



HAL
open science

Relationship Between Maximum Principle and Dynamic Programming in presence of Intermediate and Final State Constraints

Olivier Bokanowski, Anya Désilles, Housnaa Zidani

► **To cite this version:**

Olivier Bokanowski, Anya Désilles, Housnaa Zidani. Relationship Between Maximum Principle and Dynamic Programming in presence of Intermediate and Final State Constraints. ESAIM: Control, Optimisation and Calculus of Variations, In press, 10.1051/cocv/2021084 . hal-03079781v1

HAL Id: hal-03079781

<https://hal.science/hal-03079781v1>

Submitted on 17 Dec 2020 (v1), last revised 30 Aug 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Relationship Between Maximum Principle and Dynamic Programming in presence of Intermediate and Final State Constraints*

Olivier Bokanowski[†], Anya Désilles[‡] and Hasnaa Zidani[§]

September 16, 2020

Abstract

In this paper we consider a class of optimal control problems governed by a differential system. We analyse the link between the co-state arc of the Pontryagin Maximum Principle and the value function that associates the optimal value of the control problem to the initial time and state. Such a relationship has been already investigated for state-constrained problems under some controllability assumptions to guarantee Lipschitz regularity property of the value function. Here, we consider the case with intermediate and final state constraints, without any controllability assumption on the system, and without Lipschitz regularity of the value function. Furthermore, our analysis covers the case of normal optimal solutions, and abnormal solutions as well.

AMS Classification: 49K15, 49L20, 34H05

Key words: Optimal Control problems, final and/or intermediate state constraints, Maximum Principle, Hamilton-Jacobi-Bellman equation, Sensitivity analysis

1 Introduction

The aim of this paper is to study the relationship between the Maximum Principle and Dynamic Programming for control problems in presence of finale and/or intermediate state constraints. Let $T > 0$ be a fixed final time horizon, let $t_0 = 0 < t_1 < \dots < t_m = T$ be real numbers (with a fixed integer $m \geq 1$), and U be a compact subset of \mathbb{R}^r (with $r \geq 1$). Our control problem is

*This work has been partially supported by a public grant overseen by the French National Research Agency (ANR) through the "iCODE Institute project" funded by the IDEX Paris-Saclay ANR-11-IDEX-0003-02, and by the DGA under grant 0660037.

[†]Université de Paris, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France and Sorbonne-Université, CNRS, LJLL, F-75005 Paris, France olivier.bokanowski@math.univ-paris-diderot.fr

[‡]Unité de Mathématiques Appliquées (UMA), Ensta ParisTech, 828 Bd des Maréchaux, 91762 Palaiseau Cedex, France anya.desilles@gmail.com

[§]Unité de Mathématiques Appliquées (UMA), Ensta ParisTech, 828 Bd des Maréchaux, 91762 Palaiseau Cedex
Email: hasnaa.zidani@ensta-paris.fr

formulated as follows:

$$\begin{aligned}
\text{Minimize} \quad & \varphi(\mathbf{x}(T)) + \int_t^T \ell(s, \mathbf{x}(s), \mathbf{u}(s)) ds \\
\text{subject to} \quad & \dot{\mathbf{x}}(s) = f(s, \mathbf{x}(s), \mathbf{u}(s)), \text{ a.e. } s \in (t, T), & (1a) \\
& \mathbf{x}(t) = x, & (1b) \\
& \mathbf{u}(s) \in U, \text{ a.e. } s \in (t, T), & (1c) \\
& g_i(\mathbf{x}(t_i)) \leq 0, \quad i = 1, \dots, m & (1d)
\end{aligned}$$

where the dynamics $f : [0, T] \times \mathbb{R}^d \times U \rightarrow \mathbb{R}^d$, the distributed cost $\ell : [0, T] \times \mathbb{R}^d \times U \rightarrow \mathbb{R}$, the final cost $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$, and the constraints $g_i : \mathbb{R}^d \rightarrow \mathbb{R}$ are smooth functions (the precise nature of these functions is unimportant at this juncture. The assumptions will be made precise in the next section). An admissible control function is a measurable function $\mathbf{u} : [0, T] \rightarrow \mathbb{R}^m$ that satisfies (1c). An admissible process is a pair of functions (\mathbf{x}, \mathbf{u}) where \mathbf{u} is an admissible control function and \mathbf{x} is an absolutely continuous solution of (1a)-(1b) that satisfies the constraints (1d).

The study of unconstrained continuous non-linear control problems (i.e. when $g_i \equiv 0$) has provided the foundation of numerous advances in control theory over many decades. In that setting it is very well known by now (see [17, 6, 20] and the references therein) that any optimal process (\mathbf{x}, \mathbf{u}) satisfies the so-called Pontryagin Maximum principle (PMP) stating the existence of a dual arc \mathbf{p} satisfying the following relations:

$$\begin{aligned}
\dot{\mathbf{x}}(s) &= -\nabla_p H(s, \mathbf{x}(s), \mathbf{u}(s), \mathbf{p}(s)) \\
\dot{\mathbf{p}}(s) &= \nabla_x H(s, \mathbf{x}(s), \mathbf{u}(s), \mathbf{p}(s)) \\
H(s, \mathbf{x}(s), \mathbf{u}(s), \mathbf{p}(s)) &= \max_{u \in U} H(s, \mathbf{x}(s), u, \mathbf{p}(s)),
\end{aligned}$$

where $H(s, x, u, p) := -f(s, x, u) \cdot p - \ell(s, x, u)$, for every $(s, x, u, p) \in [0, T] \times \mathbb{R}^d \times U \times \mathbb{R}^d$. The PMP provides a necessary condition for any optimal process. On the other hand, the dynamic programming approach provides valuable information on the value function V that associates to every initial condition (t, x) , the optimal value $V(t, x)$ of the control problem. Under standing assumptions, the value function is guaranteed to be Lipschitz continuous everywhere, see for instance [3, 20]. It is also well known that the value function is characterized as the unique *viscosity solution* of a corresponding partial differential equation (PDE), called Hamilton-Jacobi-Bellman (HJB) equation:

$$-\partial_t V(t, x) + \mathcal{H}(t, x, D_x V(t, x)) = 0 \quad t \in [0, T], \quad x \in \mathbb{R}^d, \quad (3a)$$

$$V(T, x) = \varphi(x) \quad x \in \mathbb{R}^d, \quad (3b)$$

where $\partial_t V$ and $D_x V$ represent respectively the derivative with respect to time and to space, and where $\mathcal{H}(t, x, p) := \max_{u \in U} H(t, x, u, p)$ for every $(t, x, p) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$. Semi-smooth analysis and viscosity theory provide adequate tools to analyse the properties of solutions for HJB equations, we refer to [20, 3] and the references therein.

The connection between the dynamic programming approach and the Pontryagin maximum principle is known since the early work of Bellman and is described in introductory works of optimal control theory with the following relation :

$$\mathbf{p}(t) = \nabla_x V(t, \mathbf{x}(t)), \quad (4)$$

emphasizing the fact that the adjoint arc is a measure of the sensitivity of the optimal cost with respect to perturbations around the optimal trajectory. Combining (3a) and (4), we obtain:

$$(\mathcal{H}(t, \mathbf{x}(t), \mathbf{p}(t)), \mathbf{p}(t)) = \nabla_{t,x} V(t, \mathbf{x}(t)). \quad (5)$$

When the value function is C^1 , and all the data are smooth enough, the proof of the relations (4)-(5) is quite straightforward. However, the assumption on the regularity of the value function does not hold in general even for simple cases [4, 11, 17].

When the value function is only Lipschitz continuous, and in situations where the costate is unique, relations (4)-(5) can be extended by using the framework of the super-differential and the sub-differential [21] (we refer also to [3] where different proofs for relations (4)-(5) are provided).

If the functions f , ℓ and φ are not supposed continuously differentiable in x -variable, the co-state satisfying the PMP may not be unique. In these situations, a natural question arises regarding the existence of some co-state arc that satisfies an extension of the sensitivity relations (4)-(5).

In [7], Clarke and Vinter proved that, for a very large class of non-smooth free endpoint problems, it is possible to choose a co-state function \mathbf{p} such that

$$\mathbf{p}(t) \in \partial_x V(t, \mathbf{x}(t)) \quad \text{a. e. on } (0, T), \quad \mathbf{p}(0) \in \partial_x V(0, \mathbf{x}(0)), \quad \text{and} \quad \mathbf{p}(T) \in \partial_x V(T, \mathbf{x}(T)). \quad (6)$$

These inclusions hold for control problems with a Lipschitz continuous value function, and extend the relation (4) by using the generalised gradient of the value function (which is well defined for Lipschitz continuous functions). In [19], Vinter established also a generalization of (5) stating the existence of a co-state satisfying:

$$(\mathcal{H}(t, \mathbf{x}(t), \mathbf{p}(t)), \mathbf{p}(t)) \in \partial_{t,x} V(t, \mathbf{x}(t)) \quad \text{for all } t \in [0, T]. \quad (7)$$

The link (7) involves the generalized gradient of V in both variables and asserts a property which holds everywhere on the time interval $[0, T]$. Relations (6) and (7) do not entail each other. Moreover, some examples are given in [7, 20] showing that, in some cases, there are a number of possible choices of co-state arcs associated with the same optimal control problem, but not all of them satisfy both sensitivity relations (6) or (7) (some of the co-state arcs do not satisfy any of the two relations, while some other co-state satisfy one relation and not the other).

The proof for (6) (in [7]) and the one for (7) (in [19]) are both based on the construction of a new optimal control problem, related to the original one, but involving additional control variables, and for which the original minimizer, is still a minimizer of the new problem. Necessary conditions for optimality are then applied to the new optimal control problem which has, by construction, a richer class of variations. New information emerge from these necessary conditions and lead to the generalized sensitivity relation. It should be noticed also that the original proof of (6) (in [7]) is much complex than that of (7) and requires the approximation of the dynamic equation by a new ‘impulsive’ system.

In [12, 5], the generalization of sensitivity relations are analyzed for a general nonsmooth differential inclusion setting. In particular in [5], the authors establish the existence of a co-state that satisfies the two sensitivity relations (6)-(7). Note that in [5] the problem is in presence of state constraints but some controllability assumptions are considered so that the value function is locally Lipschitz continuous. Moreover, the proof of (7) given in [5] is based on the approximation of the dynamic equation by an impulsive system.

The introduction of state constraints into non-linear control problems fundamentally impacts their analysis. Indeed, unless some strong controllability assumptions are satisfied, the value function is no longer guaranteed to be Lipschitz continuous. The connection between the value function of a given control problem and corresponding Hamilton-Jacobi-Bellman equation, in the presence of state constraints, is a central issue in control theory, which has been attacked with methods coming from viscosity solutions theory as well as from non-smooth analysis and viability theory. It is well known that, unless some compatibility condition between constraints and dynamics does hold, the value function has not enough regularity, or can fail to be the unique constrained viscosity solution of a HJB equation. Several works have been devoted to the analysis of regularity of the value function when the control problem is in presence of state constraints, we refer for instance to [18, 13, 12, 16, 15, 14] and the references therein.

The link between dynamic programming principle and Pontryagin maximum principle has been established for pointwise state constrained control problems in [12, 5] under some controllability assumptions under which the value function is Lipschitz continuous. The case with final state constraints has been also investigated in [19, 7] under the assumption of Lipschitz regularity of the value function in a "tube" around the trajectory.

In the present work, we are interested in the case when the control problem is in presence of final and/or intermediate constraints. Such problems have already been considered in the literature and optimality conditions have been derived in the form of Pontryagin maximum principle, see for instance [9, 8, 10]. Moreover, in this setting, and without controllability assumption, the value function is in general merely lower semicontinuous, it may have infinite values and might even take finite values only on a set having empty interior. So, it is not clear how to extend the results of [7, 19, 5] (as mentioned in [7], such extension must involve generalized gradients of possibly non-Lipschitz continuous, extended valued functions).

Instead of assuming regularity assumptions on the value function, we follow an idea introduced in [1] and consider an auxiliary control problem without state constraints. The value function of this auxiliary control problem is an interesting tool that can be used to characterize the epigraph of the function ϑ . Moreover, it turns out that the value function of the auxiliary control problem can be used to establish the link between the PMP and the HJB approach along an optimal path, without assuming any controllability assumption and for general optimal solutions (including the abnormal optimal trajectories).

The contributions of this paper are threefold. First, we derive new sensitivity relations between the PMP and the dynamic programming principle for a class of control problems with endpoint and intermediate state constraints without assuming any controllability assumptions. The sensitivity relations hold for normal optimal trajectories and abnormal trajectories as well. Furthermore, we simplify the arguments introduced in [7, 19] to establish the sensitivity relations and we show that both relations can be obtained by considering a same perturbed control problem (without adding complex approximations of the dynamic equation by impulse systems).

The paper is organized as follows. Section 2 introduces the control problem with endpoint and intermediate constraints, and recall some known results in non-smooth analysis. The main results of the paper are given in section 3. Section 4 is devoted to the particular case of endpoint constrained problem. In section 5, we derive the sensitivity relations for a control problem with intermediate costs but without any constraint. This result is also a contribution of the paper and constitutes an important step for the proof of the main results of the paper. Finally, section 6 presents the proof of the main results.

2 Statement of the Problem. Hypotheses

We recall in this section some useful notions in non-smooth analysis. Then we formulate the control problem and lay out the assumptions that will be used in the paper.

2.1 Notations and some preliminaries.

In all this paper, \mathbb{R} denotes the set of real numbers, $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the Euclidean inner product and norm on \mathbb{R}^N (for any $N \geq 1$), \mathbb{B}_N is the unit closed ball $\{x \in \mathbb{R}^N : \|x\| \leq 1\}$ (also denoted \mathbb{B} if there is no ambiguity) and $\mathbb{B}(x; r) = x + r\mathbb{B}$. For any set $S \subseteq \mathbb{R}^N$, $\overset{\circ}{S}, \bar{S}, \partial S, \text{co}S$ denote its interior, closure, boundary, and convex envelope, respectively. For any $a, b \in \mathbb{R}$, we define

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

Similarly, for $a_1, \dots, a_m \in \mathbb{R}$, the notation $\bigvee_{i=1}^m a_i$ is defined as:

$$\bigvee_{i=1}^m a_i := \max(a_1, \dots, a_m).$$

The notation $W^{1,1}([a, b])$ is used for the usual Sobolev space $\{f \in L^1(a, b), f' \in L^1(a, b)\}$. Finally, abbreviation "w.r.t." stands for "with respect to", and "a.e." means "almost everywhere".

