1}([0,T];\mathbb{R}^n)$, $\lambda \in [0,1]$, a finite regular (nonnegative) Borel measure μ on $[0,T]$ and a Borel measurable function $\gamma : [0,T] \to \mathbb{R}^N$ such that:

($SC\text{-}MP_i$) Non-triviality: $\lambda + \mu([0,T]) = 1$ with $\lambda = 1$ if $\max_{t\in[0,T]} \varphi(t,\bar{\mathbf{x}}(t)) < 0$;

($SC\text{-}MP_{ii}$) $\gamma(t) \in \partial_x \varphi(t,\bar{\mathbf{x}}(t))$ for μ-a.e. $t \in [0,T]$;

($SC\text{-}MP_{iii}$) $\mathrm{supp}(\mu) \subseteq \{t \in [0,T] \mid \varphi(t,\bar{\mathbf{x}}(t)) = 0\}$;

($SC\text{-}MP_{iv}$) Costate equation and transversality condition:

$$\begin{cases} -\dot{\mathbf{p}}(t) = \partial_x H_\lambda(t,\bar{\mathbf{x}}(t),\bar{\mathbf{u}}(t),\mathbf{q}(t)) & \text{for a.e. } t \in [0,T], \\ -\mathbf{q}(T) \in \lambda \partial \psi(\bar{\mathbf{x}}(T)), \end{cases}$$

where

$$\mathbf{q}(t) = \begin{cases} \mathbf{p}(t) + \int_{[0,t[} \gamma(s)\mu(ds) & \text{if } t \in [0,T[\\ \mathbf{p}(T) + \int_{[0,T]} \gamma(s)\mu(ds) & \text{if } t = T \end{cases}$$

(SC-MP$_v$) *Maximality condition:*

$$H_\lambda(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \mathbf{q}(t)) = \max_{u \in U(t)} H_\lambda(t, \bar{\mathbf{x}}(t), u, \mathbf{q}(t)), \qquad \text{for a.e. } t \in [0,T];$$

(SC-MP$_{vi}$) *Sensitivity relations:* $(-\mathbf{p}(0), -\lambda) \in \partial_{(x,z)} \mathcal{W}(0, x_0, z_0)$, *where* $z_0 = \vartheta_0(x_0)$, *and for a.e.* $t \in [0,T]$

$$(-\mathbf{q}(t), -\lambda) \in \nu(t)\partial_{(x,z)}\mathcal{W}(t, \bar{\mathbf{x}}(t), \bar{\mathbf{z}}(t)),$$
$$(H_\lambda(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \mathbf{q}(t)), -\mathbf{q}(t), -\lambda) \in \nu(t)\partial\mathcal{W}(t, \bar{\mathbf{x}}(t), \bar{\mathbf{z}}(t)),$$

where $\nu(s) = (1 - \mu([0,s[)$ *for any* $s \in [0,T]$.

Proof. Let $r \in \mathbb{R}$ such that

$$r > e^{c_f T}(\|x_0\| + 1)$$

and let us define

$$\mathbf{F}(t,(x,z),u) := \begin{cases} (f(t,x,u), -\ell(t,x,u)) & \text{if } \|x\| < r \\ \left((f(t,x,u), -\ell\left(t, \frac{rx}{\|x\|}, u\right)\right) & \text{if } \|x\| \geq r \end{cases} \tag{30}$$

Using the fact that the mapping

$$x \mapsto \begin{cases} x & \text{if } \|x\| < r \\ \frac{rx}{\|x\|} & \text{if } \|x\| \geq r \end{cases}$$

is Lipschitz continuous on $\mathbb{R}^n$, it is not difficult to see that $(\mathbf{C}_1)$ implies that $\mathbf{F}$ satisfies $(\mathbf{H}_1)$, with $y = (x,z)$. Let $\mathbf{y} = (\mathbf{x}, \mathbf{z})$ be a solution of (8) defined on $[0,T]$ such that $\mathbf{y}(0) = y_0$, where $y_0 = (x_0, z_0)$. By Gronwall's inequality we have

$$\|\mathbf{x}(t)\| < \|\mathbf{x}(t)\| + 1 \leq e^{c_f T}(\|x_0\| + 1) < r, \qquad \forall t \in [0,T].$$

This means in particular that

$$\mathbf{z}(t) = z_0 - \int_0^t \ell(s, \mathbf{x}(s), \mathbf{u}(s))ds, \qquad \forall t \in [0,T].$$

It follows then that $\bar{\mathbf{y}} = (\bar{\mathbf{x}}, \bar{\mathbf{z}})$ is a minimizer of (P) with the dynamics $\mathbf{F}$ given by (30) and the cost

$$\Phi(t,(x,z)) = \varphi(t,x) \quad \text{and} \quad \Psi(x,z) = \psi(x) - z, \qquad \forall (x,z) \in \mathbb{R}^{n+1}.$$

