OPTIMAL CONTROL OF THE SWEEPING PROCESS WITH A NON-SMOOTH MOVING SET [*]

CRISTOPHER HERMOSILLA[†] AND MICHELE PALLADINO[‡‡]

Abstract. In this paper we prove a fully nonsmooth Pontryagin Maximum Principle for optimal control problems driven by a sweeping process with drift $\dot{x} \in f(t,x,u) - \mathcal{N}_{C(t)}(x)$. The setting we study is an optimal control problem of Mayer type in which the optimization procedure is carried out by choosing a control function $u(t)$ from a class of admissible controls $\mathcal{U}$. The choice of $u \in \mathcal{U}$ modifies the drift f and the related solution $x(t)$ to the perturbed sweeping process. Here, for the first time, we are able to prove a Pontryagin Maximum Principle in the case in which the moving set $C(t)$ is both nonsmooth and non-convex by using a novel exact penalization technique which is able to exploit the controllability properties of the dynamics.

Key words. Sweeping process, Optimal Control, Pontryagin Maximum Principle, State Constraints.

AMS subject classifications. 49K15 (Optimality conditions for problems involving ordinary differential equations), 49J45 (Methods involving semicontinuity and convergence; relaxation), 93C10 (nonlinear systems in control theory)

1. Introduction. Moreau's sweeping process is a quasi-static model introduced in the '70s by J. J. Moreau to describe the dynamical interaction between a moving constraint $C(t)$ and a point-like object $x(t)$. The time evolution of $x(t)$ is represented by the dynamic equation

$$\dot{x}(t) \in -\mathcal{N}_{C(t)}(x(t)), \qquad \text{a.e. } t \in [0,1], \qquad x(0) = x_0 \in C(0), \tag{1.1}$$

satisfying the constraint $x(t) \in C(t)$ for all $t \in [0,1]$. Here, $t \rightsquigarrow C(t)$ is a Lipschitz continuous multifunction with closed convex values, and $\mathcal{N}_{C(t)}(x)$ is the standard normal cone used in convex analysis. In the seminal paper [32], it is shown the existence and the uniqueness of the solution $x(t)$ subject to the constraint $x(t) \in C(t)$ for all $t \in [0,1]$. After that seminal contribution, several papers appeared (see, e.g. [17, 13, 21] and references therein) dealing with existence and uniqueness results associated with various generalizations of the initial value problem (1.1). In particular, there has been a strong effort for obtaining existence and uniqueness results for the, so called, perturbed sweeping process

$$\begin{cases} \dot{x}(t) \in g(t,x(t)) - \mathcal{N}^P_{C(t)}(x(t)), & \text{a.e. on } [0,1], \\ x(0) = x_0 \in C(0), \qquad x(t) \in C(t), & \forall\, t \in [0,1], \end{cases} \tag{1.2}$$

where $g : [0,1] \times \mathbb{R}^N \to \mathbb{R}^N$ is a vector field measurable with respect to t and Lipschitz continuous with respect to x, while the moving set $t \rightsquigarrow C(t)$ is a mildly non-convex set for each $t \in [0,1]$ (see, e.g. [13, 36] and references therein). Hereinafter $\mathcal{N}^P_{C(t)}$ stands for the proximal normal cone.

These existence and uniqueness results pave the way to the study of control

[*]Submitted to the editors April 29, 2022.

Funding: This research is partially supported by FONDECYT grant N 11190456 (C. Hermosilla) and by Conicyt Project Redes REDI170200.

[†]Universidad Tecnica Federico Santa Maria, Chile (cristopher.hermosill@usm.cl).

[‡]University of L'Aquila, Department of Information Engineering, Computer Science and Mathematics (DISIM), L'Aquila, Italy (michele.palladino@univaq.it).

systems such as

$$
\text{(CS)} \qquad \begin{cases} \dot{x}(t) \in f(t, x(t), u(t)) - \mathcal{N}^P_{C(t)}(x(t)), & \text{a.e. on } [0,1], \\ u \in \mathcal{U} \\ x(0) = x_0 \in C(0), \qquad x(t) \in C(t), & \forall\, t \in [0,1]. \end{cases}
$$

where $f : [0,1] \times \mathbb{R}^N \times \mathbb{R}^m \to \mathbb{R}^N$ is a controlled vector field, $C(t)$ is the moving set and $\mathcal{U}$ is a set of measurable functions. It turns out that, under general conditions on the controlled vector field and on the moving set, the control system (CS) is *well-posed*, in the sense that, for each $u \in \mathcal{U}$ and $x_0 \in C(0)$, there exists a unique, absolutely continuous arc $x : [0,1] \to \mathbb{R}^N$ solution to (CS); see for instance [20].

The study of control systems like (CS) has received an increasing attention in the last decade. Such a kind of control systems has been used to provide models for some electric circuits [2], for crowd motion [29], for hysteresis [27], for studying the evolution of a robotic intruder in soil [35] and as a tool for identification of parameters [8] for some mechanical systems subject to unilateral constraints. In particular, there has been an increasing attention for what concerns the study of optimal control problems driven by a dynamic equation as in (CS). The characterization of the value function as a unique generalized solution of a Hamilton-Jacobi-Bellman equation was derived for several kinds of optimal control problems driven by control systems like (CS) (see, e.g., [28, 34, 16, 35] and references therein).

There has also been an extensive effort to derive versions of the Pontryagin Maximum Principle for optimal control problems driven by (CS). Here, the main challenge is to derive the adjoint equation, in particular at those points in which a discontinuity of the dynamical system (CS) arises. To tackle such a problem, two main proof methodologies have been employed so far. The first one, which we refer to as *discrete approximation method* [30], consists in discretizing the optimal control problem driven by (CS), with a finer and finer time step. This leads to an optimization problem for which one can use well-known necessary conditions for constrained optimization problems. The challenging parts of the method consist in passing to the limit, both in the primal discretized problem and in the dual set of necessary conditions, when the time step vanishes. To successfully carry out these steps, several advanced techniques of variational analysis are used (see, e.g. [31, 33]). Examples of this approach can be found in [11, 12, 14, 15] and the references therein. This approach has also been able to tackle the case in which the moving set depends on a Lipschitz continuous control (see, e.g. [10]). However, for the sweeping process dynamics (CS), the state-of-the-art results obtained by using the discrete approximation technique cover the case in which the vector field satisfies mild regularity assumptions, while the constraint $C(t)$, although possibly nonsmooth and non-convex [11], has to satisfy assumptions stronger than the one employed in this paper (cfr. Hypotheses 2.1, 2.6 and 2.7 below). Furthermore, unless assuming some convexity on the control set [15], all the results achieved so far by using the discrete approximation technique, present merely a local/weak maximality condition instead of the standard, well-known *maximality condition* of the Pontryagin Maximum Principle; cfr. Condition (iv) in Theorem 2.12.

A different approach is instead based on approximating the normal cone using a differentiable (or merely Lipschitz continuous) approximation and considering the related optimal control problem for which well-known necessary conditions can be applied. Here, the main issue concerns the choice of the approximating function whose gradient approximates the normal cone. Indeed, roughly speaking, if the normal cone can be approximated by the gradient of a sequence of smooth functions, then, in the necessary conditions, one has to deal with the Hessians of these smooth functions which, in general, do not have good compactness properties. As a matter of fact, this approach was successfully applied to strictly convex, smooth moving sets [7] or to moving sets with smooth boundary [4, 5, 18, 38].

In this paper, for the first time, we are able to prove the Pontryagin Maximum Principle for a quite general, non-convex and nonsmooth moving set and on minimal hypothesis on the drift f. Indeed, in this paper we will assume that the moving set $C(t)$ is described by l differentiable functions $h_1, \ldots, h_l$ as

$$C(t) = \left\{ x \in \mathbb{R}^N \mid h_i(t,x) \leq 0, \quad i = 1, \ldots, l \right\}.$$

and such that satisfy a positive linear independence constraint qualification (see Hypothesis 2.7 for a detailed definition). To achieve such a result, we will make use of a novel exact penalization technique based on the idea that the term $-\mathcal{N}_{C(t)}(x)$ *actively* constraints the optimal state trajectory in $C(t)$. The effect of such a force is then captured by the dynamics

$$\dot{x}(t) = f(t, x(t), u(t)) - \sum_{i=1}^{l} \alpha_i(t) \nabla_x h_i(t, x(t)), \qquad x(0) = x_0, \tag{1.3}$$

where $\alpha = (\alpha_1, \ldots, \alpha_l)$ is a new control representing the action of the normal cone on the state trajectory. To obtain a solution to (CS), we add a term in the cost functional which penalizes all those trajectories of (1.3) which are not also solutions of (CS). Such a novel approach permits to exploit the small-time local controllability relation between the dynamics (1.3) and the moving constraint $C(t)$, by notably simplifying the proof of the necessary conditions.

The following notations will be used throughout the paper. For vector a $x \in \mathbb{R}^n$, $\|x\|$ denotes its Euclidean length, while $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^n$. Given $\delta > 0$, we use $\mathbb{B}(x, \delta)$ to denote the closed ball in $\mathbb{R}^n$ centered at x and with radius δ. For a multifunction $\Gamma : \mathbb{R}^n \rightsquigarrow \mathbb{R}^k$, the graph is defined as

$$\operatorname{Gr} \Gamma := \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^k \mid v \in \Gamma(x)\}.$$

Given a closed set $A \subset \mathbb{R}^n$, the distance from A is defined as

$$\operatorname{dist}_A(x) = \inf_{y \in A} \|x - y\|, \qquad x \in \mathbb{R}^n.$$

For any $x \in A$, the proximal normal cone $\mathcal{N}_A^P(x)$ at x to A is defined as the set of $p \in \mathbb{R}^n$ such that there exists $M > 0$ satisfying the relation

$$\langle p, y - x \rangle \leq M \|x - y\|^2, \qquad \text{for all } y \in A.$$

$W^{1,1}([0,1]; \mathbb{R}^n)$ is the set of absolutely continuous functions from $[0,1]$ to $\mathbb{R}^n$, while $\mathcal{M}([0,1]; \mathbb{R}^m)$ is the set of Lebesgue-measurable functions from $[0,1]$ to $\mathbb{R}^m$ and $\mathcal{BV}([0,1]; \mathbb{R}^n)$ is the set of bounded variation functions from $[0,1]$ to $\mathbb{R}^n$. Given any positive, vector-valued Borel measure μ defined on $[0,1]$, we use $\operatorname{supp}(\mu)$ for denoting its support.

Take a lower semicontinuous function $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $\bar{x} \in \operatorname{dom} f := \{x \in \mathbb{R}^k \mid f(x) < +\infty\}$. The *limiting subdifferential of f at $\bar{x}$* (also known as Mordukhovich's subdifferential) is defined a:

$$\partial f(\bar{x}) := \Big\{\xi \mid \exists\, \xi_i \to \xi \text{ and } x_i \xrightarrow{\text{dom} f} \bar{x} \text{ such that}$$

$$\limsup_{x \to x_i} \frac{\langle \xi_i, x - x_i\rangle - f(x) + f(x_i)}{|x - x_i|} \le 0 \text{ for all } i \in \mathbb{N}\Big\}.$$

If f is Lipschitz continuous in a neighborhood of $\bar{x}$, then the Clarke's subdifferential of f at $\bar{x}$ is $\operatorname{co}\partial f(\bar{x})$.

2. Setting of the problem. Given an end-point cost $g : \mathbb{R}^N \to \mathbb{R}$ we are concerned with the Mayer problem associated with a controlled sweeping process, that is,

$$\text{(P)} \qquad \begin{cases} \text{Minimize} & g(x(1)) \\ \text{over all} & x \in W^{1,1}\left([0,1];\mathbb{R}^N\right) \text{ and } u \in \mathcal{M}\left([0,1];\mathbb{R}^m\right) \\ \text{such that} & \dot{x}(t) \in f(t, x(t), u(t)) - \mathcal{N}^P_{C(t)}(x(t)), \quad \text{for a.e. } t \in [0,1], \\ & u(t) \in U(t), \qquad \text{for a.e. } t \in [0,1], \\ & x(0) = x_0 \in C(0), \end{cases}$$

In this context any process $(x,u) \in W^{1,1}\left([0,1];\mathbb{R}^N\right) \times \mathcal{M}\left([0,1];\mathbb{R}^m\right)$ satisfying the conditions of problem (P) will be called feasible. In particular, feasible processes satisfy the Cauchy problem

$$\text{(2.1)} \qquad \dot{x}(t) \in f(t, x(t), u(t)) - \mathcal{N}^P_{C(t)}(x(t)), \qquad \text{a.e. on } [0,1], \quad x(0) = x_0,$$

as well as the state constraints

$$x(t) \in C(t), \qquad \forall\, t \in [0,1],$$

under the convention that $\mathcal{N}^P_{C(t)}(x) = \emptyset$ whenever $x \notin C(t)$ for any $t \in [0,1]$. For the sake of simplicity, the initial condition x_0 remains fixed along the paper.