Let us recall some definitions and properties from sub-differential calculus that will be used throught the paper. We refer to [6, Chapter 10] or [20, Chapter 5] for more details on non-smooth analysis.

Let $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$ (with $N \geq 1$) be a locally Lipschitz function in a neighborhood of a point $x \in \mathbb{R}^N$. The function ϕ is not necessary differentiable at x , but we can define a *generalized directional derivative* at x , for any direction $v \in \mathbb{R}^N$, as follows:

$$D^\circ \phi(x; v) = \limsup_{y \rightarrow x, h \downarrow 0} \frac{\phi(y + hv) - \phi(y)}{h}. \quad (8)$$

With this notion of directional derivative, we can define the Clarke's sub-differential of ϕ at x by:

$$\partial \phi(x) = \{\xi \in \mathbb{R}^N \mid D^\circ \phi(x; v) \geq \xi \cdot v, \quad \forall v \in \mathbb{R}^d\}. \quad (9)$$

The generalized directional derivative satisfies also:

$$D^\circ \phi(x; v) = \max_{\xi \in \partial \phi(x)} \xi \cdot v, \quad \forall v \in \mathbb{R}^N. \quad (10)$$

It turns out that the sub-differential of ϕ at x is a compact convex set of \mathbb{R}^N , and if $L > 0$ is the Lipschitz constant of ϕ in a neighborhood of x then $\partial \phi(x) \subset \mathbb{B}(0, L)$. The sub-differential satisfies some properties reminiscent of the classical differential calculus; for instance for any $\alpha \in \mathbb{R}$, we have: $\partial(\alpha \phi(x)) = \alpha \partial \phi(x)$. Furthermore, we have: $\partial(\phi + \psi)(x) \subset \partial \phi(x) + \partial \psi(x)$ for ϕ and ψ Lipschitz functions in a neighborhood of x , and when ψ is continuously differentiable (in a neighborhood of x) then

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

Some other properties satisfied by the sub-differential have no smooth counterpart: consider ϕ_1, \dots, ϕ_n continuously differentiable functions on an open neighborhood Ω of x and ϕ the pointwise maximum given by $\phi(y) := \max(\phi_1(y), \dots, \phi_n(y))$ for every $y \in \mathbb{R}^N$, is a Lipschitz continuous function. Moreover, if we denote $\mathcal{I}(x) = \{i \in \{1, \dots, n\}, \mid \phi_i(x) = \phi(x)\}$. Then the sub-differential of ϕ at x is given by:

$$\partial\phi(x) = \text{co}\{\nabla\phi_i(x), i \in \mathcal{I}(x)\}. \quad (11)$$

In the sequel, for a function $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$, we will also use the lower *Dini directional derivative*, in the direction $v \in \mathbb{R}^N$, defined as:

$$D^\uparrow\phi(x; v) = \liminf_{w \rightarrow v, h \downarrow 0} \frac{\phi(x + hw) - \phi(x)}{h}, \quad (12)$$

and when ϕ is Lipschitz in a neighborhood of x then we have the following more simple expression:

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

Some useful properties, that will be used later in the paper, are given in the next Lemma:

Lemma 2.1. *Let $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$ be a Lipschitz continuous function on a neighborhood of $x \in \mathbb{R}^N$, and let $u, v \in \mathbb{R}^N$. Then*

$$D^\uparrow\phi(x; u + v) \geq D^{0,-}\phi(x; u) + D^\uparrow\phi(x; v), \quad (13)$$

where

$$D^{0,-}\phi(x; u) := -D^0(-\phi)(x; u) = \liminf_{\substack{y \rightarrow x, \\ h \downarrow 0}} \frac{\phi(y + hu) - \phi(y)}{h}. \quad (14)$$

Proof. The following decomposition holds:

$$\frac{1}{h}(\phi(x + h(u + v)) - \phi(x)) = \frac{1}{h}(\phi(x + hv + hu) - \phi(x + hv)) + \frac{1}{h}(\phi(x + hv) - \phi(x)) \quad (15)$$

Notice that $y_h = x + hv \rightarrow x$ as $h \rightarrow 0$, hence

$$\begin{aligned} \liminf_{h \downarrow 0} \frac{1}{h}(\phi(x + hv + hu) - \phi(x + hv)) &= \liminf_{h \downarrow 0} \frac{1}{h}(\phi(y_h + hu) - \phi(y_h)) \\ &\geq D^{0,-}\phi(x; u). \end{aligned}$$

Therefore, taking the $\liminf_{h \downarrow 0}$ in (15), we obtain the inequality (13). \square

As direct consequence of Lemma 2.1, for $u, v, w \in \mathbb{R}^N$, the following inequality holds:

$$D^\uparrow\phi(x; u + v + w) \geq D^{0,-}\phi(x; u) + D^{0,-}\phi(x; v) + D^\uparrow\phi(x; w). \quad (16)$$

2.2 Formulation of the problem

Let $T > 0$ be a finite time horizon and consider the following dynamical system (for $0 \leq t < T$):

$$\dot{\mathbf{x}}(s) = f(s, \mathbf{x}(s), \mathbf{u}(s)), \text{ a.e. } s \in (t, T), \quad (17)$$

where the control input $\mathbf{u} : [0, T] \rightarrow \mathbb{R}^r$ ($r \geq 1$) is a measurable function such that $u(s) \in U$ for almost every $s \in [0, T]$. Throughout this paper, we assume that

$$(\mathbf{H}_0) \quad U \text{ is a compact subset of } \mathbb{R}^r$$

We denote by \mathcal{U} the set of all admissible controls:

$$\mathcal{U} := \{\mathbf{u} : [0, T] \rightarrow \mathbb{R}^r \text{ measurable, and } \mathbf{u}(s) \in U \text{ a.e.}\}.$$

The dynamics $f : [0, T] \times \mathbb{R}^d \times U \rightarrow \mathbb{R}^d$ and the distributed cost $\ell : [0, T] \times \mathbb{R}^d \times U \rightarrow \mathbb{R}$ are given functions satisfying the following assumptions:

$$(\mathbf{H}_{1a}) \quad \begin{cases} (i) & f \text{ is continuous on } [0, T] \times \mathbb{R}^d \times U \\ (ii) & x \rightarrow f(t, x, u) \text{ is locally Lipschitz continuous in the following sense:} \\ & \forall R > 0, \exists k_R \geq 0, \forall (x, y) \in (\mathbb{B}_d(0; R))^2, \forall (t, u) \in [0, T] \times U : \\ & \quad \|f(t, x, u) - f(t, y, u)\| \leq k_R \|x - y\| \\ (iii) & \exists c_f > 0 \text{ such that } \|f(t, x, u)\| \leq c_f(1 + \|x\|) \quad \forall (t, x, u) \in [0, T] \times \mathbb{R}^d \times U. \end{cases}$$

and

$$(\mathbf{H}_{1b}) \quad \begin{cases} (i) & \ell \text{ is continuous on } [0, T] \times \mathbb{R}^d \times U \\ (ii) & x \rightarrow \ell(t, x, u) \text{ is locally Lipschitz continuous in the following sense:} \\ & \forall R > 0, \exists k_R \geq 0, \forall (x, y) \in (\mathbb{B}_d(0; R))^2, \forall (t, u) \in [0, T] \times U : \\ & \quad |\ell(t, x, u) - \ell(t, y, u)| \leq k_R \|x - y\| \\ (iii) & \exists c_\ell > 0 \text{ such that } |\ell(t, x, u)| \leq c_\ell(1 + \|x\|) \quad \forall (t, x, u) \in [0, T] \times \mathbb{R}^d \times U. \end{cases}$$

In the light of the standard theory of differential equations, assumptions (ii)-(iii) of (\mathbf{H}_{1a}) guarantee, for every $\mathbf{u} \in \mathcal{U}$ and $(t, x) \in [0, T] \times \mathbb{R}^d$, the existence of an absolutely continuous curve $\mathbf{x} : [t, T] \rightarrow \mathbb{R}^d$ which satisfies (17) and the initial data $\mathbf{x}(t) = x$. Gronwall's Lemma leads also to

$$1 + \|\mathbf{x}(s)\| \leq (1 + \|x\|)e^{c_f(s-t)}, \quad \forall s \in [t, T]. \quad (18)$$

In the sequel, we shall denote the set of all admissible pairs control-and-trajectories starting at time t from the position x as:

$$\mathbb{X}_{[t, T]}(x) := \{(\mathbf{x}, \mathbf{u}) \in W^{1,1}(t, T) \times \mathcal{U} \mid (\mathbf{x}, \mathbf{u}) \text{ satisfies (17) with } \mathbf{x}(t) = x\}.$$

Let $t_0 = 0 < t_1 < \dots < t_m = T$ be given real numbers in $[0, T]$ (for some given integer $m \geq 1$), and consider the control problem

$$\begin{cases} \text{minimize } \varphi(\mathbf{x}(T)) + \int_0^T \ell(t, \mathbf{x}(t), \mathbf{u}(t)) dt, \\ (\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[0, T]}(x_0), \\ g_i(\mathbf{x}(t_i)) \leq 0 \quad \text{for } i = 1, \dots, m, \end{cases} \quad (19)$$

with the standard convention that $\inf \emptyset = +\infty$. The final cost $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ and the constraint functions g_i are given functions satisfying:

(H₂) $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ is locally Lipschitz continuous and there exists a constant $c_\varphi \geq 0$ such that

$$|\varphi(y)| \leq c_\varphi(1 + \|y\|).$$

(H₃) For $i \in \{1, \dots, m\}$ the constraint function $g_i : \mathbb{R}^d \rightarrow \mathbb{R}$ is locally Lipschitz continuous and there exists a constant $c_g \geq 0$ such that, for all i :

$$|g_i(y)| \leq c_g(1 + \|y\|).$$

Remark 2.1. *To simplify the presentation, we assume here that every intermediate constraint is scalar. In the case where there are multiple constraints at an intermediate time t_i :*

$$g_i^j(\mathbf{x}(t_i)) \leq 0, \quad 1 \leq j \leq q_i,$$

it suffices to consider the function $g_i(\mathbf{x}(t_i)) \leq 0 := \min(g_i^1, \dots, g_i^{q_i})$ and replace the multiple constraints by a single one.

We consider the value function $\vartheta : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, associated to this control problem, as follows:

$$\vartheta(t, x) := \inf \left\{ \varphi(\mathbf{x}(T)) + \int_t^T \ell(t, \mathbf{x}(t), \mathbf{u}(t)) dt \mid \right. \\ \left. (\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[t, T]}(x), g_j(\mathbf{x}(t_j)) \leq 0 \quad \text{for all } j \text{ s.t. } t_j \geq t \right\}. \quad (20)$$

It is known that when the control problem does not include any state constraint (for instance if $g_i \equiv 0, \forall i$), under assumptions **(H₀)**-**(H₃)**, the value function ϑ is Lipschitz continuous and can be characterized as unique solution of a Hamilton-Jacobi equation, see [3, Chapter III]. Furthermore, a relation between the Pontryagin's Maximum Principle and the Hamilton-Jacobi approach for problem (19) has been established in [19]. This relation gives a link between an adjoint state of (19) and the sub-differential of the value function ϑ along an optimal trajectory starting from x_0 at $t = 0$.

In this work, we are interested in the case when the control problem is in presence of final or intermediate constraints ($m \geq 1$). In this case the value function fails to be Lipschitz continuous unless some controllability assumptions are satisfied (see [3, Chapter V]). Actually, in general the value function is discontinuous and may have infinite values. So, it is not clear how to define the sub-differential of the value function along a trajectory. A direct extension of the work [19] requires some additional assumptions on the Lipschitz regularity of ϑ in at least a *tube* around an optimal trajectory, see [7] for more details. Instead of assuming such regularity assumptions, we follow an idea introduced in [1] and consider an auxiliary control problem without state constraints. It turns out that the value function of this auxiliary control problem is an interesting tool that can be used to obtain the values of the function ϑ and also to establish the link between the PMP and the HJB approach along an optimal path.

Let us first consider the augmented dynamical system:

$$\dot{\mathbf{x}}(s) = f(x, \mathbf{x}(s), \mathbf{u}(s)) \quad \text{a.e. } s \in (t, T), \quad (21a)$$

$$\dot{\mathbf{z}}(s) = -\ell(s, \mathbf{x}(s), \mathbf{u}(s)) \quad \text{a.e. } s \in (t, T). \quad (21b)$$

This augmented dynamical system is usually used as a classical technique in control theory to recast the Bolza problem in the Mayer form:

$$\vartheta(t, x) = \inf \left\{ \varphi(\mathbf{x}(T)) - \mathbf{z}(T) \mid \begin{array}{l} (\mathbf{x}, \mathbf{z}, \mathbf{u}) \text{ satisfies (21) with } \mathbf{x}(t) = x, \mathbf{z}(t) = 0, \text{ and } g_j(\mathbf{x}(t_j)) \leq 0 \text{ for all } j \text{ s.t. } t_j \geq t \end{array} \right\}.$$

In what follows, we will assume that the augmented dynamics satisfies the following assumption:

(H₄) For any $s \in [0, T]$ and every $x \in \mathbb{R}^d$,

$$\left\{ \left(\begin{array}{c} f(s, x, u) \\ -\ell(s, x, u) + \eta \end{array} \right), u \in U, -c_\ell(1 + \|x\|) + \ell(s, x, u) \leq \eta \leq 0 \right\} \text{ is a convex set.}$$

Remark 2.2. Note that if $\ell \equiv 0$ (Mayer problem), **(H₄)** reduces to

$$f(t, x, U) \text{ convex for all } (t, x) \in [0, T] \times \mathbb{R}^d.$$

Remark 2.3. For $0 \leq a < b \leq T$ and $(x, z) \in \mathbb{R}^d \times \mathbb{R}$, consider the set of all trajectories satisfying (21) on the time interval $[a, b]$, for a control input $u \in \mathcal{U}$, and starting from a position (x, z) at time a :

$$\mathcal{S}_{[a,b]}(x, z) := \{(\mathbf{x}, \mathbf{z}) \in W^{1,1}([a, b]) \mid \exists \mathbf{u} \in \mathcal{U} \text{ such that } (\mathbf{x}, \mathbf{z}, \mathbf{u}) \text{ satisfies (21) with } \mathbf{x}(a) = x, \mathbf{z}(a) = z\}.$$

Under assumptions **(H₀)**, **(H_{1a})**, **(H_{1b})** and **(H₄)**, and by Filippov's Theorem, the multi-valued application $(t, x, z) \rightrightarrows \mathcal{S}_{[t,T]}(x, z)$ is locally Lipschitz continuous from $[0, T] \times \mathbb{R}^{d+1}$ into $W^{1,1}([0, T])$. Moreover, for every $(x, z) \in \mathbb{R}^{d+1}$ and every $t \in [0, T]$, $\mathcal{S}_{[t,T]}(x, z)$ is a compact set in $W^{1,1}([0, T])$ endowed with the $C^0([0, T])$ -topology, see [2, Theorem 1.4.1]. Therefore, for a given $(t, x) \in [0, T] \times \mathbb{R}^d$, if there exists a trajectory $\mathbf{x} \in \mathcal{S}_{[t,T]}(x)$ that satisfies the intermediate and final constraints $g_j(\mathbf{x}(t_j)) \leq 0$ for all j such that $t_j \geq t$, then the value $\vartheta(t, x)$ is finite and the control problem (20) admits a minimal solution (see Proposition 3.1 for a precise statement).