Notice too that for every $(t,u) \in \mathrm{Gr}(U)$, $(x,z) \mapsto \mathbf{F}(t,(x,z),u)$ is continuously differentiable on a tube around $\bar{\mathbf{y}} = (\bar{\mathbf{x}}, \bar{\mathbf{z}})$, because $\mathbf{F}$ agrees with $(f, -\ell)$ on $\mathbb{B}_N(0,r)$ and $x \mapsto (f(t,x,u), \ell(t,x,u))$ is continuously differentiable on a tube around $\bar{x}$. Therefore, by Theorem 3.1 we can deduce that there exist $\mathbf{p}_x \in W^{1,1}([0,T];\mathbb{R}^n)$, $\mathbf{p}_z \in W^{1,1}([0,T];\mathbb{R})$, $\lambda \in [0,1]$, a finite regular (nonnegative) Borel measure μ on $[0,T]$ and measurable functions $\gamma_x : [0,T] \to \mathbb{R}^n$ and $\gamma_z : [0,T] \to \mathbb{R}$ such that

(I) Non-triviality: $\lambda + \mu([0,T]) = 1$ with

$$\begin{cases} \lambda = 1 & \text{if } \psi(\bar{\mathbf{x}}(T)) - \vartheta(0,x_0) + \displaystyle\int_0^T \ell(t,\bar{\mathbf{x}}(t),\bar{\mathbf{u}}(t))dt > \max_{t\in[0,T]} \varphi(t,\bar{\mathbf{x}}(t)) \\ \lambda = 0 & \text{if } \psi(\bar{\mathbf{x}}(T)) - \vartheta(0,x_0) + \displaystyle\int_0^T \ell(t,\bar{\mathbf{x}}(t),\bar{\mathbf{u}}(t))dt < \max_{t\in[0,T]} \varphi(t,\bar{\mathbf{x}}(t)) \end{cases}$$

(II) $(\gamma_x(t),\gamma_z(t)) \in \partial\varphi_x(t,\bar{\mathbf{x}}(t)) \times \{0\}$ for μ-a.e. $t \in [0,T]$;

(III) $\operatorname{supp}(\mu) \subseteq \left\{ t \in [0,T] \mid \varphi(t,\bar{\mathbf{x}}(t)) = \max_{s\in[0,T]} \varphi(s,\bar{\mathbf{x}}(s)) \right\}$

(IV) Costate equation and transversality condition:

$$\begin{cases} -\dot{\mathbf{p}}_x(t) = \nabla_x H^{\mathbf{F}}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t),\bar{\mathbf{u}}(t),\mathbf{q}_x(t),\mathbf{q}_z(t)) & \text{for a.e. } t\in[0,T], \\ -\dot{\mathbf{p}}_z(t) = \nabla_z H^{\mathbf{F}}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t),\bar{\mathbf{u}}(t),\mathbf{q}_x(t),\mathbf{q}_z(t)) & \text{for a.e. } t\in[0,T], \\ (-\mathbf{q}_x(T),-\mathbf{q}_z(T)) \in \lambda\partial\psi(\bar{\mathbf{x}}(T)) \times \{-1\}, \end{cases}$$

where

$$(\mathbf{q}_x(t),\mathbf{q}_z(t)) = \begin{cases} (\mathbf{p}_x(t),\mathbf{p}_z(t)) + \displaystyle\int_{[0,t[} (\gamma_x(t),\gamma_z(t))\mu(ds) & \text{if } t\in[0,T[, \\ (\mathbf{p}_x(T),\mathbf{p}_z(T)) + \displaystyle\int_{[0,T]} (\gamma_x(t),\gamma_z(t))\mu(ds) & \text{if } t = T; \end{cases}$$

(V) Maximality condition: for a.e. $t \in [0,T]$

$$H^{\mathbf{F}}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t),\bar{\mathbf{u}}(t),\mathbf{q}_x(t),\mathbf{q}_z(t)) = \max_{u\in U(t)} H^{\mathbf{F}}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t),\bar{\mathbf{u}}(t),\mathbf{q}_x(t),\mathbf{q}_z(t)).$$

(VI) Sensitivity relations: $-(\mathbf{p}_x(0),\mathbf{p}_z(0)) \in \partial_{(x,z)}\mathcal{W}(0,x_0,z_0)$ and for a.e. $t \in [0,T]$

$$(-\mathbf{p}_x(t),-\mathbf{p}_z(t)) \in \nu(t)\partial_{(x,z)}\mathcal{W}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t)),$$
$$(H^{\mathbf{F}}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t),\bar{\mathbf{u}}(t),\mathbf{q}_x(t),\mathbf{q}_z(t)),-\mathbf{q}_x(t),-\mathbf{q}_z(t)) \in \nu(t)\partial\mathcal{W}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t)).$$