Given $\delta > 0$, we say that a feasible process $(\bar{x}, \bar{u})$ is a strong δ-local minimizer for problem (P) if

$$g(\bar{x}(1)) \le g(x(1))$$

for every feasible process (x,u) such that

$$\|x - \bar{x}\|_{L^\infty} \le \delta.$$

Our task in this paper is to exhibit optimality conditions for strong δ-local minimizers in a framework where the moving set is nonsmooth. Indeed, we are mainly concerned with the case in which $C : [0,1] \rightsquigarrow \mathbb{R}^N$ is a set-valued map determined by a given continuous mapping $h : [0,1] \times \mathbb{R}^N \to \mathbb{R}^l$ in the following way

$$C(t) = \left\{x \in \mathbb{R}^N \mid h_i(t,x) \le 0, \qquad i = 1, \dots, l\right\}.$$

As usual in the literature of sweeping processes, we assume that the moving set is uniformly prox-regular and depends continuously on the time variable. The following basic assumptions are enforced along the paper:

Hypothesis 2.1. There is $\rho > 0$ such that for each $t \in [0,1]$, the set $C(t)$ is a nonempty closed uniformly ρ-prox-regular set[1] and there is $L_C > 0$ such that

$$|\mathrm{dist}_{C(t)}(x) - \mathrm{dist}_{C(s)}(x)| \le L_C|t-s|, \qquad \forall x \in \mathbb{R}^N,\ \forall t, s \in [0,1].$$

To avoid measurability issues, in our setting the multifunction $U : [0,1] \rightsquigarrow \mathbb{R}^m$ and the dynamics $f : [0,1] \times \mathbb{R}^N \times \mathbb{R}^m \to \mathbb{R}^N$ are assumed to satisfy

Hypothesis 2.2. $\mathrm{Gr}\, U$ is a nonempty $\mathcal{L} \times \mathcal{B}^m$ measurable set.

Hypothesis 2.3. f is $\mathcal{L} \times \mathcal{B}^N \times \mathcal{B}^m$ measurable on $[0,1] \times \mathbb{R}^N \times \mathbb{R}^m$.

We use $\mathcal{U}$ to denote the set of all measurable selections of $U : [0,1] \rightsquigarrow \mathbb{R}^m$. Notice that under the assumptions we have imposed up to this point, by standard arguments of the theory of sweeping processes with drift (see for instance [20]), it follows that, given $u \in \mathcal{U}$, any process (x, u), solution of the Cauchy problem (2.1) satisfies

$$\|\dot{x}(t) - f(t, x(t), u(t))\| \le \|f(t, x(t), u(t))\| + L_C, \quad \text{for a.e. } t \in [0,1]. \tag{2.2}$$

2.1. Standing assumptions and parametric representation of the normal cone. The technique we present in the next sections for obtaining optimality conditions requires some technical assumptions, which we proceed to introduce. In particular, for the analysis we propose such hypotheses are only needed in a tube around a reference trajectory. Consequently, we assume that for a given feasible process $(\bar{x}, \bar{u})$, there is $\delta > 0$ (which remains fixed from now on), for which the following conditions hold:

Hypothesis 2.4. There exists a $\mathcal{L} \times \mathcal{B}^m$ measurable function $K_f : [0,1] \times \mathbb{R}^m \to \mathbb{R}$ such that $t \mapsto K_f(t, \bar{u}(t))$ is integrable and

$$\|f(t,x,u) - f(t,y,u)\| \le K_f(t,u)\|x - y\|,$$

for every $x, y \in \mathbb{B}(\bar{x}(t), \delta)$, $u \in U(t)$ and for a.e. $t \in [0,1]$.

Hypothesis 2.5. there exists a positive constant $c > 0$ such that

$$\|f(t,x,u)\| \le c$$

for every $u \in U(t)$, $x \in \mathbb{B}(\bar{x}(t), \delta)$ and for a.e $t \in [0,1]$.

Let us point out that by [33, Theorem 14.26] the mapping $t \rightsquigarrow \mathcal{N}^P_{C(t)}(\bar{x}(t))$ is measurable and thus, since it also has closed images, [33, Theorem 14.16] implies that there is a measurable selection[2] $\bar{\eta} \in \mathcal{M}\left([0,1]; \mathbb{R}^N\right)$ of $t \rightsquigarrow \mathcal{N}^P_{C(t)}(\bar{x}(t))$, such that

$$\dot{\bar{x}}(t) = f(t, \bar{x}(t), \bar{u}(t)) - \bar{\eta}(t), \quad \text{for a.e. } t \in [0,1], \qquad x(0) = x_0.$$

In our approach, it is important that this measurable selection can be represented by means of the partial derivatives of the scalar functions $h_1, \ldots, h_l$, in the sense that

[1] This means that

$$\langle \eta, y - x \rangle \le \frac{\|\eta\|}{2\rho}\|y - x\|^2, \qquad \forall t \in [0,1],\ x, y \in C(t),\ \eta \in \mathcal{N}^P_{C(t)}(x).$$

[2] To see this, according to the notation of [33, Theorem 14.16], it is enough to take $D(t) = \{0\}$, $X(t) = \mathcal{N}^P_{C(t)}(\bar{x}(t))$ and the Carathéodory map $F(t, \eta) = \dot{\bar{x}}(t) - f(t, \bar{x}(t), \bar{u}(t)) + \eta$.

there is $\bar{\alpha} \in \mathcal{M}\left([0,1];\mathbb{R}^l\right)$ such that

$$(2.3) \quad 0 \leq \bar{\alpha}(t) \perp h(t,\bar{x}(t)) \quad \text{and} \quad \bar{\eta}(t) = \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t,\bar{x}(t)), \quad \text{for a.e. } t \in [0,1].$$

To justify this representation, we impose the following conditions:

Hypothesis 2.6. Each h_i is continuous with $x \mapsto h_i(t,x)$ being continuously differentiable on $\mathbb{B}(\bar{x}(t),\delta)$ for all $t \in [0,1]$. Moreover, there exists a constant $L_h > 0$ such that

$$\max\left\{|h_i(t,x) - h_i(s,y)|, \|\nabla_x h_i(t,x) - \nabla_x h_i(s,y)\|\right\} \leq L_h\left(|t-s| + \|x-y\|\right)$$

for every $i = 1,\dots,l$, $(t,x),(s,y) \in (\operatorname{Gr}\bar{x} + \delta\mathbb{B})$ with $t,s \in [0,1]$.

Hypothesis 2.7. The following positive linear independence constraint qualification is satisfied: for every $t \in [0,1]$, it follows that

$$\sum_{i \in I(t,\bar{x}(t))} \alpha_i \nabla_x h_i(t,\bar{x}(t)) = 0, \text{ with each } \alpha_i \geq 0 \quad \Longrightarrow \quad \alpha_i = 0,\ \forall i \in I(t,\bar{x}(t)).$$

Here $I(t,x)$ stands for the set of active indexes, that is,

$$I(t,x) := \{i \in \{1,\dots,l\} \mid \quad h_i(t,x) = 0\}, \qquad \forall (t,x) \in [0,1] \times \mathbb{R}^N.$$

Remark 2.8. It is important to observe that the prox-regularity assumption appearing in Hypothesis 2.1 is distinct from the positive linear constraint qualification (Hypothesis 2.7). In general, these two assumptions are not related to each other. Indeed, the line $C = \{(x,y) \in \mathbb{R}^2 \mid x = 0\}$ (equivalent to $\{(x,y) \in \mathbb{R}^2 \mid x \leq 0, \ \ -x \leq 0\}$) is prox-regular, but the gradients of the constraint functions are not positively linearly independent. On the other hand, in $\mathbb{R}^2$ the functions $h_1(t,x,y) := x - \sqrt[3]{y^5}$ and $h_2(t,x,y) := -y$ satisfy Hypothesis 2.7, however, the (nonsmooth) set

$$C = \{(x,y) \in \mathbb{R}^2 \mid x - \sqrt[3]{y^5} \leq 0, y \geq 0\}$$

is not prox-regular because, on the one hand $\mathcal{N}_C^P((x,y)) = \{\lambda(1, -\frac{5}{3}\sqrt[3]{y^2}) \mid \lambda \geq 0\}$ for any $(x,y) \in \partial C$ with $y > 0$ and, on the other hand, since $(0,0) \in C$ if the set were prox-regular there would be some $\rho > 0$ such that

$$-x + \frac{5}{3}\sqrt[3]{y^5} \leq \frac{\sqrt{1 + \frac{25}{9}\sqrt[3]{y^4}}}{2\rho}(x^2 + y^2)$$

for any $(x,y) \in \partial C$ such that $y > 0$; in particular $x = \sqrt[3]{y^5}$. From here, dividing the inequality by $\sqrt[3]{y^5}$, one gets the inequality

$$\frac{2}{3} \leq \frac{\sqrt{1 + \frac{25}{9}\sqrt[3]{y^4}}}{2\rho}(\sqrt[3]{y^5} + \sqrt[3]{y})$$

which leads to $\frac{2}{3} \leq 0$ when letting $y \to 0$. Therefore, C cannot be prox-regular.

Let us also mention that positive linear independence of the gradients plus some additional qualification conditions do imply prox-regularity; see [3, Theorem 3.5].

Lemma 2.9. *Assume that Hypotheses 2.1 to 2.7 hold. Then there is* $\bar{\alpha} \in L^\infty\left([0,1];\mathbb{R}^l\right)$ *such that* (2.3) *holds.*

Proof. By [9, Corollary 10.44 and Theorem 11.36] the proximal normal cone in this setting can be written as

$$\mathcal{N}^P_{C(t)}(x) = \left\{ \sum_{i=1}^{l} \alpha_i \nabla_x h_i(t,x) \mid \alpha_i \geq 0, \text{ such that } \alpha_i h_i(t,\bar{x}(t)) = 0, \ i = 1,\ldots,l \right\}. \tag{2.4}$$

Moreover, the set-valued map

$$t \rightsquigarrow \left\{ \alpha \in \mathbb{R}^l \mid \alpha_i \geq 0, \ \alpha_i h_i(t,\bar{x}(t)) = 0, \ i = 1,\ldots,l \right\}$$

is measurable thanks to [33, Example 14.15]. Therefore, the existence of a measurable function $\bar{\alpha} : [0,1] \to \mathbb{R}^l$ that satisfies (2.3) is guaranteed by [33, Theorem 14.16]. It remains to show that $\bar{\alpha}$ is essentially bounded.

Suppose by contradiction that $\max\limits_{i=1,\ldots,l} \|\bar{\alpha}_i\|_{L^\infty} = +\infty$. Then there is a sequence $\{t_k\}_{k\in\mathbb{N}}$ of Lebesgue points of $\bar{\alpha}$ such that $r_k := \max\limits_{i=1,\ldots,l} \bar{\alpha}_i(t_k) \to +\infty$ as $k \to +\infty$. Without lost of generality, we assume that the sequence $\{t_k\}_{k\in\mathbb{N}}$ converges to some $\tau \in [0,1]$ and that $I(t_k,\bar{x}(t_k)) \subseteq I(\tau,\bar{x}(\tau))$ for any $k \in \mathbb{N}$; this last fact is justified by the continuity of the scalar functions h_i.

Notice that, thanks to (2.2) and Hypothesis 2.5, it follows that

$$\left\| \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t,\bar{x}(t)) \right\| = \|\dot{\bar{x}}(t) - f(t,\bar{x}(t),\bar{u}(t))\| \leq c + L_C. \tag{2.5}$$

Let $a_k^i := \frac{\bar{\alpha}_i(t_k)}{r_k}$ for any $k \in \mathbb{N}$ and $i = 1,\ldots,l$. Remark that each sequence $\{a_k^i\}_{k\in\mathbb{N}}$ is contained in the interval $[0,1]$. Thus, passing into a subsequence if necessary, we can assume that each sequence converges to some $a^i \in [0,1]$, with at least one a^i nonzero for some $i \in I(\tau,\bar{x}(\tau))$ (equal to 1 actually). Since, $r_k \to +\infty$, from (2.5) it follows that $\sum\limits_{i\in I(\tau,\bar{x}(\tau))} a^i \nabla_x h_i(\tau,\bar{x}(\tau)) = 0$. Therefore, by Hypothesis 2.7 each a^i must equal zero, which is a contradiction. ∎

We will also enforce the following standard condition on the cost function:

Hypothesis 2.10. The mapping $x \mapsto g(x)$ is Lipschitz continuous on $\mathbb{B}(\bar{x}(1),\delta)$.

