Improved Sensitivity Relations in State Constrained Optimal Control

Piernicola Bettiol*, H Frankowska[] and Richard B. Vinter[]

September 12, 2016

Abstract

Sensitivity relations in optimal control provide an interpretation of the costate trajectory and the Hamiltonian, evaluated along an optimal trajectory, in terms of gradients of the value function. While sensitivity relations are a straightforward consequence of standard transversality conditions for state constraint free optimal control problems formulated in terms of control-dependent differential equations with smooth data, their verification for problems with either pathwise state constraints, nonsmooth data, or for problems where the dynamic constraint takes the form of a differential inclusion, requires careful analysis. In this paper we establish validity of both ‘full’ and ‘partial’ sensitivity relations for an adjoint state of the maximum principle, for optimal control problems with pathwise state constraints, where the underlying control system is described by a differential inclusion. The partial sensitivity relation interprets the costate in terms of partial Clarke subgradients of the value function with respect to the state variable, while the full sensitivity relation interprets the couple, comprising the costate and Hamiltonian, as the Clarke subgradient of the value function with respect to both time and state variables. These relations are distinct because, for nonsmooth data, the partial Clarke subdifferential does not coincide with the projection of the (full) Clarke subdifferential on the relevant coordinate space. We show for the first time (even for problems without state constraints) that a costate trajectory can be chosen to satisfy the partial and full sensitivity relations simultaneously. The partial sensitivity relation in this paper is new for state constraint problems, while the full sensitivity relation improves on earlier results in the literature (for optimal control problems formulated in terms of Lipschitz continuous multifunctions), because a less restrictive inward pointing hypothesis is invoked in the proof, and because it is validated for a stronger set of necessary conditions.

Keywords: Optimal Control, Differential Inclusions, State Constraints, Sensitivity.

*Laboratoire de Mathmatiques, Universit de Bretagne Occidentale, 6 Avenue Victor Le Gorgeu, 29200 Brest, France, e-mail: piernicola.bettiol@univ-brest.fr

[]CNRS and Institut de Mathmatiques de Jussieu, Universit Pierre et Marie Curie (Paris 6), case 247, 4 place Jussieu, 75252 Paris cedex 05, e-mail: frankowska@math.jussieu.fr

[]Department of Electrical and Electronic Engineering, Imperial College London, Exhibition Road, London, SW7 2BT, UK, e-mail: r.vinter@imperial.ac.uk

1 Introduction

Consider the state constrained optimal control problem:

$$(P) \begin{cases} \text{Minimize } g(x(T)) \\ \text{over } x(\cdot) \in W^{1,1}([S, T]; \mathbb{R}^n) \text{ satisfying} \\ \dot{x}(t) \in F(t, x(t)) \text{ a.e. } t \in [S, T], \\ x(t) \in A \text{ for all } t \in [S, T], \\ x(S) = x_0, \end{cases}$$

the data for which comprise: an interval $[S, T]$ ($T > S$), a function $g(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$, a multifunction $F(\cdot, \cdot) : [S, T] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$, a point $x_0 \in \mathbb{R}^n$ and a closed set $A \subset \mathbb{R}^n$.

An absolutely continuous function $x(\cdot) : [t_0, t_1] \rightarrow \mathbb{R}^n$ that satisfies $\dot{x}(t) \in F(t, x(t))$, a.e. $t \in [t_0, t_1]$, is called a state trajectory (on $[t_0, t_1]$). Reference to the underlying time interval $[t_0, t_1]$ is omitted if this is clear from context. State trajectories (on $[S, T]$) that satisfy the constraints $x(S) = x_0$ and $x(t) \in A$ for all $t \in [S, T]$ are called feasible state trajectories. A feasible state trajectory that minimizes $g(x(T))$ over all feasible state trajectories is called a minimizer. The infimum of $g(x(T))$ over all feasible state trajectories $x(\cdot)$, written $\inf\{(P)\}$, is the infimum cost for (P) . If no feasible state trajectories exist, we define the infimum cost to be $+\infty$.

For each $(t, x) \in [S, T] \times \mathbb{R}^n$ denote by $(P_{t,x})$ the variant of the optimal control problem (P) in which the initial data (S, x_0) is replaced by (t, x) . We define $V(\cdot, \cdot) : [S, T] \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ to be the function

$$V(t, x) = \inf\{(P_{t,x})\} \quad \text{for } (t, x) \in [S, T] \times \mathbb{R}^n. \quad (1)$$

Write $H(\cdot, \cdot, \cdot) : [S, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ for the Hamiltonian:

$$H(t, x, p) = \max_{v \in F(t, x)} p \cdot v. \quad (2)$$

Consider initially the case $A = \mathbb{R}^n$ (no state constraints). Let $\bar{x}(\cdot)$ be a minimizer. In certain circumstances when, for example, the differential inclusion is associated with a smooth, control-dependent differential equation, the cost function is continuously differentiable and the value function is of class C^1 , it is well known that the couple comprising the Hamiltonian and the (negative) costate trajectory $p(\cdot)$ (appearing in standard first order necessary conditions of optimality) coincides with the gradient of the value function, evaluated along the graph of $\bar{x}(\cdot)$:

$$(H(t, \bar{x}(t), p(t)), -p(t)) = \nabla V(t, \bar{x}(t)) \quad \text{a.e. } t \in [S, T]. \quad (3)$$

The sensitivity relation (3) can, in these circumstances, be simply derived by noting that, for each $t \in (S, T]$, the restriction of $\bar{x}(\cdot)$ to $[t, T]$ solves the free initial time and initial state problem

$$\begin{cases} \text{Minimize } g(x(T)) - V(t, x(t)) \\ \text{over } S \leq t \leq T \text{ and } x(\cdot) \in W^{1,1}([t, T]; \mathbb{R}^n) \text{ satisfying} \\ \dot{x}(s) \in F(s, x(s)) \text{ a.e. } s \in [t, T]. \end{cases}$$

Applying necessary optimality conditions to this problem yields a costate trajectory $p^t(\cdot) : [t, T] \rightarrow \mathbb{R}^n$ (it will depend on $t \in (S, T]$). A version of the relation (3) is thereby recovered from the free-time left-endpoint transversality condition, in which $p^t(t)$ replaces $p(t)$. Since the differential inclusion arises from a smooth, control-dependent differential equation and the cost function is smooth, the costate $p(\cdot)$ restricted to $[t, T]$ is the unique solution to the costate equation on this interval satisfying the right transversality condition. It follows that $p(t)$ coincides

with $p^t(t)$; the proof is then completed by replacing $p^t(t)$ by $p(t)$. This analysis breaks down if, as is typically the case when either the data is nonsmooth or state constraints are present, it is not known in advance that the costate trajectory $p(\cdot)$ is unique. Here the sensitivity relation is in essence a strengthened necessary condition, asserting that it is possible to choose a costate trajectory to satisfy the sensitivity relation. In illustration of the pathologies that can arise, [7] provides an example of a problem where the costate trajectory is not unique, and some choices of costate trajectory violate the sensitivity relation.

Property (3) of the costate trajectory is also a sufficient optimality condition, which remains true when V is locally Lipschitz continuous and when the gradient ∇V is replaced by the (Fréchet) superdifferential of V (see, e.g., [4]).

Generalizations of this sensitivity relation, which remain valid when $V(\cdot, \cdot)$ is no longer of class C^1 but merely locally Lipschitz continuous, take the form of a subdifferential inclusion:

$$(H(t, \bar{x}(t), p(t)), -p(t)) \in \overline{\text{co}} \partial V(t, \bar{x}(t)) \text{ a.e. } t \in [S, T], \quad (4)$$

where $\partial V(t, x)$ is the limiting subdifferential of V at (t, x) (see below for summary of notation and terminology in this paper) were proved in [15], in a nonsmooth, control-dependent differential equations setting, and in [14], in the differential inclusions setting. The proof (for control-dependent differential equations), which clearly needed to be of a quite different nature to that employing the simple ‘uniqueness’ argument above, was based on the construction of a new optimal control problem, related to the original one, but involving additional variables $(a(\cdot), b(\cdot))$ and an extra integral term in the cost, which vanishes when $(a(\cdot), b(\cdot)) \equiv (0, 0)$, and for which the original minimizer $\bar{x}(\cdot)$, coupled with $(a(\cdot), b(\cdot)) \equiv (0, 0)$, is a minimizer. Necessary conditions for optimality are then applied to the new optimal control problem. Implicit in the new problem is a richer class of variations, and so one might expect new information would emerge from application of necessary conditions of optimality: the extra information is the generalized sensitivity relation.

Earlier, the following partial sensitivity relation had been proved in [7]:

$$-p(t) \in \overline{\text{co}} \partial_x V(t, \bar{x}(t)) \text{ a.e. } t \in [S, T]. \quad (5)$$

This would appear to be a consequence of (4), but it is not because, in general,

$$\text{Proj}_x \{ \overline{\text{co}} \partial V(t, \bar{x}(t)) \} \not\subset \overline{\text{co}} \partial_x V(t, \bar{x}(t)). \quad (6)$$

(Here Proj_x denotes the projection onto the x vector coordinates.) The proof of (5) also involved construction of a new optimal control problem, but now comprising only the additional variable $a(\cdot)$, together with a different additional integral cost term. Establishing that the original minimizer, coupled with $a(\cdot) \equiv 0$, was a minimizer for the new problem involved the approximation of the dynamic constraint by a new ‘impulsive’ dynamic constraint.

Now suppose the state constraint is present. Recently, in [2] and [9] (in a controlled-dependent differential equation and a differential inclusion setting respectively), analogous sensitivity relations have been derived for optimal control problems with state constraints, involving velocity sets which are measurable with respect to t and Lipschitz with respect to x . Now, they take the form

$$\left(H(t, \bar{x}(t), p(t) + \int_{[S,t]} d\eta(s)), -(p(t) + \int_{[S,t]} d\eta(s)) \right) \in \partial^0 V(t, \bar{x}(t)) \text{ a.e. } t \in (S, T].$$

Here, $\eta(\cdot)$ is a measure multiplier, associated with the state constraint and $\partial^0 V(t, x)$ is

$$\partial^0 V(t, x) := \limsup_{(t', x') \xrightarrow{E} (t, x)} \text{co } \partial V(t', x'), \quad (7)$$

where $E = [S, T] \times \text{int } A$. (The definition of the ‘lim sup’ operation is given below.) Notice that, if $x \in \text{int } A$ and V is Lipschitz continuous on a neighborhood of (t, x) , then $\partial^0 V(t, x)$ coincides with $\overline{\text{co}} \partial V(t, x)$, so (8) subsumes the earlier full sensitivity relation (4) in the state constraint-free case. The analysis is based, once again, on enriching the class of control variables. The technical challenges associated with the presence of the state constraint were overcome by invoking recent distance estimates, to approximate arbitrary feasible state trajectories for the new problems by ones which are interior to the state constraint set. Preliminary results of this nature appeared in [5].

