

A MAXIMUM PRINCIPLE FOR OPTIMAL CONTROL PROBLEMS WITH STATE AND MIXED CONSTRAINTS *

MD. HAIDER ALI BISWAS¹ AND MARIA DO ROSARIO DE PINHO²

Abstract. Here we derive a variant of the nonsmooth maximum principle for optimal control problems with both pure state and mixed state and control constraints. Our necessary conditions include a Weierstrass condition together with an Euler adjoint inclusion involving the joint subdifferentials with respect to both state and control, generalizing previous results in [16]. A notable feature is that our main results are derived combining old techniques with recent results. We use a well known penalization technique for state constrained problem together with an appeal to a recent nonsmooth maximum principle for problems with mixed constraints.

1991 Mathematics Subject Classification. 49K15, 34A60.

...

INTRODUCTION

The birth of optimal control theory, commonly associated with the publication of the seminal book [28], cannot be separated from the statement and proof of the Pontryagin maximum principle (PMP), first proved by Pontryagin and collaborators in 1956 (we refer the reader to the survey [27] for an interesting historic account of the pioneering results).

Generalizations of the Pontryagin maximum principle for problems with nonsmooth data appeared in the 1970's. In this respect the work of F. Clarke was of special relevance (see, for example, [8]). Initial versions of the nonsmooth maximum principle (NMP) were extended and refined by a number of authors (see, for example, [9, 10, 26, 31] and references therein).

Necessary optimality conditions similar to the NMP were proposed in [14]. Such conditions featured a coupled adjoint inclusion involving subdifferentials with respect to the state and control jointly and, when specialized to normal linear convex problems, they are also sufficient conditions, a feature that the PMP had whereas earlier versions of the NMP did not have. Extensions of the optimality conditions developed in [14] to cover state constrained problems were first derived in [15] for problems with convex velocity sets and later extended to the nonconvex case in [16]. However, the maximization condition or Weierstrass condition was not presented in

Keywords and phrases: optimal control; state and mixed constraints; maximum principle

* This work has been supported by the European Union Seventh Framework Programme [FP7-PEOPLE-2010-ITN] under grant agreement n. 64735-SADCO. The financial support of FCT Projects PTDC/EEA-CRO/116014/2009, PTDC/EEA-ELC/122203/2010 and PTDC/EEI-AUT/1450/2012[FCOMP-01-0124-FEDER-028894] is also gratefully acknowledged. The first author was supported while at University of Porto by PhD grant reference SFRH/BD/63707/2009, FCT, Portugal.

¹ Mathematics Discipline, Science Engineering and Technology School, Khulna University, Khulna-9208, Bangladesh

² Universidade do Porto, Faculdade de Engenharia, DEEC, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

such conditions, a feature remedied in [11] (see also Theorem 3.1 in [12]). For results similar to those in [14], see also [1], where a nonsmooth maximum principle is derived using “compatible” feedback controls.

The literature on state constrained optimal control problems is quite rich. NMP for these problems have been the focus of extended research since the publication of the seminal paper [30]. Recently, special attention has also been paid to phenomena associated with such problems like nondegeneracy, normality and regularity of minimizers; see, for example [2], [4], [21], [22], [23], to name but a few. There is also a vast literature on smooth maximum principles for mixed constrained problems; see, for example [2], [3], [18], [19], [20], and references within. Probably the most successful attempt to cover nonsmooth mixed constrained problems appeared recently in [12] (see also the accompanying paper [13]). However, the literature on nonsmooth problems with both mixed and pure state constraints is sparse.

In this paper we concentrate on problems with state and mixed constraints. We apply “old” techniques developed in [30] and a recent result obtained in [12] to derive nonsmooth necessary conditions for our problem of interest. In so doing we obtain a new NMP for state and mixed constrained problems inheriting from [12] the special feature of being a sufficient condition for normal, linear convex problems. Our approach can be viewed as an extension of [15] and [16]. Preliminary results, similar to the present work but in the absence of mixed constraints, has been announced in [6].

Bringing together “old” techniques and new results comes with a price: our nonsmooth maximum principle, when applied to mixed constrained problems, does not retain all the generality of [12] since it holds under stronger conditions. However, it still covers a large class of problems appearing in applications. An alternative approach, and possibly covering more general cases, could follow the lines of [5].

This paper is organized in the following way. In the last part of this introduction we give a brief summary of the notation used in this paper. In the next section we describe the problem of interest and present our assumptions. For the sake of completeness we present a simplified version of Theorem 7.1 in [12] in section 2. Section 3 contains the statement and discussion of our main results. The proofs of our main result, Theorem 2 below, appear in section 5.

Notation: If $g \in \mathbb{R}^m$, the inequality $g \leq 0$ is interpreted component-wise. We write $g \in \mathbb{R}_+^n$ (or $g \in \mathbb{R}^n$), if $g \geq 0$ (or $g \leq 0$). We will denote by B the closed unit ball centered at the origin regardless of the dimension of the underlying space. Also $|\cdot|$ is the Euclidean norm or the induced matrix norm on $\mathbb{R}^{p \times q}$.

Take any $A \subset \mathbb{R}^n$. Then Ψ_A is the indicator function of A and the *Euclidean distance function* with respect to A is defined as

$$d_A : \mathbb{R}^k \rightarrow \mathbb{R}, \quad y \rightarrow d_A(y) = \inf \{|y - x| : x \in A\}.$$

If $\Omega \subset \mathbb{R}^p$ and $F : \Omega \rightarrow \mathbb{R}^q$ is a multifunction, then the graph of F is defined as

$$\text{Gr } F := \{(x, y) \in \Omega \times \mathbb{R}^q : y \in F(x)\}.$$

We say that a set $S \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ is $\mathcal{L} \times \mathcal{B}$ -measurable when we refer to measurability relative to the σ -field generated by the products of Lebesgue measurable subsets in \mathbb{R} and Borel measurable subsets in $\mathbb{R}^n \times \mathbb{R}^m$.

Consider now a function $h : [t_0, t_1] \rightarrow \mathbb{R}^p$. We say that $h \in W^{1,1}([t_0, t_1]; \mathbb{R}^p)$ if and only if h is absolutely continuous; that $h \in L^1([t_0, t_1]; \mathbb{R}^p)$ iff h is integrable; and that $h \in L^\infty([t_0, t_1]; \mathbb{R}^p)$ iff h is essentially bounded. The norm of $L^1([t_0, t_1]; \mathbb{R}^p)$ is denoted by $\|\cdot\|_1$ and the norm of $L^\infty([t_0, t_1]; \mathbb{R}^p)$ is $\|\cdot\|_\infty$. Let $C^*([t_0, t_1]; \mathbb{R})$ be the dual space of the continuous functions defined from $[t_0, t_1]$ to \mathbb{R} , denoted by $C([t_0, t_1]; \mathbb{R})$, with supremum norm. The norm of $C^*([t_0, t_1]; \mathbb{R})$ is denoted by $\|\mu\|_{TV}$. The set of elements in $C^*([t_0, t_1]; \mathbb{R})$ which takes nonnegative values on nonnegative valued functions in $C([t_0, t_1]; \mathbb{R})$ is $C^\oplus([t_0, t_1]; \mathbb{R})$. For $\mu \in C^\oplus([t_0, t_1]; \mathbb{R})$, we have $\|\mu\|_{TV} = \int_{[t_0, t_1]} \mu(dt)$.

As it is clear from the introduction we shall make use of standard concepts from nonsmooth analysis. The basic concepts of nonsmooth analysis are well known so we refrain from stating them here. Instead we refer the reader to [8], [9], [26], [29] and [31], for example. Concerning nonsmooth analysis we use the following notation. For a closed set $A \subset \mathbb{R}^n$ and $x^* \in A$, the *limiting normal cone to A at x^** (also known as *Mordukhovich normal cone*) is denoted by $N_A^L(x^*)$ while the *Clarke normal cone* is $N_A^C(x^*)$.

If $f : \mathbb{R}^k \rightarrow \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function and $x^* \in \mathbb{R}^k$ a point such that $f(x^*) < +\infty$, the *limiting subdifferential* of f at x^* is denoted by $\partial^L f(x^*)$. Recall that when the function f is Lipschitz continuous near x , the convex hull of the limiting subdifferential, $\text{co } \partial^L f(x)$, coincides with the (*Clarke*) *subdifferential*, denoted here by $\partial^C f(x)$. If, however, $f : \mathbb{R}^k \rightarrow \mathbb{R}^p$, then $\partial^C f(x)$ denotes the generalized Jacobian of f (for the definition see [8]).

1. PROBLEM AND ASSUMPTIONS

Consider the problem (P) of minimizing the cost function

$$l(x(t_0), x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), u(t)) dt$$

subject to the differential equation

$$\dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e.},$$

the state constraints

$$h(t, x(t)) \leq 0 \quad \text{for all } t \in [t_0, t_1],$$

the mixed state-control constraints in equality and inequality form

$$b(t, x(t), u(t)) = 0, \quad g(t, x(t), u(t)) \leq 0 \quad \text{a.e.},$$

the control constraints

$$u(t) \in U \quad \text{a.e.},$$

and the boundary conditions

$$(x(t_0), x(t_1)) \in E.$$

Here the interval $[t_0, t_1]$ is fixed, $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^k$. The function f describing the dynamics is $f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^n$ and the functions describing the mixed constraints are $b : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^{m_b}$ and $g : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^{m_g}$. Moreover h and L are scalar functions $h : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}$, $L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}$, U is a compact set and $E \subset \mathbb{R}^n \times \mathbb{R}^n$.

Problem (P) reduces to a *standard optimal control problem* whenever the state and mixed constraints are absent and we will denote such problem by (S) . If, however, the state constraint is absent (so only mixed constraints are imposed), then (P) will be denoted by (P_M) .

For (P) (or (S) or (P_M)) a pair (x, u) comprising an absolutely continuous function x (state or trajectory) and a measurable function u (control), is called an *admissible process* if it satisfies all the constraints. In this paper, the pair (x^*, u^*) always denotes the solution of the optimal control problem under consideration.

Definition 1. We call an admissible process (x^*, u^*) a *strong local minimum* of (P) if (x^*, u^*) minimizes the cost over all admissible processes (x, u) such that

$$|x(t) - x^*(t)| \leq \varepsilon \quad \text{for all } t \in [t_0, t_1], \quad (1)$$

for some $\varepsilon > 0$. The process (x^*, u^*) is called a *local $W^{1,1}$ -minimum* if (x^*, u^*) minimizes the cost over all admissible processes (x, u) satisfying (1) and

$$\int_{t_0}^{t_1} |\dot{x}(t) - \dot{x}^*(t)| dt \leq \varepsilon.$$

To simplify the notation we set

$$X_\varepsilon(t) = \{x \in \mathbb{R}^n : |x - x^*(t)| \leq \varepsilon\}.$$

The following basic assumptions are in force throughout:

- $(t, (x, u)) \rightarrow (L(t, (x, u)), f(t, (x, u)), b(t, (x, u)), g(t, (x, u)))$ is $\mathcal{L} \times \mathcal{B}$ -measurable,
- the set E is closed,
- l is locally Lipschitz.