Finally, in order to handle nondegeneracy and normality of the maximum principle for an auxiliary optimal control problem which we will introduce later on, the next condition is also required:

Hypothesis 2.11. There exist $r' > 0$ and $K' > 0$ such that

$$\|f(t,x,u) - f(t,y,u)\| \leq K'\|x - y\|,$$

for every $x, y \in \mathbb{B}(\bar{x}(0),\delta)$, $u \in U(t)$ and for a.e. $t \in [0,r'[$.

2.2. Statement of the main result. Recall that Hypotheses 2.2 to 2.7, 2.10, and 2.11 remain always in force, and that $\bar{\alpha} \in L^\infty\left([0,1];\mathbb{R}^l\right)$ is given by Lemma 2.9 and satisfying the condition (2.3) is a reaction term introduced by the sweeping process on the control system.

THEOREM 2.12. *Suppose that* $(\bar{x}, \bar{u})$ *is a strong* δ*-local minimizer for problem* (P). *Then, there exist* $\sigma, \xi \geq 0$, $p \in W^{1,1}\left([0,1];\mathbb{R}^N\right)$, *some positive definite Borel measures* $\mu_1, \ldots, \mu_l$ *satisfying the following conditions:*

(i) $\sigma + \xi \sum_{i=1}^{l} \int_0^1 \bar{\alpha}_i(t)dt + \sum_{i=1}^{l} \mu_i([0,1]) = 1$*;*

(ii) for a.e. $t \in [0,1]$,

$$-\dot{p}(t) \in \operatorname{co} \partial_x \left\{ q(t) \cdot \left(f(t, \bar{x}(t), \bar{u}(t)) - \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t)) \right) \right\} + \xi \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t))$$

where

$$q(t) = \begin{cases} p(t) + \displaystyle\sum_{i=1}^{l} \int_{[0,t[} \nabla_x h_i(s, \bar{x}(s)) \mu_i(ds) & \text{if } t \in [0,1[, \\ p(1) + \displaystyle\sum_{i=1}^{l} \int_{[0,1]} \nabla_x h_i(s, \bar{x}(s)) \mu_i(ds) & \text{if } t = 1; \end{cases} \tag{2.6}$$

(iii) $-q(1) \in \sigma \partial g(\bar{x}(1))$*;*

(iv) for a.e. $t \in [0,1]$

$$\langle q(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle = \max_{u \in U(t)} \left\{ \langle q(t), f(t, \bar{x}(t), u) \rangle \right\};$$

(v) for each $i = 1, \ldots, l$ *we have* $\operatorname{supp}(\mu_i) \subseteq \{t \in [0,1] \mid h_i(t, \bar{x}(t)) = 0\}$*;*

(vi) $\langle q(t), \nabla_x h_i(t, \bar{x}(t)) \rangle = 0$ *for a.e.* $t \in \{s \in [0,1] \mid \bar{\alpha}_i(s) > 0\}$.

Remark 2.13. Notice that Theorem 2.12 is an extension of the Pontryagin maximum principle for problems with state constraints, and as such, degeneracy and abnormality are issues that can arise. In particular, the new multiplier ξ, which is intrinsically associated with the normal cone term of the dynamical system, needs to be understood more thoroughly. Particular degeneracy issues may appear if one has $\sigma = 0$ and

$$\xi \sum_{i=1}^{l} \int_0^1 \bar{\alpha}_i(t)dt = \sum_{i=1}^{l} \mu_i([0,1]).$$

Additional qualification conditions for avoiding degeneracy and abnormality, as the ones in [25], are planned to be studied elsewhere.

Remark 2.14. The Pontryagin maximum principle for problem (P) stated in Theorem 2.12 has some formal resemblance with the state constrained Pontryagin maximum principle in Gamkrelidze form [26, 6] in which the adjoint equation reads as

$$-\dot{p}(t) \in \partial_x H^e(t, \bar{x}(t), \bar{u}(t), p(t), \mu(t)), \qquad t \in [0,1],$$

where

$$H^e(t, x, u, p, \mu(t)) = \sup_{u \in U(t)} \left\{ \left\langle p, f(t, x, u) - \sum_{i=1}^{l} \mu_i(t) \nabla_x h_i(t, x) \right\rangle \right\}$$

and $\mu_i(t)$ is a non-negative function of bounded variation with support in the set $\{t \in [0,1] : \ h_i(t,\bar{x}(t)) = 0\}$. In particular, the analogy between Theorem 2.12 and the state constrained Pontryagin maximum principle in Gamkrelidze form arises in view of the presence of a "second order derivative" of the functional inequality describing the state constraint. However, if on one hand the second order derivatives do appear in the adjoint equation of Theorem 2.12, on the other the multiplier q is defined as in (2.6), which is the well-known Dubovitskii-Milyutin formulation.

To prove this result we introduce a novel exact penalization technique. The core of the proof of Theorem 2.12 can be found in section 4. In the next section we describe the exact penalization technique we introduce in this paper.

3. Auxiliary optimal control problem. We will divide the proof of Theorem 2.12 in two separate cases: when the optimal process $(\bar{x},\bar{u})$ is such that the normal cone is active or when the normal cone is inactive. We will start by considering the former case and, in what follows, we will work under the assumption:

$$\sum_{i=1}^{l} \int_0^1 \bar{\alpha}_i(t)dt > 0, \tag{3.1}$$

that is, the process $(\bar{x},\bar{u})$ has active normal cone.

Let us define the set of control values

$$A := \prod_{i=1}^{l} A_i \subseteq \mathbb{R}^l, \quad \text{where each } A_i := [0, \|\bar{\alpha}_i\|_{L^\infty} + 1],$$

where we recall that $\bar{\alpha}_i$ is given by Lemma 2.9 and satisfies the condition (2.3).

For $\gamma > 0$, consider the running cost $\Lambda : [0,1] \times \mathbb{R}^l \to \mathbb{R}$ given by

$$\Lambda(t,\alpha) = \frac{\gamma}{2}\|\alpha - \bar{\alpha}(t)\|^2,$$

and the final time cost $\tilde{g} : \mathbb{R}^{N+1} \to \mathbb{R}$ defined as

$$\tilde{g}(x,y) := \max\{g(x) - g(\bar{x}(1)),\, y\}.$$

We introduce the following auxiliary optimal control problem with state constraints

$$(\mathrm{P}^\gamma) \quad \begin{cases} \text{Minimize} & J(x,y,u,\alpha) := \tilde{g}(x(1),y(1)) + \int_0^1 \Lambda(t,\alpha(t))dt \\ \text{over all} & (x,y) \in W^{1,1}\left([0,1];\mathbb{R}^{N+1}\right) \text{ and } (u,\alpha) \in \mathcal{M}\left([0,1];\mathbb{R}^{m+l}\right) \\ \text{such that} & \dot{x}(t) = f(t,x(t),u(t)) - \sum_{i=1}^{l} \alpha_i(t)\nabla_x h_i(t,x(t)), \quad \text{a.e. on } [0,1], \\ & \dot{y}(t) = -\sum_{i=1}^{l} \alpha_i(t) h_i(t,x(t)), \quad \text{a.e. on } [0,1], \\ & (u(t),\alpha(t)) \in U(t) \times A \qquad \text{for a.e. } t \in [0,1], \\ & h_i(t,x(t)) \le 0, \qquad \text{for any } t \in [0,1],\ i = 1,\dots,l \\ & x(0) = x_0, \qquad y(0) = 0, \\ & \|(x,y) - (\bar{x},\bar{y})\|_{L^\infty} + \|\alpha - \bar{\alpha}\|^2_{L^2} \le \delta. \end{cases}$$

Here we regard $t \mapsto (u(t), \alpha(t))$ as a new control variable with values in the space $\mathbb{R}^{m+l} \cong \mathbb{R}^m \times \mathbb{R}^l$. Let us call X the subset of $W^{1,1}\left([0,1];\mathbb{R}^{N+1}\right) \times \mathcal{M}\left([0,1];\mathbb{R}^{m+l}\right)$ consisting of all feasible processes (x, y, u, α) for the auxiliary problem (P^γ).

Then it is not difficult to check that $(\bar{x}, \bar{y}, \bar{u}, \bar{\alpha})$ is a minimizer for (P^γ), where $\bar{y} = 0$. Indeed, let us first observe that, since $y(0) = 0$ and $\alpha_i(t)h_i(t, x(t)) \le 0$ for all $i = 1, \dots, l$, for a.e. $t \in [0,1]$, then any process $(x, y, u, \alpha) \in X$ is such that $y(t) \ge 0$ for all $t \in [0,1]$. This in particular implies that

$$\tilde{g}(x(1), y(1)) \ge 0, \qquad \forall\, (x, y, u, \alpha) \in X.$$

Since $(\bar{x}, \bar{y}, \bar{u}, \bar{\alpha}) \in X$ and $J(\bar{x}, \bar{y}, \bar{u}, \bar{\alpha}) = 0$, then this shows that $(\bar{x}, \bar{y}, \bar{u}, \bar{\alpha})$ is a minimizer for (P^γ).

4. Necessary conditions. Let us begin by recalling some auxiliary results from the literature. Consider a general optimal control problem of Bolza type with state constraints:

$$(\mathsf{P}_\mathrm{B}) \qquad \begin{cases} \text{Minimize} & \varphi(\mathsf{x}(1)) + \displaystyle\int_0^1 \mathsf{L}(t, \mathsf{x}(t), \omega(t))dt \\ \text{over all} & \mathsf{x} \in W^{1,1}([0,1];\mathbb{R}^\mathsf{n}) \text{ and } \omega \in \mathcal{M}([0,1];\mathbb{R}^\mathsf{m}) \\ \text{such that} & \dot{\mathsf{x}}(t) = \Psi(t, \mathsf{x}(t), \omega(t)), \quad \text{a.e. on } [0,1], \\ & \omega(t) \in \Omega(t), \qquad \text{for any } t \in [0,1], \\ & \mathsf{h}_i(t, \mathsf{x}(t)) \le 0, \qquad \text{for any } t \in [0,1],\ i = 1, \dots, \mathsf{l}, \\ & \mathsf{x}(0) = \mathsf{x}_0. \end{cases}$$

Let us recall that a feasible process $(\bar{\mathsf{x}}, \bar{\omega})$ for problem (P_B) is said to be a $W^{1,1}$ *local minimizer* if there exists $\varrho > 0$ such that

$$\varphi(\bar{\mathsf{x}}(1)) + \int_0^1 \mathsf{L}(t, \bar{\mathsf{x}}(t), \bar{\omega}(t))dt \le \varphi(\mathsf{x}(1)) + \int_0^1 \mathsf{L}(t, \mathsf{x}(t), \omega(t))dt$$

for any feasible process (x, ω) for problem (P_B) such that $\|\mathsf{x} - \bar{\mathsf{x}}\|_{W^{1,1}} \le \varrho$.

Optimality conditions for state constrained (and free end-point) problems such as (P_B) in form of a maximum principle are well-known nowadays. The following is a suitable version for the case we are considering (see for instance [37, Chapter 9]).

LEMMA 4.1. *Let $(\bar{\mathsf{x}}, \bar{\omega})$ be a $W^{1,1}$ local minimizer for the problem (P_B). Assume that for some $\varrho > 0$ the following holds:*

(A1) *the mappings $(t, \omega) \mapsto \Psi(t, \mathsf{x}, \omega)$ and $(t, \omega) \mapsto \mathsf{L}(t, \mathsf{x}, \omega)$ are $\mathcal{L} \times \mathcal{B}^\mathsf{m}$ measurable on $[0,1] \times \mathbb{R}^\mathsf{m}$ for $\mathsf{x} \in \mathbb{R}^\mathsf{n}$ fixed, and there is a Borel measurable function $\kappa : [0,1] \times \mathbb{R}^\mathsf{m} \to \mathbb{R}$ such that $t \mapsto \kappa(t, \bar{\omega}(t))$ is integrable and*

$$\|\Psi(t, \mathsf{x}, \omega) - \Psi(t, \mathsf{y}, \omega)\| + |\mathsf{L}(t, \mathsf{x}, \omega) - \mathsf{L}(t, \mathsf{y}, \omega)| \le \kappa(t, \omega)\|\mathsf{x} - \mathsf{y}\|$$

for all $\mathsf{x}, \mathsf{y} \in \mathbb{B}(\bar{\mathsf{x}}(t), \varrho)$, $\omega \in \Omega(t)$ and for a.e. $t \in [0,1]$.