The contributions of this paper are threefold. First, we show that it is possible to choose a costate trajectory to satisfy both the full and partial sensitivity relations simultaneously. This is a novel feature of our results, even for nonsmooth optimal control problems without state constraints since, in earlier work, it has been shown that the costate trajectory can be chosen to satisfy only one or the other of the relations. Second, we establish for the first time the validity of the partial sensitivity relation for optimal control problems formulated in terms of a differential inclusion, for general state constraint sets. Third, we improve on earlier full sensitivity relations (for problems involving a Lipschitz continuous $F(\cdot, \cdot)$ or, more generally, $F(\cdot, \cdot)$ ’s which are absolutely continuous from the left with respect to the time variable), by invoking, for the class of control systems considered, a less restrictive inward pointing condition than that of [9], and not requiring A to be a set with smooth boundary, as in [2]. ([2] and [9] permit $F(\cdot, \cdot)$ to be merely measurable with respect to the time variable, however.) We also observe that the sensitivity relations in this paper are established with reference to a stronger set of necessary conditions (namely the partially convexified Euler Lagrange condition) than those featuring in [9] (namely the Hamiltonian inclusion). The novel idea in our proofs is to replace the Mayer optimal control problem by a Bolza problem with three additional state variables (see problem (Q^ϵ) in Section 4). The Lagrangian of this Bolza problem is defined via directional derivatives of the value function in the directions of the new state variables. The application of known necessary optimality conditions of Euler Lagrange type provides not only a version of the maximum principle, but also the two sensitivity relations.

Terminology and Notation: The closed unit ball in Euclidean space is written $\mathbb{B}$. Euclidean distance is written $|\cdot|$. The interior of a set C in Euclidean space is written $\text{int } C$. We denote by $d_D(x)$ the Euclidean distance of a point x from a non-empty set D in Euclidean space, namely $\inf_{y \in D} |x - y|$. Given a non-empty subset of $E \subset \mathbb{R}^n$, we write $\text{co } E$ for the convex hull of E , while $\overline{\text{co}} E$ is the convex closure of E . E^* denotes the polar cone of E , namely

$$E^* := \{\xi \in \mathbb{R}^n : \xi \cdot x \leq 0 \text{ for all } x \in E\}.$$

Take a multifunction $G(\cdot) : E \rightsquigarrow \mathbb{R}^k$. The graph of $G(\cdot)$ is the set

$$\text{Gr } \{G(\cdot)\} := \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^k \mid x \in E \text{ and } v \in G(x)\}.$$

Take a set $E \subset \mathbb{R}^k$, a point $\bar{x} \in \mathbb{R}^k$, $\epsilon > 0$ and a multifunction $G(\cdot) : \bar{x} + \epsilon \mathbb{B} \rightsquigarrow \mathbb{R}^n$. The limit inferior and the limit superior of $G(\cdot)$ at a point $x \in \mathbb{R}^k$ along E (in the Kuratowski sense) (see

[1] or [14]) are the sets

$$\begin{aligned}\liminf_{x' \xrightarrow{E} x} G(x') &:= \{v \in \mathbb{R}^n : \limsup_{x' \xrightarrow{E} x} d_{G(x')}(v) = 0\}, \\ \limsup_{x' \xrightarrow{E} x} G(x') &:= \{v \in \mathbb{R}^n : \liminf_{x' \xrightarrow{E} x} d_{G(x')}(v) = 0\}.\end{aligned}$$

The notation $x' \xrightarrow{E} x$ indicates consideration of convergent sequences $\{x_i\}$ such that $x_i \rightarrow x$, all elements of which belong to E .

For a given interval $[t_0, t_1] \subset \mathbb{R}$, the spaces comprising equivalence classes of a.e. equal, integrable functions and a.e. equal, essentially bounded functions $L^p([t_0, t_1]; \mathbb{R}^n)$, $p = 1$ or $p = \infty$, are written briefly $L^p([t_0, t_1])$ or L^p . $\mathcal{L}^\infty([t_0, t_1]; \mathbb{R}^n)$ is the space bounded measurable functions. We write $C([t_0, t_1]; \mathbb{R}^n)$ and $W^{1,1}([t_0, t_1]; \mathbb{R}^n)$ for the set of continuous and, respectively, absolutely continuous $\mathbb{R}^n$ valued functions on $[t_0, t_1]$.

Given a set $R \subset \mathbb{R}^n$ and a multifunction $F(.,.) : [t_0, t_1] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ taking values non-empty sets, we say that $F(., x)$ is *absolutely continuous from the left, uniformly over R* (cf. [3]) if the following condition is satisfied: given any $\epsilon > 0$ we may find $\delta > 0$ such that, for any finite partition of $[t_0, t_1]$ and for any $x \in R$

$$t_0 \leq s_1 < \tau_1 \leq s_2 < \tau_2 \leq \dots \leq s_m < \tau_m \leq t_1$$

satisfying $\sum_{i=1}^m (\tau_i - s_i) < \delta$, we have

$$\sum_{i=1}^m d_{F(\tau_i, x)}(F(s_i, x)) < \epsilon.$$

in which, for arbitrary non-empty closed sets in D' , $D \in \mathbb{R}^n$,

$$d_D(D') := \inf\{\beta > 0 \mid D' \subset D + \beta\mathbb{B}\}.$$

The following constructs from nonsmooth analysis are standard. Definitions and discussion can be found in a number of texts, examples of which include [1], [8], [12] and [14].

Take a closed set $D \subset \mathbb{R}^k$ and a point $\bar{x} \in D$. The strict normal cone $\hat{N}_D(\bar{x})$ to D at $\bar{x}$ and the *limiting normal cone* $N_D(\bar{x})$ of D at $\bar{x}$ are defined as follows

$$\hat{N}_D(\bar{x}) := \{p : \limsup_{x \xrightarrow{D} \bar{x}} |x - \bar{x}|^{-1} p \cdot (x - \bar{x}) \leq 0\},$$

$$N_D(\bar{x}) := \limsup_{x \xrightarrow{D} \bar{x}} \hat{N}_D(x).$$

The *Clarke tangent cone* $T_D^C(\bar{x})$ of D at $\bar{x}$ is

$$T_D^C(\bar{x}) := (N_D^C(\bar{x}))^*.$$

Take a lower semicontinuous, extended-valued function $f : \mathbb{R}^k \rightarrow \mathbb{R} \cup \{+\infty\}$. The effective domain of f , written $\text{dom}\{f\}$, is the set $\{x \in \mathbb{R}^k \mid f(x) < \infty\}$. The epigraph of f , written $\text{epi } f$, is the set $\{(x, \alpha) \in \mathbb{R}^k \times \mathbb{R} \mid f(x) \leq \alpha\}$. The *limiting subdifferential* of f at a point $\bar{x} \in \mathbb{R}^k$ such that $f(\bar{x}) < \infty$ is

$$\partial f(\bar{x}) := \{\xi \in \mathbb{R}^k \mid (\xi, -1) \in N_{\text{epi } f}(\bar{x}, f(\bar{x}))\}.$$

The limiting subdifferential has the following useful characterization:

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

$$\limsup_{x \rightarrow x_i} |x - x_i|^{-1} (\xi_i \cdot (x - x_i) - (f(x) - f(x_i))) \leq 0 \text{ for all } i \in \mathbb{N}\}.$$

We refer to $\overline{\text{co}} \partial f(\bar{x})$ as the Clarke subdifferential of f at $\bar{x}$. If f is Lipschitz continuous on a neighborhood of $\bar{x}$ then the Clarke subdifferential has the following characterization in terms of the Fréchet derivatives of f : for any null set $\mathcal{N} \subset \mathbb{R}^k$

$$\overline{\text{co}} \partial f(\bar{x}) = \text{co}\{\xi : \exists x_i \rightarrow \bar{x} \text{ such that } x_i \notin \mathcal{N}, \nabla f(x_i) \text{ exists for each } i \text{ and } \nabla f(x_i) \rightarrow \xi\}.$$

The contingent derivative $D_{\uparrow} f(\bar{x})(d)$ of f at $\bar{x}$ in the direction $d \in \mathbb{R}^k$ is written

$$D_{\uparrow} f(\bar{x})(d) := \liminf_{h \downarrow 0, e \rightarrow d} h^{-1} [f(\bar{x} + he) - f(\bar{x})].$$

2 Sensitivity Relations

In this section we state sensitivity relations for problem (P). Reference is made to the following hypotheses, in which k is a positive number and $k_F(\cdot) \in L^1([S, T], \mathbb{R}_+)$.

(H1): $F(\cdot, \cdot)$ takes values in the space of closed, non-empty sets. $F(\cdot, x)$ is Lebesgue measurable for each $x \in \mathbb{R}^n$. For any bounded set $R \subset \mathbb{R}^n$ there exists $r > 0$ such that $F(\cdot, x)$ is absolutely continuous from the left, uniformly over $x \in (\partial A + r\mathbb{B}) \cap R$,

(H2): $F(t, x) \subset k(1 + |x|)\mathbb{B}$ and $F(t, x) \subset F(t, x') + k_F(t)|x - x'|\mathbb{B}$

for all $x, x' \in \mathbb{R}^n$, a.e. $t \in [S, T]$,

(H3): $g(\cdot)$ is locally Lipschitz continuous,

(H4): A is a nonempty closed set,

and also :

(IPC): (*Inward Pointing Condition*) for each $(t, x) \in [S, T] \times \partial A$,

$$\left(\liminf_{(t', x') \xrightarrow{D} (t, x)} \text{co } F(t', x') \right) \cap \text{int } T_A^C(x) \neq \emptyset,$$

where $D = [S, T] \times A$.