Before stating our additional assumptions we write

$$S := \{(t, x, u) \in [t_0, t_1] \times \mathbb{R}^n \times U : b(t, x, u) = 0, g(t, x, u) \leq 0\}, \quad (2)$$

$$S(t) := \{(x, u) \in \mathbb{R}^n \times U : (t, x, u) \in S\}, \quad (3)$$

$$S(t, x) := \{u \in U : (t, x, u) \in S\}, \quad (4)$$

$$S_\varepsilon^*(t) := \{(x, u) \in S(t) : x \in X_\varepsilon(t)\}. \quad (5)$$

We are now in position to state our main assumptions.

[H1] For $\phi = f, g, b$ and L , there exist integrable functions k_x^ϕ and k_u^ϕ with the following property: For almost every $t \in [t_0, t_1]$, every (x_i, u_i) for which $x_i \in X_\varepsilon(t)$ ($i = 1, 2$), $u_i \in U$ we have

$$|\phi(t, x_1, u_1) - \phi(t, x_2, u_2)| \leq k_x^\phi(t)|x_1 - x_2| + k_u^\phi(t)|u_1 - u_2|.$$

[H2] The set $U \subset \mathbb{R}^k$ is compact.

[H3] There exists an integrable function M such that, for almost every t , all $(x, u) \in S_\varepsilon^*(t)$, $\gamma^b \in \mathbb{R}^{m_b}$, $\gamma^g \in \mathbb{R}_+^{m_g}$, $\langle \gamma, g(t, x, u) \rangle = 0$ and $\eta \in N_U^L(u)$:

$$(\alpha, \beta - \eta) \in \partial_{x,u}^L \{ \langle \gamma^b, b(t, x, u) \rangle + \langle \gamma^g, g(t, x, u) \rangle \} \implies |(\gamma^b, \gamma^g)| \leq M(t)|\beta|.$$

[N] For each $t \in [t_0, t_1]$ and $x \in \mathbb{R}^n$, there exists $u \in U$ such that $b(t, x, u) = 0$ and $g(t, x, u) \leq 0$.

Assumption [H2] may be seen as quite strong. Although it could be weakened, we opt to keep it since it simplifies the proofs of the forthcoming results when limits of sequence of controls are taken. On the other hand, many problems of interest in engineering and other areas of application satisfy [H2].

Assumption [H3] is a Mangasarian Fromowitz type condition. Its smooth version (when continuous differentiability of $(x, u) \rightarrow (b(t, x, u), g(t, x, u))$ is assumed) is equivalent to well known regularity assumptions on the mixed constraints under which Maximum Principles for mixed constrained problems were proved. Those include linearly independence of the gradients $\nabla_u b_i(t, x, u)$ and linear positively independence of the gradients $\nabla_u g_j(t, x, u)$ (see for example [19]). [H3] implies, in our case, the *bounded slope* condition for $S(t)$ (see [12] for such a discussion) and it plays an important role in our setting. As for [N], it guarantees that $S(t, x) \neq \emptyset$ for each $(t, x) \in [t_0, t_1] \times \mathbb{R}^n$.

For future use, observe that the assumptions imply the existence of an integral function k_f such that:

$$|f(t, x, u)| \leq k_f(t) \text{ for all } x \in X_\varepsilon(t) \text{ and all } u \in U \text{ a.e..} \quad (6)$$

Since (x^*, u^*) solves (P) , the function $t \rightarrow L(t, x^*(t), u^*(t))$ is integrable. Thus a condition analogous to (6) holds for L : there exists an integrable function k_L such that

$$|L(t, x, u)| \leq k_L(t) \text{ for all } x \in X_\varepsilon(t) \text{ and all } u \in U \text{ a.e..} \quad (7)$$

Moreover, it is a simple matter to see that the sets

$$f(t, x, U), b(t, x, U), g(t, x, U) \text{ and } L(t, x, U) \text{ are compact for all } x \in X_\varepsilon(t). \quad (8)$$

2. AUXILIARY RESULTS

In this section we consider the problem (P) above when the state constraint is absent, i.e., we consider problem (P_M) defined above. The focus is on an adaptation of Theorem 7.1 in [12] that will be essential to our analysis. We call it “an adaptation” because it is stated under stronger assumptions than those in [12].

Theorem 1. *Let (x^*, u^*) be a local $W^{1,1}$ minimum for problem (P_M) ((P) in the absence of state constraints). Assume that the basic assumptions as well as [H2] and [H3] hold and that f, g and L satisfy [H1]. Assume also that the function $M(k_x^b + k_x^g)(k_u^f + k_u^L)$ is integrable. Then, there exist $p \in W^{1,1}([t_0, t_1]; \mathbb{R}^n)$ and a scalar $\lambda_0 \geq 0$ satisfying the nontriviality condition:*

$$\|p\|_\infty + \lambda_0 > 0, \quad (9)$$

the Euler adjoint inclusion:

$$(-\dot{p}(t), 0) \in \partial_{x,u}^C \left(\langle p(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) - K(t) |p(t), \lambda_0| d_{S(t)}(x^*(t), u^*(t)) \right) \text{ a.e.}, \quad (10)$$

the global Weierstrass condition: for all $u \in S(t, x^*(t))$,

$$\langle p(t), f(t, x^*(t), u) \rangle - \lambda_0 L(t, x^*(t), u) \leq \langle p(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) \text{ a.e.}, \quad (11)$$

and the transversality condition:

$$(p(t_0), -p(t_1)) \in N_E^L(x^*(t_0), x^*(t_1)) + \lambda_0 \partial^L l(x^*(t_0), x^*(t_1)). \quad (12)$$

Above K is an integrable function defined in terms of the Lipschitz parameters and M in [H3].

Remark 1. Before proceeding it is important to note that although Theorem 7.1 in [12] holds with the Euler adjoint inclusion written as in (10), it is nevertheless stated with (10) replaced

$$(-\dot{p}(t), 0) \in \partial_{x,u}^C \left(\langle p(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) \right) - N_{S(t)}^C(x^*(t), u^*(t)). \quad (13)$$

Observe that, appealing to the calculus rules of subdifferentials (see [31]), we deduce from (10) that

$$(-\dot{p}(t), 0) \in \partial_{x,u}^C \left(\langle p(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) \right) - \partial_{x,u}^C K(t) |p(t), \lambda_0| d_{S(t)}(x^*(t), u^*(t)) \text{ a.e.} \quad (14)$$

This, together with the fact that $\partial_{x,u}^C d_{S(t)} \subset N_{S(t)}^C$, leads to (13) (see also remark on page 4522 in [12]). Here, and for reasons that will be clear later on, we use the above sharper version of the Euler adjoint inclusion (10). Finally, it is worth to mention that when $L = 0$, the term $|p(t), \lambda_0|$ in (10) is reduced to $|p(t)|$.

A remarkable feature of Theorem 7.1 in [12] is the Euler Inclusion (10). Hypothesis [H1] is essential for the establishment of (10). To the best of our knowledge such Euler- Lagrange inclusion (with the joint subdifferentials with respect to both the state and control) was first introduced for nonsmooth problems in [14]. A main

setback of the necessary conditions in [14] dwells in the fact that (10) is not coupled with (11), a situation later remedied in [11]. Observe that in the classical NMP the inclusion

$$-\dot{p}(t) \in \partial_x^C \left(\langle p(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) \right),$$

holds instead of (10).

3. NONSMOOTH MAXIMUM PRINCIPLES FOR (P)

We now turn to problem (P) in its full generality and we assume the following assumption on the function h describing the state constraint:

[H4] For all $x \in X_\varepsilon(t)$ the function $t \rightarrow h(t, x)$ is continuous. Furthermore, there exists a constant $k_h > 0$ such that the function $x \rightarrow h(t, x)$ is Lipschitz of rank k_h for all $t \in [t_0, t_1]$.

The need to impose continuity of $t \rightarrow h$ instead of merely upper semicontinuity is discussed in [15].

We also consider the subdifferential $\partial_x^> h$ defined as

$$\partial_x^> h(t, x) := \text{co} \{ \xi : \exists (t_i, x_i) \xrightarrow{h} (t, x) : h(t_i, x_i) > 0 \ \forall i, \ \partial_x h(t_i, x_i) \rightarrow \xi \}. \quad (15)$$

Theorem 2. Let (x^*, u^*) be a strong local minimum for problem (P). Assume that f , L and g satisfy [H1], h satisfies [H4] and that [H2], [H3] and [N], as well as the basic assumptions, hold. Suppose also that the functions

$$M(k_x^b + k_x^g) \quad \text{and} \quad M(k_x^b + k_x^g)(\max\{k_u^f, k_f\} + \max\{k_u^L, k_L\}),$$

where k_f and k_L are as in (6) and (7), are integrable.

Then there exist $p \in W^{1,1}([t_0, t_1]; \mathbb{R}^n)$, $\gamma \in L^1([t_0, t_1]; \mathbb{R})$, a measure $\mu \in C^\oplus([t_0, t_1]; \mathbb{R})$ and a scalar $\lambda_0 \geq 0$ satisfying

- (i) $\mu\{[t_0, t_1]\} + \|p\|_\infty + \lambda_0 > 0$;
- (ii) $(-\dot{p}(t), 0) \in \partial_{x,u}^C(\langle q(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t))) - N_{S(t)}^C(x^*(t), u^*(t)) \quad \text{a.e.};$
- (iii) $\forall u \in S(t, x^*(t)),$
 $\langle q(t), f(t, x^*(t), u) \rangle - \lambda_0 L(t, x^*(t), u) \leq \langle q(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) \quad \text{a.e.};$
- (iv) $(p(t_0), -q(t_1)) \in N_E^I(x^*(t_0), x^*(t_1)) + \lambda_0 \partial^L l(x^*(t_0), x^*(t_1));$
- (v) $\gamma(t) \in \partial^> h(t, x^*(t)) \quad \mu\text{-a.e.};$
- (vi) $\text{supp}\{\mu\} \subset \{t \in [t_0, t_1] : h(t, x^*(t)) = 0\}.$

Here

$$q(t) = \begin{cases} p(t) + \int_{[t_0, t)} \gamma(s) \mu(ds) & t \in [t_0, t_1), \\ p(t) + \int_{[t_0, t_1]} \gamma(s) \mu(ds) & t = t_1, \end{cases} \quad (16)$$

where K is a constant defined in terms of the Lipschitz constants and M in [H3].