Now, we define an auxiliary control problem and its value function $w : [0, T] \times \mathbb{R}^d \times \mathbb{R}$, for $(x, z) \in \mathbb{R}^d \times \mathbb{R}$ and $t \in [0, T]$ as follows:

$$w(t, x, z) = \min \left\{ \max \left(\varphi(\mathbf{x}(T)) - \mathbf{z}(T), \bigvee_{i \geq 1 \text{ s.t. } t_i \geq t} g_i(\mathbf{x}(t_i)) \right) \mid (\mathbf{x}, \mathbf{z}) \in \mathcal{S}_{[t,T]}(x, z) \right\}. \quad (22)$$

We are interested in characterizing the function w as a solution to a HJB equation and also in expressing the link between the HJB equation and the Pontryagin's principle for problem (19).

In the sequel, we will use the following *Hamiltonian* function $H : [0, T] \times \mathbb{R}^d \times U \times \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ defined, for $(s, x, u, p_x, p_z) \in [0, T] \times \mathbb{R}^d \times U \times \mathbb{R}^d \times \mathbb{R}$, as:

$$H(s, x, u, p_x, p_z) := -f(s, x, u) \cdot p_x + \ell(s, x, u) p_z. \quad (23a)$$

We will also use the *maximized Hamiltonian* function $\mathcal{H} : [0, T] \times \mathbb{R}^d \times U \times \mathbb{R}^{d+1} \times \mathbb{R}$ defined, for $(s, x, u, (p_x, p_z)) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$, as:

$$\mathcal{H}(s, x, p_x, p_z) := \max_{u \in U} \left(-f(s, x, u) \cdot p_x + \ell(s, x, u) p_z \right). \quad (23b)$$

3 Main Results

The optimal control problem (19) might admit no solution if there is no trajectory that can realise all the intermediate constraints. However, under assumptions (\mathbf{H}_0) - (\mathbf{H}_4) , the auxiliary control problem does always achieved a minimum $(\mathbf{x}, \mathbf{z}) \in \mathcal{S}_{0,T]}(x, z)$. Moreover, every optimal solution of the original control problem (19) is related to an optimal solution of the auxiliary control problem (??). More precisely, we have the following result.

Proposition 3.1. *Assume that (\mathbf{H}_0) - (\mathbf{H}_4) are satisfied. Then the following assertions hold:*

(i) *For every $(t, x, z) \in [0, T] \times \mathbb{R}^{d+1}$, the auxiliary control problem (22) admits an optimal solution.*

(ii) *For every $(t, x) \in [0, T] \times \mathbb{R}^d$, if $\vartheta(t, x) < \infty$, then the optimal control problem (20) admits a solution. Moreover, the function $\vartheta : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous.*

(iii) *For every $x \in \mathbb{R}^d$ and $t \in [0, T]$, we have:*

$$\vartheta(t, x) = \min\{z \in \mathbb{R} \mid w(t, x, z) \leq 0\}$$

(by convention, the minimum is $+\infty$ if the set $\{z \in \mathbb{R} \mid w(t, x, z) \leq 0\}$ is empty).

(iv) *Let $(t, x) \in [0, T] \times \mathbb{R}^d$ such that $z_{t,x} := \vartheta(t, x) < \infty$. If $(\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[t,T]}(x)$ is an optimal solution of (20), then $(\mathbf{x}, \mathbf{z}, \mathbf{u})$ is also an optimal solution of the auxiliary problem (22), where $\mathbf{z}(s) := z_{t,x} + \int_s^T \ell(\tau, \mathbf{x}(\tau), \mathbf{u}(\tau)) d\tau$ for every $s \in [t, T]$.*

Conversely, any optimal solution to the auxiliary control problem (22) with $z = \vartheta(t, x) < \infty$ is also an optimal solution to the original problem (20).

Proof. Under assumptions (\mathbf{H}_0) - (\mathbf{H}_4) , the set of trajectories $\mathcal{S}_{[t,T]}(x, z)$ (as defined in Remark 2.3) is a non-empty compact set of $W^{1,1}([t, T])$ endowed with the C^0 -topology. This means that there exists a minimizing sequence $(\mathbf{x}_n, \mathbf{y}_n) \in \mathcal{S}_{[t,T]}(x, z)$ that converges in $C^0([t, T]; \mathbb{R}^{d+1})$ to some $(\bar{\mathbf{x}}, \bar{\mathbf{z}}) \in \mathcal{S}_{[t,T]}(x, z)$, see [2, Theorem 0.3.4 and Theorem 1.4.1]. Moreover the sequence of derivatives $(\dot{\mathbf{x}}_n, \dot{\mathbf{y}}_n)$ converges weakly in $L^1(0, T)$ to $(\dot{\bar{\mathbf{x}}}, \dot{\bar{\mathbf{z}}})$. On the other hand, for any $t \in [0, T]$ and $(x, z) \in \mathbb{R}^d \times \mathbb{R}$, the functional:

$$(\mathbf{x}, \mathbf{z}) \mapsto \max \left(\varphi(\mathbf{x}(T)) - \mathbf{z}(T), \bigvee_{i \geq 1 \text{ s.t. } t_i \geq t} g_i(\mathbf{x}(t_i)) \right)$$

is a continuous function from $C^0([t, T]; \mathbb{R}^{d+1})$ to \mathbb{R} . Therefore, problem (22) achieves a minimum, which proves assertion (i).

Assertion (ii) is also a classical result whose proof is based again on the compactness of the set of trajectories, see [15, Proposition 2.1].

To prove (iii), notice that if $w(t, x, z) \leq 0$ then there exists $(\mathbf{x}, \mathbf{z}) \in \mathcal{S}_{[t, T]}(x, z)$ such that

$$\varphi(\mathbf{x}(T)) - \mathbf{z}(T) \leq 0, \quad \text{and} \quad g_j(\mathbf{x}(t_j)) \leq 0 \text{ for all } j \text{ such that } t_j \geq t,$$

Hence, there exists an admissible control $\mathbf{u} \in \mathcal{U}$ corresponding to the trajectory (\mathbf{x}, \mathbf{z}) such that $(\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[t, T]}(x, z)$ and

$$\varphi(\mathbf{x}(T)) - \int_t^T \ell(s, \mathbf{x}(s), \mathbf{u}(s)) ds \leq z, \quad \text{and} \quad g_j(\mathbf{x}(t_j)) \leq 0 \text{ for all } j \text{ such that } t_j \geq t,$$

Hence, by definition of ϑ , we conclude that $\vartheta(t, x) \leq z$. Furthermore, if there is no z such that $w(t, x, z) \leq 0$, then it means that there is no trajectory \mathbf{x} satisfying the constraints $g_j(\mathbf{x}(t_j)) \leq 0$ ($\forall j \geq 1$ s.t. $t_j \geq t$), hence $\vartheta(t, x) = +\infty$.

Conversely, if $z_{t,x} := \vartheta(t, x) < +\infty$, then by assertion (i), there exists an admissible pair $(\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[t, T]}(x)$ such that $g_j(\mathbf{x}(t_j)) \leq 0$ for every j such that $t_j \geq t$, and such that $0 = \varphi(\mathbf{x}(T)) - \mathbf{z}(T)$, where $\mathbf{z}(T) = z_{t,x} - \int_t^T \ell(s, \mathbf{x}(s), \mathbf{u}(s)) ds$. This implies, that $w(t, x, \bar{z}) \leq 0$. If $\vartheta(t, x) = +\infty$, this means there is no admissible trajectory \mathbf{x} satisfying the constraints $g_j(\mathbf{x}(t_j)) \leq 0$ ($\forall j \geq 1$ s.t. $t_j \geq t$), so $w(t, x, z) > 0$, $\forall z$, and $\inf\{z \in \mathbb{R}, w(t, x, z) \leq 0\} = +\infty$. This concludes the proof of assertion (iii).

Finally assertion (iv) can be proved by using the same arguments as for (iii). \square

Proposition 3.1 states that the auxiliary problem admits an optimal solution even when the original problem doesn't have any solution. It is the sign of the optimal values $w(t, x, \cdot)$ that indicates if the original problem has a solution or not.

The next Theorem gives preliminary results on the auxiliary value function w , and a characterisation of w as solution of an HJB equation (assumption (\mathbf{H}_4) is not needed in this theorem).

Let us denote

$$I_1 := [0, t_1], \quad \text{and} \quad I_j :=]t_{j-1}, t_j], \quad \text{for } j = 2, \dots, m.$$

In particular, $I_1 \cup I_2 \cup \dots \cup I_m$ is a partition of $[0, T]$, and for any $t \in [0, T]$ there is a unique $j \in \{1, \dots, m\}$ such that $t \in I_j$. In the sequel, we denote by \hat{j} the application from $[0, T]$ into $\{1, \dots, m\}$ that associates to every $t \in [0, T]$ the index j such that $t \in I_j$:

$$\hat{j}(t) = j \iff t \in I_j. \tag{24}$$

Proposition 3.2. *Assume that (\mathbf{H}_0) - (\mathbf{H}_3) are verified. Then the value function w is locally Lipschitz continuous on each $I_j \times \mathbb{R}^d \times \mathbb{R}$, $j = 1, \dots, m$, with linear growth (i.e., $|w(t, x, z)| \leq C(1 + \|x\| + |z|)$, $\forall t, x, z$, for some constant $C \geq 0$). Moreover, w is the unique viscosity solution of the following HJB system (with \mathcal{H} defined in (23b)):*

$$-\partial_t w(t, x, z) + \mathcal{H}(t, x, D_x w(t, x, z), D_z w(t, x, z)) = 0, \tag{25a}$$

$$\forall t \in I_j, \quad \forall j = 1, \dots, m, \quad x \in \mathbb{R}^d, \quad z \in \mathbb{R}$$

$$w(t_j^-, x, z) = w(t_j^+, x, z) \bigvee g_j(x) \quad \forall j = 1, \dots, m-1, \quad x \in \mathbb{R}^d, \quad z \in \mathbb{R} \tag{25b}$$

$$w(T, x, z) = (\varphi(x) - z) \bigvee g_m(x) \quad x \in \mathbb{R}^d, \quad z \in \mathbb{R}. \tag{25c}$$

Proof. The result of this Theorem is quite straightforward and uses classical viscosity theory. Indeed, for $t \in]t_{m-1}, t_m] =]t_{m-1}, T]$, the value function w is given by:

$$w(t, x, z) = \min\{(\varphi(\mathbf{x}(T)) - \mathbf{z}(T)) \bigvee g_m(\mathbf{x}(T)), (\mathbf{x}, \mathbf{z}) \in \mathcal{S}_{[t, T]}(x, z)\}.$$

As a value function of a Mayer control problem, with assumptions **(H₂)**-**(H₃)**, the function w is locally Lipschitz continuous on $]t_{m-1}, T] \times \mathbb{R}^d \times \mathbb{R}$ and is the unique viscosity solution of the equation (see [3, Chapter III]):

$$\begin{aligned} -\partial_t w(t, x, z) + \mathcal{H}(t, x, D_x w(t, x, z), D_z w(t, x, z)) &= 0, \quad t \in]t_{m-1}, T], \quad x \in \mathbb{R}^d, \quad z \in \mathbb{R} \\ w(T, x, z) &= (\varphi(x) - z) \bigvee g_m(x) \quad x \in \mathbb{R}^d, \quad z \in \mathbb{R}. \end{aligned}$$

Now, by using the Dynamic Programming principle, for $i = 1, \dots, m-1$ and for every $t \in I_i$, we have for any $h > 0$ such that $t_i < t + h \leq t_{i+1}$:

$$w(t, x, z) = \min\{w(t+h, \mathbf{x}(t+h), \mathbf{z}(t+h)) \bigvee g_i(\mathbf{x}(t_i)) \mid (\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[t, t+h]}(x)\}.$$

Therefore, for $t+h \rightarrow t_i^+$, we obtain

$$w(t, x, z) = \min\{w(t_i^+, \mathbf{x}(t_i), \mathbf{z}(t_i)) \bigvee g_i(\mathbf{x}(t_i)) \mid (\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[t, t_i]}(x)\}.$$

In particular, for $t = t_i^-$, we deduce the relation (25b).

Again by classical arguments, for any $i = 1, \dots, m-1$, the function w is Lipschitz continuous on I_i and is the unique viscosity solution of:

$$-\partial_t w(t, x, z) + \mathcal{H}(t, x, D_x w(t, x, z), D_z w(t, x, z)) = 0 \quad \forall (t, x, z) \in I_i \times \mathbb{R}^d \times \mathbb{R} \quad (26a)$$

$$w(t_i^-, x, z) = w(t_i^+, x, z) \bigvee g_i(x) \quad \forall (x, z) \in \mathbb{R}^d \times \mathbb{R}. \quad (26b)$$

More precisely, it is a classical result that there is a unique locally Lipschitz continuous viscosity solution of (26a) with the given terminal data at $t = t_i$ ($w(t_i, x, z) = w(t_i^+, x, z) \bigvee g_i(x)$). Let us denote by w_i this solution. Then $w(t, x, z) = w_i(t, x, z)$ for any $t_{i-1} < t \leq t_i$. While w_i is Lipschitz regular up to the left boundary $t = t_{i-1}$ of $[t_{i-1}, t_i]$, the function w may differ from w_i at this boundary because of the jump condition (25b). Uniqueness then follows by recursion on $i = m-1, \dots, 1$. This concludes the proof. \square

So far, Propositions 3.1 and 3.2 establish the link between the original control problem and the auxiliary one. One main feature of the function w consists in the fact that it is more regular than ϑ , it takes finite values everywhere, and it is characterized as the viscosity solution of a Hamilton-Jacobi equation (which is convenient for numerical approximations). Now, we are interested in linking the optimality conditions for an optimal solution of problem (19) with the auxiliary value function w . This link is the main result of this paper and it is stated precisely in the following Theorem.