Define the extended valued function $\widehat{V} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ given by

$$\widehat{V}(y) = \begin{cases} -V(S, y) & \text{if } y \in A \\ +\infty & \text{otherwise.} \end{cases}$$

in which $V(\cdot, \cdot)$ is the value function (1). $H(\cdot, \cdot, \cdot)$ is the Hamiltonian function (2).

Theorem 2.1 Assume (H1)-(H4) and (IPC). Then $V(\cdot, \cdot)$ is locally Lipschitz continuous on $[S, T] \times A$.

Let $\bar{x}(\cdot)$ be a minimizer for (P). Then there exists $p(\cdot) \in W^{1,1}([S, T]; \mathbb{R}^n)$ and a function of bounded variation $\eta(\cdot) : [S, T] \rightarrow \mathbb{R}^n$, continuous from the right on (S, T) , such that

(i): for some positive Borel measure μ on $[S, T]$ and Borel measurable selection

$$\nu(t) \in (\overline{\text{co}} N_A(\bar{x}(t))) \cap \mathbb{B} \quad \mu - \text{a.e.} \quad t \in [S, T]$$

we have

$$\eta(t) = \int_{[S, t]} \nu(s) d\mu(s), \quad \text{for all } t \in (S, T],$$

(ii): $\dot{p}(t) \in \text{co} \{r : (r, q(t)) \in N_{Gr\{F(t, \cdot)\}}(\bar{x}(t), \dot{\bar{x}}(t))\}$ for a.e. $t \in [S, T]$,

(iii): $-q(T) \in \partial g(\bar{x}(T))$, $q(S) \in \partial \widehat{V}(\bar{x}(S))$ and

(iv): $q(t) \cdot \dot{\bar{x}}(t) = \max_{v \in F(t, \bar{x}(t))} q(t) \cdot v$ for a.e. $t \in [S, T]$,

where

$$q(t) := \begin{cases} p(S) & \text{for } t = S, \\ p(t) + \eta(t) & \text{for } t \in (S, T], \end{cases}$$

and the following ‘partial’ and ‘full’ sensitivity relations are satisfied:

(v) $-q(t) \in \partial_x^0 V(t, \bar{x}(t))$ for a.e. $t \in (S, T]$,

in which, for $(t, x) \in [S, T] \times A$,

$$\partial_x^0 V(t, x) := \limsup_{x' \xrightarrow{G} x} \text{co } \partial_x V(t, x'), \quad (8)$$

where $G = \text{int } A$.

(vi) $(H(t, \bar{x}(t), q(t)), -q(t)) \in \partial^0 V(t, \bar{x}(t))$ for a.e. $t \in (S, T]$,

in which, for $(t, x) \in [S, T] \times A$,

$$\partial^0 V(t, x) := \limsup_{(t', x') \xrightarrow{E} (t, x)} \text{co } \partial V(t', x'),$$

where $E = [S, T] \times \text{int } A$.

The hypotheses under which the assertions of the above theorem are valid include the requirement that $F(\cdot, x)$ is absolutely continuous from the left. Can this regularity hypothesis be relaxed? The following theorem addresses this question; it tells us that the assertions remain true under the weaker hypothesis that $F(\cdot, x)$ is merely measurable, provided that the inward pointing hypothesis (IPC) is replaced by:

(IPC)': (Modified Inward Pointing Condition) take any $t \in [S, T]$, $x \in \partial A$. Then there exists a null-set $A_{t,x} \subset [S, T]$ such that, given any $v \in \mathbb{R}^n$ such that

$$v \in \limsup_{\substack{(s,y) \rightarrow (t,x) \\ s \notin A_{t,x}}} F(s, y) \quad \text{and} \quad \max_{n \in N_A^1(x)} n \cdot v \geq 0,$$

there exists $w \in \mathbb{R}^n$ such that

$$w \in \liminf_{\substack{(s,y) \rightarrow (t,x) \\ s \notin A_{t,x}}} \text{co} F(s, y) \quad \text{and} \quad \max_{n \in N_A^1(x)} n \cdot (w - v) < 0.$$

Here $N_A^1(x) = N_A(x) \cap \{\xi \mid |\xi| = 1\}$. It is shown in [9, Lemma 3.5] that (IPC)' is a stronger hypothesis than (IPC). So the theorem below illustrates how the severity of the regularity hypothesis regarding $F(\cdot, x)$ and the inward pointing condition can be traded against one another, while retaining the sensitivity relations. A special case of interest in which (IPC)' is satisfied is the following: A has C^1 smooth boundary and (IPC) is satisfied.

Theorem 2.2 *The assertions of Thm. 2.1 remain valid when hypothesis (H1) is replaced by*

(H1)': $F(.,.)$ takes values in the space of closed, non-empty sets and $F(.,x)$ is Lebesgue measurable for all $x \in \mathbb{R}^n$,

and (IPC) is replaced by (IPC)'.

Comments.

- (a): *Simpler forms of the hypotheses:* (H1) is satisfied if $F(.,.)$ has non-empty compact values and is Lipschitz continuous. If $F(.,.)$ is continuous, then the non-empty intersection condition in (IPC) reduces to the simpler equivalent form:

$$\text{co } F(t, x) \cap \text{int } T_A^C(x) \neq \emptyset \text{ for each } (t, x) \in [S, T] \times \partial A.$$

- (b): The conditions (v) and (vi) are more precise sensitivity relations than those expressed in terms of the Clarke subdifferential $\overline{\text{co}} \partial V(t, \bar{x}(t))$ and partial Clarke subdifferential $\overline{\text{co}} \partial_x V(t, \bar{x}(t))$, because the definitions of $\partial^0 V(t, \bar{x}(t))$ and $\partial_x^0 V(t, \bar{x}(t))$ involve limit taking from within the interior of the state constraint set. Indeed

$$\partial_x^0 V(t, x) = \text{co } \partial_x V(t, x) \quad \text{and} \quad \partial^0 V(t, x) = \text{co } \partial V(t, x)$$

for $(t, x) \in [S, T] \times \text{int } A$ and

$$\partial_x^0 V(t, x) \stackrel{\text{strict}}{\subset} \overline{\text{co}} \partial_x V(t, x) \quad \text{and} \quad \partial^0 V(t, x) \stackrel{\text{strict}}{\subset} \overline{\text{co}} \partial V(t, x)$$

for $(t, x) \in [S, T] \times \partial A$.

- (c): We emphasize that conditions (v) and (vi) are distinct necessary conditions since, as we have already remarked, it is possible that

$$\text{Proj}_x \{ \partial^0 V(t, x) \} \not\subset \partial_x^0 V(t, x).$$

For instance for $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ given by $f(t, x) = |x + 1 - t| - |x|$ we have $f(1, x) = 0$ for all x and therefore $\partial_x^0 f(1, 0) = \{0\}$. It is not difficult to verify that $\partial^0 f(1, 0)$ is equal to the convex hull of the set $\{\pm(1, 0), \pm(1, -2)\}$. Thus $\text{Proj}_x \{ \partial^0 f(1, 0) \} \not\subset \partial_x^0 f(1, 0)$.

The two theorems above assert that a single costate trajectory can be chosen, with respect to which *both* the full and partial sensitivity relation are satisfied. Earlier derived sensitivity relations assert only that a costate can be chosen to satisfy either the full or the partial sensitivity relation.

- (d): The partial sensitivity relation has previously been derived only for nonsmooth optimal control problems not involving state constraints, in which the dynamic constraint is formulated in terms of a control-dependent differential equation. The partial sensitivity is derived here, for the first time, in relation to differential inclusion problems with state constraints.
- (e): The assertions of Thms. 2.1 and 2.2 remain valid (with possibly different $p(.)$ and $\eta(.)$) if we replace the Euler-Lagrange condition (ii) with the following convexified Euler-Lagrange inclusion:

$$(ii)' \quad \dot{p}(t) \in \text{co} \{ r \mid (r, q(t)) \in N_{\text{Gr}\{F(t, \cdot)\}}(\bar{x}(t), \dot{\bar{x}}(t)) \} \quad \text{for a.e. } t \in [S, T].$$

This is true because if $F(.,.)$ is replaced by its convex hull $\text{co } F(.,.)$ the hypotheses invoked in Thms. 2.1 and 2.2 are still valid. Furthermore, $\bar{x}(.)$ remains a minimizer and the value function remains the same. See [3, Proof of Lemma 5.2] and [9, proof of Thm. 5.1]. (The inward pointing conditions (IPC) or (IPC)' have a key role in establishing these facts.) Now apply Thms. 2.1 and 2.2 with $\text{co } F(.,.)$ in place of $F(.,.)$

As is well-known, cf. [14], (ii)' implies the partially convexified Hamiltonian inclusion

$$(ii)'' \quad \dot{p}(t) \in \text{co} \{r \mid (-r, \dot{\bar{x}}(t)) \in \partial_{x,p} H(t, \bar{x}(t), p(t))\} \quad \text{for a.e. } t \in [S, T],$$

and also, therefore, the weaker Hamiltonian inclusion

$$(ii)''' \quad (-\dot{p}(t), \dot{\bar{x}}(t)) \in \text{co } \partial_{x,p} H(t, \bar{x}(t), p(t)) \quad \text{for a.e. } t \in [S, T].$$

- (f): The full sensitivity relation in Thm. 2.2 was previously derived in [9] under the modified inward pointing condition (IPC)'. (See also [2] for such a relation, in the context of state constrained optimal control problems involving control dependent differential equations, when (IPC) is strengthened to require A to have smooth boundary.) Thm 2.1 improves on [9] because, in addition to adjoining the partial sensitivity relation, it asserts the co-state trajectory can be chosen to satisfy the partially convexified Euler Lagrange in place of the weaker Hamiltonian inclusion. (See comment (e).) The full sensitivity relation of Thm 2.1 improves on that in [9], when the multifunction F is absolutely continuous from the left (and, in particular, when it is Lipschitz continuous), also because it invokes the less restrictive inward pointing hypothesis (IPC) in place of (IPC)'.