If, furthermore, the functions $(x, u) \rightarrow b(t, x, u)$ and $(x, u) \rightarrow g(t, x, u)$ are strictly differentiable at $(x^*(t), u^*(t))$ a.e., then there exist measurable functions

$$\gamma^b : [t_0, t_1] \rightarrow \mathbb{R}^{m_b}, \quad \gamma^g : [t_0, t_1] \rightarrow \mathbb{R}_+^{m_g} \quad \text{with} \quad \langle \gamma^g(t), g(t, x^*(t), u^*(t)) \rangle = 0 \quad \text{a.e.}$$

and

$$|\gamma^b(t), \gamma^g(t)| \leq M(t) \{ |p(t)| k_u^f(t) + \lambda_0 k_u^L(t) \} \quad \text{a.e.},$$

such that the inclusion (ii) above is expressible in the explicit multiplier form

$$\begin{aligned} (p(t), \nu(t)) \in \\ \partial_{x,u}^C \{ \langle p(t), f(t, x^*(t), u^*(t)) \rangle - \lambda_0 L(t, x^*(t), u^*(t)) \} - \\ \nabla_{x,u} \{ \langle \gamma^b(t), b(t, x^*(t), u^*(t)) \rangle - \langle \gamma^g(t), g(t, x^*(t), u^*(t)) \rangle \} \quad a.e., \end{aligned}$$

where ν is a measurable function satisfying $\nu(t) \in N_U^C(u^*(t))$ a.e..

The integrability of the functions $M(k_x^b + k_x^g)$ and $M(k_x^b + k_x^g)(\max\{k_u^f, k_f\} + \max\{k_u^L, k_L\})$, is assumed in Theorem 2 in contrast with Theorem 1 where merely integrability of $M(k_x^b + k_x^g)(k_u^f + k_u^L)$ is imposed. The reason for strengthening such assumptions will be clear in the proof of Proposition 4 in section 4.

Theorem 2 can also be extended to deal with a local $W^{1,1}$ -minimum for (P) . This is accomplished as in the last steps of the proof of Lemma 9.4.1 in [31]. We omit the proof but for the sake of completeness state the result.

Theorem 3. *Let (x^*, u^*) be merely a local $W^{1,1}$ -minimum for problem (P) . Then the conclusions of Theorem 2 hold.*

Imposing Lipschitz continuity with respect to the control (which can be partially relaxed as discussed in [12]) contrasts with the usual hypothesis of measurability. This may be seen as a disadvantage. However, (10) is valid because [H1] and (10) is responsible for the very fact that this variant of the nonsmooth Maximum Principle is also sufficient for linear-convex problems. In this respect, the nonsmoothness is essential. Let us consider the following linear convex problem with mixed and state constraints:

$$(LC) \quad \left\{ \begin{array}{ll} \text{Minimize } l(x(t_0), x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), u(t)) \, dt \\ \text{subject to} \\ \quad \dot{x}(t) = A_1(t)x(t) + B_1(t)u(t) & \text{a.e.} \\ \quad 0 = A_2(t)x(t) + B_2(t)u(t) & \text{a.e.} \\ \quad 0 \geq A_3(t)x(t) + B_3(t)u(t) & \text{a.e.} \\ \quad u(t) \in U & \text{a.e.} \\ \quad D(t)x(t) \leq 0 & \text{for all } t \in [t_0, t_1] \\ \quad (x(t_0), x(t_1)) \in E, \end{array} \right.$$

Here we assume that U and E are convex sets, $(x, u) \rightarrow L(t, x, u)$ convex, $A_1 : [t_0, t_1] \rightarrow \mathbb{R}^{n \times n}$ and $B_1 : [t_0, t_1] \rightarrow \mathbb{R}^{n \times k}$ integrable, $A_2 : [t_0, t_1] \rightarrow \mathbb{R}^{m_b \times n}$, $A_3 : [t_0, t_1] \rightarrow \mathbb{R}^{m_g \times n}$, $B_2 : [t_0, t_1] \rightarrow \mathbb{R}^{m_b \times k}$ and $B_3 : [t_0, t_1] \rightarrow \mathbb{R}^{m_g \times k}$ measurable and $D : [t_0, t_1] \rightarrow \mathbb{R}^{1 \times n}$ continuous. If the conclusions of Theorem 2 hold, then subdifferentials and normal cones reduce to those of convex analysis.

Proposition 1. *Suppose that (x^*, u^*) is an admissible process for (LC) and that the conclusions of Theorem 2 hold for (x^*, u^*) with $\lambda = 1$. Then (x^*, u^*) is a local minimizer.*

Here we do not present the proof of the above Proposition since it is a simple adaptation of the proof of Proposition 4.1 in [15].

4. PROOF OF THEOREM 2

We now dedicate this section to the proof of Theorem 2. We first prove it in the case where $L \equiv 0$. Thus, the assumption that $M(k_x^b + k_x^g)(\max\{k_u^f, k_f\} + \max\{k_u^L, k_L\})$ is integrable reduces to the integrability of $M(k_x^b + k_x^g) \max\{k_u^f, k_f\}$.

Before engaging in the proof, however, we introduce a multifunction F^- related to our problem and we state some well known properties of F^- that will be of importance in the forthcoming analysis. We then prove the

first part of the Theorem assuming convexity of the velocity set. Such requirement will be removed next. In the last stage of the proof we consider a nonzero L and the case in which $(x, u) \rightarrow b(t, x, u)$ and $(x, u) \rightarrow g(t, x, u)$ are strictly differentiable. The proof is built up in several steps, summarized as lemmas or propositions.

We consider the multifunction $F^- : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, defined as

$$F^-(t, x) := \{f(t, x, u) : u \in S(t, x)\}, \quad (17)$$

Proposition 2. *Assume that our basic assumptions as well as [H1]–[H3] and [N] hold. Then*

- a) *the multifunctions $(t, x) \rightarrow S(t, x)$, $F^-(t, x)$ are non-empty and compact valued;*
- b) *the multifunction F^- is $\mathcal{L} \times \mathcal{B}$ measurable;*
- c) *there exists an integrable function c such that, for almost every $t \in [t_0, t_1]$, we have $|\gamma| \leq c(t)$ for all $x \in X_\varepsilon(t)$ and $\gamma \in F^-(t, x)$;*
- d) *there exist a scalar $\rho > 0$ and an integrable function K_F such that, for almost every $t \in [a, b]$, we have*

$$F^-(t, x) \subset F^-(t, x') + K_F(t)|x - x'|B$$

for all $x, x' \in \{x : |x - x^(t)| \leq \rho\}$.*

For a discussion and proof of this Proposition we refer the reader to [25].

Set $\bar{\varepsilon} = \min\{\varepsilon, \rho\}$, where ρ is as defined in Proposition 2. There is no loss of generality in assuming that from now on the parameter defining the sets $X_\varepsilon(t)$ and $S_\varepsilon^*(t)$ is $\bar{\varepsilon}$. However, and to simplify the notation, we do not redesignate those sets in what follows.

Now we consider is a special case of (P) in the form

$$(Q) \quad \left\{ \begin{array}{ll} \text{Minimize } l(x(t_0), x(t_1)) \\ \text{subject to} \\ \dot{x}(t) = f(t, x(t), u(t)) & \text{a.e.} \\ 0 = b(t, x(t), u(t)) & \text{a.e.} \\ 0 \geq g(t, x(t), u(t)) & \text{a.e.} \\ u(t) \in U & \text{a.e.} \\ h(t, x(t)) \leq 0 & \text{for all } t \in [t_0, t_1] \\ (x(t_0), x(t_1)) \in E_0 \times \mathbb{R}^n, \end{array} \right.$$

when the velocity set is convex and the necessary conditions are expressed in terms of a (possibly) larger subdifferential of $x \rightarrow h(t, x)$ than $\partial_x^> h$:

$$\bar{\partial}_x h(t, x) := \text{co} \{ \lim \xi_i : \xi_i \in \partial_x h(t_i, x_i), (t_i, x_i) \rightarrow (t, x) \}. \quad (18)$$

Observe that (Q) differs from (P) (recall we are assuming here that $L \equiv 0$) since $E = E_0 \times \mathbb{R}^n$. Here we assume that $E_0 \subset \mathbb{R}^n$ is closed.

Proposition 3. *We position the hypotheses of Theorem 3 and [C] below.*

[C] *For all $(t, x) \in [t_0, t_1] \times \mathbb{R}^n$, the set $F^-(t, x)$ is convex.*

Then all the conclusions (i)–(vi) of Theorem 2 hold with $\bar{\partial}_x h(t, x)$ replacing $\partial_x^> h$, (ii) and (iv) expressible in the form

$$\begin{aligned} (-\dot{p}(t), 0) &\in \partial_{x,u}^C \left(\langle q(t), f(t, x^*(t), u^*(t)) \rangle - K(t)|(q(t), \lambda_0)|d_{S(t)}(x^*(t), u^*(t)) \right) \\ (p(t_0), -q(t_1)) &\in N_{E_0}^L(x^*(t_0)) \times \{0\} + \lambda_0 \partial^L l(x^*(t_0), x^*(t_1)) \end{aligned}$$

Proof. The main stages of the proof consist on simple adaptation of previous work (see [30] and [15]). Thus, we refrain from presenting all the details. A complete report can be found in [7].

Set $h^+(t, x) := \max\{0, h(t, x)\}$ and define a sequence of problems, called (Q_i) , where the cost is

$$l(x(t_0), x(t_1)) + i \int_{t_0}^{t_1} h^+(t, x(t)) dt,$$

subject to the constraints

$$\dot{x}(t) = f(t, x(t), u(t)), \quad b(t, x(t), u(t)) = 0, \quad g(t, x(t), u(t)) \leq 0, \quad u(t) \in U, \quad \text{a.e.}$$

and $(x(t_0), x(t_1)) \in E_0 \times \mathbb{R}^n$. Assume that the following interim hypothesis holds:

$$\mathbf{[IH]} \quad \liminf_{i \rightarrow \infty} \{Q_i\} = \inf\{Q\}.$$

In the final step of the proof we show that $\mathbf{[C]}$ implies $\mathbf{[IH]}$.