Theorem 3.1. *Assume that **(H₀)**-**(H₄)** are satisfied. Let $x_0 \in \mathbb{R}^n$ be such that $z_0 = \vartheta(0, x_0) < +\infty$, and assume that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{X}_{[0, T]}(x)$ is a solution to the control problem (19). Let $\bar{\mathbf{z}}(\cdot)$ be defined by:*

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

Then, there exist multipliers $(\lambda_i)_{0 \leq i \leq m}$ and a co-state $\bar{\mathbf{p}}(\cdot)$, absolutely continuous on each sub-interval (t_{j-1}, t_j) : $\bar{\mathbf{p}}|_{(t_{j-1}, t_j)} \in W^{1,1}((t_{j-1}, t_j), \mathbb{R}^d)$, such that the following assertions hold:

(i) non-triviality of the multipliers:

$$\sum_{j=0}^m \lambda_j = 1, \quad \lambda_j \geq 0 \quad \text{for } j \geq 0, \quad \text{and} \quad \lambda_j g_j(\bar{\mathbf{x}}(t_j)) = 0 \quad \forall j \geq 1; \quad (27)$$

(ii) co-state equation:

$$\dot{\bar{\mathbf{p}}}(s) \in \partial_x H(s, \bar{\mathbf{x}}(s), \bar{\mathbf{u}}(s), \bar{\mathbf{p}}_x(s), -\lambda_0), \quad \text{a.e. } s \in [0, T], \quad (28a)$$

$$\bar{\mathbf{p}}(t_i^-) \in \bar{\mathbf{p}}(t_i^+) + \lambda_i \partial g_i(\bar{\mathbf{x}}(t_i)) \quad \forall i = 1, \dots, m-1, \quad (28b)$$

$$\bar{\mathbf{p}}(T) \in \lambda_0 \partial \varphi(\bar{\mathbf{x}}(T)) + \lambda_m \partial g_m(\bar{\mathbf{x}}(T)); \quad (28c)$$

(iii) Pontryagin's Principle (with H defined as in (23a)):

$$H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \bar{\mathbf{p}}(t), -\lambda_0) = \max_{u \in U} H(t, \bar{\mathbf{x}}(t), u, \bar{\mathbf{p}}(t), -\lambda_0), \quad \text{for a.e. } t \in [0, T]; \quad (29)$$

(iv) general sensitivity relation:

$$\left(\mathcal{H}(t, \bar{\mathbf{x}}(t), \bar{\mathbf{p}}(t), -\lambda_0), \bar{\mathbf{p}}(t), -\lambda_0 \right) \in q(t) \partial_{t,x,z} w(t, \bar{\mathbf{x}}(t), \bar{\mathbf{z}}(t)) \quad \forall t \in I_j, \quad j = 1, \dots, m \quad (30)$$

where the function $q(t)$ is given by:

$$q(t) := \lambda_0 + \sum_{j=\hat{j}(t)}^m \lambda_j, \quad \text{with } \hat{j}(\cdot) \text{ defined as in (24);}$$

(v) sensitivity relations:

$$(\bar{\mathbf{p}}(t), -\lambda_0) \in q(t) \partial_{x,z} w(t, \bar{\mathbf{x}}(t), \bar{\mathbf{z}}(t)) \quad \text{for a.e. } t \in [0, T], \quad (31a)$$

$$(\bar{\mathbf{p}}(0), -\lambda_0) \in \partial_{x,z} w(0, x_0, z_0), \quad (31b)$$

$$(\bar{\mathbf{p}}(t_i^\pm), -\lambda_0) \in q(t_i^\pm) \partial_{x,z} w(t_i^\pm, \bar{\mathbf{x}}(t_i), \bar{\mathbf{z}}(t_i)) \quad \forall i = 1, \dots, m-1. \quad (31c)$$

Remark 3.1. In the case when $\lambda_0 \neq 0$, the co-state $\bar{\mathbf{p}}$ and the multipliers $(\lambda_i)_{0 \leq i \leq m}$ can be rescaled, and the statement (i) can be replaced by:

$$\lambda_0 = 1, \quad \lambda_j \geq 0 \quad \text{and} \quad \lambda_j g_j(\bar{\mathbf{x}}(t_j)) = 0 \quad \forall j \geq 1.$$

In the case of a Mayer problem (i.e., when there is no distributed cost, $\ell \equiv 0$), the previous theorem holds with $\bar{z}(t) = \text{const} = z_0$. Furthermore, in this case, we can establish the sensitivity relations in a slightly different way. The precise statement is given in the next theorem.

Theorem 3.2. Assume that (\mathbf{H}_0) - (\mathbf{H}_4) are satisfied with $\underline{\ell} \equiv 0$. Let $x_0 \in \mathbb{R}^n$ be such that $z_0 = \vartheta(0, x_0) < +\infty$, and assume that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{X}_{[0,T]}(x)$ is a solution to the control problem (19). Then, there exist multipliers $(\lambda_i)_{0 \leq i \leq m}$ and a co-state $\bar{\mathbf{p}}(\cdot)$, absolutely continuous on each sub-interval (t_{j-1}, t_j) : $\bar{\mathbf{p}}|_{(t_{j-1}, t_j)} \in W^{1,1}((t_{j-1}, t_j), \mathbb{R}^d)$, and such that the relations (i)-(iii) of Theorem 3.1 hold true. Furthermore, we have the following sensitivity relations.

(iv') *General sensitivity relation:*

$$\left(\mathcal{H}(t, \bar{\mathbf{x}}(t), \bar{\mathbf{p}}(t), -\lambda_0), \bar{\mathbf{p}}(t) \right) \in q(t) \partial_{t,x} w(t, \bar{\mathbf{x}}(t), z_0) \quad \forall t \in I_j, \quad j = 1, \dots, m, \quad (32)$$

where the function $q(t)$ is given by:

$$q(t) := \lambda_0 + \sum_{\hat{j}(t)}^m \lambda_j, \quad \text{with } \hat{j}(\cdot) \text{ defined as in (24);}$$

(v') *sensitivity relations:*

$$\bar{\mathbf{p}}(t) \in q(t) \partial_x w(t, \bar{\mathbf{x}}(t), z_0) \quad \text{for a.e. } t \in [0, T], \quad (33a)$$

$$\bar{\mathbf{p}}(0) \in \partial_x w(0, x_0, z_0), \quad (33b)$$

$$\bar{\mathbf{p}}(t_i^\pm) \in q(t_i^\pm) \partial_x w(t_i^\pm, \bar{\mathbf{x}}(t_i), z_0) \quad \forall i = 1, \dots, m-1. \quad (33c)$$

Theorem 3.2 provides sensitivity relations related the the adjoint state $\bar{\mathbf{p}}$ while Theorem 3.1 expresses the sensitivity relations with $(\bar{\mathbf{p}}, \lambda_0)$ and thus gives an additional information on the multiplier λ_0 . Note also that conditions (31) do not subsume (33) since the assertion $(a, b) \in \partial_{x,z} w(t, x, z)$ does not imply that $a \in \partial_x w(t, x, z)$.

The proof of Theorem 3.1 in the case when $m = 1$ (no intermediate state constraints) will be first given in section 4. The proof of Theorem 3.1 (as well as Theorem 3.2), for the general case $m \geq 1$, will be the focus of sections 5 and 6.

Comments. First of all, let us notice that when there are no state constraints (for instance when $m = 1$ and $g_1(y) \equiv C$, where C is a sufficiently large negative constant), then

$$\vartheta(t, x) - z = w(t, x, z). \quad (34)$$

Moreover, in this case, the transversality condition (27) implies that $\lambda_1 = 0$ and $\lambda_0 = 1$.

Therefore, Theorem 3.1 gives the existence of a co-state $\bar{\mathbf{p}}(\cdot)$, absolutely continuous on $[0, T]$, satisfying

$$\begin{aligned} \dot{\bar{\mathbf{p}}}(s) &\in \partial_x H(s, \bar{\mathbf{x}}(s), \bar{\mathbf{u}}(s), \bar{\mathbf{p}}(s), -1), \quad \text{a.e. } s \in [0, T] \\ \bar{\mathbf{p}}(T) &= \nabla \varphi(\mathbf{x}(T)), \end{aligned}$$

the Pontryagin's Maximum principle (PMP): a.e. $t \in [0, T]$:

$$H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \bar{\mathbf{p}}(t), -1) = \max_{u \in U} H(t, \bar{\mathbf{x}}(t), u, \bar{\mathbf{p}}(t), -1)$$

and the link between PMP and the function ϑ (which can be deduced from (30) and using (34)):

$$\left(\mathcal{H}(t, \bar{\mathbf{x}}(t), \bar{\mathbf{p}}(t), -1), \bar{\mathbf{p}}(t) \right) \in \partial_{t,x} \vartheta(t, \bar{\mathbf{x}}(t)) \quad \text{for every } t \in [0, T].$$

This result is exactly the one already established in [19] (see also [7]).

Theorems 3.1 & 3.2 establish the relation between the PMP and the HJB approach by using the auxiliary value function w and not the value function ϑ . Let us stress on the fact that because the function ϑ may be discontinuous, a relation between the PMP and the HJB approach cannot be expressed by using only ϑ . Indeed, if the value function is not locally Lipschitz then it is not even clear which sub-differential should be used, see [7].

Another feature of Theorems 3.1 & 3.2 is that no qualification constraint is required and so the extremal $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{p}})$ may be abnormal (i.e., $\lambda_0 = 0$). In all cases, the multiplier λ_0 is related to the sensitivity of the auxiliary value function w w.r.t the variable z .

4 The case of Mayer problem with final state constraints

In this section, we prove most of the assertions of Theorem 3.1 in the particular case where the control problem is a Mayer problem with only one end-point constraint (case $m = 1$, $t_m = t_1 = T$, $\ell \equiv 0$), and we furthermore denote $g = g_1$ throughout this section. The problem in this simplified setting is :

$$\begin{aligned} & \text{minimize} && \varphi(\mathbf{x}(T)) \\ & \text{subject to} && g(\mathbf{x}(T)) \leq 0 \quad \text{and} \quad (\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[0,T]}(x) \end{aligned} \quad (35)$$

where g and φ are given locally Lipschitz continuous functions.

In this setting the auxiliary optimal control problem becomes

$$w(t, x, z) = \min_{(\mathbf{x}, \mathbf{u}) \in \mathbb{X}_{[t,T]}(x)} (\varphi(\mathbf{x}(T)) - z) \bigvee g(\mathbf{x}(T)). \quad (36)$$

This problem corresponds to an augmented dynamics where the dynamics for the z variable is zero. We recall that if $z_0 = \inf\{z \in \mathbb{R}, w(0, x, z) \leq 0\}$ is finite, then any optimal trajectory (\mathbf{x}, \mathbf{u}) for the auxiliary OCP associated with the value $w(0, x, z_0)$ is also an optimal solution of the initial problem (20), and $z_0 = \vartheta(0, x)$.

Theorem 4.1. *Assume that (\mathbf{H}_0) - (\mathbf{H}_4) are satisfied. Let $x_0 \in \mathbb{R}^d$ and assume $z_0 := \vartheta(0, x)$ is finite. Let $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{X}_{[0,T]}(x)$ be an optimal solution of the control problem (35). Then there exist multipliers $(\lambda_0, \lambda_1) \in \mathbb{R} \times \mathbb{R}$ and a co-state $\bar{\mathbf{p}} \in W^{1,1}((0, T), \mathbb{R}^d)$ such that:*

- (i) *non-triviality of the multipliers: $\lambda_0 + \lambda_1 = 1$, $\lambda_0, \lambda_1 \geq 0$ and $\lambda_1 g(\bar{\mathbf{x}}(T)) = 0$,*
- (ii) *co-state equation:*

$$\dot{\bar{\mathbf{p}}}(t) \in \partial_x H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \bar{\mathbf{p}}(t), -\lambda_0) \quad \text{a.e. } t \in [0, T], \quad (37a)$$

$$\bar{\mathbf{p}}(T) \in \lambda_0 \partial \varphi(\bar{\mathbf{x}}(T)) + \lambda_1 \partial g(\bar{\mathbf{x}}(T)), \quad (37b)$$

- (iii) *Pontryagin's principle:*

$$H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \bar{\mathbf{p}}(t), -\lambda_0) = \max_{u \in U} H(t, \bar{\mathbf{x}}(t), u, \bar{\mathbf{p}}(t), -\lambda_0) \quad \text{a.e. } t \in [0, T], \quad (37c)$$

- (iv) *link between PMP and the function w :*

$$(\mathcal{H}(t, \bar{\mathbf{x}}(t), \bar{\mathbf{p}}(t), -\lambda_0), \bar{\mathbf{p}}(t), -\lambda_0) \in \partial_{t,x,z} w(t, \bar{\mathbf{x}}(t), z_0) \quad \text{for every } t \in [0, T], \quad (37d)$$

$$(\bar{\mathbf{p}}(0), -\lambda_0) \in \partial_{x,z} w(0, x_0, z_0), \quad (37e)$$

$$(\bar{\mathbf{p}}(T), -\lambda_0) \in \partial_{x,z} w(T, \bar{\mathbf{x}}(T), z_0), \quad (37f)$$

$$(\bar{\mathbf{p}}(t), -\lambda_0) \in \partial_{x,z} w(t, \bar{\mathbf{x}}(t), z_0) \quad \text{a.e. } t \in [0, T]. \quad (37g)$$

Remark 4.1. *In the setting of Mayer problem with final state constraint (and without intermediate state constraints), we notice that the function $q(t)$ in Theorem 3.1-(iv) satisfies $q(t) = \lambda_0 + \lambda_1 \equiv 1$. Therefore the function $q(\cdot)$ does not appear in Theorem 4.1.*