(A2) $\mathrm{Gr}\,\Omega \subseteq \mathbb{R}^{\mathsf{m}+1}$ *is a nonempty $\mathcal{L} \times \mathcal{B}^\mathsf{m}$ measurable set.*

(A3) $\mathsf{x} \mapsto \varphi(\mathsf{x})$ *is Lipschitz continuous on $\mathbb{B}(\bar{\mathsf{x}}(1), \varrho)$.*

(A4) *Each h_i is continuous on $[0,1] \times \mathbb{R}^\mathsf{n}$ with $\mathsf{x} \mapsto \mathsf{h}_i(t, \mathsf{x})$ being differentiable on $\mathbb{B}(\bar{\mathsf{x}}(t), \varrho)$ for all $t \in [0,1]$ fixed such that $\nabla_\mathsf{x}\mathsf{h}_i$ is continuous at any $(t, \mathsf{x}) \in [0,1] \times \mathbb{R}^\mathsf{n}$ that satisfies $\mathsf{x} \in \mathbb{B}(\bar{\mathsf{x}}(t), \varrho)$.*

Then there exist $\mathsf{p} \in W^{1,1}([0,1];\mathbb{R}^\mathsf{n})$, $\lambda \ge 0$ and (positive) Borel measures $\mu_1, \dots, \mu_\mathsf{l}$ on $[0,1]$ such that

(i) $(\lambda, \mu_1, \ldots, \mu_{\mathsf{l}}, \mathsf{p}) \neq (0, 0, \ldots, 0, 0)$;
(ii) $-\dot{\mathsf{p}}(t) \in \operatorname{co} \partial_x \mathsf{H}_\lambda\left(t, \bar{\mathsf{x}}(t), \mathsf{q}(t), \bar{\omega}(t)\right)$ *for a.e.* $t \in [0, 1]$;
(iii) $-\mathsf{q}(1) \in \lambda \partial \varphi(\bar{\mathsf{x}}(1))$;
(iv) $\mathsf{H}_\lambda\left(t, \bar{\mathsf{x}}(t), \mathsf{q}(t), \bar{\omega}(t)\right) = \max_{\omega \in \Omega} \mathsf{H}_\lambda\left(t, \bar{\mathsf{x}}(t), \mathsf{q}(t), \omega\right)$ *for a.e.* $t \in [0, 1]$;
(v) $\operatorname{supp}(\mu_i) \subseteq \{t \in [0, 1] \mid \mathsf{h}_i(t, \bar{\mathsf{x}}(t)) = 0\}$ *for all* $i = 1, \ldots, \mathsf{l}$.

Here $\mathsf{H}_\lambda\left(t, \mathsf{x}, \mathsf{q}, \omega\right) = \langle \mathsf{q}, \Psi(t, \mathsf{x}, \omega)\rangle - \lambda \mathsf{L}(t, \mathsf{x}, \omega)$ *and*

$$\mathsf{q}(t) = \begin{cases} \mathsf{p}(t) + \displaystyle\sum_{i=1}^{\mathsf{l}} \int_{[0,t[} \nabla_{\mathsf{x}} \mathsf{h}_i(s, \bar{\mathsf{x}}(s)) \mu_i(ds) & \text{if } t \in [0, 1[\\ \mathsf{p}(1) + \displaystyle\sum_{i=1}^{\mathsf{l}} \int_{[0,1]} \nabla_{\mathsf{x}} \mathsf{h}_i(s, \bar{\mathsf{x}}(s)) \mu_i(ds) & \text{if } t = 1 \end{cases}$$

Furthermore, if stronger conditions are satisfied then degenerate multipliers can be ruled out from Lemma 4.1, and normality ($\lambda = 1$) can be ensured.

LEMMA 4.2. *Let* $(\bar{\mathsf{x}}, \bar{\omega})$ *be a* $W^{1,1}$ *local minimizer for the problem* $(\mathsf{P}_{\mathrm{B}})$. *Assume that for some* $\varrho > 0$ *the assumptions in Lemma 4.1 hold. Then the conditions of Lemma 4.1 are satisfied with* $\lambda = 1$ *provided that there exist* $r, \eta, \mathbf{k}_\Psi, \mathbf{k}_\omega > 0$ *such that*

(CQ_1) *for all* $\mathsf{x}, \mathsf{y} \in \mathbb{B}(\bar{\mathsf{x}}(0), \varrho)$, $\omega \in \Omega(t)$ *and for a.e.* $t \in [0, r[$

$$\|\Psi(t, \mathsf{x}, \omega) - \Psi(t, \mathsf{y}, \omega)\| + |\mathsf{L}(t, \mathsf{x}, \omega) - \mathsf{L}(t, \mathsf{y}, \omega)| \leq \mathbf{k}_\Psi \|\mathsf{x} - \mathsf{y}\|$$

(CQ_2) *if* $\max\limits_{i=1,\ldots,\mathsf{l}} \mathsf{h}_i(0, \mathsf{x}_0) = 0$, *then there exists a measurable selection (a control)* ω_1 *of* $\Omega : [0, 1] \to \mathbb{R}^{\mathsf{m}}$ *satisfying that for a.e.* $t \in [0, r[$

$$\max\left\{\|\Psi(t, \mathsf{x}_0, \bar{\omega}(t))\|, |\mathsf{L}(t, \mathsf{x}_0, \bar{\omega}(t))|, \|\Psi(t, \mathsf{x}_0, \omega_1(t))\|, |\mathsf{L}(t, \mathsf{x}_0, \omega_1(t))|\right\} \leq \mathbf{k}_\omega$$

and

$$\langle \nabla_{\mathsf{x}} \mathsf{h}_i(s, \mathsf{x}), \Psi(t, \mathsf{x}_0, \omega_1(t)) - \Psi(t, \mathsf{x}_0, \bar{\omega}(t))\rangle < -\eta$$

for any $i \in \{1, \ldots, \mathsf{l}\}$ *with* $\mathsf{h}_i(s, \mathsf{x}) \geq 0$, $\mathsf{x} \in \mathbb{B}(\mathsf{x}_0, r)$ *and* $s \in [0, r[$.

(CQ_3) *there exists a measurable selection (a control)* ω_2 *of* $\Omega : [0, 1] \to \mathbb{R}^{\mathsf{m}}$ *satisfying that*

$$\|\Psi(t, \bar{\mathsf{x}}(t), \bar{\omega}(t)) - \Psi(t, \bar{\mathsf{x}}(t), \omega_2(t))\| + |\mathsf{L}(t, \bar{\mathsf{x}}(t), \bar{\omega}(t)) - \mathsf{L}(t, \bar{\mathsf{x}}(t), \omega_2(t))| \leq \mathbf{k}_\omega$$

and

$$\langle \nabla_{\mathsf{x}} \mathsf{h}_i(s, \bar{\mathsf{x}}(s)), \Psi(t, \bar{\mathsf{x}}(t), \omega_2(t)) - \Psi(t, \bar{\mathsf{x}}(t), \bar{\omega}(t))\rangle < -\eta$$

for any $i \in \{1, \ldots, \mathsf{l}\}$ *with* $\mathsf{h}_i(s, \bar{\mathsf{x}}(s)) = 0$, *for a.e.* $t, s \in (\tau - r, \tau] \cap [0, 1]$, *where* $\tau := \inf\left\{t \in [0, 1] \mid \exists i \in I(t, \bar{\mathsf{x}}(t)), \displaystyle\int_{[t,1]} \mu_i(ds) = 0\right\}$.

Proof. This results is a direct consequence of [25, Theorem 4.2]. It is enough to remark that, under the assumptions we have considered in this paper, it follows that

$$\partial_x^{>} h(t, x) = \begin{cases} \operatorname{co}\left\{\nabla_x \mathsf{h}_i(t, x) \mid \mathsf{h}_i(t, x) = \mathsf{h}(t, x)\right\} & \text{if } \mathsf{h}(t, x) \geq 0, \\ \emptyset & \text{otherwise,} \end{cases}$$

where $\partial_x^{>}\mathsf{h}(t,x)$ stands for the hybrid partial subdifferential of $\mathsf{h}(t,x) = \max_{i=1,\ldots,l} \mathsf{h}_i(t,x)$. Therefore, it is not difficult to check that (CQ_2) and (CQ_3) imply (CQ) and (CQ_n) in [25, Theorem 4.2], respectively. ∎

Remark 4.3. Let us point out that the constraints qualifications in [25, Theorem 4.2] were for the first time proposed in [24] and were as well stated in terms of the hybrid partial subdifferential.

4.1. Optimality conditions for the auxiliary problem. It is not difficult to check that, for any $\gamma > 0$ the assumptions in Lemma 4.1 hold for the state variables $\mathsf{x} = (x,y) \in \mathbb{R}^{N+1}$, the control $\omega = (u,a) \in \mathbb{R}^{m+l}$,

$$\varphi(x,y) = \tilde{g}(x,y), \quad \mathsf{h}_i(t,x,y) = h_i(t,x), \quad \Omega(t) = U(t) \times A,$$

$$\Psi(t,(x,y),(u,a)) = \left(f(t,x,u) - \sum_{i=1}^{l} a_i \nabla_x h_i(t,x),\ -\sum_{i=1}^{l} a_i h_i(t,x) \right)$$

and

$$\mathsf{L}(t,(x,y),(u,a)) = \Lambda(t,a).$$

One then can apply Lemma 4.1 to (P^γ), implying the existence of some $\lambda^\gamma \geq 0$, $(p_x^\gamma, p_y^\gamma) \in W^{1,1}\left([0,1];\mathbb{R}^{N+1}\right)$ and (positive) Borel measures $\mu_1^\gamma, \ldots, \mu_l^\gamma$ such that

(i) $\left(\lambda^\gamma, \mu_1^\gamma, \ldots, \mu_l^\gamma, p_x^\gamma, p_y^\gamma\right) \neq (0,0,\ldots,0,0,0)$;

(ii) for a.e. $t \in [0,1]$,

$$\begin{aligned} -\dot{p}_x^\gamma(t) \in& \operatorname{co} \partial_x \left\langle q_x^\gamma(t), f(t,\cdot,\bar{u}(t)) - \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t,\cdot) \right\rangle (\bar{x}(t)) \\ & - p_y^\gamma(t) \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t,\bar{x}(t)), \\ -\dot{p}_y^\gamma(t) =& 0; \end{aligned}$$

(iii) $-(q_x^\gamma(1), p_y^\gamma(1)) \in \lambda^\gamma \partial \tilde{g}(\bar{x}(1), \bar{y}(1))$;

(iv) for a.e. $t \in [0,1]$

$$\begin{aligned} &\left\langle q_x^\gamma(t), f(t,\bar{x}(t),\bar{u}(t)) - \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t,\bar{x}(t)) \right\rangle - p_y^\gamma(t) \sum_{i=1}^{l} \bar{\alpha}_i(t) h_i(t,\bar{x}(t)) \\ &= \max_{u \in U(t)} \left\{ \langle q_x^\gamma(t), f(t,\bar{x}(t),u) \rangle \right\} - \min_{a \in A} \sum_{i=1}^{l} \left\{ a_i b_i^\gamma(t) + \frac{\gamma \lambda^\gamma}{2} (a_i - \bar{\alpha}_i(t))^2 \right\} \end{aligned}$$

(v) $\operatorname{supp}(\mu_i^\gamma) \subseteq \{t \in [0,1] \mid h_i(t,\bar{x}(t)) = 0\}$ for all $i = 1,\ldots,l$.

where

$$b_i^\gamma(t) := \langle q_x^\gamma(t), \nabla_x h_i(t,\bar{x}(t)) \rangle + p_y^\gamma(t) h_i(t,\bar{x}(t)), \qquad \forall t \in [0,1]$$

and

$$q_x^\gamma(t) = \begin{cases} p_x^\gamma(t) + \displaystyle\sum_{i=1}^{l} \int_{[0,t[} \nabla_x h_i(s,\bar{x}(s))\mu_i^\gamma(ds) & \text{if } t \in [0,1[, \\ p_x^\gamma(1) + \displaystyle\sum_{i=1}^{l} \int_{[0,1]} \nabla_x h_i(s,\bar{x}(s))\mu_i^\gamma(ds) & \text{if } t = 1. \end{cases}$$

4.1.1. Constraint Qualifications. We will now show that the optimal control problem (P^γ) satisfies certain constraint qualification which guarantees that Lemma 4.1 is satisfied in normal form. This in turn will mean that Lemma 4.2 can be applied. To be more precise, we will prove that $\lambda^\gamma = 1$.