We now define \mathcal{W} to be the set of pairs (u, s) , where $u : [t_0, t_1] \rightarrow \mathbb{R}^k$ is a measurable function, such that

$$\begin{aligned} \mathcal{W} := \{ & (u, s) : \exists x \in W^{1,1}([t_0, t_1]; \mathbb{R}^n) \text{ such that} \\ & \dot{x}(t) = f(t, x(t), u(t)), \quad b(t, x(t), u(t)) = 0, \quad g(t, x(t), u(t)) \leq 0, \quad u(t) \in U \text{ a.e.} \\ & x(t) \in X_\varepsilon(t) \text{ for all } t \in [t_0, t_1], \quad x(t_0) = s, \quad s \in E_0 \}. \end{aligned}$$

Equip \mathcal{W} with the metric

$$\delta((u, s), (v, s')) := \int_{t_0}^{t_1} |u(t) - v(t)| dt + |s - s'|$$

and set

$$J_i(u, s) := l(x(t_0), x(t_1)) + i \int_{t_0}^{t_1} h^+(t, x(t)) dt,$$

where x is the trajectory corresponding to (u, s) . Under our assumptions, the set (\mathcal{W}, δ) is a complete metric space and $J_i : \mathcal{W} \rightarrow \mathbb{R}$ is continuous.

Denote by (O_i) the problem

$$\text{Min } \{J_i(u, s) : (u, s) \in \mathcal{W}\}.$$

For any i , $(u^*, x^*(t_0))$ is an admissible solution of (O_i) with

$$J_i(u^*, x^*(t_0)) = l(x^*(t_0), x^*(t_1)) = \inf Q.$$

Set $\rho_i := J_i(u^*, x^*(t_0)) - \inf Q_i$. Since $J_i(u^*, x^*(t_0)) \geq \inf Q_i$ we have $\rho_i \geq 0$. By \mathbf{IH} , we have $\lim_{i \rightarrow \infty} \rho_i = 0$.

Ekeland's Theorem (see [31]) applies to (O_i) . Following the approach in [15] (see also [30]) we deduce the existence of a sequence $(u_i, s_i) \in \mathcal{W}$ such that

$$\int_{t_0}^{t_1} |u_i(t) - u^*(t)| dt + |s_i - x^*(t_0)| \leq \sqrt{\rho_i}.$$

Let x_i be the trajectory associated with u_i . It follows from the above that u_i converges strongly to u^* . Then there exists a subsequence $\{u_i\}$ (we do not relabel) converging to u^* for almost every $t \in [t_0, t_1]$. Along the corresponding subsequence, x_i converges uniformly to x^* . Discarding initial terms of this sequence, if necessary, we guarantee that $x_i(t) \in X_{\varepsilon/2}(t)$ for all $t \in [t_0, t_1]$. In control terms, we have that, for each i , (x_i, u_i) solves

the problem:

$$(\bar{P}_i) \quad \left\{ \begin{array}{ll} \text{Minimize} & l(x(t_0), x(t_1)) + i \int_{t_0}^{t_1} h^+(t, x(t)) dt + \sqrt{\rho_i} \int_{t_0}^{t_1} |u(t) - u_i(t)| dt \\ \text{subject to} & \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [t_0, t_1], \\ & b(t, x(t), u(t)) = 0 \quad \text{a.e. } t \in [t_0, t_1], \\ & g(t, x(t), u(t)) \leq 0 \quad \text{a.e. } t \in [t_0, t_1], \\ & u(t) \in U \quad \text{a.e. } t \in [t_0, t_1], \\ & x(t_0) \in E_0. \end{array} \right.$$

Here the integrand cost is $L_i(t, x, u) = ih^+(t, x) + \sqrt{\rho_i}|u - u_i(t)|$. Under our conditions, this function is Lipschitz with Lipschitz rank $k_u^{L_i} = \sqrt{\rho_i}$. The integrability of $M(k_x^b + k_x^g)$ and $M(k_x^b + k_x^g) \max\{k_u^f, k_f\}$ implies that $M(k_x^b + k_x^g)k_u^f$ and $M(k_x^b + k_x^g)(k_u^f + k_u^{L_i})$ are integrable functions. We are in position to apply Theorem 1 to (\bar{P}_i) . This yields the existence of an absolutely continuous function p_i and a scalar $\lambda_i \geq 0$ such that

$$(p_i(t), \lambda_i) \neq 0 \text{ for all } t, \quad (19)$$

$$(-\dot{p}_i(t), 0) \in \partial_{x,u}^C \hat{H}_\lambda(t, x_i(t), p_i(t), u_i(t)), \quad (20)$$

$$(x_i(t), u) \in S(t) \implies \langle p_i(t), f(t, x_i(t), u) \rangle - \sqrt{\rho_i} \lambda_i |u - u_i(t)| \leq \langle p_i(t), f(t, x_i(t), u_i(t)) \rangle \text{ a.e.}, \quad (21)$$

$$(p_i(t_0), -p_i(t_1)) \in N_{E_0}^L(x_i(t_0)) \times \{0\} + \lambda_i \partial^L l(x_i(t_0), x_i(t_1)), \quad (22)$$

where

$$\hat{H}_\lambda(t, x, p, u) = \langle p, f(t, x, u) \rangle - i\lambda h^+(t, x) - \lambda \sqrt{\rho_i} |u - u_i| - K(t) |(p, \lambda)| d_{S(t)}(x, u).$$

Appealing to the sum rule of subdifferentials (see [8] or [31]), we have

$$\begin{aligned} & \partial_{x,u}^C \hat{H}_\lambda(t, x, p, u) \subset \\ & \partial_{x,u}^C H_\lambda(t, x, p, u) - i\lambda \partial_{x,u}^C h^+(t, x) - \lambda \sqrt{\rho_i} \partial_{x,u}^C |u - u_i| - K(t) |(p, \lambda)| \partial_{x,u}^C d_{S(t)}(x, u), \end{aligned}$$

By the max rule (Proposition 2.3.12 in [8]) we have

$$\partial_{x,u}^C h^+(t, s) \subset \begin{cases} \{\sigma \Gamma : \sigma \in [0, 1], \Gamma = (\gamma, 0) \in \partial_{x,u}^C h(t, s)\} & \text{if } h(t, s) = 0, \\ \{0\} & \text{otherwise.} \end{cases}$$

We also have

$$\partial_{x,u}^C |u - u_i| \subset \{(0, e) : e \in B\}$$

and, by Theorem 2.6.2 in [8],

$$\partial_{x,u}^C \langle p, f(t, x, u) \rangle \subset p \partial_{x,u}^C f(t, x, u)$$

where $\partial_{x,u}^C f(t, x, u)$ is the generalized gradient of f .

Set $\beta = (\Phi, \Gamma, \sigma, e, \Delta)$ and define the multifunctions

$$\Sigma_i(t) = \begin{cases} [0, 1] & \text{if } h(t, x_i(t)) = 0, \\ \{0\} & \text{otherwise,} \end{cases}$$

$$G_i(t, \beta) = p_i(t) \Phi - i\lambda_i \sigma \Gamma - i\sqrt{\rho_i} e - K(t) |(p_i(t), \lambda_i)| \Delta$$

and

$$\Omega_i(t) = \partial_{x,u}^C f(t, x_i(t), u_i(t)) \times \partial_{x,u}^C h(t, x_i(t)) \times \Sigma_i(t) \times \partial_{x,u}^C d_{S(t)}(x_i(t), u_i(t)).$$

We can now rewrite (20) as $(-\dot{p}_i(t), 0) \in \{G_i(t, \beta) : \beta \in \Omega_i(t)\}$. For each i , G_i is measurable in t and continuous in β . Appealing to properties of subdifferentials and generalized Jacobians, it is straightforward to verify that Ω_i is compact valued and measurable. It follows from Theorem 2.3.11 in [31] that there exists a measurable function $\beta_i = (\Phi_i, \Gamma_i, \sigma_i, e_i, \delta_i)$ such that $\beta_i(t) \in \Omega_i(t)$ a.e. and

$$(-\dot{p}_i(t), 0) \in G_i(t, \beta_i(t)) \text{ a.e..} \quad (23)$$

Observe that, for almost every $t \in [t_0, t_1]$, we have

$$\begin{aligned} \Phi_i(t) &= (\phi_i^1(t), \phi_i^2(t) \in \partial_{x,u}^C f(t, x_i(t), u_i(t)), \quad \Gamma_i(t) = (\gamma_i(t), 0) \in \partial_{x,u}^C h(t, x_i(t)), \quad \sigma_i(t) \in \Sigma_i(t), \\ e_i(t) &\in B, \quad \Delta_i(t) = (\delta_i^1(t), \delta_i^2(t)) \in \partial_{x,u}^C d_{S(t)}(x_i(t), u_i(t)). \end{aligned}$$

We now introduce the measure $\mu_i \in C^\oplus([t_0, t_1]; \mathbb{R})$ such that, for any Borel set $\mathcal{B} \subset [t_0, t_1]$,

$$\int_{\mathcal{B}} d\mu_i(t) = \int_{\mathcal{B}} i\lambda_i \sigma_i(t) dt.$$

Then

$$\int_{[t_0, t]} i\lambda_i \gamma_i(t) \sigma_i(t) dt = \int_{[t_0, t]} \gamma_i(t) d\mu_i(t).$$

Define the measure $\pi_i \in C^*([t_0, t_1]; \mathbb{R}^n)$ by $d\pi_i(t) = \dot{p}_i(t) dt$. Let $b_i \in \mathbb{R}^n$. Now (19)–(22) can be expressed as

$$\begin{aligned} p_i(t) &= b_i + \int_{[t_0, t]} d\pi_i(t) \quad \text{for all } t \in [t_0, t_1], \\ \int_{[t_0, t]} d\pi_i(t) &= \int_{[t_0, t]} \left(p_i(t) \phi_i^1(t) - K(t) |p_i(t), \lambda_i| \delta_i^1(t) \right) dt - \int_{[t_0, t]} \gamma_i(t) d\mu_i(t) \quad \text{for all } t \in [t_0, t_1], \\ 0 &= \int_{[t_0, t]} \left(p_i(t) \phi_i^2(t) - K(t) |p_i(t), \lambda_i| \delta_i^2(t) \right) dt - \lambda_i \sqrt{\rho_i} e_i(t) \quad \text{for all } t \in [t_0, t_1], \\ (b_i, -b_i - \int_{[t_0, t_1]} d\pi_i(t)) &\in N_{E_0}(x_i(t_0)) \times \{0\} + \lambda_i \partial^L l(x_i(t_0), x_i(t_1)), \\ (x_i(t), u) \in S(t) &\implies \langle p_i(t), f(t, x_i(t), u) \rangle - \sqrt{\rho_i} \lambda_i |u - u_i(t)| \leq \langle p_i(t), f(t, x_i(t), u_i(t)) \rangle \text{ a.e.,} \\ &|b_i| + |\lambda_i| + |\mu_i| = 1. \end{aligned}$$

We normalise the multipliers to get the last equality.