Now, since

$$\vartheta(0,x_0) = \psi(\bar{\mathbf{x}}(T)) + \int_0^T \ell(t,\bar{\mathbf{x}}(t),\bar{\mathbf{u}}(t))dt \quad \text{and} \quad \max_{t\in[0,T]} \varphi(t,\bar{\mathbf{x}}(t)) \le 0,$$

from (I) and (III), we obtain (SC-MP$_i$) and (SC-MP$_{iii}$). Also, from (II), (IV) and the fact that the dynamics doesn't depend on the z variable, we have that $\mathbf{p}_z = \mathbf{q}_z \equiv \lambda$, and consequently,

$$H^{\mathbf{F}}(t,\bar{\mathbf{x}}(t),\bar{\mathbf{z}}(t),\bar{\mathbf{u}}(t),\mathbf{q}_x(t),\mathbf{q}_z(t)) = H_\lambda(t,\bar{\mathbf{x}}(t),\bar{\mathbf{u}}(t),\mathbf{q}_x(t)), \quad \text{a.e. on } [0,T].$$

Therefore, defining $\mathbf{q} := \mathbf{q}_x$, $\mathbf{p} := \mathbf{p}_x$ and $\gamma := \gamma_x$ we get from (II), (IV), (V), (VI) the remainder conditions (SC-MP$_{ii}$), (SC-MP$_{iv}$), (SC-MP$_v$), (SC-MP$_{vi}$) on the statement of the theorem. □ □

6. Discussion about the assumptions

To conclude, we provide a short discussion about the assumptions we have done in our main results. The focus will be on Theorem 3.1, however it extends in a natural way to Theorem 5.1.

The hypotheses $(\mathbf{H}_1)$ and $(\mathbf{H}_2)$ are rather standard assumptions made in the literature over the data that comprise the control system, although $(\mathbf{H}_1)$(b) can be weakened to be satisfied in a local sense. Moreover, the hypothesis $(\mathbf{H}_3)$ is somewhat essential in our setting since, otherwise, it may not be possible to ensure the Lipschitz continuous character of the value function.

One of the main contributions of this work is that we are able to provide a global and a partial sensitivity relation with a single perturbed problem, that are valid for the same costate. To accomplish this, it is fundamental that one is able to ensure the existence of continuously differentiable trajectories of the control system for any given initial state and initial velocity. At this point is where assumptions (A_2) and (A_3) start playing a major role. If one drops these assumptions, it is still possible to get the global sensitivity relation, but not the partial one. To see this one needs to set the control $\mathbf{b} \equiv 0$ in the perturbed problem $(\mathcal{P}_\varepsilon)$, use [25, Lemma 4.3] to argue the existence a subset $\Theta \subseteq [0,T]$ of full measure such that for every $(t,x) \in \Theta \times \mathbb{R}^N$ and $v \in \mathrm{co}(\mathbf{F}(t,x,U(t)))$ there is $(\mathbf{y},\mathbf{u}) \in \mathbb{X}^{\mathbf{F}}_{[t,T]}(x)$ for which

$$\lim_{h\downarrow 0} \left\| \frac{\mathbf{y}(t+h)-x}{h} - v \right\| = 0.$$

And finally, instead of (16) use

$$\frac{d}{dt}\mathcal{V}(t,\mathbf{y}) \geq -\, D^\circ\mathcal{V}((t,\mathbf{y});(\boldsymbol{\omega},-\boldsymbol{\theta})) + (1+\boldsymbol{\omega})D^\downarrow\mathcal{V}((t,\mathbf{y});(1,\mathbf{F}(t,\mathbf{y},\mathbf{u})).$$

A similar argument, but setting the control $(\boldsymbol{\omega},\boldsymbol{\theta}) \equiv (0,0)$ in the perturbed problem $(\mathcal{P}_\varepsilon)$ may be used to get the partial sensitivity relation. However, the problem in this case is that, contrary to the lower Dini derivative, in general one does not have

$$D^\circ\mathcal{V}((t,\mathbf{y}(t));(0,\mathbf{b})) = D^\circ_x\mathcal{V}(t,\cdot)(\mathbf{y}(t);\mathbf{b}),$$

a fact that is essential in the proof of Lemma 3.