Proofs of the theorems are supplied below. As in earlier related analysis, we modify the optimal control problem in such a manner that the original minimizer $\bar{x}(.)$ remains a minimizer when the space of admissible state trajectories is enlarged by addition of extra variables; this supplementary information about $\bar{x}(.)$ yields the sensitivity relations. There are however two technical innovations in the analysis to achieve the improved sensitivity relations of this paper. First, we introduce three extra variables $a(.,)$, $b(.,)$ and $c(.,)$ in order to establish the simultaneous validity of the full and partial sensitivity relations (in place of the single variable or pair of variables previously employed to generate, separately, the partial and full sensitivity relations respectively). Second, we interpret the extra variables as state variables, not control-like variables as before.

3 Discrete Approximation of Lebesgue Integrals

Consider a function $g(.) : [t_0, t_1] \rightarrow \mathbb{R}^k$. For any given $\tau \in \mathbb{R}$ and positive integer N , we associate with $g(.)$ the discrete vector measure on Borel subsets of $[t_0, t_1]$:

$$g_{N,\tau}(s) := \frac{1}{N} \sum_{j=-\infty}^{+\infty} g\left(\tau + \frac{j}{N}\right) \delta_{\{\tau + \frac{j}{N}\}}(s)$$

in which $\delta_{\{t\}}(.)$ denotes the unit measure concentrated at the point $t \in [t_0, t_1]$. (For purposes of evaluating the right side, we set $g_{N,\tau}(s) = 0$ for $s \notin [t_0, t_1]$; thus the summation contains only a finite number of terms with non-zero weights.)

Observe that

$$\int_{[t_0, t_1]} dg_{N, \tau}(s) = \sum_{\{j \mid \tau + \frac{j}{N} \in [t_0, t_1]\}} g(\tau + \frac{j}{N}) \times \frac{1}{N} \quad (9)$$

is an ‘Euler scheme’ approximation of the integral

$$\int_{[t_0, t_1]} g(s) ds .$$

A proof of the following proposition appears in ([7], Cor. 7.2).

Proposition 3.1 *Take an interval $[t_0, t_1] \subset \mathbb{R}$ and a function $g(\cdot) \in \mathcal{L}^\infty([t_0, t_1]; \mathbb{R}^k)$. Then there exists a Lebesgue subset $\mathcal{T} \subset \mathbb{R}$, of full measure, and a sequence of integers $N_i \uparrow \infty$ with the following property: for every $\tau \in \mathcal{T}$*

$$\sup_{t \in [t_0, t_1]} \left| \int_{[t_0, t]} dg_{N_i, \tau}(s) - \int_{[t_0, t]} g(s) ds \right| \rightarrow 0, \quad \text{as } i \rightarrow \infty .$$

(Notice that $g(\cdot)$ is taken to be a function, not an equivalence class of Lebesgue a.e. equal functions; the set $\mathcal{T}$ having the asserted properties will change if $g(\cdot)$ is modified on a null-set.)

Lemma 3.2 *Take a multifunction $F(\cdot, \cdot) : [S, T] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ that satisfies hypothesis (H1), a point $z_0 \in \mathbb{R}^n$, an interval $[t_0, t_1] \subset [S, T]$, a Lebesgue measurable function $a(\cdot) \in \mathcal{L}^\infty([t_0, t_1]; \mathbb{R}^n)$ and an $\mathcal{L} \times \mathcal{B}^n$ bounded measurable function $q(\cdot, \cdot) : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}$. Let $z(\cdot) \in W^{1,1}([t_0, t_1]; \mathbb{R}^n)$ satisfy the differential inclusion*

$$\begin{cases} \dot{z}(t) \in F(t, z(t)) + a(t) \\ z(t_0) = z_0 . \end{cases} \quad (10)$$

Then a sequence of integers $N_i \uparrow \infty$ and a point $\tau \in [t_0, t_1]$ can be found such that $t_0 - \tau$ and $t_1 - \tau$ are both irrational, and, for each i , there exists a solution $z_i(\cdot) \in W^{1,1}([t_0, t_1]; \mathbb{R}^n)$ to the measure-driven differential inclusion

$$\begin{cases} dz_i(t) \in F(t, z_i(t))dt + da_{N_i, \tau}(t) & \text{on } [t_0, t_1] \\ z_i(t_0) = z_0 \end{cases} \quad (11)$$

satisfying

$$\|z_i(\cdot) - z(\cdot)\|_{L^\infty} \rightarrow 0$$

and for $\vartheta(s) := q(s, a(s))$

$$\int_{[t_0, t_1]} d\vartheta_{N_i, \tau}(s) \rightarrow \int_{[t_0, t_1]} q(s, a(s)) ds, \quad (12)$$

as $i \rightarrow \infty$. (We remark that a solution to the differential inclusion (11) is taken here to mean a piecewise absolutely continuous function on a collection of contiguous subintervals of $[t_0, t_1]$ with discontinuities at the atoms of $a_{N, \tau}(\cdot)$ and jumps equal to the weights of these atoms.)

Proof. For any given $\tau \in \mathbb{R}$ and positive integer N , we consider the discrete measures on Borel subsets of $[t_0, t_1]$ associated with the bounded measurable functions $s \rightarrow a(s)$ and $s \rightarrow \vartheta(s)$ respectively:

$$a_{N, \tau}(s) = \frac{1}{N} \sum_{j=-\infty}^{+\infty} a(\tau + \frac{j}{N}) \delta_{\{\tau + \frac{j}{N}\}}(s)$$

and

$$\vartheta_{N,\tau}(s) = \frac{1}{N} \sum_{j=-\infty}^{+\infty} \vartheta\left(\tau + \frac{j}{N}\right) \delta_{\{\tau + \frac{j}{N}\}}(s) = \frac{1}{N} \sum_{j=-\infty}^{+\infty} q\left(\tau + \frac{j}{N}, a\left(\tau + \frac{j}{N}\right)\right) \delta_{\{\tau + \frac{j}{N}\}}(s). \quad (13)$$

Applying Prop. 3.1 to the $\mathbb{R}^{n+1}$ valued function $s \rightarrow (a(s), \vartheta(s))$, we are assured of the existence of $\tau \in [t_0, t_1]$ such that $t_0 - \tau$ and $t_1 - \tau$ are both irrational, and a sequence $N_i \uparrow \infty$ such that (12) holds true and

$$\int_{[t_0, t]} da_{N_i, \tau}(s) \rightarrow \int_{[t_0, t]} a(s) ds \quad \text{uniformly on } [t_0, t_1] \text{ as } i \rightarrow \infty.$$

Note next that the trajectory $z(\cdot)$ satisfies

$$z(t) := \int_{[t_0, t]} a(s) ds + y(t) \quad \text{for all } t \in [t_0, t_1],$$

where $y(\cdot)$ is a solution to the differential inclusion

$$\begin{cases} \dot{y}(t) \in F(t, y(t) + \int_{[t_0, t]} a(s) ds) \\ y(t_0) = z_0. \end{cases}$$

We deduce from the generalized Filippov existence theorem (see e.g. [1] or [14]) that, for each i , there exists a solution $y_i(\cdot)$ to the differential inclusion

$$\begin{cases} \dot{y}_i(t) \in F(t, y_i(t) + \int_{[t_0, t]} da_{N_i, \tau}(s)) \\ y_i(t_0) = z_0 \end{cases} \quad (14)$$

such that

$$\|y_i(\cdot) - y(\cdot)\|_{L^\infty} \rightarrow 0, \quad \text{as } i \rightarrow \infty.$$

We have

$$\int_{[t_0, t]} da_{N_i, \tau}(s) + y_i(t) \rightarrow \int_{[t_0, t]} a(s) ds + y(t),$$

uniformly on $[t_0, t_1]$, as $i \rightarrow \infty$. Now define, for each i ,

$$z_i(t) = \int_{[t_0, t]} da_{N_i, \tau}(s) + y_i(t).$$

The preceding property, expressed in terms of $z_i(\cdot)$ and $z(\cdot)$ is, simply,

$$\|z_i(\cdot) - z(\cdot)\|_{L^\infty} \rightarrow 0 \quad \text{as } i \rightarrow \infty.$$

We have confirmed the existence of a sequence $z_i(\cdot)$ with the stated properties.

4 Proofs of Theorems 2.1 and 2.2

Proof of Theorem 2.1: Assume the hypotheses of Theorem 2.1 are satisfied.

Lemma 4.1 (i): *The restriction of $V(\cdot, \cdot)$ to $[S, T] \times A$ is locally Lipschitz continuous.*

(ii): *There exists a subset $\mathcal{D} \subset [S, T]$ of full measure such that, for all $t \in \mathcal{D}$, $x \in \text{int } A$ and $e \in F(t, x)$,*

$$D_\uparrow V(t, x)(1, e) \geq 0.$$

Proof. Property (i) can be derived from the distance estimates in [3] (in a manner similar to that used to establish regularity properties of the value function in other contexts, involving different control system descriptions and formulations of the inward pointing condition. See, e.g., [9, Proof of Thm. 4.1]). For a proof of (ii) see [9, Lemma 5.4].

Take a minimizer $\bar{x}(\cdot)$. For given $\epsilon > 0$ we define the functions $\sigma^\epsilon(\cdot, \cdot) : [S, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ and $\tilde{\sigma}^\epsilon : [S, T] \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ by

$$\sigma^\epsilon(t, a) := \sup\{a \cdot \xi : \xi \in -\partial_x V(t, x), x \in (\bar{x}(t) + \epsilon \mathbb{B}) \cap \text{int } A\}. \quad (15)$$

$$\begin{aligned} \tilde{\sigma}^\epsilon(t, b, c) := \sup \{ (\xi, \eta) \cdot (c, -(1+c)b) : (\xi, \eta) \in \partial V(t', x'), \\ (t', x') \in ((t, \bar{x}(t)) + \epsilon \mathbb{B}) \cap [S, T] \times \text{int } A \}. \end{aligned}$$

Notice that from well known properties of support functions we have

$$\sigma^\epsilon(t, a) = \sup\{a \cdot \xi : \xi \in -\text{co } \partial_x V(t, x), x \in (\bar{x}(t) + \epsilon \mathbb{B}) \cap \text{int } A\}$$

and a similar assertion is valid also for $\tilde{\sigma}^\epsilon$.