To prove this result, we require the following intermediate result.

LEMMA 4.4. *Let $g_1, \ldots, g_k \in \mathbb{R}^N$ be positively linearly independent. Then, there are $\beta > 0$ and $a \in \Delta_k = \{x \in \mathbb{R}^k \mid x_i \geq 0,\ \sum_{i=1}^k x_i = 1\}$ such that*

$$\left\langle g_j, \sum_{i=1}^{k} a_i g_i \right\rangle > \beta, \qquad \forall j = 1, \ldots, k.$$

Proof. Define $\Phi : \Delta_k \to \mathbb{R}^k$ via the formula

$$\Phi_j(a) := \left\langle g_j, \sum_{i=1}^{k} a_i g_i \right\rangle, \qquad \forall a \in \Delta_k, j = 1, \ldots, k.$$

By contradiction, assume that $\Phi(\Delta_k) \cap \mathbb{R}^k_{++} = \emptyset$, where $\mathbb{R}^k_{++}$ is the interior of the positive orthant of $\mathbb{R}^k$. This means in particular that the nonempty convex compact set $\Phi(\Delta_k)$ can be separated from $\mathbb{R}^k_{++}$, which means that for some $s_1, \ldots, s_k \in \mathbb{R}$, not all zero at the same time, we have that

$$\left\langle \sum_{j=1}^{k} s_j g_j, \sum_{i=1}^{k} a_i g_i \right\rangle \leq \sum_{j=1}^{k} s_j x_j, \qquad \forall a \in \Delta_k,\ x_1, \ldots, x_k \geq 0.$$

From here, it follows that $s_1, \ldots, s_k \geq 0$, and so, without lost of generality, we may assume that $s = (s_1, \ldots, s_k) \in \Delta_k$. Thus, evaluating the preceding inequality at $a = s$ and $x = 0$ we get $\sum_{j=1}^k s_j g_j = 0$, which contradicts the fact that $g_1, \ldots, g_k \in \mathbb{R}^N$ are positively linearly independent vectors. ∎

We are now ready to prove that normality can be ensured for problem (P^γ).

PROPOSITION 4.5. *For each $\gamma > 0$, one can choose $\lambda^\gamma = 1$, where λ^γ is the multiplier given by Lemma 4.1 applied to (P^γ).*

Proof. It is enough to check that constraints qualification conditions (CQ_1), (CQ_2) and (CQ_3) in Lemma 4.2 hold for $\mathsf{x} = (\bar{x}, \bar{y})$ and $\bar{\omega} = (\bar{u}, \bar{\alpha})$, where

$$\bar{y}(t) = -\sum_{i=1}^{l} \int_0^t -\bar{\alpha}_i(s) h_i(s, \bar{x}(s)) ds.$$

First of all, (CQ_1) in Lemma 4.2 holds because of Hypothesis 2.6 and Hypothesis 2.11. Indeed, for any $(x,y), (\tilde{x},\tilde{y}) \in \mathbb{B}((x_0,0),\delta)$, $u \in U(t)$ and $a \in A$ and for a.e.

$t \in [0, r'[$ we have

$$\|\Psi(t,(x,y),(u,a)) - \Psi(t,(\tilde{x},\tilde{y}),(u,a))\| \leq \left(K' + 2L_h \sum_{i=1}^{l} (\|\bar{\alpha}_i\|_{L^\infty} + 1)\right) \|x - \tilde{x}\|.$$

Let us now check (CQ_2) in Lemma 4.2 holds. Notice that, thanks to Hypothesis 2.5 and Lemma 2.9, for any feasible control (u, α) for problem (P^γ) we have that $\|\Psi(t,(x_0,0),(u(t),\alpha(t)))\|$ and $|\mathsf{L}(t,(x_0,0),(u(t),\alpha(t)))|$ are uniformly bounded for $t \in [0,1]$. Consequently, to check (CQ_2) we only need to focus on the inequality with the inner product.

Suppose that $\max_{i=1,\dots,l} h_i(0,x_0) = 0$. Let $a^0 \in \Delta_l$ and $\beta^0 > 0$ be given by Lemma 4.4 with $g_i = \nabla_x h_i(0,x_0)$ for any $i \in I(0,x_0)$. Notice that thanks to Hypothesis 2.7 the family $\{\nabla_x h_i(0,x_0)\}_{i \in I(0,x_0)}$ is positively linearly independent, thus it is enough to set $a^0_j = 0$ if $j \notin I(0,x_0)$. In the light of Hypothesis 2.6, it follows then that for any $j \in I(0,x_0)$ we have

$$\begin{aligned}
\left\langle \nabla_x h_j(s,x), \sum_{i=1}^{l} a_i^0 \nabla_x h_i(t,x_0) \right\rangle &= \left\langle \nabla_x h_j(0,x_0), \sum_{i=1}^{l} a_i^0 \nabla_x h_i(0,x_0) \right\rangle \\
&\quad + \left\langle \nabla_x h_j(s,x) - \nabla_x h_j(0,x_0), \sum_{i=1}^{l} a_i^0 \nabla_x h_i(0,x_0) \right\rangle \\
&\quad + \left\langle \nabla_x h_j(s,x), \sum_{i=1}^{l} a_i^0 \left[\nabla_x h_i(t,x_0) - \nabla_x h_i(0,x_0)\right] \right\rangle \\
&> \beta^0 - L_h^2 (s + |x - x_0| + t)
\end{aligned}$$

Let $(s,x) \in [0,1] \times \mathbb{R}^N$ such that $h_j(s,x) \geq 0$. Notice that, thanks to Hypothesis 2.6, there is $r^0 > 0$ such that if $s, t \in [0, r^0)$ and $x \in \mathbb{B}(x_0, r^0)$, then we may assume that $j \in I(0,x_0)$. Therefore, it is not difficult to check that (CQ_2) holds for $\bar{\omega} = (\bar{u}, \bar{\alpha})$ with $\omega_1 = (\bar{u}, a^0 + \bar{\alpha})$, $r = \min\{r^0, \frac{\beta^0}{6L_h^2}\}$ and $\eta = \frac{\beta^0}{2}$.

Let us finally verify that (CQ_3) in Lemma 4.2 holds for $\bar{\mathsf{x}} = (\bar{x}, \bar{y})$ and $\bar{\omega} = (\bar{u}, \bar{\alpha})$. Let $\tau^\gamma \in [0,1]$ be given by

$$\tau^\gamma := \inf \left\{ t \in [0,1] \mid \exists i \in I(t, \bar{x}(t)),\ \int_{[t,1]} \mu_i^\gamma(ds) = 0 \right\}.$$

Let $a^\tau \in \Delta_l$ and $\beta^\tau > 0$ be given by Lemma 4.4 with $g_i = \nabla_x h_i(\tau, \bar{x}(\tau))$ for any $i \in I(\tau, \bar{x}(\tau))$. Similarly as done for checking (CQ_2), we set $a_j^\tau = 0$ whenever $j \notin I(\tau, \bar{x}(\tau))$. It follows then that for any $j \in I(\tau^\gamma, \bar{x}(\tau^\gamma))$ we have

$$
\begin{aligned}
\left\langle \nabla_x h_j(s, \bar{x}^\gamma(s)), \sum_{i=1}^{l} a_i^\tau \nabla_x h_i(t, \bar{x}^\gamma(t)) \right\rangle &= \left\langle \nabla_x h_j(\tau, \bar{x}(\tau)), \sum_{i=1}^{l} a_i^\tau \nabla_x h_i(\tau, \bar{x}(\tau)) \right\rangle \\
&+ \left\langle \nabla_x h_j(s, \bar{x}(s)) - \nabla_x h_j(\tau, \bar{x}(\tau)), \sum_{i=1}^{l} a_i^\tau \nabla_x h_i(\tau, \bar{x}(\tau)) \right\rangle \\
&+ \left\langle \nabla_x h_j(s, \bar{x}(s)), \sum_{i=1}^{l} a_i^\tau \left[\nabla_x h_i(t, \bar{x}(t)) - \nabla_x h_i(\tau, \bar{x}(\tau))\right] \right\rangle \\
&> \beta^\tau - L_h^2(|s-\tau| + |\bar{x}(s) - \bar{x}(\tau)| + |\bar{x}(t) - \bar{x}(\tau)| + |t-\tau|)
\end{aligned}
$$

Hence, since $\bar{x}$ is continuous, there is $r^\tau > 0$ such that

$$
|\bar{x}(s) - \bar{x}(\tau)| + |\bar{x}(t) - \bar{x}(\tau)| \leq \frac{\beta^\tau}{6L_h^2}, \qquad \forall s, t \in (\tau^\gamma - r^\tau, \tau^\gamma] \cap [0,1].
$$

Therefore, it is not difficult to check that (CQ_3) holds for $\bar{\mathsf{x}} = (\bar{x}, \bar{y})$ and $\bar{\omega} = (\bar{u}, \bar{\alpha})$ with $\omega_2 = (\bar{u}, a^\tau + \bar{\alpha})$, $r = \min\{r^\tau, \frac{\eta^\tau}{6L_h^2}\}$ and $\eta = \frac{\beta^\tau}{2}$, which completes the proof. $\square$

As a direct consequence of Lemma 4.1 and Lemma 4.2, we get the following optimality conditions for the auxiliary optimal control problem (P^γ).

Proposition 4.6. *For any given $\gamma > 0$, recall that the assumption (3.1) still holds. Then there exist $p^\gamma \in W^{1,1}([0,1]; \mathbb{R}^N)$, $\sigma^\gamma, \xi^\gamma \geq 0$, positive definite Borel measures $\mu_1^\gamma, \dots, \mu_l^\gamma$ such that:*

(a) $\sigma^\gamma + \xi^\gamma + \sum_{i=1}^{l} \mu_i^\gamma([0,1]) = 1$;

(b) for a.e. $t \in [0,1]$

$$
\begin{aligned}
-\dot{p}^\gamma(t) \in &\operatorname{co} \partial_x \left\langle q^\gamma(t), f(t, \cdot, \bar{u}(t)) - \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \cdot) \right\rangle (\bar{x}(t)) \\
&+ \xi^\gamma \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t)),
\end{aligned} \tag{4.1}
$$

where

$$
q^\gamma(t) = \begin{cases} p^\gamma(t) + \displaystyle\sum_{i=1}^{l} \int_{[0,t[} \nabla_x h_i(s, \bar{x}(s)) \mu_i^\gamma(ds) & \text{if } t \in [0,1[, \\ p^\gamma(1) + \displaystyle\sum_{i=1}^{l} \int_{[0,1]} \nabla_x h_i(s, \bar{x}(s)) \mu_i^\gamma(ds) & \text{if } t = 1, \end{cases};
$$

(c) $-q^\gamma(1) \in \sigma^\gamma \partial g(\bar{x}(1))$ *;*

(d) for a.e. $t \in [0,1]$

$$
\max_{u \in U(t)} \{\langle q^\gamma(t), f(t, \bar{x}(t), u)\rangle\} = \langle q^\gamma(t), f(t, \bar{x}(t), \bar{u}(t))\rangle, \tag{4.2}
$$

and for each $i = 1, \dots, l$

$$
\min_{a \in [0, A_i]} \left\{ a b_i^\gamma(t) + \frac{\gamma(a - \bar{\alpha}_i(t))^2}{2} \right\} = \bar{\alpha}_i(t) b_i^\gamma(t) \tag{4.3}
$$

where

$$b_i^\gamma(t) := \langle q^\gamma(t), \nabla_x h_i(t,\bar{x}(t))\rangle - \xi^\gamma h_i(t,\bar{x}(t)), \qquad \forall t \in [0,1]$$

(e) for each $i = 1,\dots,l$ we have $\operatorname{supp}(\mu_i^\gamma) \subseteq \{t \in [0,1] \mid h_i(t,\bar{x}(t)) = 0\}$.
Moreover, there exists $c > 0$ such that for any $\gamma > 0$ and $n \in \mathbb{N}$ it follows that

$$\|p^\gamma(1)\| \le c \quad \textit{and} \quad \|\dot{p}^\gamma(t)\| \le c(K_f(t,\bar{u}(t)) + 1), \qquad \textit{for a.e. } t \in [0,1]. \tag{4.4}$$

Proof. Let $(p_x^\gamma, p_y^\gamma) \in W^{1,1}\left([0,1];\mathbb{R}^{N+1}\right)$, $\lambda^\gamma \ge 0$ and (positive) Borel measures $\mu_1^\gamma,\dots,\mu_l^\gamma$ given by Lemma 4.1 applied to (P^γ). Recall that by Lemma 4.2 we can actually take $\lambda^\gamma = 1$.