Finally, let us point out that the smoothness assumption over the dynamics (A_1) in Theorem 3.1 has been done only for the sake of the exposition. A nonsmooth version of Theorem 3.1, would have instead of (MP_{iv}) the following differential inclusion

$$-\dot{\mathbf{p}}(t) \in \mathrm{co}\,\partial_y H^{\mathbf{F}}(t,\bar{\mathbf{y}}(t),\bar{\mathbf{u}}(t),\mathbf{q}(t)), \quad \text{for a.e. } t\in[0,T],$$

which can be obtained by considering the following perturbed optimal control problem with state constraints

$$(\mathcal{P}'_\varepsilon)\quad \begin{cases} \text{Minimize} & J'(\mathbf{y},\boldsymbol{\zeta},\mathbf{z}) \\ \text{over all} & (\mathbf{y},\boldsymbol{\zeta},\mathbf{z}) \in W^{1,1}\left([0,T];\mathbb{R}^{N+2}\right),\ \mathbf{u}\in\mathcal{U} \text{ and measurable} \\ & \ \text{functions } (\boldsymbol{\omega},\boldsymbol{\theta},\mathbf{b}) : [0,T] \to \mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^N \\ \text{such that} & \dot{\mathbf{y}}(t) = (1+\boldsymbol{\omega}(t))\mathbf{F}(t,\mathbf{y}(t),\mathbf{u}(t)) + \boldsymbol{\theta}(t) + \mathbf{b}(t), \\ & \dot{\boldsymbol{\zeta}}(t) = \sigma_0^\varepsilon(t,\boldsymbol{\omega}(t),\boldsymbol{\theta}(t)) + \sigma_1^\varepsilon(t,\mathbf{b}(t)), \\ & \dot{\mathbf{z}}(t) = 0, \\ & |\boldsymbol{\omega}(t)| \leq \varepsilon,\ \|\boldsymbol{\theta}(t)\| \leq \varepsilon,\ \|\mathbf{b}(t)\| \leq \varepsilon, \\ & \text{for a.e. } t\in[0,T], \\ & \Phi(t,\mathbf{y}(t)) + \boldsymbol{\zeta}(t) - \mathbf{z}(t) \leq 0, \quad \text{for all } t\in[0,T]. \end{cases}$$

where the perturbed cost function is now defined via the formula:

$$J'(\mathbf{y},\boldsymbol{\zeta},\mathbf{z}) := \left[\mathbf{z}(T) \bigvee (\Psi(\mathbf{y}(T)) + \boldsymbol{\zeta}(T))\right] - \mathcal{V}(0,\mathbf{y}(0)) - \boldsymbol{\zeta}(0).$$

It is not difficult to see that setting $\bar{\boldsymbol{\zeta}} \equiv 0$, $\bar{\boldsymbol{\omega}} \equiv 0$, $\bar{\boldsymbol{\theta}} \equiv 0$ and $\bar{\mathbf{b}} \equiv 0$ as before and

$$\bar{\mathbf{z}} \equiv \max_{t\in[0,T]} \Phi(t, \bar{\mathbf{y}}(t)),$$

for any given $\varepsilon \in]0,1]$, we have that $((\bar{\mathbf{y}}, \bar{\boldsymbol{\zeta}}, \bar{\mathbf{z}}), (\bar{\mathbf{u}}, \bar{\boldsymbol{\omega}}, \bar{\boldsymbol{\theta}}, \bar{\mathbf{b}}))$ is a strong local minimizer of the problem $(\mathcal{P}'_\varepsilon)$; similar arguments used to prove that the functional J is non-negative can be used to show that the functional J' is also non-negative for any feasible trajectory. Finally, instead of using Lemma 1 for getting a set of optimality conditions for the perturbed problem, one needs to use the nonsmooth maximum principle for problems with state constraints ([24, Theorem 9.3.1]). The convergence analysis is rather standard and almost the same we have provided in our proof.

Appendix A. Proof of Lemma 2

Proof. Let $r > 0$ and $L_{\Phi,\Psi} \geq 0$ be a common Lipschitz constants for $\Phi(t,\cdot)$ and Ψ on the closed ball of radius $\tilde{r} = (1+r)e^{c_{\mathbf{F}}T} - 1$. Let us define $e(k_{\mathbf{F}}) := \exp(k_{\mathbf{F}}T)$, where $k_{\mathbf{F}} > 0$ is the Lipschitz modulus of $\mathbf{F}$ given by $(\mathbf{H}_1)$.