Up to now we have worked with a fix i . Next, we take limits obtaining the conclusions of Proposition 3. This is done following the steps of the proof of Theorem 3.1 in [15]. We omit the details.

Finally, we claim that [C] implies [IH]. The proof of our claim, presented next, follows closely the approach used in the end step of the proof of Theorem 3.1 in [15].

For each i , choose a feasible process (x_i, u_i) for problem (Q_i) such that $x_i(t) \in X_\varepsilon(t)$ and

$$l(x_i(t_0), x_i(t_1)) + i \int_{t_0}^{t_1} h^+(t, x_i(t)) dt \leq \inf(Q_i) + \frac{1}{i}. \quad (24)$$

Recall the definition of F^- in (17). Taking into account that F^- is convex valued by [C] and the properties of F^- stated in Lemma 2, we invoke Theorem 2.5.3 in [31] to deduce that $x_i \rightarrow x$ uniformly for some $x \in W^{1,1}$ such that $x(t) \in X_\varepsilon(t)$ and

$$\begin{cases} \dot{x}(t) &\in F^-(t, x(t)), \\ x(t_0) &\in E_0. \end{cases}$$

By Theorem 2.3.11 in [31] there exists a measurable function u such that

$$\begin{cases} \dot{x}(t) &= f(t, x(t), u(t)), \\ 0 &= b(t, x(t), u(t)), \\ 0 &\geq g(t, x(t), u(t)), \\ u &\in U, \\ x(t_0) &\in E_0. \end{cases}$$

Since $l(x_i(t_0), x_i(t_1)) + i \int_{t_0}^{t_1} h^+(t, x_i(t)) dt$ is bounded by (24) and $t \rightarrow h^+(t, x(t))$ is continuous, we get

$$\lim_{i \rightarrow \infty} \int_{t_0}^{t_1} h^+(t, x_i(t)) dt = \int_{t_0}^{t_1} h^+(t, x(t)) dt.$$

Continuity of $t \rightarrow h^+(t, x(t))$ allow us to conclude that $h(t, x(t)) \leq 0$ for all $t \in [t_0, t_1]$ (see the proof of Theorem 3.1 in [15] for details). It follows that (x, u) is admissible for (Q) . Since (x^*, u^*) is optimal for (Q) , we have

$$l(x^*(t_0), x^*(t_1)) \leq l(x(t_0), x(t_1)). \quad (25)$$

On the other hand, by (24) we have

$$l(x(t_0), x(t_1)) = \lim_{i \rightarrow \infty} l(x_i(t_0), x_i(t_1)) + i \int_{t_0}^{t_1} h^+(t, x_i(t)) dt \leq \liminf_{i \rightarrow \infty} (Q_i).$$

This, together with (25), yields

$$l(x^*(t_0), x^*(t_1)) = \inf(Q) \leq l(x(t_0), x(t_1)) \leq \inf(Q_i),$$

that is, $\inf(Q) \leq \liminf_{i \rightarrow \infty} (Q_i)$. Since we also know that (x^*, u^*) is admissible for each (Q_i) , we conclude that $\inf(Q) \geq \liminf_{i \rightarrow \infty} (Q_i)$. Thus $\inf(Q) = \liminf_{i \rightarrow \infty} (Q_i)$, completing our proof. \square

Our next task is to cover problems with non-convex velocity sets. We start by validating the conclusions of Theorem 2 for a mini-max optimal control problem with $E = E_0 \times \mathbb{R}^n$ (here we follow an approach developed in Chapter 9 of [31]).

$$(\tilde{R}) \quad \begin{cases} \text{Minimize } \tilde{l}(x(t_0), x(t_1), \max_{t \in [t_0, t_1]} h(t, x(t))), \\ \text{over } x \in W^{1,1} \text{ and measurable functions } u \text{ satisfying} \\ \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e.}, \\ b(t, x(t), u(t)) = 0 \quad \text{a.e.}, \\ g(t, x(t), u(t)) \geq 0 \quad \text{a.e.}, \\ u(t) \in U \quad \text{a.e.}, \\ x(t_0) \in E_0. \end{cases}$$

where $\tilde{l} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ is a given function and $E_0 \subset \mathbb{R}^n$ is a given closed set.

Proposition 4. *Suppose that (x^*, u^*) is a strong local minimum for problem (\tilde{R}) . Assume that the data of problem (\tilde{R}) satisfy the assumptions of Theorem 2 and [H5] below.*

[H5] *The function \tilde{l} is Lipschitz continuous on a neighbourhood of $(x^*(t_0), x^*(t_1), \max_{t \in [t_0, t_1]} h(t, x^*(t)))$ and if $z' \geq z$, then we have $\tilde{l}(y, x, z') \geq \tilde{l}(y, x, z)$, for all $(y, x) \in \mathbb{R}^n \times \mathbb{R}^n$.*

Then there exist an absolutely continuous function $p : [t_0, t_1] \rightarrow \mathbb{R}^n$, an integrable function $\gamma : [t_0, t_1] \rightarrow \mathbb{R}^n$ and a measure $\mu \in C^\oplus([t_0, t_1]; \mathbb{R})$, such that

$$(-\dot{p}(t), 0) \in \partial_{x,u}^C \langle q(t), f(t, x^*(t), u^*(t)) \rangle - K(t)|(q(t), 1)| \partial_{x,u}^C d_{S(t)}(x^*(t), u^*(t)) \quad a.e. \quad (26)$$

$$(p(t_0), -q(t_1), \int_{[t_0, t_1]} \mu(ds)) \in N_{E_0}^L(x^*(t_0)) \times \{0, 0\} + \partial^L \tilde{l}(x^*(t_0), x^*(t_1), \max_{t \in [t_0, t_1]} h(t, x^*(t))), \quad (27)$$

$$\gamma(t) \in \bar{\partial} h(t, x^*(t)) \quad \mu\text{-a.e.}, \quad (28)$$

$$\forall u \in S(t, x^*(t)), \quad \langle q(t), f(t, x^*(t), u) \rangle \leq \langle q(t), f(t, x^*(t), u^*(t)) \rangle \quad a.e., \quad (29)$$

$$\text{supp}\{\mu\} \subset \left\{ t \in [t_0, t_1] : h(t, x^*(t)) = \max_{s \in [t_0, t_1]} h(s, x^*(s)) \right\}, \quad (30)$$

where q is defined as in (16).

Remark 2. Before presenting the proof, a remark on (27) is called for. This inclusion is not as sharp as the Euler adjoint inclusion in Theorem 1 (in this regard see Remark 1). However, (27) is enough for our purpose.

Proof. The proof is an adaptation of an approach in [31] (pages 352-353) and so we keep the details to the essential.

Adjusting ε , if necessary, it is a simple task to see that x^* minimizes

$$\tilde{l}(x(t_0), x(t_1), \max_{t \in [t_0, t_1]} h(t, x(t)))$$

over all the trajectories of the differential inclusion $\dot{x}(t) \in F^-(t, x(t))$ such that $x(t_0) \in E_0$ and $x \in X_\varepsilon(t)$. Now, Theorem 2.7.2 in [31] guarantees that, for any trajectory x satisfying

$$\dot{x}(t) \in \text{co } F^-(t, x(t)) \quad \text{a.e.}, \quad x(t_0) \in E_0, \quad x(t) \in X_\varepsilon(t), \quad (31)$$

there exists a F^- trajectory w such that $w(t_0) = x(t_0)$ and $w(t) \in X_\varepsilon(t)$. In view of Proposition 2 and the continuity of \tilde{l} , Theorem 2.7.3 in [31] asserts that x^* minimizes also $\tilde{l}(x(t_0), x(t_1), \max_{t \in [t_0, t_1]} h(t, x(t)))$ over all $x \in W^{1,1}$ satisfying (31).

Let $x \in W^{1,1}$ be such that $\dot{x}(t) \in \text{co } F^-(t, x(t))$ a.e.. Then, by Proposition 2 and Carathéodory's Theorem, there exist $(u_1(t), u_2(t), \dots, u_{n+1}(t))$ and $(\lambda_1(t), \dots, \lambda_{n+1}(t)) \in \Lambda$, where $u_i(t) \in U$ for $i = 1, \dots, n+1$, such that

$$(x, (u_1(t), u_2(t), \dots, u_{n+1}(t)), (\lambda_1(t), \dots, \lambda_{n+1}(t)))$$

is a solution to the system

$$\begin{cases} \dot{x}(t) = \sum_{i=1}^{n+1} \lambda_i(t) f(t, x(t), u_i(t)), & \text{a.e.}, \\ b(t, x(t), u_i(t)) = 0, & i = 1, \dots, n+1, \quad \text{a.e.}, \\ g(t, x(t), u_i(t)) \leq 0, & i = 1, \dots, n+1, \quad \text{a.e.}, \\ (\lambda_1(t), \dots, \lambda_{n+1}(t)) \in \Lambda, & \text{a.e.}, \\ u_i(t) \in U, & \text{a.e. for } i = 1, \dots, n+1, \end{cases} \quad (32)$$

where

$$\Lambda := \{\lambda' \in \mathbb{R}^{n+1} : \lambda'_i \geq 0 \text{ and } \sum_{i=1}^{n+1} \lambda'_i = 1\}.$$

For a discussion on relaxation of F^- see [25] or [24].