Proof of Theorem 4.1. Let $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ be a solution of the control problem (35), then $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, z_0)$ is an optimal solution for the auxiliary control problem (36), for $t = 0$, $x = x_0$ and $z = z_0$. By using the results of [5], we get the existence of an adjoint state $\mathbf{p} \equiv (\mathbf{p}_x, \mathbf{p}_z) \in W^{1,1}([0, T], \mathbb{R}^{d+1})$ satisfying the PMP:

$$\dot{\mathbf{p}}_x(t) \in \partial_x H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \mathbf{p}_x(t), \mathbf{p}_z(t)), \quad \text{a.e. } t \in [0, T], \quad (38a)$$

$$\dot{\mathbf{p}}_z(t) = 0, \quad \text{a.e. } t \in [0, T], \quad (38b)$$

$$H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \mathbf{p}_x(t), \mathbf{p}_z(t)) = \max_{u \in U} H(t, \bar{\mathbf{x}}(t), u, \mathbf{p}_x(t), \mathbf{p}_z(t)), \quad \text{a.e. } t \in [0, T], \quad (38c)$$

and the following sensitivity relations:

$$(\mathcal{H}(t, \bar{\mathbf{x}}(t), \mathbf{p}_x(t), \mathbf{p}_z(t)), \mathbf{p}_x(t), \mathbf{p}_z(t)) \in \partial_{t,x,z} w(t, \bar{\mathbf{x}}(t), z_0) \quad \text{for a.e. } t \in [0, T], \quad (39a)$$

$$(\mathbf{p}_x(0), \mathbf{p}_z(0)) \in \partial_{x,z} w(0, x_0, z_0), \quad (39b)$$

$$(\mathbf{p}_x(T), \mathbf{p}_z(T)) \in \partial_{x,z} w(T, \bar{\mathbf{x}}(T), z_0), \quad (39c)$$

$$(\mathbf{p}_x(t), \mathbf{p}_z(t)) \in \partial_{x,z} w(t, \bar{\mathbf{x}}(t), z_0) \quad \text{for a.e. } t \in [0, T], \quad (39d)$$

where the Hamiltonian functions H and \mathcal{H} are the same as (23a) with $\ell \equiv 0$. If we set $\bar{\mathbf{p}} := \mathbf{p}_x$, then equations (38a), (38c) lead immediately to (37a) and (37c). From (38b) we deduce also that $\mathbf{p}_z(t) = \text{const}$.

On the other hand, relation (39c) gives

$$\begin{pmatrix} \mathbf{p}_x(T) \\ \mathbf{p}_z(T) \end{pmatrix} \in \partial_{x,z} \left[(\varphi(\bar{\mathbf{x}}(T)) - z_0) \bigvee g(\bar{\mathbf{x}}(T)) \right].$$

By using (11), we have

$$\partial_{x,z} \left[(\varphi(y) - z) \bigvee g(y) \right] \subset \left\{ \lambda_0 \begin{pmatrix} \xi_\varphi \\ -1 \end{pmatrix} + \lambda_1 \begin{pmatrix} \xi_g \\ 0 \end{pmatrix}, \lambda_0, \lambda_1 \geq 0, \lambda_0 + \lambda_1 = 1, \right. \\ \left. \xi_\varphi \in \partial\varphi(y), \xi_g \in \partial g(y) \right\}.$$

Therefore we obtain the existence of a couple (λ_0, λ_1) such that $\lambda_0, \lambda_1 \geq 0$, $\lambda_0 + \lambda_1 = 1$, and

$$\begin{aligned} \mathbf{p}_x(T) &\in \lambda_0 \partial\varphi(\bar{\mathbf{x}}(T)) + \lambda_1 \partial g(\bar{\mathbf{x}}(T)) \\ \mathbf{p}_z(T) &= -\lambda_0. \end{aligned}$$

In the case when $g(\bar{\mathbf{x}}(T)) < 0$, since \bar{z} is the minimal value such that $w(0, x, \bar{z}) = 0$, we have $\varphi(\bar{\mathbf{x}}(T)) - \bar{z} = 0$ and therefore $g(\bar{\mathbf{x}}(T)) < \varphi(\bar{\mathbf{x}}(T)) - \bar{z}$. In that case, $\lambda_1 = 0$ and $\lambda_0 = 1$. We conclude that the nontriviality of the multipliers λ_0, λ_1 .

Now, the general sensitivity relation (39a) leads to:

$$(\mathcal{H}(t, \bar{\mathbf{x}}(t), \bar{\mathbf{p}}(t), -\lambda_0), \bar{\mathbf{p}}(t), -\lambda_0) \in \partial_{t,x,z} w(t, \bar{\mathbf{x}}(t), z_0) \quad \text{for a.e. } t \in [0, T]$$

which is still true for every $t \in [0, T]$, since \mathcal{H} , $\bar{\mathbf{p}}$, $\bar{\mathbf{x}}$ are continuous functions and the generalized gradient has closed graph. The other sensitivity relations (37e)-(37g) can be directly deduced from (39b)-(39d). □

As it can be seen in the proof of Theorem 4.1, $\mathbf{p}_z(t) = \text{const} = -\lambda_0$. Also \mathbf{p}_z represents a sensitivity of the cost function w with respect to the z variable.

The case when $\lambda_0 = 0$ means that there is no sensitivity with respect to the z variable and the cost only depends on the constraint $g(\mathbf{x}(T))$.

5 Sensitivity relations for a control problem with intermediate costs

In order to prove Theorem 3.1, we need to proceed in several steps. The first step aims to establish sensitivity relations for a general control problem where the cost function is in the form of a maximum of intermediate costs at given intermediate times $0 < t_1 < t_2 < \dots < t_m = T$. In this section we consider functions $(\Phi_j)_{1 \leq j \leq m}$ with the following regularity assumption:

(H $_{\Phi}$) For each $j = 1, \dots, m$, $\Phi_j : \mathbb{R}^d \rightarrow \mathbb{R}$ is locally Lipschitz continuous.

Consider the dynamics $F : [0, T] \times \mathbb{R}^N \times U$ satisfying the following assumption: are given functions satisfying the following assumptions:

$$\mathbf{(H_F)} \quad \left\{ \begin{array}{l} (i) \quad F \text{ is continuous on } [0, T] \times \mathbb{R}^N \times U \\ (ii) \quad x \rightarrow F(t, x, u) \text{ is locally Lipschitz continuous in the following sense:} \\ \quad \forall R > 0, \exists k_R \geq 0, \forall (x, y) \in (\mathbb{B}_N(0; R))^2, \forall (t, u) \in [0, T] \times U : \\ \quad \quad \|F(t, x, u) - F(t, y, u)\| \leq k_R \|x - y\| \\ (iii) \quad \exists c_F > 0 \text{ such that} \\ \quad \quad \|F(t, x, u)\| \leq c_F(1 + \|x\|), \quad \forall (t, x, u) \in [0, T] \times \mathbb{R}^N \times U. \end{array} \right.$$

The control problem that is considered here is the following:

$$\begin{aligned} & \text{minimize } \bigvee_{j=1}^m \Phi_j(\mathbf{y}(t_j)) & (40) \\ & \dot{\mathbf{y}}(s) = F(s, \mathbf{y}(s), \mathbf{u}(s)), \text{ a.e. on } [0, T], \\ & \mathbf{y}(0) = \mathbf{y}_0, \\ & \mathbf{u} \in \mathcal{U}. \end{aligned}$$

Note that the above control problem has no state constraint. Assumption **(H $_{\mathbf{F}}$)** guarantee that for every $\mathbf{u} \in \mathcal{U}$ and $(t, x) \in [0, T] \times \mathbb{R}^N$, the existence of an absolutely continuous curve $\mathbf{y} : [t, T] \rightarrow \mathbb{R}^N$ which satisfies the dynamic equation and the initial data $\mathbf{y}(t) = y$. We define the set of all admissible pairs control-and-trajectories starting at time t from the position y as:

$$\mathbb{X}_{[t, T]}^F(y) := \{(\mathbf{y}, \mathbf{u}) \in W^{1,1}(t, T) \times \mathcal{U} \mid (\mathbf{y}, \mathbf{u}) \text{ satisfies } \dot{\mathbf{y}}(s) = F(s, \mathbf{y}(s), \mathbf{u}(s)) \text{ with } \mathbf{y}(t) = y\}.$$

The value function associated to the control problem (40) is defined as follows. For every $x \in \mathbb{R}^d$, for every $t \in [0, T]$ and $j \in \{1, \dots, m\}$ such that $t \in I_j$:

$$\mathcal{V}(t, y) := \min \left\{ \bigvee_{k=j}^m \Phi_j(\mathbf{y}(t_k)), (\mathbf{y}, \mathbf{u}) \in \mathbb{X}_{[t, T]}^F(y) \right\}.$$

By using the same arguments as in the proof of Proposition 3.2, we can show that the value function \mathcal{V} is locally Lipschitz continuous on each $I_j \times \mathbb{R}^d$ ($\forall j$), and that \mathcal{V} is a viscosity solution of the following system of HJB equations:

$$\begin{cases} -\frac{\partial \mathcal{V}}{\partial t}(t, y) + \max_{u \in U} (-F(t, y, u) \cdot D_y \mathcal{V}(t, y)) = 0 & t \in I_j, j = 1, \dots, m, y \in \mathbb{R}^d \\ \mathcal{V}(t_j^-, y) = \mathcal{V}(t_j^+, y) \vee \Phi_j(y) & j = 1, \dots, m-1, y \in \mathbb{R}^d \\ \mathcal{V}(T, y) = \Phi_m(y) & y \in \mathbb{R}^d. \end{cases} \quad (41)$$

Furthermore, if we assume that all (Φ_j) have linear growth (i.e., $\exists c_\Phi > 0$ such that $|\Phi_j(y)| \leq c_\Phi(1 + |y|) \forall y \in \mathbb{R}^N$ and $j = 1, \dots, m$), then \mathcal{V} has also a linear growth on $[0, T] \times \mathbb{R}^d$ and \mathcal{V} is the unique viscosity solution of (41) with linear growth.

Optimality conditions for the problem (40) can be deduced as particular cases of the problems considered in [9] or in [10]. The main result of this section aims at deriving the relation between the co-state associated to an optimal solution and the generalized gradient of the value function along that optimal solution.

It is worth noticing that the sensitivity relations cannot be deduced directly from the existing literature, however the proof is inspired by the arguments in [7, 19, 5]. We give here the complete proof for two main reasons: first, we show how the arguments can be adapted in case of intermediate costs and discontinuities of the value function at times t_i , for $i = 1, \dots, m-1$. Secondly, we give a simplified proof and show that all the sensitivity relations can be obtained by a simple perturbed control problem without using an approximation of the dynamic equation by impulse systems (as required in [7, 5]). In this section, we introduce the Hamiltonian corresponding to problem (40):

$$H^F(t, y, u, p) = -F(t, y, u) \cdot p, \quad \text{for } t \in [0, T], y \in \mathbb{R}^N, u \in U, p \in \mathbb{R}^N.$$

Theorem 5.1. *Assume that the hypothesis (\mathbf{H}_0) , (\mathbf{H}_F) and (\mathbf{H}_Φ) hold. Let (\bar{y}, \bar{u}) be a local minimizer for problem (40). Set*

$$\mathcal{J}(\bar{y}) := \{j \in \{1, \dots, m\} \mid \Phi_j(\bar{y}(t_j)) = \bigvee_{k=1}^m \Phi_k(\bar{y}(t_k))\}. \quad (42)$$

Then there exist $\nu \in \mathbb{R}^m$ and $\mathbf{p}(\cdot) : [0, T] \rightarrow \mathbb{R}^d$, with $\mathbf{p}(\cdot) \in W^{1,1}(I_j, \mathbb{R}^N)$ for every $j = 1, \dots, m$, and such that

(i) Normality:

$$\sum_{j=1}^m \nu_j = 1, \quad \nu_j \geq 0, \quad \text{and } \nu_j = 0 \text{ for } j \notin \mathcal{J}(\bar{y}). \quad (43a)$$

(ii) Adjoint state equation. For every $(t, y, u, p) \in [0, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N$, then we have:

$$\begin{cases} \dot{\mathbf{p}}(t) \in \partial_y H^F(t, \bar{y}(t), \bar{u}(t), \mathbf{p}(t)) & \text{a.e. on } [0, T], \\ \mathbf{p}(t_j^-) - \mathbf{p}(t_j^+) \in \nu_j \partial \Phi_j(\bar{y}(t_j)) & j = 1, \dots, m-1, \\ \mathbf{p}(T) \in \nu_m \partial \Phi_m(\bar{y}(T)). \end{cases} \quad (43b)$$

(iii) *Maximality:*

$$H^F(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{p}(t)) = \max_{u \in U} H^F(t, \bar{\mathbf{y}}(t), u, \mathbf{p}(t)) \quad (43c)$$

(iv) *General sensitivity relation:* $\forall t \in [0, T] \setminus \{t_1, \dots, t_{m-1}\}$

$$(H^F(t, \bar{\mathbf{y}}(t), \bar{\mathbf{u}}(t), \mathbf{p}(t)), \mathbf{p}(t)) \in q(t) \partial \mathcal{V}(t, \bar{\mathbf{y}}(t)), \quad \text{where } q(t) := \sum_{\hat{j}(t)}^m \nu_j; \quad (43d)$$

(v) *Sensitivity relations:* for a.e. $t \in [0, T]$:

$$\mathbf{p}(t) \in q(t) \partial_y \mathcal{V}(t, \bar{\mathbf{y}}(t)), \quad (43e)$$

$$\mathbf{p}(0) \in \partial_y \mathcal{V}(0, \bar{\mathbf{y}}(0)), \quad (43f)$$

$$\mathbf{p}(t_i^\pm) \in q(t_i^\pm) \partial_y \mathcal{V}(t_i^\pm, \bar{\mathbf{y}}(t_i)) \quad \forall i = 1, \dots, m-1, \quad (43g)$$

Remark 5.1. *We have also, analogously to (43g), the relation*

$$\mathbf{p}(T) \in q(T) \partial_x \mathcal{V}(T, \bar{\mathbf{y}}(T)),$$

but since $q(T) = \nu_m$ and $\mathcal{V}(T, x) = \Phi_m(x)$ this leads to the last relation already given in (43b).