By a use of the max rule for limiting subdifferential (see [37, Theorem 5.5.2]), it follows from Lemma 4.1, condition iii), that there exist $\sigma^\gamma, \xi^\gamma \ge 0$, such that

$$\partial\tilde{g}(\bar{x}(1),\bar{y}(1)) \subseteq \sigma^\gamma\left(\partial g(\bar{x}(1)) \times \{0\}\right) + \xi^\gamma\left(\{0\} \times \{1\}\right) \tag{4.5}$$

and $\sigma^\gamma + \xi^\gamma = 1$. In view of the latter relation, one has that $1 \le \sigma^\gamma + \xi^\gamma + \sum_{i=1}^l \mu_i^\gamma([0,1])$ and in particular Item (a) follows (up to a rescaling, if necessary). Observe that, in view of Lemma 4.1, condition ii) and of (4.5), one has that $p_y^\gamma(t) \equiv -\xi^\gamma$ for any $t \in [0,1]$. Furthermore, by defining $p^\gamma(t) = p_x^\gamma(t)$, we see that Item (b) is satisfied and, by observing again (4.5), also Item (c) is satisfied. Notice that at this stage that $\mu_i^\gamma(ds)$ is a positive definite measure such that

$$\operatorname{supp}(\mu_i^\gamma) \subseteq \{t \in [0,1] \mid h_i(t,\bar{x}(t)) = 0\}$$

and therefore Item (e) also holds.

Let us now show that the maximum condition turns into two independent maximum conditions (Item (d) in the statement of the proposition), one for the standard control u and another for the auxiliary control α.

Notice that the maximality condition given by Lemma 4.1 can be written as

$$\begin{aligned}
&\max_{u \in U(t)} \{\langle q^\gamma(t), f(t,\bar{x}(t),u) - f(t,\bar{x}(t),\bar{u}(t))\rangle\} \\
&= \min_{a \in A} \sum_{i=1}^l \left\{ a_i b_i^\gamma(t) + \frac{\gamma(a_i - \bar{\alpha}_i(t))^2}{2} \right\} - \sum_{i=1}^l \bar{\alpha}_i(t) b_i^\gamma(t).
\end{aligned}$$

Since the right-hand side term in the equality is nonpositive and the left-hand side term is nonnegative, we get that maximality condition splits in two conditions, namely, for a.e. $t \in [0,1]$, one gets (4.2) and

$$\min_{a \in A} \sum_{i=1}^l \left\{ a_i b_i^\gamma(t) + \frac{\gamma(a_i - \bar{\alpha}_i(t))^2}{2} \right\} = \sum_{i=1}^l \bar{\alpha}_i(t) b_i^\gamma(t). \tag{4.6}$$

By the definition of A, it is not difficult to see that (4.3) follows from (4.6).

Notice that thanks to the definition of q^γ, Item (a) and the Hypothesis 2.6 we get

$$\|q^\gamma(t)\| \le \|p^\gamma(t)\| + L_h \sum_{i=1}^l \mu_i^\gamma([0,1]) \le \|p^\gamma(t)\| + L_h, \qquad \forall t \in [0,1].$$

Furthermore, in view of Item (c) one obtains the following estimate:

$$\|p^\gamma(1)\| \leq \sigma^\gamma L_g + L_h \sum_{i=1}^{l} \mu_i^\gamma([0,1]) \leq L_g + L_h.$$

Therefore, the end point $\|p^\gamma(1)\|$ is uniformly bounded. Moreover, by the adjoint equation Item (b), Hypothesis 2.4, Hypothesis 2.6 and Lemma 2.9 it follows that for for a.e. $t \in [0,1]$ we also have

$$\|\dot{p}^\gamma(t)\| \leq \left(K_f(t, \bar{u}(t)) + L_h \sum_{i=1}^{l} (\|\bar{\alpha}_i\|_{L^\infty} + 1) \right) (\|p^\gamma(t)\| + L_h) + \xi^\gamma L_h \sum_{i=1}^{l} \|\bar{\alpha}_i\|_{L^\infty}.$$

Hence, in view of the previous estimates and an application of the Grönwall's Lemma, the relation (4.4) follows. This completes the proof. □

Remark 4.7. It is worth noticing at this stage that the necessary conditions provided by Proposition 4.6 are non-trivial even when σ^γ and $\sum_{i=1}^{l} \mu_i^\gamma([0,1])$ are both vanishing, as long as condition (3.1) holds. Indeed, for any given matrix-valued measurable selection

$$X(t) \in D_x f(t, \bar{x}(t), \bar{u}(t)), \qquad \text{a.e. } t \in [0,1],$$

one has that p^γ is the unique solution of the initial value problem

$$-\dot{p}(t) = p(t)X(t) + \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t)), \quad p(1) = 0.$$

By calling $M(t)$ the fundamental matrix solution associated to the linear system $\dot{v}(t) = X(t)v(t)$, then

$$p^\gamma(t) = M(t) \int_t^1 M^{-1}(\tau) \cdot \left(\sum_{i=1}^{l} \bar{\alpha}_i(\tau) \nabla_x h_i(\tau, \bar{x}(\tau)) \right) d\tau.$$

Furthermore in view of condition (3.1), one has that there exists a positive measure set I such that

$$\sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t)) \neq 0, \quad \text{a.e. } t \in I.$$

Hence, since $M(t)$ is a full rank matrix for each $t \in [0,1]$, one has that $p^\gamma(t)$ is an absolutely continuous arc such that $\sup_{t\in[0,1]} \|p^\gamma(t)\| \neq 0$.

4.2. About the convergence of the multipliers. For a given sequence $\gamma_n \to 0$ as $n \to +\infty$, let us study the convergence of the multipliers and the conclusions that can be gathered from Proposition 4.6. At this stage, we remind that condition (3.1) still holds.

Notice that from Proposition 4.6 it can be deduced that the sequence $\{p^{\gamma_n}\}_{n\in\mathbb{N}}$ is relatively compact in $W^{1,1}\left([0,1];\mathbb{R}^N\right)$ (thanks to [1, Theorem 0.3.4]), in the sense that one may assume that there is $p \in W^{1,1}\left([0,1];\mathbb{R}^N\right)$ such that $p^{\gamma_n} \to p$ uniformly on $[0,1]$ and $\dot{p}^{\gamma_n} \rightharpoonup \dot{p}$ weakly in $L^1\left([0,1];\mathbb{R}^N\right)$, along a subsequence, which we do not relabel.

Using now a routine convergence procedure, one can extract a subsequence (again we do not relabel) such that $\sigma^{\gamma_n} \to \sigma$, $\xi^{\gamma_n} \to \xi$, $q^{\gamma_n} \to q$ pointwise on $[0,1]$, $\mu_i^{\gamma_n} \rightharpoonup \mu_i$ weakly* for all $i = 1, \dots, l$ as $n \to +\infty$, where $\sigma, \xi \geq 0$, each μ_i is a positive definite Borel measure and $q \in \mathcal{BV}\left([0,1]; \mathbb{R}^N\right)$ is such that

$$q(t) = \begin{cases} p(t) + \displaystyle\sum_{i=1}^{l} \int_{[0,t[} \nabla_x h_i(s, \bar{x}(s))\mu_i(ds) & \text{if } t \in [0,1[, \\ p(1) + \displaystyle\sum_{i=1}^{l} \int_{[0,1]} \nabla_x h_i(s, \bar{x}(s))\mu_i(ds) & \text{if } t = 1. \end{cases}$$

It is not difficult to see that the weak-$\star$ convergence of the sequences $\{\mu_i^{\gamma_n}\}_{n\in\mathbb{N}}$ to μ_i implies that $\mu_i^{\gamma_n}([0,1]) \to \mu_i([0,1])$ as $n \to +\infty$, for any $i = 1, \dots, l$.

Therefore, it follows from Item (a) that $\sigma + \xi + \sum_{i=1}^{l} \mu_i([0,1]) = 1$ and, in view of the condition (3.1), one also has $\sigma + \xi \sum_{i=1}^{l} \int_0^1 \bar{\alpha}_i(t)dt + \sum_{i=1}^{l} \mu_i([0,1]) > 0$. Then, by further rescaling (if necessary), one obtains the condition

(I) $\sigma + \xi \displaystyle\sum_{i=1}^{l} \int_0^1 \bar{\alpha}_i(t)dt + \sum_{i=1}^{l} \mu_i([0,1]) = 1;$

Furthermore, the right-hand side in Item (b) converges a.e. in $[0,1]$ to

$$\operatorname{co} \partial_x \left\langle q(t), f(t, \cdot, \bar{u}(t)) - \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \cdot) \right\rangle (\bar{x}(t)) + \xi \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t)),$$

and we deduce that Item (b) yields to

(II) for a.e. $t \in [0,1]$

$$\begin{aligned} -\dot{p}(t) \in & \operatorname{co} \partial_x \left\{ q(t) \cdot \left(f(t, \bar{x}(t), \bar{u}(t)) - \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t)) \right) \right\} \\ & + \xi \sum_{i=1}^{l} \bar{\alpha}_i(t) \nabla_x h_i(t, \bar{x}(t)) \end{aligned}$$

Notice that the fact that the limiting subdifferential of a Lipschitz continuous function has closed graph ([9, Proposition 10.10]) combined with Item (c) leads to

(III) $-q(1) \in \sigma \partial g(\bar{x}(1));$

Note as well that (4.2) in Item (d) leads to the following maximality condition

(IV) for a.e. $t \in [0,1]$,

$$\langle q(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle = \max_{u \in U(t)} \left\{ \langle q(t), f(t, \bar{x}(t), u) \rangle \right\}.$$

In view of the properties of the weak convergence of measures, one has that the supports of the measures μ_i^γ for $i = 1, \dots, l$ satisfy

(V) for each $i = 1, \dots, l$ we have $\operatorname{supp}(\mu_i) \subseteq \{t \in [0,1] \mid h_i(t, \bar{x}(t)) = 0\}$.

Finally, let $I_i \subseteq [0,1]$ be a measurable set of full measure in $[0,1]$ such that (4.3) is well-defined for any $t \in I_i$. It is worth noticing that $\bar{\alpha}_i(t) < A_i$ and so is either 0 or an internal point of $[0, A_i]$, for every $t \in I_i$, $i = 1, \dots, l$. Furthermore, when $\bar{\alpha}_i(t) > 0$, then it satisfies the first order necessary condition

$$b_i^{\gamma_n}(t) = 0. \tag{4.7}$$

Since $b_i^{\gamma_n}(t) = \langle q^{\gamma_n}(t), \nabla_x h_i(t, \bar{x}(t))\rangle - \xi^{\gamma_n} h_i(t, \bar{x}(t))$ and by observing that, if $\bar{\alpha}_i(t) > 0$ then $h_i(t, \bar{x}(t)) = 0$, it follows from (4.7) that

$$\langle q^{\gamma_n}(t), \nabla_x h_i(t, \bar{x}(t))\rangle = 0, \qquad \text{when } \ \bar{\alpha}_i(t) > 0.$$

Consequently, by passing to the limit for $n \to \infty$ (4.3) in Item (d) yields to the following orthogonality condition.

(VI) $\langle q(t), \nabla_x h_i(t, \bar{x}(t))\rangle = 0$ for a.e. $t \in \{s \in [0,1] \mid \bar{\alpha}_i(s) > 0\}$.

This concludes the proof of Theorem 2.12 when condition (3.1) holds.