Lemma 4.2

(i): The functions $\sigma^\epsilon(\cdot, a)$ and $\tilde{\sigma}^\epsilon(\cdot, b, c)$ are $\mathcal{L}$ measurable for all $a, b \in \mathbb{R}^n$, $c \in \mathbb{R}$, and there exists $K > 0$ such that for all $t \in [S, T]$, $\sigma^\epsilon(t, \cdot)$ and $\tilde{\sigma}^\epsilon(t, \cdot, \cdot)$ are Lipschitz continuous with Lipschitz constant K (independent from time).

(ii): Fix $\epsilon > 0$ and bounded measurable functions $b : [S, T] \rightarrow \mathbb{R}^n$, $c : [S, T] \rightarrow \mathbb{R}$. Take any $t_0 < t_1$ in the interval $[S, T]$. Let $z(\cdot) : [t_0, t_1] \rightarrow \text{int } A$ be a piecewise absolutely continuous function, continuous from the right, with a single point of discontinuity at $s \in (t_0, t_1)$. Assume that the restrictions of $z(\cdot)$ to $[t_0, s)$ and $[s, t_1]$ satisfy $\dot{z}(t) \in (1 + c(t))(b(t) + F(t, z(t)))$ a.e. on these intervals and $\|\bar{x} - z\|_{L^\infty} < \epsilon$. Assume, furthermore, that the line segment $[z(s^-), z(s)] \subset \text{int } A$. Then

$$V(t_1, z(t_1)) - V(t_0, z(t_0)) \geq -\sigma^\epsilon(s, z(s) - z(s^-)) - \int_{t_1}^{t_2} \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt.$$

Proof.

(i): Using results from [1, Chapter 8] we can show that the functions $\sigma^\epsilon(\cdot, x)$ and $\tilde{\sigma}^\epsilon(\cdot, y, r)$ are Lebesgue measurable for all $x, y \in \mathbb{R}^n$, $r \in \mathbb{R}$. Since, by Lemma 4.1, $V(\cdot, \cdot)$ is locally Lipschitz continuous on $[S, T] \times A$, by known properties of limiting subdifferentials and polar functions, $\sigma^\epsilon(t, \cdot)$ and $\tilde{\sigma}^\epsilon(t, \cdot, \cdot)$ are Lipschitz continuous with Lipschitz constant K .

(ii): By adding and subtracting terms, we can write

$$\begin{aligned} V(t_1, z(t_1)) - V(t_0, z(t_0)) &= V(t_1, z(t_1)) - V(s, z(s)) + V(s, z(s)) - V(s, z(s^-)) \\ &\quad + V(s, z(s^-)) - V(t_0, z(t_0)). \end{aligned} \quad (16)$$

(Here, we make use of the fact that $V(\cdot, \cdot)$ is finite valued on $[S, T] \times A$, so that the added and subtracted terms giving rise to the above identity are finite.)

Claim: $\frac{d}{dt} V(t, z(t)) + \tilde{\sigma}^\epsilon(t, b(t), c(t)) \geq 0$ a.e. $t \in [t_1, t_2]$.

To verify this, consider the subset $\mathcal{D} \subset [S, T]$ of Lemma 4.1 and also a measurable selection $v(t) \in F(t, z(t))$ such that

$$\dot{z}(t) = (1 + c(t))(b(t) + v(t)) \text{ a.e.} \quad (17)$$

Take $t \in \mathcal{D}$ at which the Lipschitz continuous function $V(\cdot, z(\cdot))$ is differentiable, $v(t) \in F(t, z(t))$ and (17) is satisfied (such t 's have full measure). Then

$$\frac{d}{dt}V(t, z(t)) = \lim_{h \downarrow 0} h^{-1} (V(t + h, z(t) + h(1 + c(t))(v(t) + b(t))) - V(t, z(t))) .$$

By Lebourg's mean value theorem for Clarke subdifferentials [14, Thm. 4.5.3], for all sufficiently small $h > 0$ and some $\theta_h \in [0, 1]$, we can find

$$p_h \in \text{co } \partial V(t + h + h\theta_h c(t), z(t) + h(1 + c(t))(v(t) + b(t)) - h\theta_h(1 + c(t))b(t))$$

such that

$$h p_h \cdot (c(t), -(1 + c(t))b(t)) = \Delta V(h) ,$$

in which

$$\Delta V(h) := V(t + h(1 + c(t)), z(t) + h(1 + c(t))v(t)) - V(t + h, z(t) + h(1 + c(t))(b(t) + v(t))) .$$

By Lemma 4.1 however, $D_{\uparrow}V(t, z(t))(1, v(t)) \geq 0$. Since $D_{\uparrow}V(t, z(t))(\cdot)$ is positively homogeneous, it follows that $D_{\uparrow}V(t, z(t))(1 + c(t), (1 + c(t))v(t)) \geq 0$. But then

$$\begin{aligned} 0 &\leq \liminf_{h \downarrow 0} h^{-1} (V(t + h(1 + c(t)), z(t) + h(1 + c(t))v(t)) - V(t, z(t))) \\ &\leq \limsup_{h \downarrow 0} h^{-1} (V(t + h, z(t) + h(1 + c(t))(v(t) + b(t))) - V(t, z(t))) \\ &\quad + \limsup_{h \downarrow 0} h^{-1} \Delta V(h) \\ &\leq \frac{d}{dt}V(t, z(t)) + \limsup_{h \downarrow 0} p_h \cdot (c(t), -(1 + c(t))b(t)) . \end{aligned}$$

But then $\frac{d}{dt}V(t, z(t)) + \tilde{\sigma}^\epsilon(t, b(t), c(t)) \geq 0$, by definition of $\tilde{\sigma}^\epsilon(\cdot, \cdot, \cdot)$. The claim is verified.

It follows that

$$V(t_1, z(t_1)) - V(s, z(s)) \geq - \int_s^{t_1} \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt$$

and

$$V(s, z(s^-)) - V(t_0, z(t_0)) \geq - \int_{t_0}^s \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt .$$

We now examine the middle term on the right side of (16). Recalling that $[z(s), z(s^-)] \subset (\bar{x}(s) + \epsilon\mathbb{B}) \cap \text{int } A$, we deduce from the mean value theorem that

$$V(s, z(s)) - V(s, z(s^-)) = \xi \cdot (z(s) - z(s^-))$$

for some $\xi \in \text{co } \partial_x V(s, x)$ in which x is a point in $[z(s^-), z(s)]$. This implies that

$$V(s, z(s)) - V(s, z(s^-)) \geq -\sigma^\epsilon(s, z(s) - z(s^-)) .$$

Combining these relations yields the desired inequality.

For arbitrary $\epsilon > 0$ consider the optimal control problem:

$$(Q^\epsilon) \left\{ \begin{array}{l} \text{Minimize } g(x(T)) - V(S, x(S)) + \int_S^T \sigma^\epsilon(t, a(t))dt + \int_S^T \tilde{\sigma}^\epsilon(t, b(t), c(t))dt \\ \text{over } (x(\cdot), a(\cdot), b(\cdot)) \in (W^{1,1}([S, T]; \mathbb{R}^n))^3 \text{ and } c(\cdot) \in W^{1,1}([S, T]; \mathbb{R}) \text{ satisfying} \\ \dot{x}(t) \in a(t) + (1 + c(t))(b(t) + F(t, x(t))) \quad \text{a.e. } t \in [S, T], \\ (\dot{a}(t), \dot{b}(t), \dot{c}(t)) \in (\epsilon\mathbb{B}) \times (\epsilon\mathbb{B}) \times (\epsilon\mathbb{B}) \quad \text{a.e. } t \in [S, T], \\ x(t) \in A \cap (\bar{x}(t) + (\epsilon/4)\mathbb{B}) \quad \text{for all } t \in [S, T], \\ (a(S), b(S), c(S)) = 0. \end{array} \right.$$

Lemma 4.3 *Let $\bar{x}(\cdot)$ be a minimizer for (P) . Then $(\bar{x}(\cdot), \bar{a}(\cdot) \equiv 0, \bar{b}(\cdot) \equiv 0, \bar{c}(\cdot) \equiv 0)$ is a minimizer for (Q^ϵ) for all sufficiently small $\epsilon > 0$.*

Proof. Take $\epsilon > 0$ (we shall place restrictions on its value presently). Let $(x(\cdot), a(\cdot), b(\cdot), c(\cdot))$ be an arbitrary feasible state trajectory for (Q^ϵ) . For fixed $(a(\cdot), b(\cdot), c(\cdot))$, $x(\cdot)$ can be interpreted as a state trajectory (on $[S, T]$) for the differential inclusion

$$\dot{x}(t) \in a(t) + (1 + c(t))(b(t) + F(t, x(t))) .$$

$x(\cdot)$ satisfies the state constraint $x(t) \in A$ for $t \in [S, T]$. Take any bounded set $G \subset \mathbb{R}^n$. It may be deduced from the characterization of the interior of the Clarke tangent cone provided in [13] and a routine compactness argument that, for suitably small $\epsilon > 0$ and $D = [S, T] \times A$,

$$\text{int } T_A^C(z) \cap \left(\liminf_{(t', z') \xrightarrow{D} (t, z)} (a(t') + \text{co}(1 + c(t'))(b(t') + F(t', z'))) \right) \neq \emptyset$$

for all $(t, z) \in [S, T] \times (G \cap \partial A)$.