Set $v(t) = (u_1(t), u_2(t), \dots, u_{n+1}(t))$, $V = U \times \dots \times U$,

$$\begin{aligned} \tilde{f}(t, x, v, \lambda) &= \sum_{i=1}^{n+1} \lambda_i f(t, x, u_i), & \tilde{b}(t, x, v, \lambda) &= (b(t, x, u_1), \dots, b(t, x, u_{n+1})), \\ \tilde{h}(t, x, y, z) &= \tilde{h}(t, x) - z, & \tilde{g}(t, x, v, \lambda) &= (g(t, x, u_1), \dots, g(t, x, u_{n+1})) \end{aligned}$$

and

$$\tilde{S}(t) = \left\{ (x, v, \lambda) \in \mathbb{R}^n \times V \times \Lambda : \tilde{b}(t, x, v, \lambda) = 0, \tilde{g}(t, x, v, \lambda) \leq 0 \right\}.$$

Write $y^* = \tilde{l}(x^*(t_0), x^*(t_1), z^*)$ and $z^* = \max_{t \in [t_0, t_1]} h(t, x^*(t))$. Proposition 2, Theorem 2.3.13 in [31], Carathéodory's Theorem and [H5] allow us to deduce that

$$\{x^*, y^*, z^*, (u_1^*, \dots, u_{n+1}^*) \equiv (u^*, \dots, u^*), (\lambda_1^*, \lambda_2^*, \dots, \lambda_{n+1}^*) \equiv (1, 0, \dots, 0)\}$$

is a strong minimizer for the optimal control problem

$$(O) \left\{ \begin{array}{l} \text{Minimize } y(t_1) \\ \text{over } x \in W^{1,1} \text{ and measurable functions } u_1, \dots, u_{n+1}, \lambda_1, \dots, \lambda_{n+1} \text{ satisfying} \\ \dot{x}(t) = \tilde{f}(t, x(t), v(t), \lambda(t)), \dot{y}(t) = 0, \dot{z}(t) = 0 \quad \text{a.e.}, \\ \tilde{b}(t, x(t), v(t), \lambda(t)) = 0 \quad \text{a.e.}, \\ \tilde{g}(t, x(t), v(t), \lambda(t)) \leq 0 \quad \text{a.e.}, \\ (v(t), \lambda(t)) \in V \times \Lambda \quad \text{a.e.}, \\ \tilde{h}(t, x(t), y(t), z(t)) \leq 0 \quad \text{for all } t \in [t_0, t_1], \\ (x(t_0), x(t_1), y(t_0), z(t_0)) \in \text{epi}\{\tilde{l} + \Psi_{E_0 \times \mathbb{R}^n \times \mathbb{R}}\}. \end{array} \right.$$

Here $(\lambda_1, \dots, \lambda_{n+1})$, (u_1, \dots, u_{n+1}) are regarded as control variables.

We now show that the conditions under which Proposition 3 holds are satisfied by the data of (O). Since this is a problem with convex velocity set, [C] holds. Hypothesis [H2] also holds since the set $V \times \Lambda$ is compact. It is a simple matter to see that \tilde{h} satisfies [H4]. Since b and g satisfy [N], we know that, for each t and each $x \in \mathbb{R}^n$, there exists a $\hat{u} \in U$ such that $b(t, x, \hat{u}) = 0$ and $g(t, x, \hat{u}) \leq 0$. Take $\hat{v} = (\hat{u}, \hat{u}, \dots, \hat{u})$. Then we have $\tilde{b}(t, x, \hat{v}, \hat{\lambda}) = 0$ and $\tilde{g}(t, x, \hat{v}, \hat{\lambda}) \leq 0$, i.e., the data of (O) satisfy [N]. We now need to prove that \tilde{f} , \tilde{b} and \tilde{g} satisfy [H1] and that [H3] hold. Starting with [H1], take any $x, x' \in X_\varepsilon(t)$, $v, v' \in V$ and $\lambda, \lambda' \in \Lambda$. Since f satisfies [H1] and (6), we have

$$\begin{aligned} & |\tilde{f}(t, x, v, \lambda) - \tilde{f}(t, x', v', \lambda')| = \left| \sum_{i=1}^{n+1} \lambda_i f(t, x, u_i) - \sum_{i=1}^{n+1} \lambda'_i f(t, x', u'_i) \right| \\ & \leq \left| \sum_{i=1}^{n+1} \lambda_i f(t, x, u_i) - \sum_{i=1}^{n+1} \lambda_i f(t, x', u'_i) \right| + \left| \sum_{i=1}^{n+1} \lambda_i f(t, x', u'_i) - \sum_{i=1}^{n+1} \lambda'_i f(t, x', u'_i) \right| \\ & \leq k_x^f(t) \sum_{i=1}^{n+1} \lambda_i |x - x'| + k_u^f(t) \sum_{i=1}^{n+1} \lambda_i |u - u'_i| + k_f(t) \sum_{i=1}^{n+1} |\lambda_i - \lambda'_i| \\ & \leq k_x^f(t) |x - x'| + k_u^f(t) (n+1) |v - v'| + k_f(t) (n+1) |\lambda - \lambda'| \\ & \leq k_x^f(t) |x - x'| + 2 \max\{k_u^f(t), k_f(t)\} (n+1) |(v, \lambda) - (v', \lambda')|. \end{aligned}$$

Here we use the fact that $\sum_{i=1}^m |x_i| \leq m|(x_1, \dots, x_m)|$. We get analogous relations for \tilde{b} and \tilde{g} . Summarizing, we have

[HS1] For almost every $t \in [a, b]$, for all $x, x' \in X_\varepsilon(t)$, $v, v' \in V$ and $\lambda, \lambda' \in \Lambda$ we have

$$\begin{aligned} |\tilde{f}(t, x, v, \lambda) - \tilde{f}(t, x', v', \lambda')| &\leq k_x^f(t)|x - x'| + 2 \max\{k_u^f(t), k_f(t)\}(n+1)|(v, \lambda) - (v', \lambda')|, \\ |\tilde{b}(t, x, v, \lambda) - \tilde{b}(t, x', v', \lambda')| &\leq k_x^b(t)(n+1)|x - x'| + k_u^b(t)(n+1)|(v, \lambda) - (v', \lambda')|, \\ |\tilde{g}(t, x, v, \lambda) - \tilde{g}(t, x', v', \lambda')| &\leq k_x^g(t)(n+1)|x - x'| + k_u^g(t)(n+1)|(v, \lambda) - (v', \lambda')|. \end{aligned}$$

This means that \tilde{f} , \tilde{b} and \tilde{g} satisfy [H1].

Next we explore the consequences of [H3] to our new data. Consider any $x \in X_\varepsilon(t)$, (v, λ) such that $(x, v, \lambda) \in \tilde{S}(t)$ and $(\eta, \xi) = (\eta_1, \dots, \eta_{n+1}, \xi) \in N_{V \times \Lambda}^L(v, \lambda)$. Then, we have

$$(x, u_i) \in S(t), \quad \eta_i \in N_U^L(u_i) \text{ for } i = 1, \dots, n+1 \text{ and } \xi = (\xi_1, \dots, \xi_{n+1}) \in N_\Lambda^L(\lambda).$$

Take

$$\tilde{\gamma}^b = (\gamma_1^b, \dots, \gamma_{n+1}^b), \quad \gamma_i^b \in \mathbb{R}^{m_b}, \quad \tilde{\gamma}^g = (\gamma_1^g, \dots, \gamma_{n+1}^g), \quad \gamma_i^g \in \mathbb{R}_+^{m_g}$$

such that $\langle \tilde{\gamma}^g, \tilde{g}(t, x, v, \lambda) \rangle = 0$. Since $\langle \tilde{\gamma}^g, \tilde{g}(t, x, v, \lambda) \rangle = \sum_{i=1}^{n+1} \langle \tilde{\gamma}_i^g, g(t, x, u_i) \rangle = 0$ and $\langle \tilde{\gamma}_i^g, g(t, x, u_i) \rangle \leq 0$, we deduce that $\langle \tilde{\gamma}_i^g, g(t, x, u_i) \rangle = 0$ for $i = 1, \dots, n+1$.

By the sum rule (see [31]), we have the following estimate for $\partial_{x,v,\lambda}^L \left\{ \langle \tilde{\gamma}^b, \tilde{b} \rangle + \langle \tilde{\gamma}^g, \tilde{g} \rangle \right\}$:

$$\begin{aligned} \partial_{x,v,\lambda}^L \left\{ \langle \tilde{\gamma}^b, \tilde{b}(t, x, v, \lambda) \rangle + \langle \tilde{\gamma}^g, \tilde{g}(t, x, v, \lambda) \rangle \right\} \subset \\ \left\{ \left(\sum_{i=1}^{n+1} \alpha_i, \beta_1 - \eta_1, \dots, \beta_{n+1} - \eta_{n+1}, 0 \right) : (\alpha_i, \beta_i - \eta_i) \in \partial_{x,u_i}^L \left\{ \langle \gamma_i^b, b(t, x, u_i) \rangle + \langle \gamma_i^g, g(t, x, u_i) \rangle \right\} \right\}. \end{aligned}$$

Suppose now that

$$(\alpha, \beta - \eta, \chi - \xi) \in \partial_{x,v,\lambda}^L \left\{ \langle \tilde{\gamma}^b, \tilde{b}(t, x, v, \lambda) \rangle + \langle \tilde{\gamma}^g, \tilde{g}(t, x, v, \lambda) \rangle \right\}.$$

Then, for each $i = 1, \dots, n+1$, there exists $(\alpha_i, \beta_i - \eta_i) \in \partial_{x,u_i}^L \left\{ \langle \gamma_i^b, b(t, x, u_i) \rangle + \langle \gamma_i^g, g(t, x, u_i) \rangle \right\}$ such that

$$(\alpha, \beta - \eta, \chi - \xi) = \left(\sum_{i=1}^{n+1} \alpha_i, \beta_1 - \eta_1, \dots, \beta_{n+1} - \eta_{n+1}, 0 \right).$$

Since, by [H3], we have $|\langle \gamma_i^b, \gamma_i^g \rangle| \leq M(t)|\beta_i|$ for each $i = 1, \dots, n+1$, we deduce that

$$|\langle \tilde{\gamma}^b, \tilde{\gamma}^g \rangle| \leq \sum_{i=1}^{n+1} |\langle \gamma_i^b, \gamma_i^g \rangle| \leq M(t) \sum_{i=1}^{n+1} |\beta_i| \leq M(t) \sum_{i=1}^{n+1} |(\beta_i, \xi_i)| \leq M(t)(n+1)|(\beta, \xi)|.$$

Thus the following condition is satisfied:

[HS3] For almost every $t \in [a, b]$, all $x \in X_\varepsilon(t)$ and all (v, λ) such that $(x, v, \lambda) \in \tilde{S}(t)$, $(\eta, \xi) \in N_{V \times \Lambda}^L(v, \lambda)$ and any $\tilde{\gamma} \in \mathbb{R}_+^{m \times (n+1)}$ such that $\langle \tilde{\gamma}, \tilde{g}(t, x, v, \lambda) \rangle = 0$ we have

$$(\alpha, \beta - \eta, \chi - \xi) \in \partial_{x,v,\lambda}^L \left\{ \langle \tilde{\gamma}^b, \tilde{b}(t, x, v, \lambda) \rangle + \langle \tilde{\gamma}^g, \tilde{g}(t, x, v, \lambda) \rangle \right\} \implies |\langle \tilde{\gamma}^b, \tilde{\gamma}^g \rangle| \leq (n+1)M(t)|(\beta, \eta)|.$$

Finally, observe that the function $\tilde{M}(k_x^{\tilde{b}} + k_x^{\tilde{g}})k_u^{\tilde{f}}$, where $\tilde{M} = M(n+1)$, $k_x^{\tilde{b}} = k_x^b(n+1)$, $k_x^{\tilde{g}} = k_x^g(n+1)$ and $k_u^{\tilde{f}} = 2 \max\{k_u^f, k_f\}(n+1)$, is integrable. This is because

$$\tilde{M}(k_x^{\tilde{b}} + k_x^{\tilde{g}})k_u^{\tilde{f}} = 2(n+1)^3 M(k_x^b + k_x^g) \max\{k_u^f, k_f\}$$

and we are working under the assumption that $M(k_x^b + k_x^g) \max\{k_u^f, k_f\}$ is integrable.