4.3. Proof of the main result. So far, we have shown Theorem 2.12 in the case in which the normal cone related to the process $(\bar{x}, \bar{u})$ is active (namely, when condition (3.1) holds). It remains to show Theorem 2.12 when $\bar{\alpha}_i(t) = 0$, a.e. $t \in [0,1]$ and for all $i = 1, \dots, l$. However in this latter case, $(\bar{x}, \bar{u})$ is also a minimizer of the standard optimal control problem with state constraints (and no end-point constraint)

$$(\bar{P}) \qquad \begin{cases} \text{Minimize} & g(x(1)) \\ \text{over all} & x \in W^{1,1}\left([0,1]; \mathbb{R}^N\right) \ \text{ and } u \in \mathcal{M}\left([0,1]; \mathbb{R}^m\right) \\ \text{such that} & \dot{x}(t) = f(t, x(t), u(t)), \quad \text{a.e. on } [0,1], \\ & u(t) \in U(t) \qquad \text{for a.e. } t \in [0,1], \\ & h_i(t, x(t)) \le 0, \qquad \text{for any } t \in [0,1],\ i = 1, \dots, l \\ & x(0) = x_0, \\ & \|x - \bar{x}\|_{L^\infty} \le \delta. \end{cases}$$

Again, it is not difficult to check that the assumptions in Lemma 4.1 hold for the state variables $\mathsf{x} = x \in \mathbb{R}^N$, the control $\omega = u \in \mathbb{R}^m$,

$$\varphi = g, \quad \mathsf{h}_i = h_i, \quad \Omega = U, \quad \Psi = f \quad \text{ and } \quad \mathsf{L} = 0.$$

One then can apply Lemma 4.1 to $(\bar{P})$, implying the existence of some $\sigma \ge 0$, $p \in W^{1,1}\left([0,1]; \mathbb{R}^N\right)$ and (positive) Borel measures $\mu_1, \dots, \mu_l$ such that

(i) $(\sigma, \mu_1, \dots, \mu_l) \ne (0, 0, \dots, 0)$;

(ii) for a.e. $t \in [0,1]$,

$$-\dot{p}(t) \in \operatorname{co} \partial_x \langle q(t), f(t, \bar{x}(t), \bar{u}(t))\rangle\,;$$

where

$$q(t) = \begin{cases} p(t) + \displaystyle\sum_{i=1}^{l} \int_{[0,t[} \nabla_x h_i(s, \bar{x}(s)) \mu_i(ds) & \text{if } t \in [0,1[, \\ p(1) + \displaystyle\sum_{i=1}^{l} \int_{[0,1]} \nabla_x h_i(s, \bar{x}(s)) \mu_i(ds) & \text{if } t = 1. \end{cases}$$

(iii) $-q(1) \in \sigma \partial g(\bar{x}(1))$;

(iv) for a.e. $t \in [0,1]$

$$\langle q(t), f(t, \bar{x}(t), \bar{u}(t))\rangle = \max_{u \in U(t)} \{\langle q(t), f(t, \bar{x}(t), u)\rangle\}\,;$$

(v) $\operatorname{supp}(\mu_i) \subseteq \{t \in [0,1] \mid h_i(t, \bar{x}(t)) = 0\}$ for all $i = 1, \dots, l$.

By rescaling the relation (i), one can easily obtains the non-triviality condition

$$\sigma + \sum_{i=1}^{l} \mu_i([0,1]) = 1.$$

Hence, by observing again conditions (ii)-(v) above, one has that the statement of Theorem 2.12 is verified also in the case in which the normal cone is not active. This concludes the proof of Theorem 2.12.

5. Examples.

5.1. A smooth boundary case. In this section we will apply Theorem 2.12 to a slight modification of an example presented in [5]. Consider the optimal control problem

$$\text{(E)} \qquad \begin{cases} \text{Minimize} & g(x(1)) \\ \text{over all} & x \in W^{1,1}([0,1];\mathbb{R}^2) \text{ and } u \in \mathcal{M}([0,1];\mathbb{R}^2) \\ \text{such that} & \dot{x}(t) \in u(t) - \mathcal{N}_C^P(x(t)), \quad \text{for a.e. } t \in [0,1], \\ & u(t) \in [-1,1]^2, \qquad \text{for a.e. } t \in [0,1], \\ & x(t) \in C, \qquad \text{for any } t \in [0,1]. \\ & x(0) = (0, x_2^0) \in C, \end{cases}$$

in which $C = \{(x_1,x_2) \in \mathbb{R}^2 \mid -x_2 \leq 0\}$, $x = (x_1,x_2)$, $g(x) = x_1 + x_2$ and $0 < x_2^0 < 1$. This is clearly an example of (P) in which the constraint consists in a fixed half plane described by the function $h(t,x) = -x_2$; here, since we are in the case $l = 1$, we suppress the subindex $i = 1$ in the data of the problem. Given the simple structure of the constraint, the normal cone can easily be written as

$$\mathcal{N}_C^P(x) = \left\{(0,-\alpha) \in \mathbb{R}^2 \mid \alpha \geq 0\right\}.$$

Since the contour lines of g are defined by $\nabla g(x) = (1,1)$, it is reasonable to expect that any optimal solution will reach the point $(-1,0)$. Indeed, it can be accomplished with the optimal control $\hat{u} = (-1,-1)$, which generates the optimal trajectory $\hat{x}$ touching ∂C at time $\hat{t} = x_2^0$ and sliding along ∂C in the interval $[\hat{t},1]$. In this case

$$\hat{\alpha}(t) = \begin{cases} 0 & \text{if } t \in [0,\hat{t}] \\ 1 & \text{if } t \in]\hat{t},1] \end{cases}$$

At this stage, it is worth noticing that such an optimal trajectory is not unique. Indeed, for instance, if $\tilde{t} = 1 - x_2^0$, then the control

$$\tilde{u}(t) = \begin{cases} (-1,0), & \text{if } t \in [0,\tilde{t}[\\ (-1,-1), & \text{if } t \in]\tilde{t},1] \end{cases}$$

generates a different optimal trajectory $\tilde{x}$ which touches ∂C just at $t = 1$.

Let $\bar{x}$ an optimal trajectory and $\bar{u}$ a corresponding optimal control. Let $\bar{t} \in]0,1]$ the first time at which $\bar{x}(\bar{t}) \in \partial C$, that is, $\bar{x}_2(\bar{t}) = 0$. Let $\bar{\alpha}$ be the reaction terms associated with the normal cone satisfying

$$\dot{x}(t) = u(t) - \bar{\alpha}(t) \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \text{for a.e. } t \in [0,1].$$

By the complementary condition (2.3), it follows that $\bar{\alpha}(t) = 0$ a.e. on $[0, \bar{t}]$. Moreover, if $\bar{x}_2(t) = 0$ for $t \in (a, b)$, then $\dot{\bar{x}}_2(t) = 0$ for any $t \in (a, b)$, which implies

$$\bar{u}_2(t) = -\bar{\alpha}(t) \leq 0, \qquad \text{a.e. whenever } \bar{x} \in \partial C.$$

Therefore, any optimal control must be such that $\bar{u}_2(t) \in [-1, 0]$ whenever $\bar{x} \in \partial C$.

Bearing the previous considerations in mind, we now apply Theorem 2.12. The adjoint equation (ii) reads as $\dot{p} = (0, \xi\bar{\alpha}(t))$ a.e. $t \in [0, 1]$, while the transversality condition (iii) implies $-q(1) = \sigma(1, 1)$, where $\sigma \geq 0$, $q_1(t) = p_1(t)$ for any $t \in [0, 1]$ and

$$q_2(t) = \begin{cases} p_2(t) - \mu([0, t[) & \text{if } t \in [0, 1[, \\ p_2(1) - \mu([0, 1]) & \text{if } t = 1, \end{cases}$$

where μ is a positive measure satisfying the conditions (i) and (v). Notice that $q_1(t) = p_1(t) = -\sigma$ for any $t \in [0, 1]$. The maximality condition (iii) leads then to

$$\sigma(\bar{u}_1(t) + 1) = q_2(t)\bar{u}_2(t) - |q_2(t)|, \qquad \text{a.e. on } [0, 1]. \tag{5.1}$$

The fact that $\sigma(\bar{u}_1(t) + 1) \geq 0$ and $q_2(t)\bar{u}_2(t) - |q_2(t)| \leq 0$, implies that both terms in (5.1) are zero. Let us also point out that by (vi) it follows that

$$q_2(t) = 0, \qquad \text{a.e. on } \{s \in [0, 1] \mid \bar{\alpha}_i(s) > 0\}.$$

Thus the maximality condition holds immediately on the set $\{s \in [0, 1] \mid \bar{\alpha}_i(s) > 0\}$.

Combining these facts we obtain

$$-\sigma = q_2(1) = p_2(1) - \mu([0, 1]) = p_2(t) + \xi \int_t^1 \bar{\alpha}(s)ds - \mu([0, 1]),$$

and so, by the non-triviality condition we get

$$p_2(t) = 2\mu([0, 1]) - 1 + \xi \int_0^t \bar{\alpha}(s)ds.$$

Therefore, since $\bar{\alpha}(t) = 0$ if $t \in [0, \bar{t}[$, we get that $p_2(t) = 2\mu([0, 1]) - 1$ for any $t \in [0, \bar{t}[$. By (v), it also follows that $q_2(t) = 2\mu([0, 1]) - 1$ for any $t \in [0, \bar{t}[$.

Notice that if $\mu([0, 1]) > \frac{1}{2}$, then $q_2(t) > 0$ on $[0, \bar{t}[$ and so $\bar{u}_2(t) = 1$ on $[0, \bar{t}[$, which is clearly a contradiction with the fact that $\bar{x}_2(\bar{t}) = 0$. If on the other hand $\mu([0, 1]) < \frac{1}{2}$, by continuity, there exists $\tau \in]\bar{t}, 1]$ such that

$$\xi \int_{\Theta_\tau} \bar{\alpha}(s)ds < \frac{1}{2} - \mu([0, 1]), \qquad \text{where } \Theta_\tau = [\bar{t}, \tau] \cap \{s \in [0, 1] \mid \bar{\alpha}_i(s) > 0\}.$$

Since

$$q_2(t) = 2\mu([0, 1]) - 1 + \xi \int_{\bar{t}}^t \bar{\alpha}(s)ds - \mu([0, t[), \qquad \forall t \in [\bar{t}, t[,$$

the latter implies that $q_2(t) < 0$ for on Θ_τ, which contradicts (vi). Therefore, we must have $\mu([0, 1]) = \frac{1}{2}$, and so $p_2(t) = q_2(t) = 0$ for any $t \in [0, \bar{t}[$.

We now distinguish between the cases $\sigma = 0$ and $\sigma > 0$. Let us consider first the abnormal case. If $\sigma = 0$, one has that $q_1 = p_1 \equiv 0$ on $[0, 1]$. Notice that the non-triviality condition implies that $\xi \int_0^1 \bar{\alpha}(s)ds = \frac{1}{2}$, in particular the normal cone must be active on a measurable set with positive measure. Let us point out that the multiplier

$$\mu([x_2^0, t[) = \mu([x_2^0, t]) = \frac{(t - x_2^0)}{2(1 - x_2^0)}, \qquad \forall t \in [x_2^0, 1] \qquad \text{and} \qquad \xi = \frac{1}{2(1 - x_2^0)}$$

with $\mathrm{supp}(\mu) = [x_2^0, 1]$, produce an admissible multiplier for the optimal trajectory $\hat{x}$ described above.

Consider now the normal case, that is, $\sigma > 0$ and so, one has $p_1 \equiv -\sigma$. From the maximality condition (iii) and the structure of the control set, we get then that $\bar{u}_1 \equiv -1$ on $[0, 1]$.

If $p_2 \equiv 0$, then $\sigma = \frac{1}{2}$, $\xi = 0$ and $\mu = \frac{1}{2}\delta_1$ produce an admissible multiplier for *any* optimal trajectory which satisfies all of the conditions of Theorem 2.12. This multiplier is exactly the one identified by the necessary conditions in [5, Example 1] and is the only multiplier related, for instance, to the optimal trajectory $\tilde{x}$ described above, which doesn't not satisfy the condition for the abnormal case.