All the other hypotheses are satisfied for the application of the distance estimate ([3], Thm. 2.3) with reference to the multifunction $(t, x) \rightsquigarrow a(t) + (1 + c(t))(b(t) + F(t, x))$ and state constraint $x(t) \in A$. (Note in particular that the requirement the multifunction be absolutely continuous from the left is met, because of $(H1)$ and the fact that $a(\cdot), b(\cdot), c(\cdot)$ are Lipschitz continuous functions.) Accordingly, we can therefore find a sequence of absolutely continuous functions $x_i(\cdot) : [S, T] \rightarrow \mathbb{R}^n$ such that, for each i , $x_i(S) = x(S)$ and

$$\dot{x}_i(t) \in a(t) + (1 + c(t))(b(t) + F(t, x_i(t))) \quad \text{a.e.}$$

$$x_i(t) \in \text{int } A \quad \text{for all } t \in (S, T] , \tag{18}$$

$$\|x_i(\cdot) - x(\cdot)\|_{L^\infty} \rightarrow 0 \quad \text{as } i \rightarrow \infty . \tag{19}$$

Since $\|\bar{x}(\cdot) - x(\cdot)\|_{L^\infty} \leq \epsilon/4$, we can arrange, by discarding initial terms in the sequence, that

$$\|\bar{x}(\cdot) - x_i(\cdot)\|_{L^\infty} \leq \epsilon/3 \quad \text{for all } i .$$

Take a sequence $\delta_i \downarrow 0$. Now fix the value of the index i , and take $\rho > 0$ such that

$$x_i(t) + \rho\mathbb{B} \subset \text{int } A \quad \text{for all } t \in [S + \delta_i, T] . \tag{20}$$

According to Lemma 3.2, applied to the set valued map $(t, x) \rightsquigarrow (1 + c(t))(b(t) + F(t, x))$ we may choose a sequence of integers $N_j \uparrow \infty$ and $\tau \in \mathbb{R}$ such that $S + \delta_i - \tau$ and $T - \tau$ are irrational and for each j , there exists a solution $z_j(\cdot)$ to the measure-driven differential inclusion

$$\begin{cases} dz_j(t) \in (1 + c(t))(b(t) + F(t, z_j(t)))dt + da_{N_j, \tau}(t) & \text{on } [S + \delta_i, T] \\ z_j(S + \delta_i) = x_i(S + \delta_i) , \end{cases}$$

such that

$$\|z_j(\cdot) - x_i(\cdot)\|_{L^\infty} \rightarrow 0, \quad \text{as } j \rightarrow \infty, \quad (21)$$

and for $\vartheta^\epsilon(s) := \sigma^\epsilon(s, a(s))$

$$\int_{[S+\delta_i, T]} d\vartheta_{N_j, \tau}^\epsilon(s) \rightarrow \int_{[S+\delta_i, T]} \sigma^\epsilon(s, a(s)) ds \quad \text{as } j \rightarrow \infty. \quad (22)$$

But then, by (20) and (21), we have that, for all j sufficiently large and $t \in [S + \delta_i, T]$,

$$z_j(t) + (1/2)\rho\mathbb{B} \subset \text{int } A \quad \text{and} \quad z_j(t) \in \bar{x}(t) + (1/2)\epsilon\mathbb{B}.$$

Write $\{s_1, \dots, s_{M_j}\}$ for the points of discontinuity of z_j in $[S + \delta_i, T]$. (The s_k 's are all located in the open set $(S + \delta_i, T)$ because $S + \delta_i - \tau$ and $T - \tau$ are irrational). Write also $t_k := S + \delta_i + \frac{k}{N}$, for all $k = 0, 1, \dots, M_j - 1$, and $t_{M_j} := T$. Notice that, for each $k \in \{1, \dots, M_j\}$, the function $z_j(\cdot) : [t_{k-1}, t_k] \rightarrow \mathbb{R}^n$ is piecewise absolutely continuous with at most one point of discontinuity at $s_k \in (t_{k-1}, t_k)$. We have

$$\begin{aligned} g(z_j(T)) - V(S + \delta_i, z_j(S + \delta_i)) &= V(T, z_j(T)) - V(S + \delta_i, z_j(S + \delta_i)) \\ &= \sum_{k=1}^{M_j} [V(t_k, z_j(t_k)) - V(t_{k-1}, z_j(t_{k-1}))], \end{aligned}$$

and applying Lemma 4.2, we arrive at

$$\begin{aligned} g(z_j(T)) - V(S + \delta_i, z_j(S + \delta_i)) &\geq - \sum_{k=1}^{M_j} \sigma^\epsilon(s_k, z_j(s_k) - z_j(s_k^-)) - \int_{S+\delta_i}^T \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt \\ &= - \frac{1}{N_j} \sum_{k=1}^{M_j} \sigma^\epsilon(s_k, a(s_k)) - \int_{S+\delta_i}^T \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt \\ &= - \int_{[S+\delta_i, T]} d\vartheta_{N_j, \tau}^\epsilon(s) - \int_{S+\delta_i}^T \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt. \end{aligned}$$

(In the last two equalities we have used (9) and (13).) Passing to the limit as $j \rightarrow \infty$ with the help of (21) and (22) we obtain

$$g(x_i(T)) - V(S + \delta_i, x_i(S + \delta_i)) \geq - \int_{S+\delta_i}^T \sigma^\epsilon(t, a(t)) dt - \int_{S+\delta_i}^T \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt.$$

In view of (19), and since $V(\cdot, \cdot)$ is continuous on $[S, T] \times A$, we obtain from this last relation, in the limit as $i \rightarrow \infty$,

$$g(x(T)) - V(S, x(S)) + \int_S^T \sigma^\epsilon(t, a(t)) dt + \int_S^T \tilde{\sigma}^\epsilon(t, b(t), c(t)) dt \geq 0.$$

On the other hand, since $\bar{x}(\cdot)$ is a minimizer for (P) and $\sigma^\epsilon(t, 0) = 0$, $\tilde{\sigma}^\epsilon(t, 0, 0) = 0$, we have

$$g(\bar{x}(T)) - V(S, \bar{x}(S)) + \int_S^T \sigma^\epsilon(t, \bar{a}(t) = 0) dt + \int_S^T \tilde{\sigma}^\epsilon(t, \bar{b}(t) = 0, \bar{c}(t) = 0) dt = 0.$$

This confirms that $(\bar{x}(\cdot), \bar{a}(\cdot) \equiv 0, \bar{b}(\cdot) \equiv 0, \bar{c}(\cdot) \equiv 0)$ is a minimizer for (Q^ϵ) .

Therefore for given $\epsilon > 0$ sufficiently small,

$$(\bar{x}(\cdot), \bar{a}(\cdot) \equiv 0, \bar{b}(\cdot) \equiv 0, \bar{c}(\cdot) \equiv 0, \bar{d}(\cdot) \equiv 0, \bar{y}(\cdot) \equiv \widehat{V}(\bar{x}(S))) \quad (23)$$

is a strong local minimizer for the optimal control problem

$$\begin{cases} \text{Minimize } g(x(T)) + y(T) + d(T) \\ \text{over } (x(\cdot), a(\cdot), b(\cdot)) \in (W^{1,1}([S, T]; \mathbb{R}^n))^3, (c(\cdot), d(\cdot), y(\cdot)) \in (W^{1,1}([S, T]; \mathbb{R}))^3 \text{ satisfying} \\ (\dot{x}(t), \dot{a}(t), \dot{b}(t), \dot{c}(t), \dot{d}(t), \dot{y}(t)) \in F_1(t, x(t), a(t), b(t), c(t), d(t), y(t)) \text{ , a.e. ,} \\ (x(t), a(t), b(t), c(t), d(t), y(t)) \in A \times (\mathbb{R}^n)^2 \times (\mathbb{R})^3, \text{ for all } t \in [S, T] \text{ ,} \\ y(S) \geq \widehat{V}(x(S)), (a(S), b(S), c(S), d(S)) = 0, \end{cases}$$

in which $F_1 : [S, T] \times (\mathbb{R}^n)^3 \times (\mathbb{R})^3 \rightsquigarrow (\mathbb{R}^n)^3 \times (\mathbb{R})^3$ is the multifunction

$$F_1(t, x, a, b, c, d, y) := \{(a + (1 + c)(b + v), e_1, e_2, e_3, \sigma^\epsilon(t, a) + \bar{\sigma}^\epsilon(t, b, c), 0) \mid v \in F(t, x), e_1, e_2 \in \epsilon\mathbb{B}, e_3 \in [-\epsilon, +\epsilon]\}.$$

Here, ‘strong local minimizer’ means that the state trajectory (23) minimizes the cost over feasible state trajectories in some L^∞ neighborhood of (23). Carrying out a similar analysis to that in [2] or [9], based on an application of the distance estimate from [3] to the differential inclusion

$$\dot{x}(t) \in a(t) + (1 + c(t))(b(t) + F(t, x(t)))$$

for fixed Lipschitz continuous functions $a(\cdot), b(\cdot), c(\cdot)$, satisfying $\|a(\cdot)\|_{L^\infty} + \|b(\cdot)\|_{L^\infty} + \|c(\cdot)\|_{L^\infty} \leq \delta$ for $\delta > 0$ sufficiently small, and also on Lemma 4.3, we can conclude that that

$$(\bar{x}(\cdot), \bar{a}(\cdot) \equiv 0, \bar{b}(\cdot) \equiv 0, \bar{c}(\cdot) \equiv 0, \bar{d}(\cdot) \equiv 0, \bar{y}(\cdot) \equiv \widehat{V}(\bar{x}(S)))$$

remains a strong local solution, when the state constraint ‘ $x(t) \in A$ ’ is removed and the cost function is replaced by

$$g(x(T)) + y(T) + d(T) + C \max_{t \in [S, T]} d_A(x(t)) \text{ ,} \quad (24)$$

for some suitably large number C . Since at this point of the analysis we invoke distance estimates for a left absolutely continuous multifunction, it is important to underline the fact that $a(\cdot), b(\cdot), c(\cdot)$ are restricted to the class of suitably bounded Lipschitz continuous function, not of measurable functions (as in the similar analysis employed in [7]).