Remark 5.2. *The cost function being a maximum of intermediate costs, the sensitivity relation that gives a link between the adjoint arc and the value function holds actually along the optimal trajectory until the final time t_{m_0} that corresponds to the time where the maximal intermediate cost is reached.*

Proof. Let $(\bar{\mathbf{y}}, \bar{\mathbf{u}})$ be a local minimizer of problem (40). The proof of the optimality conditions and sensitivity relations will be split into three steps.

Step 1. For every $j = 1, \dots, m$, define the function $\mathcal{V}_j : [t_{j-1}, t_j] \times \mathbb{R}^N \rightarrow \mathbb{R}$ by:

$$\mathcal{V}_j(t, y) := \min \left\{ \bigvee_{k=j}^m \Phi_k(\mathbf{y}(t_k)), (\mathbf{y}, \mathbf{u}) \in \mathbb{X}_{[t, T]}^F(y) \right\}.$$

Then \mathcal{V}_j is locally Lipschitz on $[t_{j-1}, t_j] \times \mathbb{R}^d \times \mathbb{R}$. Moreover, on the interval $t \in [t_{j-1}, t_j]$, the function $\mathcal{V}(t, \cdot)$ differs from \mathcal{V}_j only at $t = t_{j-1}$. More precisely,

$$\begin{aligned} \mathcal{V}(t, x) &= \mathcal{V}_j(t, x) \text{ for any } t \in I_j, \\ \mathcal{V}(t_{j-1}^-, x) &= \max(\mathcal{V}_j(t_{j-1}, x), \Phi_{j-1}(x)). \end{aligned} \quad (44)$$

For a given $\varepsilon \in]0, 1]$, we define the following subset of \mathbb{R}^{d+1} , for any $s \in [0, T]$:

$$G_{\varepsilon, 0}(s) := \left\{ (\alpha, \beta) \in \partial \mathcal{V}_j(\tau, \xi) \mid \tau \in [t_{j-1}, t_j], (\tau, \xi) \in (s, \bar{\mathbf{y}}(s)) + \varepsilon \mathbb{B}_{N+1} \right\}, \quad (45)$$

where j is such that $s \in I_j$ (recall that \mathbb{B}_{N+1} denotes the closed unit ball of \mathbb{R}^{N+1}). We define also the set:

$$G_{\varepsilon, 1}(s) := \left\{ \beta \in \partial_x \mathcal{V}_j(s, \xi) \mid \xi \in \bar{\mathbf{y}}(s) + \varepsilon \mathbb{B}_N \right\}. \quad (46)$$

For any bounded set S of \mathbb{R}^N , $\partial\mathcal{V}_j$ is bounded on $I_j \times S$, hence $G_{\varepsilon,0}(s)$ and $G_{\varepsilon,1}(s)$ are uniformly bounded sets.

Now, introduce the *support* functions $\sigma_{\varepsilon,0} : [0, T] \times \mathbb{R} \times \mathbb{R}^N$ and $\sigma_{\varepsilon,1} : [0, T] \times \mathbb{R}^N$ defined as:

$$\sigma_{\varepsilon,0}(s, \omega, \theta) := \sup_{(\alpha, \beta) \in G_{\varepsilon,0}(s)} (\alpha, \beta) \cdot (\omega, -\theta) \quad \forall s \in [0, T], \quad \omega \in \mathbb{R}, \quad \text{and } \theta \in \mathbb{R}^N; \quad (47)$$

$$\sigma_{\varepsilon,1}(s, b) := \sup_{\beta \in G_{\varepsilon,1}(s)} \beta \cdot (-b) \quad \forall s \in [0, T], \quad \text{and } b \in \mathbb{R}^N \quad (48)$$

The quantities $G_{\varepsilon,1}(\cdot)$ and $\sigma_{\varepsilon,1}(\cdot)$ will be useful in order to analyze the sensitivity relations w.r.t the state variable.

By similar arguments as in [19, Lemma 3.1], we can show that, for any $j = 1, \dots, m$, the function $\sigma_{\varepsilon,0}$ is upper semi-continuous on $[t_{j-1}, t_j] \times \mathbb{R}^N \times \mathbb{R}$. Moreover, for any $s \in (t_{j-1}, t_j)$, $\sigma_{\varepsilon,0}(s, \cdot, \cdot)$ is continuous on $\mathbb{R} \times \mathbb{R}^N$. On the other hand, the function $\sigma_{\varepsilon,1}$ is bounded on any bounded set of $[0, T] \times \mathbb{R} \times \mathbb{R}^N$. In particular, for any essentially bounded and measurable functions $(\omega, \theta) : [0, T] \rightarrow \mathbb{R} \times \mathbb{R}^d$, the function $\sigma_{\varepsilon,0}(\cdot, \omega(\cdot), \theta(\cdot))$ is integrable. Similarly, for any measurable $\mathbf{b} : [0, T] \rightarrow \mathbb{R}^d$, the function $\sigma_{\varepsilon,1}(\cdot, \mathbf{b}(\cdot))$ is measurable and bounded on $[0, T]$. Therefore we can define the following perturbed optimal control problem:

$$(\mathcal{P}_\varepsilon) : \text{minimize } J_\varepsilon(\mathbf{y}, \zeta, \mathbf{u}, \omega, \theta, \mathbf{b}) = \bigvee_{j=1}^m (\Phi_j(\mathbf{y}(t_j)) + \zeta(t_j)) - \mathcal{V}(0, \mathbf{y}(0)) - \zeta(0) \quad (49a)$$

such that

$$\begin{cases} \dot{\mathbf{y}}(t) = (1 + \omega(t))F(t, \mathbf{y}(t), \mathbf{u}(t)) + \theta(t) + \mathbf{b}(t) & \text{a.e. } t \in (0, T), \\ \dot{\zeta}(t) = \sigma_{\varepsilon,0}(t, \omega(t), \theta(t)) + \sigma_{\varepsilon,1}(t, \mathbf{b}(t)) & \text{a.e. } t \in (0, T), \end{cases} \quad (49b)$$

$$\mathbf{u}(t) \in U, \quad |\omega(t)| \leq \varepsilon, \quad \|\theta(t)\| \leq \varepsilon, \quad \|\mathbf{b}(t)\| \leq \varepsilon \quad \text{a.e. } t \in (0, T), \quad (49c)$$

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

Step 2. For any given $\varepsilon \in]0, 1]$, let us show that $(\bar{\mathbf{y}}(\cdot), \zeta \equiv 0, \bar{\mathbf{u}}, 0, 0, 0)$ is a local minimizer of the problem $(\mathcal{P}_\varepsilon)$. For this, we first remark that $(\bar{\mathbf{y}}(\cdot), \zeta \equiv 0)$ is a solution of (49b) with the control law $(\bar{\mathbf{u}}(\cdot), \omega \equiv 0, \theta \equiv 0, \mathbf{b} \equiv 0)$ and

$$J_\varepsilon(\bar{\mathbf{y}}, 0, \bar{\mathbf{u}}(t), 0, 0, 0) = \bigvee_{j=1}^m \Phi_j(\bar{\mathbf{y}}(t_j)) - \mathcal{V}(0, \bar{\mathbf{y}}(0)) = 0.$$

Now, consider any solution $(\mathbf{y}(\cdot), \zeta(\cdot), \mathbf{u}(\cdot), \omega(\cdot), \theta(\cdot), \mathbf{b}(\cdot))$ of (49b)-(49c) and such that \mathbf{y} lies in an ε -tube around $\bar{\mathbf{y}}$, i.e.,

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

We will show that

$$J_\varepsilon(\mathbf{y}, \zeta, \mathbf{u}, \omega, \theta, \mathbf{b}) \geq 0, \quad (50)$$

which in turn will lead directly to the desired conclusion that is $(\bar{\mathbf{y}}(\cdot), \zeta \equiv 0, \bar{\mathbf{u}}, 0, 0, 0)$ is a local minimizer of the problem $(\mathcal{P}_\varepsilon)$.

As $\mathcal{V}(\cdot, \mathbf{y}(\cdot))$ is a Lipschitz continuous function on I_j (for any j), it is then almost everywhere differentiable on $[0, T]$. Furthermore, we have the following key result:

Lemma 5.1.

$$\frac{d}{ds} \mathcal{V}(s, \mathbf{y}(s)) + \sigma_{\varepsilon,0}(s, \boldsymbol{\omega}(s), \boldsymbol{\theta}(s)) + \sigma_{\varepsilon,1}(s, \mathbf{b}(s)) \geq 0, \quad \text{a.e. on } [0, T]. \quad (51)$$

Proof of Lemma 5.1. The arguments of this proof are similar to the ones introduced in [19]. The difference is that here we establish inequality (51) with respect to the two perturbations $(\boldsymbol{\omega}, \boldsymbol{\theta})$ and \mathbf{b} . This is a key result to derive all sensitivity relations by considering the perturbed control problem $(\mathcal{P}_\varepsilon)$ and without considering additional approximations of the state (as it was done in [7, 5]).

It is sufficient to prove that

$$\frac{d}{ds} \mathcal{V}_j(s, \mathbf{y}(s)) + \sigma_{\varepsilon,0}(s, \boldsymbol{\omega}(s), \boldsymbol{\theta}(s)) + \sigma_{\varepsilon,1}(s, \mathbf{b}(s)) \geq 0, \quad \text{a.e. on } [t_{j-1}, t_j]. \quad (52)$$

For almost every $s \in [t_{j-1}, t_j]$, it holds:

$$\begin{aligned} \frac{d}{ds} \mathcal{V}_j(s, \mathbf{y}(s)) &= \lim_{h \downarrow 0} \frac{1}{h} \left(\mathcal{V}_j(s+h, \mathbf{y}(s+h)) - \mathcal{V}_j(s, \mathbf{y}(s)) \right) \\ &= \lim_{h \downarrow 0} \frac{1}{h} \left(\mathcal{V}_j(s+h, \mathbf{y}(s) + h\dot{\mathbf{y}}(s)) - \mathcal{V}_j(s, \mathbf{y}(s)) \right) \\ &= D^\uparrow \mathcal{V}_j((s, \mathbf{y}(s)); (1, \dot{\mathbf{y}}(s))) \end{aligned} \quad (53)$$

where D^\uparrow is the upper Dini directional derivative. Furthermore (with the notation $\mathbf{F} = F(s, \mathbf{y}(s), \mathbf{u}(s))$) we have

$$\begin{pmatrix} 1 \\ \dot{\mathbf{y}}(s) \end{pmatrix} = \begin{pmatrix} 1 \\ (1 + \boldsymbol{\omega}(s))\mathbf{F} + \boldsymbol{\theta}(s) + \mathbf{b}(s) \end{pmatrix} = (1 + \boldsymbol{\omega}(s)) \begin{pmatrix} 1 \\ \mathbf{F} \end{pmatrix} + \begin{pmatrix} -\boldsymbol{\omega}(s) \\ \boldsymbol{\theta}(s) \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{b}(s) \end{pmatrix}. \quad (54)$$

By using (16) and the relation $D^{\circ,-} \phi(x; \lambda v) = \lambda D^{\circ,-} \phi(x; v)$ (for $\lambda \geq 0$), we get:

$$\begin{aligned} \frac{d}{ds} \mathcal{V}_j(s, \mathbf{y}(s)) = D^\uparrow \mathcal{V}_j((s, \mathbf{y}(s)); (1, \dot{\mathbf{y}}(s))) &\geq (1 + \boldsymbol{\omega}(s)) D^{\circ,-} \mathcal{V}_j((s, \mathbf{y}(s)); (1, F(s, \mathbf{y}(s), \mathbf{u}(s)))) \\ &\quad + D^{\circ,-} \mathcal{V}_j((s, \mathbf{y}(s)); (-\boldsymbol{\omega}(s), \boldsymbol{\theta}(s))) \\ &\quad + D^\uparrow \mathcal{V}_j((s, \mathbf{y}(s)); (0, \mathbf{b}(s))) \end{aligned} \quad (55)$$

From one hand, Lemma 5.2 (at the end of the section) says that for a.e. $s \in I_j$:

$$D^{\circ,-} \mathcal{V}_j((s, \mathbf{y}(s)); (1, F(s, \mathbf{y}(s), \mathbf{u}(s)))) \geq 0. \quad (56)$$

On the other hand, for any given values of $\boldsymbol{\omega}, \boldsymbol{\theta}$, it holds

$$\begin{aligned} -D^{\circ,-} \mathcal{V}_j((s, \mathbf{y}(s)); (-\boldsymbol{\omega}, \boldsymbol{\theta})) &= D^\circ \mathcal{V}_j((s, \mathbf{y}(s)); (\boldsymbol{\omega}, -\boldsymbol{\theta})) \\ &= \max_{(\alpha, \beta) \in \partial \mathcal{V}_j(s, \mathbf{y}(s))} (\alpha \boldsymbol{\omega} - \beta \cdot \boldsymbol{\theta}) \\ &\leq \max_{(\alpha, \beta) \in G_{\varepsilon,0}(s)} (\alpha \boldsymbol{\omega} - \beta \cdot \boldsymbol{\theta}) = \sigma_{\varepsilon,0}(s, \boldsymbol{\omega}, \boldsymbol{\theta}). \end{aligned} \quad (57)$$

We turn now to the third term in the right hand side of (55). Since $\mathbf{y}(s) \in \bar{\mathbf{y}}(s) + \varepsilon\mathbb{B}_N$, by definition of $G_{\varepsilon,1}(s)$ and $\sigma_{\varepsilon,1}(s, \cdot)$, for any $\mathbf{b} \in \mathbb{R}^N$, we have

$$\begin{aligned}
-D^\dagger \mathcal{V}_j((s, \mathbf{y}(s)); (0, \mathbf{b})) &= -\liminf_{h \downarrow 0} \frac{1}{h} (\mathcal{V}_j(s, \mathbf{y}(s) + h\mathbf{b}) - \mathcal{V}_j(s, \mathbf{y}(s))) \\
&= \limsup_{h \downarrow 0} \frac{1}{h} (\mathcal{V}_j(s, \mathbf{y}(s)) - \mathcal{V}_j(s, \mathbf{y}(s) + h\mathbf{b})) \\
&\leq \limsup_{h \downarrow 0, \xi \rightarrow \mathbf{y}(s)} \frac{1}{h} (\mathcal{V}_j(s, \xi - h\mathbf{b}) - \mathcal{V}_j(s, \xi)) \\
&\leq D_x^\circ(\mathcal{V}_j(s, \cdot))(\mathbf{y}(s); -\mathbf{b}) \\
&\leq \max_{\beta \in G_{\varepsilon,1}(s)} \beta \cdot (-\mathbf{b}) = \sigma_{\varepsilon,1}(s, -\mathbf{b})
\end{aligned} \tag{58}$$