Notice that for any $t \in [0,T]$, $x \in \mathbb{B}_N(0;r)$ and $(\mathbf{y},\mathbf{u}) \in \mathbb{X}^{\mathbf{F}}(t,x)$, we have by (9) that $\mathbf{y}(s) \in \mathbb{B}_N(0;\tilde{r})$ for any $s \in [t,T]$. Similarly, $\tilde{\mathbf{y}}(s) \in \mathbb{B}_N(0;\tilde{r})$ for any $(\tilde{\mathbf{y}},\tilde{\mathbf{u}}) \in \mathbb{X}^{\mathbf{F}}(\tilde{t},\tilde{x})$ such that $\tilde{\mathbf{y}}(t) = x$, for some $\tilde{t} \in [0,t]$ and $\tilde{x} \in \mathbb{R}^N$.

Let $t \in [0,T]$ and $x, x' \in \mathbb{B}_N(0;r)$. Notice that $\mathcal{V}(t,x) \geq \Phi(t,x)$, and so since $\mathbb{X}^{\mathbf{F}}(t,x) \neq \emptyset$, we have that $\mathcal{V}(t,x) \in \mathbb{R}$. Therefore, for any $\varepsilon > 0$ there is $(\mathbf{y}_\varepsilon,\mathbf{u}_\varepsilon) \in \mathbb{X}^{\mathbf{F}}(t,x)$ such that

$$\mathcal{V}(t,x) \geq \max_{s\in[t,T]} \Phi(s,\mathbf{y}_\varepsilon(s)) \bigvee \Psi(\mathbf{y}_\varepsilon(T)) - \varepsilon.$$

It follows then that

$$\mathcal{V}(t,x') - \mathcal{V}(t,x) \leq \max_{s\in[t,T]} (\Phi(s,\mathbf{y}'_\varepsilon(s)) - \Phi(s,\mathbf{y}_\varepsilon(s))) \bigvee (\Psi(\mathbf{y}'_\varepsilon(T)) - \Psi(\mathbf{y}_\varepsilon(T))) + \varepsilon,$$

where $\mathbf{y}'_\varepsilon$ is such that $(\mathbf{y}'_\varepsilon,\mathbf{u}_\varepsilon) \in \mathbb{X}^{\mathbf{F}}(t,x')$. Here we have also used the fact that

$$(a \vee b) - (c \vee d) \leq (a-c) \vee (b-d) \quad \text{and} \quad \sup_{a\in A} \phi_1(a) - \sup_{a\in A} \phi_2(a) \leq \sup_{a\in A} (\phi_1(a) - \phi_2(a)).$$

Using the fact that $\mathbf{y}_\varepsilon(s), \mathbf{y}'_\varepsilon(s) \in \mathbb{B}_N(0;\tilde{r})$ for any $s \in [t,T]$, we get that

$$\mathcal{V}(t,x') - \mathcal{V}(t,x) \leq L_{\Phi,\Psi} \max_{s\in[t,T]} \|\mathbf{y}'_\varepsilon(s)) - \mathbf{y}_\varepsilon(s)\| + \varepsilon,$$

Since we also have a.e. on $[t,T]$ that

$$\|\dot{\mathbf{y}}'_\varepsilon(s) - \dot{\mathbf{y}}_\varepsilon(s)\| = \|\mathbf{F}(s,\mathbf{y}'_\varepsilon(s),\mathbf{u}_\varepsilon(s)) - \mathbf{F}(s,\mathbf{y}_\varepsilon(s),\mathbf{u}_\varepsilon(s))\| \leq k_{\mathbf{F}} \|\mathbf{y}'_\varepsilon(s) - \mathbf{y}_\varepsilon(s)\|,$$

by Gronwall's inequality we have

$$\|\mathbf{y}'_\varepsilon(s) - \mathbf{y}_\varepsilon(s)\| \leq e^{k_{\mathbf{F}}(s-t)} \|\mathbf{y}'_\varepsilon(t) - \mathbf{y}_\varepsilon(t)\| \leq e(k_{\mathbf{F}}) \|x' - x\|, \qquad \forall s \in [t,T],$$

and consequently,

$$\mathcal{V}(t,x') - \mathcal{V}(t,x) \leq L_{\Phi,\Psi} e(k_{\mathbf{F}}) \|x' - x\| + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, by switching the roles of x and x' we get

$$|\mathcal{V}(t,x') - \mathcal{V}(t,x)| \leq L_{\Phi,\Psi} e(k_{\mathbf{F}}) \|x' - x\|, \quad \forall t \in [0,T], x, x' \in \mathbb{B}_N(0;r). \tag{31}$$

Take now also $t' \in [0,T]$, then it follows from (31) that

$$\begin{aligned} |\mathcal{V}(t',x') - \mathcal{V}(t,x)| &\leq |\mathcal{V}(t',x') - \mathcal{V}(t',x)| + |\mathcal{V}(t',x) - \mathcal{V}(t,x)| \\ &\leq L_{\Phi,\Psi} e(k_{\mathbf{F}}) \|x' - x\| + |\mathcal{V}(t',x) - \mathcal{V}(t,x)|. \end{aligned}$$