The above implies that

$$(\bar{x}(\cdot), \bar{a}(\cdot) \equiv 0, \bar{b}(\cdot) \equiv 0, \bar{c}(\cdot) \equiv 0, \bar{d}(\cdot) \equiv 0, \bar{y}(\cdot) \equiv \widehat{V}(\bar{x}(S)), \bar{z}(\cdot) \equiv 0) \quad (25)$$

is a strong local minimizer for the following problem

$$(\tilde{Q}) \begin{cases} \text{Minimize } g(x(T)) + y(T) + d(T) + Cz(T) \\ (x(\cdot), a(\cdot), b(\cdot)) \in (W^{1,1}([S, T]; \mathbb{R}^n))^3, (c(\cdot), d(\cdot), y(\cdot), z(\cdot)) \in (W^{1,1}([S, T]; \mathbb{R}))^4, \\ (\dot{x}(t), \dot{a}(t), \dot{b}(t), \dot{c}(t), \dot{d}(t), \dot{y}(t), \dot{z}(t)) \in \tilde{F}(t, x(t), a(t), b(t), c(t), d(t), y(t), z(t)) \text{ a.e.} \\ d_A(x(t)) - z(t) \leq 0, \quad \forall t \in [S, T] \\ y(S) \geq \widehat{V}(x(S)), (a(S), b(S), c(S), d(S)) = 0, \end{cases}$$

in which $\tilde{F}(\cdot) : [S, T] \times (\mathbb{R}^n)^3 \times (\mathbb{R})^4 \rightsquigarrow (\mathbb{R}^n)^3 \times (\mathbb{R})^4$ is the multifunction defined by

$$\tilde{F}(t, x, a, b, c, d, y, z) := F_1(t, x, a, b, c, d, y) \times \{0\} \quad (26)$$

which, *de facto*, does not depend on the last three variables d, y, z .

We pause to recall the well known property of Lipschitz multifunctions (see, e.g., [6]), a proof of which is included here because of its key role in the analysis below:

Lemma 4.4 Consider a set valued map $G(\cdot) : \mathbb{R}^m \rightsquigarrow \mathbb{R}^m$ and assume that for some $y_0 \in \mathbb{R}^m$ and $\varepsilon > 0$, $c \geq 0$, $G(\cdot)$ is Lipschitz continuous on the ball $y_0 + \varepsilon\mathbb{B}$ with Lipschitz constant c . Then, for every $\bar{y} \in \mathbb{R}^m$ satisfying $|\bar{y} - y_0| < \varepsilon$ and any $\bar{z} \in G(\bar{y})$, $(r, q) \in \mathbb{R}^m \times \mathbb{R}^m$, the following implication holds true

$$(r, q) \in N_{Gr\{G(\cdot)\}}(\bar{y}, \bar{z}) \Rightarrow |r| \leq c|q|.$$

Proof. Since elements in the normal cone are limits of elements in strict normal cones at neighbouring points, it suffices to verify the implication for an arbitrary point $(r, q) \in \hat{N}_{Gr\{G\}}(\bar{y}, \bar{z})$. Take $\alpha_i \downarrow 0$ and, for each i , take $y_i = \bar{y} + \alpha_i r$. Under the Lipschitz continuity hypothesis there exists, for each i , $z_i \in G(\bar{y} + \alpha_i r)$ such that $|z_i - \bar{z}| \leq \alpha_i c|r|$. By definition of elements in the strict normal cone however, there exists $\epsilon_i \downarrow 0$ such that, for each i ,

$$r \cdot (y_i - \bar{y}) + q \cdot (z_i - \bar{z}) \leq \epsilon_i (|y_i - \bar{y}| + |z_i - \bar{z}|).$$

Since $y_i - \bar{y} = \alpha_i r$ and $|z_i - \bar{z}| \leq \alpha_i c|r|$, it follows that $\alpha_i |r|^2 - \alpha_i c|q||r| \leq \epsilon_i (1 + c)\alpha_i |r|$. Dividing by α_i we obtain, in limit as $i \rightarrow 0$, $|r|^2 - c|q||r| \leq 0$. This implies $|r| \leq c|q|$.

We now apply the necessary conditions of [14, Thm. 10.3.1] with reference to the minimizer (25). The state constraint is

$$h(x, f = (a, b, c, d, y, z)) \leq 0$$

where $h(x, f) := d_A(x) - z$. The necessary conditions are expressed in terms of a hybrid subdifferential, namely

$$\begin{aligned} \partial h^>(\bar{x}(t), \bar{f}(t)) &:= \\ \overline{\text{co}}\{\xi \mid \exists \xi_i \rightarrow \xi, (x_i, f_i) \rightarrow (\bar{x}(t), \bar{f}(t)) \text{ s.t. } \xi_i = \nabla h(x_i, f_i) \text{ and } h(x_i, f_i) > 0 \text{ for each } i\} \end{aligned}$$

in which $\bar{f}(t) = (\bar{a}(t), \bar{b}(t), \bar{c}(t), \bar{d}(t), \bar{y}(t), \bar{z}(t))$. From well-known relations between limiting subdifferentials to the distance function and standard subdifferential calculus, we know

$$\partial h^>(\bar{x}(t), \bar{f}(t)) \subset \begin{cases} (\overline{\text{co}} N_A(\bar{x}(t))) \cap \mathbb{B} \times \{(0, 0, 0, 0, 0)\} \times \{-1\} & \text{if } \bar{x}(t) \in \partial A, \\ \emptyset & \text{if } \bar{x}(t) \in \text{int } A. \end{cases}$$

We can then draw the following conclusions: there exists a costate trajectory

$$\tilde{p}(\cdot) = (p(\cdot), s(\cdot), \omega(\cdot), \xi(\cdot), \tau(\cdot), \phi(\cdot), \psi(\cdot)) \in W^{1,1}([S, T]; \mathbb{R}^{3n+4}),$$

whose components $p(\cdot)$, $s(\cdot)$, $\omega(\cdot)$, $\xi(\cdot)$, $\tau(\cdot)$, $\phi(\cdot)$ and $\psi(\cdot)$ are associated with the state trajectory components $(\bar{x}(\cdot), \bar{\zeta}(\cdot))$ in which

$$\bar{\zeta}(\cdot) := (\bar{a}(\cdot) \equiv 0, \bar{b}(\cdot) \equiv 0, \bar{c}(\cdot) \equiv 0, \bar{d}(\cdot) \equiv 0, \bar{y}(\cdot) \equiv \hat{V}(\bar{x}(S)), \bar{z}(\cdot) \equiv 0),$$

a number $\lambda \geq 0$, a positive Borel measure μ on $[S, T]$ satisfying

$$\text{supp } \mu \subset \{t \mid \bar{x}(t) \in \partial A\}, \quad (27)$$

and a Borel measurable selection

$$\nu(t) \in (\overline{\text{co}} N_A(\bar{x}(t))) \cap \mathbb{B} \quad \mu - \text{a.e.} \quad t \in [S, T] \quad (28)$$

such that $(\tilde{p}(\cdot), \mu, \lambda) \neq 0$ and

- (a): $\dot{\tilde{p}}(t) \in \text{co}\{\alpha \mid (\alpha, \tilde{q}(t)) \in N_{Gr\{\bar{F}(t, \cdot)\}}((\bar{x}(t), \bar{\zeta}(t)), (\dot{\bar{x}}(t), \dot{\bar{\zeta}}(t)))\} \text{ a.e.,}$
- (b): $-\tilde{q}(T) \in \lambda \partial g(\bar{x}(T)) \times \{0\} \times \{0\} \times \{0\} \times \{\lambda\} \times \{\lambda\} \times \{\lambda C\},$

$$(c): (p(S), \phi(S)) \in N_{\text{epi}} \hat{V}(\bar{x}(S), \hat{V}(\bar{x}(S))), \psi(S) = 0,$$

$$(d): \tilde{q}(t) \cdot (\dot{\bar{x}}(t), \dot{\bar{\zeta}}(t)) \geq \tilde{q}(t) \cdot \tilde{v}, \quad \text{for all } \tilde{v} \in \tilde{F}(t, \bar{x}(t), \bar{\zeta}(t)), \text{ a.e.}$$

in which $\tilde{q}(\cdot) : [S, T] \rightarrow \mathbb{R}^{3n+4}$ is the function defined by : $\tilde{q}(S) = \tilde{p}(S)$ and for all $t \in (S, T]$

$$\tilde{q}(t) := \tilde{p}(t) + \left(\int_{[S,t]} \nu(s) d\mu(s), 0, 0, 0, 0, 0, - \int_{[S,t]} d\mu(s) \right).$$

(Note that we have made use of a form of necessary conditions for state constrained problems, in which $\tilde{q}(\cdot)$ is a right continuous function (defined in terms of integrals over $[S, t]$. The necessary conditions are equivalent to those of [14, Thm. 10.3.1], expressed in terms of a left continuous function $\tilde{q}(\cdot)$ (defined in terms of integrals over $[S, t]$ for $t \in (S, T)$). This follows from the fact that $\int_{[S,t]} \nu(s) d\mu(s)$ is of bounded variation, and therefore differs from $\int_{[S,t]} \nu(s) d\mu(s)$ on a set of measure zero.)

To study the implications of condition (a) above, we need to estimate, for each t , limiting normal vectors to $\text{Gr } \tilde{F}(t, \cdot)$ at the point $((\bar{x}(t), \bar{\zeta}(t)), (\dot{\bar{x}}(t), \dot{\bar{\zeta}}(t)))$. Using the definition of limiting normal vectors as limits of strict normal vectors we deduce the following relation:

$$((\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7), (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7)) \in N_{\text{Gr } \tilde{F}(t, \cdot)}((\bar{x}(t), \bar{\zeta}(t)), (\dot{\bar{x}}(t), \dot{\bar{\zeta}}(t)))$$

implies that $\alpha_5 = \alpha_6 = \alpha_7 = 0$, $\beta_2 = 0$, $\beta_3 = \beta_4 = 0$,

$$(\alpha_1, \beta_1) \in N_{\text{Gr } F(t, \cdot)}(\bar{x}(t), \dot{\bar{x}}(t)), \quad -(\alpha_2 + \beta_1) \in \beta_5 \text{ co } \partial_a \sigma^\epsilon(t, \bar{a}(t) = 0) \quad (29)$$

and

$$-(\alpha_3 + \beta_1, \alpha_4 + \beta_1 \cdot \dot{\bar{x}}(t)) \in \beta_5 \text{ co } \partial_{b,c} \tilde{\sigma}^\epsilon(t, \bar{b}(t) = 0, \bar{c}(t) = 0). \quad (30)$$

It follows then from (a) that

$$s(\cdot) \equiv 0, \quad \omega(\cdot) \equiv 0, \quad \xi(\cdot) \equiv 0, \quad \dot{\tau}(\cdot) \equiv 0, \quad \dot{\phi}(\cdot) \equiv 0, \quad \dot{\psi}(\cdot) \equiv 0, \quad (31)$$

and

$$\dot{p}(t) \in \text{co} \{r \mid (r, q(t)) \in N_{\text{Gr } \{F(t, \cdot)\}}(\bar{x}(t), \dot{\bar{x}}(t))\} \text{ a.e.}, \quad (32)$$

in which

$$q(t) := p(t) + \int_{[S,t]} \nu(s) d\mu(s) \quad \text{for } t \in (S, T]. \quad (33)$$

From (b) it follows that

$$-q(T) \in \lambda \partial g(\bar{x}(T)) \quad (34)$$

and

$$\tau(T) = -\lambda, \quad \phi(T) = -\lambda, \quad \psi(T) - \int_{[S,T]} d\mu(s) = -\lambda C,$$

Since $\psi(S) = 0$, we know $\psi(\cdot) \equiv 0$. But then

$$\int_{[S,t]} d\mu(s) = \lambda C.$$

We can deduce that $\lambda > 0$. Indeed if this were not the case then, by the preceding equation $\mu = 0$. But then it can be deduced from (32) and (33), with the help of Lemma 4.4, that $p(\cdot) = 0$.