Finally, by combining inequalities (55), (56), (57), (58) and using the fact that $1 + \boldsymbol{\omega}(s) \geq 0$, we conclude that (51) is satisfied. \square

Now, by integrating (51) over each interval (t_{j-1}, t_j) , for $j = 1, \dots, m$, we get:

$$0 \leq \mathcal{V}(t_j^-, \mathbf{y}(t_j)) - \mathcal{V}(t_{j-1}^+, \mathbf{y}(t_{j-1})) + \zeta(t_j) - \zeta(t_{j-1}),$$

which combined with (44) gives:

$$\mathcal{V}(t_{j-1}^+, \mathbf{y}(t_{j-1})) + \zeta(t_{j-1}) \leq (\mathcal{V}(t_j^+, \mathbf{y}(t_j)) + \zeta(t_j)) \bigvee (\Phi_j(\mathbf{y}(t_j)) + \zeta(t_j)). \tag{59}$$

Using a recursion on (59) and the fact that $\mathcal{V}(T, \mathbf{y}(T)) = \Phi_m(\mathbf{y}(T))$, we obtain

$$\begin{aligned}
\mathcal{V}(0, \mathbf{y}(0)) + \zeta(0) &\leq (\mathcal{V}(t_1^+, \mathbf{y}(t_1)) + \zeta(t_1)) \bigvee (\Phi_1(\mathbf{y}(t_1)) + \zeta(t_1)) \\
&\leq (\mathcal{V}(t_2^+, \mathbf{y}(t_2)) + \zeta(t_2)) \bigvee (\Phi_2(\mathbf{y}(t_2)) + \zeta(t_2)) \bigvee (\Phi_1(\mathbf{y}(t_1)) + \zeta(t_1)) \\
&\leq \dots \leq \bigvee_{j=1}^m (\Phi_j(\mathbf{y}(t_j)) + \zeta(t_j)).
\end{aligned}$$

This is sufficient to justify that $J_\varepsilon(\mathbf{y}, \zeta, \mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\theta}, \mathbf{b}) \geq 0$ which is the desired claim to prove in this step.

Step 3: Now, we apply the Pontryagin maximum principle to the optimal control problem $(\mathcal{P}_\varepsilon)$ with intermediate costs. Such a problem is a particular case of optimal control problems for multi-processes in [9] and [8] (see also [10]).

Since $(\bar{\mathbf{y}}, \zeta \equiv 0)$ is a local minimizer of the optimal control problem $(\mathcal{P}_\varepsilon)$, then according to [9, 8], there exist $\nu^\varepsilon \in \mathbb{R}^m$ and $(\mathbf{p}_\varepsilon(\cdot), \mathbf{q}_\varepsilon(\cdot))$ absolutely continuous on each sub-interval I_j such

that

$$\left\{ \begin{array}{l}
(i) \quad \nu_j^\varepsilon \geq 0, \quad \sum_{j=1}^m \nu_j^\varepsilon = 1, \quad \text{and} \quad \nu_k^\varepsilon = 0 \quad \text{for } k \notin \mathcal{J}(\bar{\mathbf{y}}) \\
(ii) \quad \begin{pmatrix} \dot{\mathbf{p}}_\varepsilon(s) \\ \dot{\mathbf{q}}_\varepsilon(s) \end{pmatrix} \in \partial_{y,\zeta} H_\varepsilon(s, \bar{\mathbf{y}}(s), \bar{\boldsymbol{\zeta}}(s), \bar{\mathbf{u}}(s), 0, 0, 0, \mathbf{p}_\varepsilon(s), \mathbf{q}_\varepsilon(s)), \quad \text{a.e. on } [0, T] \\
(iii) \quad H_\varepsilon(s, \bar{\mathbf{y}}(s), \bar{\boldsymbol{\zeta}}(s), \bar{\mathbf{u}}(s), 0, 0, 0, \mathbf{p}_\varepsilon(s), \mathbf{q}_\varepsilon(s)) = \max_{\substack{u \in U, \|\theta\| \leq \varepsilon, \\ |\omega| \leq \varepsilon, \|b\| \leq \varepsilon}} H_\varepsilon(s, \bar{\mathbf{y}}(s), \bar{\boldsymbol{\zeta}}(s), u, \theta, \omega, b, \mathbf{p}_\varepsilon(s), \mathbf{q}_\varepsilon(s)) \quad \text{a.e. on } [0, T] \\
(iv) \quad - \begin{pmatrix} \mathbf{p}_\varepsilon(0) \\ \mathbf{q}_\varepsilon(0) \end{pmatrix} \in \partial_{y,\zeta} (-\mathcal{V}(0, \bar{\mathbf{y}}(0)) - \bar{\boldsymbol{\zeta}}(0)) \\
(v) \quad \begin{pmatrix} \mathbf{p}_\varepsilon(T) \\ \mathbf{q}_\varepsilon(T) \end{pmatrix} \in \nu_m^\varepsilon \partial_{y,\zeta} (\Phi_m(\bar{\mathbf{y}}(T)) + \bar{\boldsymbol{\zeta}}(T)) \\
(vi) \quad \begin{pmatrix} \mathbf{p}_\varepsilon(t_j^-) - \mathbf{p}_\varepsilon(t_j^+) \\ \mathbf{q}_\varepsilon(t_j^-) - \mathbf{q}_\varepsilon(t_j^+) \end{pmatrix} \in \nu_j^\varepsilon \partial_{x,\zeta} (\Phi_j(\bar{\mathbf{y}}(t_j)) + \bar{\boldsymbol{\zeta}}(t_j)), \quad j = 1, \dots, m-1
\end{array} \right. \quad (60)$$

where H_ε is defined by:

$$H_\varepsilon(s, y, \zeta, u, \theta, \omega, b, p, q) := -((1 + \omega)F(s, y, u) + \theta + b) \cdot p - (\sigma_{\varepsilon,0}(s, \theta, \omega) + \sigma_{\varepsilon,1}(s, b))q. \quad (61)$$

Analysis of (60)(ii). As H_ε does not depend on ζ , it comes:

$$\dot{\mathbf{q}}_\varepsilon(s) \equiv 0,$$

and

$$\dot{\mathbf{p}}_\varepsilon(s) \in \partial_y H_\varepsilon(s, \bar{\mathbf{y}}(s), \bar{\boldsymbol{\zeta}}(s), \bar{\mathbf{u}}(s), 0, 0, 0, \mathbf{p}_\varepsilon(s), \mathbf{q}_\varepsilon(s)), \quad \text{a.e. on } [0, T]$$

In addition, by definition of H_ε , we can easily check that:

$$\partial_y H_\varepsilon(s, \bar{\mathbf{y}}(s), \bar{\boldsymbol{\zeta}}(s), \bar{\mathbf{u}}(s), 0, 0, 0, \mathbf{p}_\varepsilon(s), \mathbf{q}_\varepsilon(s)) = (1 + \bar{\omega}(s)) \partial_y H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}_\varepsilon(s)).$$

As $\bar{\omega} \equiv 0$, we conclude that:

$$\left\{ \begin{array}{l}
\dot{\mathbf{p}}_\varepsilon \in \partial_y H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}_\varepsilon(s)), \quad \text{a.e. on } [0, T] \\
\dot{\mathbf{q}}_\varepsilon \equiv 0, \quad \text{a.e. on } [0, T].
\end{array} \right. \quad (62)$$

Analysis of assertions (60) (iv), (v) and (vi). The function $(y, \zeta) \mapsto \Phi_m(y) + \zeta$ is a separable sum of a locally Lipschitz function Φ_m and a C^1 function of ζ . Hence (v) can be rewritten

$$\begin{pmatrix} \mathbf{p}_\varepsilon(T) \\ \mathbf{q}_\varepsilon(T) \end{pmatrix} \in \nu_m^\varepsilon \partial \Phi_m(\bar{\mathbf{y}}(T)) \times \{1\},$$

which gives

$$\mathbf{p}_\varepsilon(T) \in \nu_m^\varepsilon \partial \Phi_m(\bar{\mathbf{y}}(T)) \quad \text{and} \quad \mathbf{q}_\varepsilon(T) = \nu_m^\varepsilon. \quad (63)$$

With similar arguments, we get also:

$$\mathbf{p}_\varepsilon(0) \in \partial_y \mathcal{V}(0, \bar{\mathbf{y}}(0)), \quad \mathbf{q}_\varepsilon(0) = 1, \quad (64)$$

as well as

$$\mathbf{p}_\varepsilon(t_j^-) - \mathbf{p}_\varepsilon(t_j^+) \in \nu_j^\varepsilon \partial \Phi_j(\bar{\mathbf{y}}(t_j)), \quad \mathbf{q}_\varepsilon(t_j^-) - \mathbf{q}_\varepsilon(t_j^+) = \nu_j^\varepsilon. \quad (65)$$

It follows from (62), (63) and (65) that $\mathbf{q}(\cdot)$ is piecewise constant, with:

$$\mathbf{q}_\varepsilon(s) = \sum_{k=j}^m \nu_k^\varepsilon \quad \text{for } s \in I_j, \quad j = 1, \dots, m. \quad (66)$$

In particular, $s \rightarrow \mathbf{q}_\varepsilon(s)$ is a decreasing function with $\mathbf{q}_\varepsilon(0) = 1$, $\mathbf{q}_\varepsilon(T) = \nu_m^\varepsilon$, and consequently $0 \leq \mathbf{q}_\varepsilon(s) \leq 1$ for a.e. $s \in [0, T]$.

Analysis of assertion (60)(iii): $\forall u \in U, \forall (\omega, \theta) \in \mathbb{R} \times \mathbb{R}^d$ such that for a.e. $s \in (0, T)$ and for every $|\omega| \leq \varepsilon, \|\theta\| \leq \varepsilon, \|b\| \leq \varepsilon$, we have:

$$-(1 + \omega)F(s, \bar{\mathbf{y}}(s), u) \cdot \mathbf{p}_\varepsilon(s) - \sigma_\varepsilon(s, \theta, \omega)\mathbf{q}_\varepsilon(s) - \sigma_{\varepsilon,1}(s, b)\mathbf{q}_\varepsilon(s) \leq -F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s)) \cdot \mathbf{p}_\varepsilon(s). \quad (67)$$

In particular, taking $\theta = 0, \omega = 0$ and $b = 0$ in (67), it comes:

$$-F(s, \bar{\mathbf{y}}(s), u) \cdot \mathbf{p}_\varepsilon(s) \leq -F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s)) \cdot \mathbf{p}_\varepsilon(s) \quad \forall u \in U,$$

that is,

$$H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}_\varepsilon(s)) = \max_{u \in U} H^F(s, \bar{\mathbf{y}}(s), u, \mathbf{p}_\varepsilon(s)) \quad \text{a.e. on } (0, T).$$

On the other hand, with $u = \bar{\mathbf{u}}(s)$ in (67), for every $(\omega, \theta) \in \mathbb{R}^{d+1}$ such that $|\omega| \leq \varepsilon, \|\theta\| \leq \varepsilon$ and $\|b\| \leq \varepsilon$:

$$\begin{aligned} & -(1 + \omega)F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s)) \cdot \mathbf{p}_\varepsilon(s) - \theta \cdot \mathbf{p}_\varepsilon(s) - b \cdot \mathbf{p}_\varepsilon(s) - \sigma_{\varepsilon,0}(s, \theta, \omega)\mathbf{q}_\varepsilon(s) - \sigma_{\varepsilon,1}(s, b)\mathbf{q}_\varepsilon(s) \\ & \leq -F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s)) \cdot \mathbf{p}_\varepsilon(s). \end{aligned} \quad (68)$$

Hence, for $b = 0$:

$$-\omega F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s)) \cdot \mathbf{p}_\varepsilon(s) - \theta \cdot \mathbf{p}_\varepsilon(s) - \sigma_{\varepsilon,0}(s, \theta, \omega)\mathbf{q}_\varepsilon(s) \leq 0$$

and

$$\omega H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}_\varepsilon(s)) - \theta \cdot \mathbf{p}_\varepsilon(s) \leq \mathbf{q}_\varepsilon(s) \sigma_{\varepsilon,0}(s, \theta, \omega) = \mathbf{q}_\varepsilon(s) \left(\max_{(\alpha, \beta) \in G_{\varepsilon,0}(s)} (\alpha, \beta) \cdot (\omega, -\theta) \right).$$

Therefore, for all $|\omega| \leq \varepsilon$ and $\|\theta\| \leq \varepsilon$,

$$(\omega, -\theta) \cdot (H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}_\varepsilon(s)), \mathbf{p}_\varepsilon(s)) \leq \mathbf{q}_\varepsilon(s) \left(\max_{(\alpha, \beta) \in G_{\varepsilon,0}(s)} (\alpha, \beta) \cdot (\omega, -\theta) \right).$$

By a separating hyperplane argument as in [19], since $\mathbf{q}_\varepsilon(s) \geq 0$, this last inequality implies:

$$(H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}_\varepsilon(s)), \mathbf{p}_\varepsilon(s)) \in \mathbf{q}_\varepsilon(s) \overline{\text{co}}G_{\varepsilon,0}(s) \quad \text{a.e. } s \in [0, T]. \quad (69)$$

On the other hand, if we set $\theta = 0$ and $\omega = 0$ in (68), with $\|b\| \leq \varepsilon$, we obtain

$$-b \cdot \mathbf{p}_\varepsilon(s) - \sigma_{\varepsilon,1}(s, b)\mathbf{q}_\varepsilon(s) \leq 0.$$

We then deduce that

$$\mathbf{p}_\varepsilon(s) \in \mathbf{q}_\varepsilon(s) \overline{\text{co}} G_{\varepsilon,1}(s), \quad \text{a.e. } s \in [0, T]. \quad (70)$$