We are now in position to apply Proposition 3 to (O). We also show that the conclusions of Proposition 3 hold with $\lambda_0 = 1$. This is done mimicking the proof of Proposition 9.5.4 in [31]. Details are omitted. Putting together all our findings, we obtain the required conditions. \square

We now prove the main result for (P) Here we follow closely the proof of Proposition 9.5.5 in [31] and so, once again, we keep the details to a minimum. Observe that we now work with a general endpoint constraint $(x(t_0), x(t_1)) \in E$.

Recall that (6), (7) and (8) hold. Consider the set V of all (x, u, e) , where $x \in W^{1,1}$, u is a measurable function and $e \in \mathbb{R}^n$, satisfying $x \in X_\varepsilon(t)$, $\dot{x}(t) = f(t, x(t), u(t))$, $(x(t), u(t)) \in S(t)$ for a.e. $t \in [t_0, t_1]$ and $(x(t_0), e) \in E$. Define now the function $d_V : V \times V \rightarrow \mathbb{R}$ by

$$d_V((x, u, e), (x', u', e')) = |x(t_0) - x'(t_0)| + |e - e'| + \int_{t_0}^{t_1} |u(t) - u'(t)| dt$$

For all i , we choose $\delta_i \downarrow 0$ and set

$$\tilde{l}_i(x, y, x', y', z) := \max\{l(x, y) - l(x^*(t_0), x^*(t_1)) + \delta_i^2, z, |x' - y'|\}.$$

The set V is nonempty since $(x^*, u^*, x^*(t_1)) \in V$. It is straightforward to verify that d_V defines a metric on V , (V, d_V) is a complete metric space,

$$(x, u, e) \rightarrow \tilde{l}_i(x(t_0), e, x(t_1), e, \max_{t \in [t_0, t_1]} h(t, x(t)))$$

is continuous on (V, d_V) and

$$(x_i, u_i, e_i) \rightarrow (x, u, e) \in (V, d_V) \implies \|x_i - x\|_{L^\infty} \rightarrow 0.$$

We refer the reader to [7] for a complete proof of these facts. We also remark that

$$\tilde{l}_i(x^*(t_0), x^*(t_1), x^*(t_1), x^*(t_1), \max_{t \in [t_0, t_1]} h(t, x^*(t))) = \delta_i^2. \quad (33)$$

Consider now the optimization problem

$$\text{Minimize } \{\tilde{l}_i(x(t_0), e, x(t_1), e, \max_{t \in [t_0, t_1]} h(t, x(t))) : (x, u, e) \in V\}.$$

Since \tilde{l}_i is non-negative, it follows from (33) that $(x^*, u^*, x^*(t_1))$ is an δ_i^2 -minimizer for the above problem. Ekeland's variational principle (Theorem 3.3.1 in [31]) applies. It asserts the existence of a sequence $\{(x_i, u_i, e_i)\}$ in V such that for each i

$$\begin{aligned} & \tilde{l}_i(x_i(t_0), e_i, x_i(t_1), e_i, \max_{t \in [t_0, t_1]} h(t, x_i(t))) \leq \\ & \tilde{l}_i(x(t_0), e, x(t_1), e, \max_{t \in [t_0, t_1]} h(t, x(t))) + \delta_i d_V((x, u, e), (x_i, u_i, e_i)) \end{aligned} \quad (34)$$

for all $(x, u, e) \in V$ and

$$d_V((x_i, u_i, e_i), (x^*, u^*, x^*(t_1))) \leq \delta_i. \quad (35)$$

The inequality (35) implies that $e_i \rightarrow x^*(t_1)$ and $u_i \rightarrow u^*$ strongly in L^1 . Then there exists a subsequence (we do not relabel) such that $u_i \rightarrow u^*$ almost everywhere and $x_i \rightarrow x^*$ uniformly. Define the arc $y_i \equiv e_i$. We have $y_i \rightarrow x^*(t_1)$ uniformly. By (34) we can now conclude that $(x_i, y_i, w_i \equiv 0, u_i)$ is a strong local minimum for the optimal control problem with mixed constraints

$$(\tilde{R}_i) \begin{cases} \text{Minimize } \tilde{l}_i(x(t_0), y(t_0), x(t_1), y(t_1), \max_{t \in [t_0, t_1]} h(t, x(t))) \\ \quad + \delta_i[|x(t_0) - x_i(t_0)| + |y(t_0) - y_i(t_0)| + w(t_1)] \\ \text{over } x, y, w \in W^{1,1} \text{ and measurable functions } u \text{ satisfying} \\ \dot{x}(t) = f(t, x(t), u(t)), \dot{y}(t) = 0, \dot{w}(t) = |u(t) - u_i(t)| \text{ a.e.,} \\ (x(t), y(t), w(t), u(t)) \in \hat{S}(t) \text{ a.e.,} \\ (x(t_0), y(t_0), w(t_0)) \in E \times \{0\}, \end{cases}$$

where

$$\hat{S}(t) = \{(x, y, w, u) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times U : b(t, x, u) = 0, g(t, x, u) \leq 0\}.$$

Observe that

$$(x, y, w, u) \in \hat{S}(t) \iff (x, u) \in S(t),$$

where $S(t)$ is defined in (3). The data of (\tilde{R}_i) satisfies all the assumptions of the Proposition 4. Applying it we deduce the existence of absolutely continuous functions $p_i^x, p_i^y, p_i^w \in W^{1,1}$, an integrable function γ_i and a non-negative measure $\mu_i \in C^\oplus([t_0, t_1]; \mathbb{R})$ satisfying

- (a) $(-\dot{p}_i^x(t), -\dot{p}_i^y(t), -\dot{p}_i^w(t), 0) \in \partial_{x,u}^C \langle q_i^x(t), f(t, x_i(t), u_i(t)) \rangle + p_i^w(t)|u_i(t) - u_i(t)| \\ - K(t)|(\langle q_i^x(t), p_i^y(t), p_i^w(t) \rangle, \lambda_i)|\partial_{x,u}^C |d_{\hat{S}(t)}(x_i(t), u_i(t)) \text{ a.e.,}$
- (b) $\forall u \in S(t, x_i(t)), \langle q_i(t), f(t, x_i(t), u) \rangle + p_i^w(t)|u - u_i(t)| \leq \langle q_i(t), f(t, x_i(t), u_i(t)) \rangle \text{ a.e.,}$
- (c) $(p_i^x(t_0), p_i^y(t_0), p_i^w(t_0), -q_i(t_1), -p_i^y(t_1), -p_i^w(t_1), \int_{[t_0, t_1]} \mu_i(ds)) \in \\ N_{E \times \{0\}}^L(x_i(t_0), y_i(t_0), w_i(t_0)) \times \{0, 0, 0, 0\} + \partial^L \{l_i(x_i(t_0), y_i(t_0), x_i(t_1), y_i(t_1), \max_{t \in [t_0, t_1]} h(t, x_i(t))) \\ + \delta_i[|x_i(t_0) - x_i(t_0)| + |y_i(t_0) - y_i(t_0)| + w_i(t_1)]\},$
- (d) $\gamma_i(t) \in \bar{\partial}_x h(t, x_i(t)) \quad \mu\text{-a.e.,}$
- (e) $\text{supp}\{\mu_i\} \subset \{t : h(t, x_i(t)) = \max_{s \in [t_0, t_1]} h(s, x_i(s))\}.$

where

$$q_i(t) := \begin{cases} p_i^x(t) + \int_{[t_0, t]} \gamma_i(s) \mu_i(ds) & \text{if } t < t_1, \\ p_i^x(t_1) + \int_{[t_0, t_1]} \gamma_i(s) \mu_i(ds) & \text{if } t = t_1, \end{cases}$$

in the above relations.

A straightforward verification shows that

$$N_{\hat{S}(t)}^C(x_i(t), y_i(t), 0, u_i(t)) = \left\{ (\eta_i, 0, 0, \xi_i) : (\eta_i, \xi_i) \in N_{S(t)}^C(x_i(t), u_i(t)) \right\}.$$

This, together with (a) above and the subdifferential Sum Rule in [8], asserts that $\dot{p}_i^y = 0$ and $\dot{p}_i^w = 0$. Thus $p_i^y(t) = p_i^y$ and $p_i^w(t) = p_i^w$. It follows that (a) leads to

$$(-\dot{p}_i(t), 0) \in \partial^C \langle q_i(t), f(t, x_i(t), u_i(t)) \rangle + (0, p_i^w \beta_i(t)) - N_{S(t)}^C(x_i(t), u_i(t)) \quad (36)$$

with $\beta_i \in \mathbb{R}^k$ and $\|\beta_i(t)\| \leq 1$. From (c), we now get $q_i^w = -\delta_i$ and

$$\begin{aligned} (p_i^x(t_0), p_i^y, -q_i(t_1), -p_i^y, \int_{[t_0, t_1]} \mu_i(dt)) &\in N_E^L(x_i(t_0), y_i(t_0)) \times \{(0, 0, 0)\} + \\ \partial^L \tilde{l}_i(x_i(t_0), y_i(t_0), x_i(t_1), y_i(t_1), \max\{h(t, x_i(t))\}) &+ \delta_i(B \times B) \times \{(0, 0, 0)\}. \end{aligned} \quad (37)$$

From (37), we conclude that $\{\|\mu_i\|_{TV}\}$, $\{p_i^y\}$ and $\{p_i(t_1)\}$ are all bounded sequences. Then from (36) we conclude that $\{p_i\}$ is uniformly bounded and $\{p_i\}$ is uniformly integrably bounded. Following subsequence extraction we get

$$p_i \rightarrow p \text{ uniformly, } p_i^y \rightarrow p^y,$$

and

$$\mu_i \rightarrow \mu, \quad \gamma_i \mu_i(dt) \rightarrow \gamma \mu(dt) \text{ weakly}^*,$$

for some $p \in W^{1,1}$, $p^y \in \mathbb{R}^n$, $\mu \in C^\oplus$ and some Borel measurable function γ . It is then a simple matter to see that $\text{supp}\{\mu\}$ is a subset of $\{t : h(t, x^*(t)) = \max_{s \in [t_0, t_1]} h(s, x^*(s))\}$ and that $\gamma(t) \in \bar{\partial}_x h(t, x^*(t))$ μ -a.e..