5.2. A non-smooth boundary case. Consider the optimal control problem

$$\text{(E)} \qquad \begin{cases} \text{Minimize} & g(x(1)) \\ \text{over all} & x \in W^{1,1}([0,1]; \mathbb{R}^2) \text{ and } u \in \mathcal{M}([0,1]; \mathbb{R}) \\ \text{such that} & \dot{x}(t) \in \begin{pmatrix} u(t) \\ 0 \end{pmatrix} - \mathcal{N}_C^P(x(t)), \quad \text{for a.e. } t \in [0,1], \\ & u(t) \in [-1, 1], \qquad \text{for a.e. } t \in [0,1], \\ & x(t) \in C, \qquad \text{for any } t \in [0,1]. \\ & x(0) = (0, \frac{1}{4}) \in C, \end{cases}$$

in which $C = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \leq \frac{1}{2},\ 2x_1 + x_2 \leq 1\}$, $x = x(x_1, x_2)$ and $g(x) = x_2 - x_1$. This is clearly an example of (P) in which the constraint consists in the intersection of two half planes described by the functions $h_1(t, x) = x_1 + x_2 - \frac{1}{2}$ and $h_2(t, x) = 2x_1 + x_2 - 1$. The normal cone in this case can be written as

$$\mathcal{N}_C^P(x) = \left\{ \begin{pmatrix} \alpha_1 + 2\alpha_2 \\ \alpha_1 + \alpha_2 \end{pmatrix} \in \mathbb{R}^2 \mid \alpha_1, \alpha_2 \geq 0,\ \alpha_1(x_1 + x_2 - \frac{1}{2}) = \alpha_2(2x_1 + x_2 - 1) = 0 \right\}.$$

It is not difficult to see that in this case the global minimum is attained with the optimal control $\hat{u}(t) = 1$ for any $t \in [0, 1]$, which generates the optimal trajectory $\hat{x}$ touching ∂C at time $\hat{t} = \frac{1}{4}$ and sliding along ∂C in the interval $[\frac{1}{4}, 1]$, switching from the constraint given by h_1 to the constraint determined by h_2 at time $t = \frac{3}{4}$. In this case

$$\hat{\alpha}_1(t) = \begin{cases} \frac{1}{2} & \text{if } t \in [\frac{1}{4}, \frac{3}{4}], \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad \hat{\alpha}_2(t) = \begin{cases} \frac{2}{5} & \text{if } t \in [\frac{3}{4}, 1], \\ 0 & \text{otherwise.} \end{cases}$$

Let $\bar{x}$ a strong local minimizer, $\bar{u}$ a corresponding optimal control and $\bar{\alpha}_1$ and $\bar{\alpha}_2$ be the reaction terms associated with the normal cone satisfying

$$\dot{x}(t) = \bar{u}(t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \bar{\alpha}_1(t) \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \bar{\alpha}_2(t) \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \text{for a.e. } t \in [0,1].$$

By Theorem 2.12 there exist $\sigma, \xi \geq 0$, $p^0 \in \mathbb{R}^2$, some positive definite Borel measures μ_1, μ_2 satisfying the following conditions:

1. $\sigma + \xi\|\bar{\alpha}_1\|_{L^1} + \xi\|\bar{\alpha}_2\|_{L^1} + \mu_1([0,1]) + \mu_2([0,1]) = 1$;
2. $p^0 + (\mu_1([0,1]) - \xi\|\bar{\alpha}_1\|_{L^1})\begin{pmatrix}1\\1\end{pmatrix} + (\mu_2([0,1]) - \xi\|\bar{\alpha}_2\|_{L^1})\begin{pmatrix}2\\1\end{pmatrix} = -\sigma\begin{pmatrix}-1\\1\end{pmatrix}$.
3. $q_1(t)\bar{u}(t) = |q_1(t)|$ for a.e. $t \in [0,1]$, where
$$q(t) = p^0 + \left(\mu_1([0,t[) - \xi\int_0^t \bar{\alpha}_1(s)ds\right)\begin{pmatrix}1\\1\end{pmatrix} + \left(\mu_2([0,t[) - \xi\int_0^t \bar{\alpha}_2(s)ds\right)\begin{pmatrix}2\\1\end{pmatrix};$$
4. $\operatorname{supp}(\mu_1) \subseteq \left\{t \in [0,1] \mid \bar{x}_1(t) + \bar{x}_2(t) = \frac{1}{2}\right\}$ and
$$\operatorname{supp}(\mu_2) \subseteq \{t \in [0,1] \mid 2\bar{x}_1(t) + \bar{x}_2(t) = 1\}.$$
5. $q_1(t) + q_2(t) = 0$ for a.e. $t \in \{s \in [0,1] \mid \bar{\alpha}_1(s) > 0\}$ and $2q_1(t) + q_2(t) = 0$ for a.e. $t \in \{s \in [0,1] \mid \bar{\alpha}_2(s) > 0\}$.

Since $x_0 \in \operatorname{int}(C)$, then $q_1(t) = p_1^0$ on some interval $[0,\varepsilon[$. It follows then that $p_1^0 \geq 0$, otherwise by condition 3, we would have $\bar{u}(t) = -1$ and so $\bar{x}(t) \in \operatorname{int}(C)$ for any $t \in [0,1]$. By condition 1 it follows that $\sigma = 1$ because the state constraint is never active, however by condition 2 we would have $p_1^0 = \sigma = 1$.

As a matter of fact, taking $p_1^0 = \sigma = 1$, ξ and $\mu_1 = \mu_2 = 0$ one gets a suitable multiplier that satisfies all the conditions stated above, and with it one recovers the global minimum of the problem.

It is not difficult to see that if the measures satisfy the conditions
$$\mu_i([0,t[) = \xi\int_0^t \bar{\alpha}_i(s)ds \quad \text{and} \quad \mu_i([0,1]) = \xi\|\bar{\alpha}_i\|_{L^1}, \qquad i \in \{1,2\},$$
then if $\sigma > 0$, one gets a family of multipliers that allows us to recover the global minimum of the problem from Theorem 2.12; in this case one needs $p_1^0 = \sigma$. Notice that if $\sigma = 0$ and $p_1^0 = 0$, we get also get a multiplier that satisfies all the conditions stated above, however in this case the optimality conditions become degenerate.

Acknowledgement. The authors would like to thank Prof. Richard Vinter and the anonymous referees for providing very helpful comments on the first version of the paper.

REFERENCES

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

[2] B. Acary, O. Bonnefon, and B. Brogliato, Nonsmooth Modeling and Simulation for Switched Circuits, Springer, Berlin, 2011

[3] S. Adly, F. Nacry, L. Thibault, *Preservation of prox-Regularity of sets with applications to constrained optimization*, SIAM Journal on Control and Optimization 26(1): 2016, 448–473.

[4] C. Arroud, G. Colombo, *Necessary conditions for a nonclassical control problem with state constraints*, IFAC-PapersOnLine, Volume 50, Issue 1, July 2017, Pages 506-511

[5] CE Arroud, G Colombo, *A maximum principle for the controlled sweeping process*, Set-Valued and Variational Analysis 26 (3), 607-629, 2018.

[6] A.V. Arutyunov, D.Y. Karamzin and F.L. Pereira *The Maximum Principle for Optimal Control Problems with State Constraints by R.V. Gamkrelidze: Revisited*, Journal of Optimization Theory and Applications, vol. 149, (2011), pp 474-493.

[7] M. Brokate and P. Krejci *Optimal control of ODE systems involving a rate independent variational inequality*, Discrete Contin. Dyn. Syst. Ser. B, 18 (2013), pp. 331-348.

[8] B. Brogliato, C. Prieur, and A. Tanwani, *Observer design for frictionless and unilaterally constrained mechanical systems: A passivity-based approach*, IEEE Trans. Automat. Control, Issue 9, Sept.-2016, Page(s) 2386 - 2401 doi:10.1109/TAC.2015.2492098.

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

[10] T. H. Cao, G. Colombo, B. S. Mordukhovich and Tr. Dao Nguyen, *Optimization of fully controlled sweeping processes*, Journal of Differential Equations, vol. 295, (2021),138-186.

[11] T.H. Cao, B. S. Mordukhovich, *Optimal control of a nonconvex perturbed sweeping process*, Journal of Differential Equations, vol. 266, (2019), pp. 1003-1050.

[12] G. Colombo, P. Gidoni, *On the optimal control of rate-independent soft crawlers.* J. Math. Pures Appl., 46 (2021), 127-157.

[13] G. Colombo, V. Goncharov, *The sweeping processes without convexity*, Set-Valued Anal. 7 (1999), 357-374.

[14] G. Colombo, R. Henrion, Nguyen D. Hoang, B. S. Mordukhovich, *Optimal control of the sweeping process over polyhedral controlled sets*, J. Differential Equations 260 (2016), 3397-3447.

[15] G. Colombo, B. S. Mordukhovich, Tr. Dao Nguyen, *Optimization of a perturbed sweeping process by constrained discontinuous controls*, SIAM Journal on Control and Optimization 58(4): January 2020, 2678-2709.

[16] G. Colombo and M. Palladino, *The minimum time function for the controlled Moreau's sweeping process*, SIAM Journal on Control and Optimization, vol. 54, no. 4, pp. 2036-2062, 2016.

[17] B. Cornet, *Existence of slow solutions for a class of differential inclusions*, J. Math. Anal. Appl. 96(1), 130-147 (1983)

[18] M.d.R. de Pinho, M.M.A. Ferreira, G.V. Smirnov, *Optimal control involving sweeping processes,* Set-Valued Var. Anal. 27 (2019), 523-548.

[19] R. M. Dudley. *Real Analysis and Probability.* Cambridge university press, 2002.

[20] J.F. Edmond and L. Thibault, *Relaxation of an optimal control problem involving a perturbed sweeping process*, Mathematical Programming - Series B, 104: 347–373, 2005.

[21] J.F. Edmond, L. Thibault, *BV solutions of nonconvex sweeping process differential inclusion with perturbation*, J. Diff. Equat. 226, 135-179 (2006)

[22] I. Ekeland, R. Témam. *Convex Analysis and Variational Problems.* SIAM, 1999.

[23] L. Evans, R. Gariepy. *Measure Theory and Fine Properties of Functions.* CRC Press, 2015.

[24] M.M.A. Ferreira, R.B. Vinter, *When is the maximum principle for state constrained problems nondegenerate?*, J. Math. Anal. Appl., Vol. 187, Issue 2, (1994), pp. 438-467.

[25] F. Fontes, S. Lopes. Normal forms of necessary conditions for dynamic optimization problems with pathwise inequality constraints. *Journal of Mathematical Analysis and Applications*, 399: 27–37, 2013.

[26] R.V. Gamkrelidze, *Optimal control processes under bounded phase coordinates*, Izv. Akad. Nauk SSSR, Ser. Mat., 24(3), 315-356, 1960.

[27] P. Krejci, *Evolution variational inequalities and multidimensional hysteresis operators*, in Nonlinear Differential Equations (Chvalatice, 1998), Res. Notes Math. 404, Chapman & Hall/CRC, London, 1999, pp. 47-110.

[28] P.-L. Lions and A.-S. Sznitman, *Stochastic differential equations with reflecting boundary conditions*, Comm. Pure Appl. Math., 37 (1984), pp. 511-537.

[29] B. Maury, A. Roudneff-Chupin, F. Santambrogio, and J. Venel, *Handling congestion in crowd motion modeling*, Netw. Heterog. Media, 6 (2011), pp. 485-519.

[30] B. S Mordukhovich, *Discrete Approximations and Refined Euler-Lagrange Conditions for Nonconvex Differential Inclusions*, SIAM Journal on Control and Optimization, Vol. 33, n.3 (1995), pp. 882-915.

[31] B. S. Mordukhovich, *Variational analysis and generalized differentiation I: Basic theory.* Vol. 330. Springer Science & Business Media, 2006.

[32] J.J. Moreau, *Evolution problem associated with a moving convex set in a Hilbert space*, Journal of Differential Equations, Volume 26, Issue 3, December 1977, Pages 347-374

[33] R. T. Rockafellar, R. J-B. Wets. Variational Analysis. Springer-Verlag Berlin Heidelberg, 1998.

[34] O.-S. Serea, *On reflecting boundary problem for optimal control*, SIAM Journal on Control and Optimization, vol. 42, no. 2, pp. 559-575, 2003.

[35] F. Tedone, M. Palladino, *Hamilton-Jacobi-Bellman Equations for Control Systems with friction*, IEEE Trans. Automat. Control, 25 November 2020, doi:10.1109/TAC.2020.3040726.

[36] L. Thibault. Sweeping process with regular and nonregular sets. *J. Differential Equations*, 193 (2003) 1–26.

[37] R. Vinter. *Optimal Control.* Birkhäuser Basel, 2010.

[38] V. Zeidan, C. Nour, H. Saoud, *A nonsmooth maximum principle for a controlled nonconvex sweeping process*, Journal of Differential Equations, Volume 269, Issue 11, 15 November 2020, Pages 9531-9582.