The rest of the proof follows the same arguments as in [19]. All the previous properties were established for a given $\varepsilon > 0$. Now let us consider a sequence $\varepsilon_i \rightarrow 0$ and denote by $\mathbf{p}^{\varepsilon_i}(\cdot)$ the corresponding adjoint states. One can find a subsequence such that $\mathbf{p}^{\varepsilon_i}(\cdot)$ converges uniformly on each sub-interval I_j to an absolutely continuous function $\mathbf{p}(\cdot)$, and ν^{ε_i} converges to some $\nu \in \mathbb{R}^m$ satisfying the normality condition. It is clear that the limit function $\mathbf{p}(\cdot)$ satisfies

$$\dot{\mathbf{p}}(s) \in \partial_y H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}(s)) \quad \text{a.e. on } [0, T]$$

and

$$-\mathbf{p}(s) \cdot F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s)) = \max_{u \in U} (-F(s, \bar{\mathbf{y}}(s), u) \cdot \mathbf{p}(s)), \quad \text{a.e. on } [0, T].$$

Moreover, all boundary conditions are preserved:

$$\begin{cases} -\mathbf{p}(0) \in -\partial_y \mathcal{V}(0, \bar{\mathbf{y}}(0)), \\ \mathbf{p}(t_j^-) - \mathbf{p}(t_j^+) \in \nu_j \partial \Phi_j(\bar{\mathbf{y}}(t_j)) \quad \text{for } j = 1, \dots, m-1, \\ \mathbf{p}(T) \in \nu_m \partial \Phi_m(\bar{\mathbf{y}}(T)). \end{cases} \quad (71)$$

Finally, by passing to the limit in (69), one has $\forall j \in \{1, \dots, m\}$, for almost every $s \in J_j$:

$$\begin{aligned} & (H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}(s)), \mathbf{p}(s)) \\ & \in \mathbf{q}(s) \bigcap_{\varepsilon > 0} \overline{\text{co}} \left\{ \partial \mathcal{V}_j(\tau, y) \mid (\tau, y) \in (s, \bar{\mathbf{y}}(s)) + \varepsilon \mathbb{B}_{d+1} \cap (I_j \times \mathbb{R}^d) \right\} \equiv \mathbf{q}(s) \partial \mathcal{V}_j(s, \bar{\mathbf{y}}(s)). \end{aligned}$$

From this relationship, using the same argument as in [19], one deduces that

$$(H^F(s, \bar{\mathbf{y}}(s), \bar{\mathbf{u}}(s), \mathbf{p}(s)), \mathbf{p}(s)) \in \mathbf{q}(s) \partial \mathcal{V}_j(s, \bar{\mathbf{y}}(s)) \quad \forall s \in (t_{j-1}, t_j).$$

Also, by passing to the limit in (70), we obtain for a.e. $s \in I_j$:

$$\mathbf{p}(s) \in \mathbf{q}(s) \bigcap_{\varepsilon > 0} \overline{\text{co}} \left\{ \partial_y \mathcal{V}_j(s, y) \mid y \in \bar{\mathbf{y}}(s) + \varepsilon \mathbb{B}_N \right\} \equiv \mathbf{q}(s) \partial_y \mathcal{V}_j(s, \bar{\mathbf{y}}(s)). \quad (72)$$

This concludes the proof. □

The proof of Lemma 5.1 is based on the following result whose proof comes directly from the dynamic principle of the value function \mathcal{V}_j (a sketch of the proof is given below for sake of completeness).

Lemma 5.2. *For any $(s, y, u) \in [t_{j-1}, t_j] \times \mathbb{R}^d \times U$,*

$$D^{\circ, -} \mathcal{V}_j((s, y), (1, F(s, y, u))) := \liminf_{h \downarrow 0, s' \rightarrow s, y' \rightarrow y} \frac{1}{h} \left(\mathcal{V}_j(s' + h, y' + hF(s, y, u)) - \mathcal{V}_j(s', y') \right) \geq 0.$$

Proof of Lemma 5.2. Consider the constant admissible control $\mathbf{u}(\theta) \equiv u$ and let $\mathbf{y}_{s,y}^u(\cdot)$ denote the absolutely continuous solution of $\dot{\mathbf{y}}(\theta) = F(\theta, \mathbf{y}(\theta), u)$ a.e. $\theta \in [0, T]$, with $\mathbf{y}(s) = y$.

Let $\varepsilon > 0$. By assumption (\mathbf{H}_F) on the dynamics F , and by standard arguments, we can show that, for h sufficiently small and for any (s', y') close enough to (s, y) , we have:

$$\|\mathbf{y}_{s',y'}^u(s' + h) - (y' + hF(s, y, u))\| \leq h\varepsilon.$$

Moreover, since \mathcal{V}_j is locally Lipschitz continuous on $[t_{j-1}, t_j] \times \mathbb{R}^N$, there exists $L > 0$ such that:

$$\mathcal{V}_j(s' + h, y' + hF(s, y, u)) \geq \mathcal{V}_j(s' + h, \mathbf{y}_{s',y'}^u(s' + h)) - Lh\varepsilon.$$

This inequality with the dynamic programming principle leads to:

$$\mathcal{V}_j(s' + h, y' + hF(s, y, u)) \geq \mathcal{V}_j(s' + h, \mathbf{y}_{s',y'}^u(s' + h)) - Lh\varepsilon \geq \mathcal{V}_j(s', y') - Lh\varepsilon.$$

Therefore,

$$\liminf_{h \downarrow 0, s' \rightarrow s, y' \rightarrow y} \frac{1}{h} (\mathcal{V}_j(s' + h, y' + hF(s, y, u)) - \mathcal{V}_j(s', y')) \geq -2L\varepsilon,$$

and since the result holds for any $\varepsilon > 0$, we deduce the desired inequality. \square

6 Proof of the main results (Theorems 3.1 and 3.2)

Proof of of Theorem 3.1. Let $x_0 \in \mathbb{R}^n$ be such that $z_0 = \vartheta(0, x_0) < +\infty$, and assume that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{X}_{[0,T]}(x)$ is a solution to the control problem (19). Let $\bar{\mathbf{z}}(\cdot)$ be defined by:

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

By proposition 3.1, we know that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{z}})$ is also an optimal solution to the augmented control problem associated to the value $w(0, x, z)$ (as defined in (22)). By using the result of Theorem 5.1, we get the existence of $\nu \in \mathbb{R}^m$ and $\mathbf{p}(\cdot) = (\mathbf{p}_x(\cdot), \mathbf{p}_z(\cdot))$ absolutely continuous on each sub-interval $I_j: \mathbf{p}_x|_{(t_{j-1}, t_j)} \in W^{1,1}([t_{j-1}, t_j], \mathbb{R}^d)$ and $\mathbf{p}_z|_{(t_{j-1}, t_j)} \in W^{1,1}([t_{j-1}, t_j], \mathbb{R})$, such that

- (i) $\sum_{j=1}^m \nu_j = 1$, $\nu_j \geq 0$ ($\forall j$), and $j \notin \mathcal{J}(\bar{\mathbf{y}}) \Rightarrow \nu_j = 0$ ($\forall j$)
- (ii) $\begin{cases} \dot{\mathbf{p}}_x(t) \in \partial_x H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \mathbf{p}_x(t), \mathbf{p}_z(t)) \\ \dot{\mathbf{p}}_z(t) = 0 \end{cases}$ a.e. on $[0, T]$
- (iii) $\begin{cases} \mathbf{p}_x(t_j^-) - \mathbf{p}_x(t_j^+) \in \nu_j \partial_x g_j(\bar{\mathbf{x}}(t_j)) & j = 1, \dots, m-1 \\ \mathbf{p}_z(t_j^-) - \mathbf{p}_z(t_j^+) = 0 & j = 1, \dots, m-1 \\ (\mathbf{p}_x(T), \mathbf{p}_z(T)) \in \nu_m \partial_{x,z} (g_m(\bar{\mathbf{x}}(T)) \vee (\varphi(\bar{\mathbf{x}}(T)) - \bar{\mathbf{z}}(T))) \end{cases}$
- (iv) $H(t, \bar{\mathbf{x}}(t), \bar{\mathbf{u}}(t), \mathbf{p}_x(t), \mathbf{p}_z(t)) = \max_{u \in U} H(t, \bar{\mathbf{x}}(t), u, (\mathbf{p}_x(t), \mathbf{p}_z(t)))$
- (v) $(\mathcal{H}(t, \bar{\mathbf{x}}(t), \mathbf{p}_x(t), \mathbf{p}_z(t)), \mathbf{p}_x(t), \mathbf{p}_z(t)) \in q(t) \partial_{(t,x,z)} w(t, \bar{\mathbf{x}}(t), \mathbf{z}(t))$,
 $\forall t \in]t_{j-1}, t_j[, \forall j = 1, \dots, m$,
 $(\mathbf{p}_x(0), \mathbf{p}_z(0)) \in \partial_{x,z} w(0, \bar{\mathbf{x}}(0), \bar{\mathbf{z}}(0))$,
 $(\mathbf{p}_x(t), \mathbf{p}_z(t)) \in q(t) \partial_{x,z} w(t, \bar{\mathbf{x}}(t), \bar{\mathbf{z}}(t))$, a.e. $t \in (0, T)$.

From (ii) we deduce that $\mathbf{p}_z(t)$ is constant on all intervals (t_{j-1}, t_j) , and since $\mathbf{p}_z(t_i^-) = \mathbf{p}_z(t_i^+)$ by (iii), we deduce that $\mathbf{p}_z(t)$ is constant on $[0, T]$.

By using the sub-differential calculus rule for the maximum of two functions, we conclude that there exist $\alpha_0 \geq 0, \alpha_1 \geq 0$ such that $\alpha_0 + \alpha_1 = 1$ and

$$\begin{pmatrix} \mathbf{p}_x(T) \\ \mathbf{p}_z(T) \end{pmatrix} \in \nu_m \begin{pmatrix} \alpha_0 \partial \varphi(\bar{\mathbf{x}}(T)) + \alpha_1 \partial g_m(\bar{\mathbf{x}}(T)) \\ -\alpha_0 \end{pmatrix}$$

with $\alpha_1 g_m(\bar{\mathbf{x}}(T)) = 0$, and $\mathbf{p}_z(T) = -\alpha_0 \nu_m$. Finally, if we set

$$\begin{cases} \lambda_0 := \nu_m \alpha_0, \\ \lambda_j := \nu_j, \quad \text{for } j = 1, \dots, m-1, \\ \lambda_m := \nu_m \alpha_1, \end{cases} \quad \text{and } \bar{\mathbf{p}} \equiv \mathbf{p}_x,$$

then all the assertions of Theorem 3.1 follow directly from the above optimality conditions. \square

Proof of Theorem 3.2. Recall here that we assume $\ell \equiv 0$. We apply Theorem 5.1 for the optimal solution $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$, with fixed parameter $z_0 = \vartheta(0, x_0)$, the optimal control problem $(t, x) \rightarrow w(t, x, z_0)$ where w is defined as usual by (22), considering the variable $x \in \mathbb{R}^d$, the dynamics $F(t, x, u) := f(t, x, u)$, and the functions $\Phi_j(x) := g_j(x)$ for $j = 1, \dots, m-1$ and $\Phi_m(x) := \max(g_m(x), \varphi(x) - z_0)$. The proof then follows the same lines as for the proof of Theorem 3.1. \square

References

- [1] A. Altarovici, O. Bokanowski, and H. Zidani. A general Hamilton-Jacobi framework for non-linear state-constrained control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(02):337–357, 2013.
- [2] J-P. Aubin and A. Cellina. *Differential inclusions. Set-valued maps and viability theory*, volume 264 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1984.
- [3] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems and Control: Foundations and Applications. Birkhäuser, Boston, 1997.
- [4] L. D. Berkovitz. *Optimal control theory*. Springer, New York, 1974.
- [5] P. Bettiol, H. Frankowska, and R. B. Vinter. Improved sensitivity relations in state constrained optimal control. *Appl. Math. Optim.*, 71(2):353–377, 2015.
- [6] F. H. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. 2013.
- [7] F. H. Clarke and R. B. Vinter. The relationship between the maximum principle and dynamic programming. *SIAM J. Control Optim.*, 25(5):1291–1311, 1987.

- [8] F. H. Clarke and R. B. Vinter. Applications of optimal multiprocesses. *SIAM Journal on Control and Optimization*, 27(5):1048–1071, 1989.
- [9] F. H. Clarke and R. B. Vinter. Optimal multiprocesses. *SIAM Journal on Control and Optimization*, 27(5):1072–1091, 1989.
- [10] A. V. Dmitruk and A. M. Kaganovich. Maximum principle for optimal control problems with intermediate constraints. *Comput. Math. Model.*, 22(2):180–215, 2011. Translation of Nelineĭnaya Din. Upr. No. 6 (2008), 101–136.
- [11] W. H. Fleming and R. W. Rischel. *Deterministic and stochastic optimal control*. Springer, New York, 1975.
- [12] H. Frankowska and M. Mazzola. On relations of the adjoint state to the value function for optimal control problems with state constraints. *Nonlinear Differential Equations and Applications NoDEA*, 20:361–383, 2013.
- [13] H. Frankowska and R. B. Vinter. Existence of neighboring feasible trajectories: Applications to dynamic programming for state-constrained optimal control problems. *Journal of Optimization Theory and Applications*, 104(1):20–40, 2000.
- [14] C. Hermosilla, R. B. Vinter, and H. Zidani. Hamilton-Jacobi-Bellman equations for optimal control processes with convex state constraints. *Systems Control Lett.*, 109:30–36, 2017.
- [15] C. Hermosilla, P. Wolenski, and H. Zidani. The mayer and minimum time problems with stratified state constraints. *Set-Valued and Variational Analysis*, 26:643662, 2018.
- [16] C. Hermosilla and H. Zidani. Infinite horizon problems on stratifiable state-constraints sets. *Journal of Differential Equations*, 258(4):1420–1460, 2015.
- [17] L.S. Pontryagin, V.G. Boltjanskii, R. V. Gramkrelidze, and E.F. Mishchenko. *Optimal control theory*. Springer, New York, 1974.
- [18] H. Soner. Optimal control with state-space constraint I. *SIAM Journal on Control and Optimization*, 24(3):552–561, 1986.
- [19] R. B. Vinter. New results on the relationship between dynamic programming and the maximum principle. *Math. Control Signals Systems*, 1(1):97–105, 1988.
- [20] R. B. Vinter. *Optimal control*. Birkhäuser, Boston, 2000.
- [21] X. Y. Zhou. Maximum principle, dynamic programming, and their connection in deterministic control. *Journal of Optimization Theory and Applications*, 65(2):363–373, 1990.