Take again $\varepsilon > 0$ arbitrary and $(\mathbf{y}_\varepsilon, \mathbf{u}_\varepsilon) \in \mathbb{X}^{\mathbf{F}}(t,x)$ such that

$$\mathcal{V}(t,x) \geq \max_{s\in[t,T]} \Phi(s,\mathbf{y}_\varepsilon(s)) \bigvee \Psi(\mathbf{y}_\varepsilon(T)) - \varepsilon.$$

Suppose in addition that $t' > t$, then

$$\mathcal{V}(t,x) \geq \max_{s\in[t',T]} \Phi(s,\mathbf{y}_\varepsilon(s)) \bigvee \Psi(\mathbf{y}_\varepsilon(T)) - \varepsilon.$$

Therefore,

$$\mathcal{V}(t',x) - \mathcal{V}(t,x) \leq \max_{s\in[t',T]} (\Phi(s,\mathbf{y}'_\varepsilon(s)) - \Phi(s,\mathbf{y}_\varepsilon(s))) \bigvee (\Psi(\mathbf{y}'_\varepsilon(T)) - \Psi(\mathbf{y}_\varepsilon(T))) + \varepsilon,$$

where $\mathbf{y}'_\varepsilon$ is such that $(\mathbf{y}'_\varepsilon, \mathbf{u}_\varepsilon|_{[t',T]}) \in \mathbb{X}^{\mathbf{F}}(t',x)$. From here, we obtain

$$\begin{aligned}
\mathcal{V}(t',x) - \mathcal{V}(t,x) &\leq L_{\Phi,\Psi} e^{k_{\mathbf{F}}(T-t')} \|x - \mathbf{y}_\varepsilon(t')\| + \varepsilon \\
&\leq L_{\Phi,\Psi} e(k_{\mathbf{F}}) \int_t^{t'} \|\mathbf{F}(s,\mathbf{y}_\varepsilon(s),\mathbf{u}_\varepsilon(s))\| ds + \varepsilon \\
&\leq L_{\Phi,\Psi} e(k_{\mathbf{F}}) c_{\mathbf{F}} (1+\tilde{r}) |t - t'| + \varepsilon.
\end{aligned}$$

Suppose now that $t' < t$. Let $\mathbf{u} \in \mathcal{U}$ be a measurable selection of U, whose existence is justified by the Aumann's Measurable Selection Theorem ([24, Theorem 2.3.12]). Let $\mathbf{y}$ be the unique solution of the Cauchy problem

$$\dot{\mathbf{y}}(s) = \mathbf{F}(s,\mathbf{y}(s),\mathbf{u}(s)), \text{ a.e. } s \in (t',t), \qquad \mathbf{y}(t) = x$$

and define

$$\mathbf{y}^\sharp_\varepsilon(s) = \begin{cases} \mathbf{y}(s) & \text{if } s \in [t',t) \\ \mathbf{y}_\varepsilon(s) & \text{if } s \in [t,T] \end{cases} \quad \text{and} \quad \mathbf{u}^\sharp_\varepsilon(s) = \begin{cases} \mathbf{u}(s) & \text{if } s \in [t',t) \\ \mathbf{u}_\varepsilon(s) & \text{if } s \in [t,T] \end{cases}$$

Notice that

$$\max_{s\in[t',T]} \Phi(s,\mathbf{y}^\sharp_\varepsilon(s)) = \max_{s\in[t',t]} \Phi(s,\mathbf{y}(s)) \bigvee \max_{s\in[t,T]} \Phi(s,\mathbf{y}_\varepsilon(s))$$

and that for any $s \in [t',t]$ we have

$$\begin{aligned}
\Phi(s,\mathbf{y}(s)) &\leq \Phi(t,x) + L_{\Phi,\Psi} (|s-t| + \|\mathbf{y}(s) - x\|) \\
&\leq \Phi(t,x) + L_{\Phi,\Psi} \left(|t'-t| + \int_{t'}^{t} \|\mathbf{F}(s,\mathbf{y}(s),\mathbf{u}(s))\| ds \right) \\
&\leq \Phi(t,x) + \tilde{L}_{\mathcal{V}} |t'-t| \leq \max_{s\in[t,T]} \Phi(s,\mathbf{y}_\varepsilon(s)) + \tilde{L}_{\mathcal{V}} |t'-t|,
\end{aligned}$$

where $\tilde{L}_{\mathcal{V}} := L_{\Phi,\Psi}(1 + c_{\mathbf{F}}(1+\tilde{r}))$. Therefore,

$$\max_{s\in[t',T]} \Phi(s,\mathbf{y}^\sharp_\varepsilon(s)) \leq \max_{s\in[t,T]} \Phi(s,\mathbf{y}_\varepsilon(s)) + \tilde{L}_{\mathcal{V}} |t'-t|$$

and so

$$\mathcal{V}(t,x) \geq \left(\max_{s\in[t',T]} \Phi(s,\mathbf{y}^\sharp_\varepsilon(s)) - \tilde{L}_{\mathcal{V}} |t'-t| \right) \bigvee \Psi(\mathbf{y}^\sharp_\varepsilon(T)) - \varepsilon.$$