We now introduce $q := p + \int \gamma \mu(ds)$. A convergence analysis along the lines of the proof of Theorem 3.1 in [15] and an appeal to the upper semi continuity properties of limiting subdifferentials and normal cones allow us to pass to the limit in relationships (36) and (c) leading to

$$(-\dot{p}(t), 0) \in \partial^C(\langle q(t), f(t, x^*(t), u^*(t)) \rangle - K(t)|\langle q(t), p^y \rangle| d_{S(t)}(x^*(t), u^*(t))) \text{ a.e. } t \in [t_0, t_1]. \quad (38)$$

From (b), we also deduce that

$$\forall u \in S(t, x^*(t)), \langle q(t), f(t, x^*(t), u) \rangle \leq \langle q(t), f(t, x^*(t), u^*(t)) \rangle \text{ a.e..}$$

We now turn to (37). Simple calculations lead us to conclude that

$$\tilde{l}_i(x_i(t_0), y_i(t_0), x_i(t_1), y_i(t_1), \max_{s \in [t_0, t_1]} h(s, x(s))) > 0. \quad (39)$$

for all sufficiently large i (for details, see the proof of Proposition 9.5.5 in [31]).

Set $z_i = \max_{s \in [t_0, t_1]} h(s, x_i(s))$. Following the steps in the proof of Proposition 9.5.5 in [31], we deduce that

$$\begin{aligned} \partial^L \tilde{l}_i(x_i(t_0), y_i(t_0), x_i(t_1), y_i(t_1), z_i) &\subset \\ \{(a, b, e, -e, c) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} : & \\ \exists \tilde{\lambda} \geq 0, \tilde{\lambda} + |e| = 1 \text{ and } (a, b, c) \in \tilde{\lambda} \partial^L \max\{l(x, y) - l(x_i(t_0), y_i(t_0)) + \delta_i^2, z\} &|_{(x_i(t_0), y_i(t_0), z_i)}\}. \end{aligned} \quad (40)$$

This estimation of the limiting subdifferential of \tilde{l} , together with (37), yields

$$\begin{cases} p_i^y = -q_i(t_1) \\ \tilde{\lambda}_i + |q_i(t_1)| = 1 \\ \|\mu_i\|_{TV} = c = \tilde{\lambda}_i(1 - \alpha_i) \\ (p_i(t_0), -q_i(t_1)) \in N_E^L(x_i(t_0), y_i(t_0)) + \alpha_i \tilde{\lambda}_i \partial^L l(x_i(t_0), y_i(t_0)) + \delta_i(b_1, b_2) \end{cases}$$

Also, by (37), we have $\mu_i = 0$ if $z_i \leq 0$. This is because $z_i \leq 0$ implies

$$\tilde{l}_i(x, y, x', y', z) := \max\{l(x, y) - l(x^*(t_0), x^*(t_1)) + \delta_i^2, |x' - y'|\}$$

for (x, y, x', y', z) near $(x_i(t_0), y_i(t_0), x_i(t_1), y_i(t_1), z_i)$, which, in turn, implies $\alpha_i = 1$.

Set $\lambda_i = \alpha_i \tilde{\lambda}_i$. Then $\|\mu_i\|_{TV} = \tilde{\lambda}_i(1 - \alpha_i) = 1 - |q_i(t_1)| - \lambda_i$, and this gives

$$\lambda_i + \|\mu_i\|_{TV} + |q_i(t_1)| = 1.$$

For a subsequence, we have $\lambda_i \rightarrow \lambda$, for some $\lambda \geq 0$. Taking limits as $i \rightarrow \infty$ we obtain

$$(p(t_0), -q(t_1)) \in \lambda \partial^L l(x^*(t_0), x^*(t_1)) + N_E^L(x^*(t_0), x^*(t_1))$$

and

$$\lambda + \|\mu\|_{TV} + |q(t_1)| = 1.$$

It remains to prove that $\bar{\partial}_x h$ will be replaced by $\partial_x^> h$. This is done along the lines of the proof of Proposition 9.5.5 in [31] (we omit the details).

Theorem 2 is now proved in the case $L \equiv 0$. The case in which L is nonzero is reducible to the previous one by adding a state y with dynamics $\dot{y}(t) = L(t, x(t), u(t))$, an additional initial condition $y(0) = 0$, and replacing the integral cost by $y(t_1)$. Applying Theorem 2 to this problem we get the required conditions.

It remains to prove the last assertion of Theorem 2. This covers the case when b and g are strictly differentiable along the optimal solution. This is done appealing to Proposition 4.1 in [12] (see also the proof of Theorem 4.3 in [12]). The proof is now complete.

Acknowledgements: The authors thank the anonymous reviewers of this paper for multiple suggestions and remarks that greatly improved the manuscript.

REFERENCES

- [1] Z. Artstein. *Pontryagin Maximum Principle revisited with feedbacks*. European Journal of Control 17 (2011) pp. 46–54.
- [2] A. V. Arutyunov. *Optimality Conditions. Abnormal and Degenerate Problems*. (1st edn) Kluwer Academic Publishers, Dordrecht (2000).
- [3] A. Arutyunov, D. Karamzin, F.L. Pereira. *Maximum principle in problems with mixed constraints under weak assumptions of regularity*. Journal of Optimization, 59 7 (2010) pp. 1067–1083.
- [4] P. Bettiol, H. Frankowska. *Lipschitz regularity of solution map of control systems with multiple state constraints*. Discrete Contin. Dynam. Systems, Series A, 32 (2012) pp. 1–26.
- [5] P. Bettiol, A. Boccia, R. B. Vinter. *Stratified Necessary Conditions for Differential Inclusions with State Constraints*. SIAM J. Control Optim. 51 (2013) pp. 3903–3917.
- [6] M. H. A. Biswas, M. d. R. de Pinho. *A Variant of Nonsmooth Maximum Principle for State Constrained Problems*. IEEE Proc. of 51th CDC (CDC12) (2012) pp.7685–7690.
- [7] M. H. A. Biswas. *Necessary Conditions for Optimal Control Problems with State Constraints: Theory and Applications*. University of Porto, Faculty of Engineering, DEEC, PDEEC, PhD Thesis, 2013.
- [8] F. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley, New York (1993).
- [9] F. Clarke, Y. Ledyev, R. J. Stern, P. R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer-Verlag, New York (1998).
- [10] F. Clarke. *Necessary conditions in dynamic optimization*. Mem. Amer. Math. Soc. 2005.
- [11] F. Clarke, M. d. R. de Pinho. *The Nonsmooth Maximum Principle*. Control Cybernetics, 38 (2009) pp. 1151–1168.
- [12] F. Clarke, M. d. R. de Pinho. *Optimal control problems with mixed constraints*. SIAM J. Control Optim. 48 (2010) pp. 4500–4524.
- [13] F. Clarke, Y. Ledyev, M. d. R. de Pinho. *An extension of the Schwarzkopf multiplier rule in optimal control*. SIAM J. Control Optim. 49 (2011) pp. 599–610.
- [14] M. d. R. de Pinho, R. B. Vinter. *An Euler-Lagrange inclusion for optimal control problems*. IEEE Trans. Automat. Control, 40 (1995) pp. 1191–1198.
- [15] M. d. R. de Pinho, M. M. A. Ferreira, F. A. C. C. Fontes. *An Euler-Lagrange inclusion for optimal control problems with state constraints*. Dynam. Control Systems 8 (2002) pp. 23–45.
- [16] M. d. R. de Pinho, M. M. A. Ferreira, F. A. C. C. Fontes. *Unmaximized inclusion necessary conditions for nonconvex constrained optimal control problems*. ESAIM Control Optim. Calc. Var., 11 (2005) pp. 614–632.
- [17] M. d. R. de Pinho, P. Loewen, G. N. Silva. *A weak maximum principle for optimal control problems with nonsmooth mixed constraints*. Set-Valued and Variational Analysis, 17 (2009) pp. 203–221.
- [18] E. N. Devdaryani, Y. S. Ledyev. *Maximum principle for implicit control systems*. Appl. Math. Optim., 40 (1999) pp. 79–103.
- [19] A. V. Dmitruk. *Maximum principle for the general optimal control problem with phase and regular mixed constraints*. Comput. Math. Model., 4 (1993) pp. 364–377.

- [20] M. R. Hestenes. *Calculus of Variations and Optimal Control Theory*. John Wiley, New York (1966).
- [21] M. M. A. Ferreira, F. A. C. C. Fontes, R. B. Vinter. *Nondegenerate necessary conditions for nonconvex optimal control problems with state constraints*. Journal of Math. Anal. and App., 233 (1999) pp. 116–129.
- [22] Fernando A. C. C. Fontes, Sofia O. Lopes. Normal Forms of Necessary Conditions for Dynamic Optimization Problems with Pathwise Inequality Constraints Journal of Mathematical Analysis and Applications, 399, 1 (2013) pp.27–37.
- [23] H. Frankowska. *Regularity of minimizers and of adjoint states for optimal control problems under state constraints*. J. Convex Analysis, 13 (2006) pp. 299 – 328.
- [24] I. Kornienko, M. d. R. de Pinho. *Differential Inclusion Approach for Mixed Constrained Problems Revisited*. Submitted (2014).
- [25] I. Kornienko, M. d. R. de Pinho. *Properties of some control systems with mixed constraints in the form of inequalities*. Report, ISR, DEEC, FEUP (2013). (<http://paginas.fe.up.pt/mrpinho/>)
- [26] B. Mordukhovich. *Variational analysis and generalized differentiation. Basic Theory*. Fundamental Principles of Mathematical Sciences 330, Springer-Verlag, Berlin (2006).
- [27] H. J. Pesch, M. Plail. *The Maximum Principle of optimal control: A history of ingenious ideas and missed opportunities*. Control and Cybernetics, 38 (2009) pp. 973–995.
- [28] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. *The Mathematical Theory of Optimal Processes*. John Wiley, New York (1962).
- [29] Rockafellar R. T. and Wets B. *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin (1998).
- [30] R. B. Vinter, G. Pappas. *A maximum principle for nonsmooth optimal-control problems with state constraints*. J. Math. Anal. Appl., 89 (1982) pp. 212–232.
- [31] R. B. Vinter. *Optimal Control*. Birkhäuser, Boston (2000).