This contradicts $(p(\cdot), \mu, \lambda) \neq 0$. We can therefore arrange, by scaling the multipliers that $\lambda = 1$.

The preceding relations now imply that $\phi(S) = -1$. So $(p(S), -1) \in N_{\text{epi } \widehat{V}}(\bar{x}(S), \widehat{V}(\bar{x}(S)))$, whence

$$p(S) \in \partial \widehat{V}(\bar{x}(S)). \quad (35)$$

Notice also that (d) implies

$$q(t) \cdot \dot{\bar{x}}(t) \geq q(t) \cdot \tilde{v}, \quad \text{for all } v \in F(t, \bar{x}(t)), \text{ a.e.} \quad (36)$$

We deduce from (31) that $\dot{s}(\cdot) \equiv 0$, $\dot{\omega}(\cdot) \equiv 0$, $\dot{\xi}(\cdot) \equiv 0$. Since $\tau(\cdot) \equiv -1$, it follows from (29) and (30)

$$q(t) \in \text{co } \partial_a \sigma^\epsilon(t, \bar{a}(t) = 0), \quad (37)$$

and

$$(q(t), q(t) \cdot \dot{\bar{x}}(t)) \in \text{co } \partial_{b,c} \tilde{\sigma}^\epsilon(t, \bar{b}(t) = 0, \bar{c}(t) = 0). \quad (38)$$

We see from (28), (32), (33), (34) (with $\lambda = 1$), (35) and (36) that conditions (i)-(iv) of Thm. 2.1 are satisfied, for $p(\cdot)$, $\nu(\cdot)$, μ and $q(\cdot)$ as above. However the sensitivity relations (v) and (vi) are satisfied only in the approximate forms (37) and (38). It remains to replace the approximate forms by (v) and (vi) themselves. For this purpose take $\epsilon_i \downarrow 0$. For each i , we repeat the preceding analysis with $\epsilon = \epsilon_i$, to obtain multipliers, which we label $p_i(\cdot)$, $\nu_i(\cdot)$, μ_i and $q_i(\cdot)$.

It can be deduced from (37) and the local Lipschitz continuity of V on $[S, T] \times A$ that $\{q_i(\cdot)\}_{i \geq 1}$ is bounded in $L^\infty([S, T]; \mathbb{R}^n)$. Since $-p_i(T) \in \partial g(\bar{x}(T))$ we also know that $\{p_i(S)\}$ is bounded. We can then conclude from Lemma 4.4 that

$$|\dot{p}_i(t)| \leq k_F(t) |q_i(t)| \quad \text{for a.e. } t \in [S, T].$$

Consequently, for some integrable function $\ell : [S, T] \rightarrow \mathbb{R}_+$ and all large i , $|\dot{p}_i(t)| \leq \ell(t)$ a.e. and the sequence $\{p_i(\cdot)\}_{i \geq 1}$ is bounded in $L^\infty([S, T]; \mathbb{R}^n)$. Define $\eta_i \in C([S, T]; \mathbb{R}^n)^*$ by $\eta_i(dt) = \nu_i(t) \mu_i(dt)$.

Extracting subsequences (without relabelling) we show that, for some absolutely continuous function $p(\cdot) : [S, T] \rightarrow \mathbb{R}^n$, some positive Borel measure μ on $[S, T]$ and functions of bounded variation $\eta(\cdot)$ and $q(\cdot)$, $p_i(\cdot) \rightarrow p(\cdot)$ uniformly, $\dot{p}_i(\cdot)$ converge to $\dot{p}(\cdot)$ weakly in $L^1([S, T]; \mathbb{R}^n)$, $\mu_i \rightarrow \mu$ weakly* in $C([S, T]; \mathbb{R})^*$, η_i converge to η in $C([S, T]; \mathbb{R}^n)^*$ and $q_i(\cdot) \rightarrow q(\cdot)$ a.e.

By [14, Proposition 9.2.1] there exists a Borel measurable selection $\nu(s) \in (\overline{\text{co}} N_A(\bar{x}(s))) \cap \mathbb{B}$ such that $\eta(dt) = \nu(t) \mu(dt)$. It can be shown, by means of a standard analysis (see [14, proof of Thm. 7.4.1]), that assertions (i)-(iv) of the theorem statement are satisfied, with reference to the limiting multipliers $p(\cdot)$, μ , $q(\cdot)$ and the selector $\nu(\cdot)$.

Since (37) and (38) hold true with q replaced by q_i , we deduce (v) and (vi) by passing to the limit and invoking the definitions of $\partial^0 V$ and $\partial_x^0 V$. All the assertions of the theorem have been confirmed.

Proof of Thm. 2.2: The proof of Thm. 2.1, supplied above, makes use of analytical tools, referred to as ‘distance estimates’, asserting the existence of a state trajectory interior to A on $(S, T]$ whose L^∞ distance from a given state trajectory is linearly estimated. The distance estimates are used in the proof of Lemma 4.3 (to construct sequences x_i of state trajectories for the extended control system satisfying (18) and (19)). They are used again to justify replacing

the state constraint $x(t) \in A$ by a penalty term in the cost function (24) for optimal control problem $(\tilde{Q})$.

The only reason why hypothesis (H1) ($F(\cdot, x)$ is absolutely continuous from the left'), and not the weaker (H1)' ($F(\cdot, x)$ is measurable'), is invoked in Thm. 2.1 is to justify the application, at these steps of the proof, of distance estimates in [3], which require (H1) and the inward pointing hypothesis (IPC) to be satisfied. Suitable distance estimates are available also from [9] to carry these steps of the proof, when (H1) and (IPC) are replaced by (H1)' and (IPC)'. The assertions of Thm. 2.1 are therefore valid under this modification of the hypotheses; we have proved Thm. 2.2.

We commented earlier on the fact that, in this paper, the auxiliary functions $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$ are treated as state trajectory components. This was to justify the application of distance estimates in [3] requiring these functions to be at least continuous. If, however the distance estimates of [9] are used (in the proof of Thm. 2.2), $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$ are permitted to be control-like variables. This fact can be used to provide an alternative, slightly simpler, proof of Thm. 2.2 alone.

Acknowledgments. This work was co-funded by the European Union under the 7th Framework Programme “FP7-PEOPLE-2010-ITN”, grant agreement number 264735-SADCO, and by EPSRC under the grant Control For Energy and Sustainability, grant reference EP/G066477/1.

References

- [1] J.-P. Aubin and H. Frankowska, SET-VALUED ANALYSIS, Birkhäuser Boston, Inc., Boston, Basel, Berlin, 1990.
- [2] P. Bettiol and R. Vinter, *Sensitivity interpretations of the co-state variable for optimal control problems with state constraints*, SIAM J. Control Optim., 48 (2010), pp. 3297–3317.
- [3] P. Bettiol, H. Frankowska and R. B. Vinter, *L^∞ estimates on trajectories confined to a closed subset*, J. of Differential Equations, 252 (2012), pp. 1912–1933.
- [4] P. Cannarsa and H. Frankowska, *Some characterizations of optimal trajectories in control theory*, SIAM J. Control Optim., 29 (1991), pp. 1322–1347.
- [5] A. Cernea and H. Frankowska, *A connection between the maximum principle and dynamic programming for constrained control problems*, SIAM J. Control Optim., 44 (2005), pp. 673–703.
- [6] F. H. Clarke, *A general theorem on necessary conditions in optimal control*, Journal of Discrete and Continuous Dynamical Systems 29 (2011), pp. 485–503.
- [7] F. H. Clarke and R. B. Vinter, *The relationship between the maximum principle and dynamic programming*, SIAM J. Control Optim., 25 (1987), pp. 1291–1311.
- [8] F. H. Clarke, Y. S. Ledyaev, R. J. Stern and P. R. Wolenski, NONSMOOTH ANALYSIS AND CONTROL THEORY, Graduate Texts in Mathematics, vol. 178, Springer Verlag, New York, 1998.
- [9] H. Frankowska and M. Mazzola, *On relations of the adjoint state to the value function for optimal control problems with state constraints*, J. Nonlinear Differential Equations and

Applications (2012), (published on line) <http://link.springer.com/article/10.10072Fs00030-012-0183-0>.

- [10] H. Frankowska and F. Rampazzo, *Filippov's and Filippov-Wazewski's theorems on closed domains*, *J. Differential Equations*, 161 (2000), pp. 449–478.
- [11] F. Rampazzo and R.B. Vinter, *Degenerate optimal control problems with state constraints*, *SIAM J. Control Optim.*, 39 (2000), pp. 989–1007.
- [12] R. T. Rockafellar and R. J.-B. Wets *VARIATIONAL ANALYSIS*, Grundlehren der Mathematischen Wissenschaften, vol. 317, Springer Verlag, New York, 1998.
- [13] R. T. Rockafellar *Clarke's tangent cones and the boundaries of closed sets in $\mathbb{R}^n$* , *Nonlinear Anal. TMA*, 3 (1979), pp. 145–154.
- [14] R. B. Vinter, *OPTIMAL CONTROL*, Birkhäuser, Boston, 2000.
- [15] R. B. Vinter, *New results on the relationship between dynamic programming and the maximum principle*, *Math. Control Signals Systems*, 1 (1988) pp. 97–105.