Let $\mathbf{y}'_\varepsilon$ be such that $(\mathbf{y}'_\varepsilon, \mathbf{u}^\sharp_\varepsilon) \in \mathbb{X}^{\mathbf{F}}(t', x)$. Then, we have

$$\begin{aligned}\mathcal{V}(t',x) - \mathcal{V}(t,x) \leq & \max_{s\in[t',T]} \Phi(s, \mathbf{y}'_\varepsilon(s)) \bigvee \Psi(\mathbf{y}'_\varepsilon(T)) \\ & - \left(\max_{s\in[t',T]} \Phi(s, \mathbf{y}^\sharp_\varepsilon(s)) - \tilde{L}_{\mathcal{V}}|t'-t|\right) \bigvee \Psi(\mathbf{y}^\sharp_\varepsilon(T)) + \varepsilon \\ \leq & \left(\max_{s\in[t',T]} \big(\Phi(s,\mathbf{y}'_\varepsilon(s)) - \Phi(s,\mathbf{y}^\sharp_\varepsilon(s))\big) + \tilde{L}_{\mathcal{V}}|t'-t|\right) \bigvee \big(\Psi(\mathbf{y}'_\varepsilon(T)) - \Psi(\mathbf{y}^\sharp_\varepsilon(T))\big) + \varepsilon\end{aligned}$$

Since we also have a.e. on $[t', T]$ that

$$\|\dot{\mathbf{y}}'_\varepsilon(s) - \dot{\mathbf{y}}^\sharp_\varepsilon(s)\| = \|\mathbf{F}(s, \mathbf{y}'_\varepsilon(s), \mathbf{u}^\sharp_\varepsilon(s)) - \mathbf{F}(s, \mathbf{y}^\sharp_\varepsilon(s), \mathbf{u}^\sharp_\varepsilon(s))\| \leq k_{\mathbf{F}} \|\mathbf{y}'_\varepsilon(s) - \mathbf{y}^\sharp_\varepsilon(s)\|,$$

by Gronwall's inequality we have

$$\|\mathbf{y}'_\varepsilon(s) - \mathbf{y}^\sharp_\varepsilon(s)\| \leq e^{k_{\mathbf{F}}(s-t)} \|\mathbf{y}'_\varepsilon(t') - \mathbf{y}^\sharp_\varepsilon(t')\| = e(k_{\mathbf{F}}) \|x - \mathbf{y}^\sharp_\varepsilon(t')\|, \qquad \forall s \in [t', T].$$

Given that

$$\|x - \mathbf{y}^\sharp_\varepsilon(t')\| \leq c_{\mathbf{F}}(1+\tilde{r})|t'-t|,$$

we get

$$\mathcal{V}(t',x) - \mathcal{V}(t,x) \leq \tilde{L}_{\mathcal{V}}|t'-t|\,(e(k_{\mathbf{F}})+1) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, by switching the roles of t and t', and using either case 1 or 2 described above, we get

$$|\mathcal{V}(t',x) - \mathcal{V}(t,x)| \leq \tilde{L}_{\mathcal{V}}|t'-t|\,(e(k_{\mathbf{F}})+1)$$

and so

$$|\mathcal{V}(t',x') - \mathcal{V}(t,x)| \leq L_{\Phi,\Psi} e(k_{\mathbf{F}}) \|x'-x\| + \tilde{L}_{\mathcal{V}}\,(e(k_{\mathbf{F}})+1)\,|t'-t|,$$

which leads to the conclusion. □ □

References

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

[2] Aubin, J.P., Cellina, A.: Differential inclusions: Set-Valued Maps and Viability Theory, *Grundlehren der mathematischen Wissenschaften*, vol. 264. Springer-Verlag, Berlin Heidelberg New York Tokyo (1984)

[3] Bagagiolo, F., Maggistro, R., Zoppello, M.: A differential game with exit costs. Dynamic Games and Applications **10**(2), 297–327 (2020)

[4] Bardi, M., Capuzzo-Dolcetta, I.: Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Systems and Control: Foundations and Applications. Birkhäuser, Boston (1997)

[5] Bardi, M., Koike, S., Soravia, P.: Pursuit-evasion games with state constraints: dynamic programming and discrete-time approximations. Discrete & Continuous Dynamical Systems **6**(2), 361 (2000)

[6] Barron, E.N.: The Pontryagin maximum principle for minimax problems of optimal control. Nonlinear Analysis: Theory, Methods & Applications **15**(12), 1155–1165 (1990)

[7] Bettiol, P., Cardaliaguet, P., Quincampoix, M.: Zero-sum state constrained differential games: existence of value for Bolza problem. International Journal of Game Theory **34**(4), 495–527 (2006)

[8] Bettiol, P., Frankowska, H., Vinter, R.B.: Improved sensitivity relations in state constrained optimal control. Applied Mathematics and Optimization **71**(2), 353–377 (2015). DOI 10.1007/s00245-014-9260-6

[9] Bettiol, P., Vinter, R.: Sensitivity interpretations of the costate variable for optimal control problems with state constraints. SIAM Journal on Control and Optimization **48**(5), 3297–3317 (2010)
[10] Bokanowski, O., Désilles, A., Zidani, H.: Relationship between maximum principle and dynamic programming principle in presence of intermediate and final state constraints. ESAIM: Control, Optimisation and Calculus of Variations **27** (2021)
[11] Clarke, F.: Functional Analysis, Calculus of Variations and Optimal Control, *Graduate Texts in Mathematics*, vol. 264. Springer–Verlag (2013)
[12] Clarke, F.H., Vinter, R.B.: The relationship between the maximum principle and dynamic programming. SIAM Journal on Control and Optimization **25**(5), 1291–1311 (1987)
[13] Frankowska, H., Mazzola, M.: On relations of the adjoint state to the value function for optimal control problems with state constraints. Nonlinear Differential Equations and Applications NoDEA **20**, 361–383 (2013)
[14] Frankowska, H., Vinter, R.: 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). DOI 10.1023/A:1004668504089
[15] Goebel, R.: Regularity of the optimal feedback and the value function in convex problems of optimal control. Set-Valued Analysis **12**(1), 127–145 (2004)
[16] Hermosilla, C., Vinter, R.B., Zidani, H.: Hamilton-Jacobi-Bellman equations for optimal control processes with convex state constraints. Systems & Control Letters **109**, 30–36 (2017). DOI 10.1016/j.sysconle.2017.09.004
[17] Hermosilla, C., Wolenski, P., Zidani, H.: The Mayer and minimum time problems with stratified state constraints. Set-Valued and Variational Analysis **26**, 643–662 (2018). DOI doi.org/10.1007/s11228-017-0413-z
[18] Hermosilla, C., Zidani, H.: Infinite horizon problems on stratifiable state-constraints sets. Journal of Differential Equations **258(4)**, 1420–1460 (2015). DOI 10.1016/j.jde.2014.11.001
[19] Rockafellar, R.T.: Conjugate convex functions in optimal control and the calculus of variations. Journal of Mathematical Analysis and Applications **32**(1), 174–222 (1970)
[20] Rockafellar, R.T.: Convex analysis, vol. 28. Princeton University Press, Princeton, N.J. (1970)
[21] Rockafellar, R.T., Wolenski, P.R.: Convexity in Hamilton–Jacobi theory I: Dynamics and duality. SIAM Journal on Control and Optimization **39**(5), 1323–1350 (2000)
[22] Soner, H.: Optimal control with state-space constraint I. SIAM Journal on Control and Optimization **24**(3), 552–561 (1986). DOI 10.1137/0324032
[23] Vinter, R.: New results on the relationship between dynamic programming and the maximum principle. Mathematics of Control, Signals, and Systems **1**(1), 97–105 (1988)
[24] Vinter, R.: Optimal control. Modern Birkhäuser Classics. Birkhäuser, Boston (2000)
[25] Vinter, R., Wolenski, P.: Hamilton-Jacobi theory for optimal control problems with data measurable in time. SIAM Journal on Control and Optimization **28**(6), 1404–1419 (1990)
[26] Zhou, X.Y.: Maximum principle, dynamic programming, and their connection in deterministic control. Journal of Optimization Theory and Applications **65**(2), 363–373 (1990)

Departamento de Matemática, Universidad Técnica Federico Santa María,, Avenida España 1680, Valparaíso, Chile
Email address: `cristopher.hermosill@usm.cl`

Laboratoire de Mathématiques, INSA Rouen Normandie,, 685 Avenue de l'Université, Saint-Étienne-du-Rouvray, France
Email address: `hasnaa.zidani@insa-rouen.fr`