SIAM J. CONTROL OPTIM.
Vol. 60, No. 3, pp. 1436–1462

UNBOUNDED CONTROL, INFIMUM GAPS, AND HIGHER ORDER NORMALITY*

MONICA MOTTA†, MICHELE PALLADINO‡, AND FRANCO RAMPAZZO†

Abstract. In optimal control theory one sometimes extends the minimization domain of a given problem, with the aim of achieving the existence of an optimal control. However, this issue is naturally confronted with the possibility of a gap between the original infimum value and the extended one. Avoiding this phenomenon is not a trivial issue, especially when the trajectories are subject to endpoint constraints. However, since the seminal works by Warga, some authors have recognized "normality" of an extended minimizer as a condition guaranteeing the absence of an infimum gap. Yet, normality is far from being necessary for this goal, a fact that makes the search for weaker assumptions a reasonable aim. In relation to a control-affine system with unbounded controls, in this paper we prove a sufficient no-gap condition based on a notion of *higher order normality*, which is less demanding than the standard normality and involves iterated Lie brackets of the vector fields defining the dynamics.

Key words. optimal control, infimum gap, impulsive control, high order conditions, lie brackets

AMS subject classifications. 49K15, 49N25, 93C10

DOI. 10.1137/21M1431692

1. Introduction.

Let an endpoint constrained optimal control problem be given, together with an extension of the minimization domain. With the expression (*local*) *infimum gap* one usually refers to the possibility that the cost of some extended sense control-trajectory pair is strictly smaller than the local infimum cost of the original problem. The occurrence of a gap is often independent of the cost functional itself and is rather a purely dynamical phenomenon. Indeed, given a local minimizing control-trajectory pair for the extended optimization problem, the occurrence of a gap is due to the fact that the endpoint constraint is locally separated, at the final point of the minimizing trajectory, from the *reachable set*, which we define as the set of the endpoints of strict sense trajectories close to the given trajectory.

Starting with Warga's pioneering works [34, 35, 36, 37], several papers have shown that the occurrence of an infimum gap is sufficient for an extended sense control-trajectory pair to be an abnormal extremal [22, 25, 26, 28, 29, 27, 12, 13, 14]. (Let us recall that an extremal is called *abnormal* provided the corresponding cost multiplier in the Maximum Principle can be chosen equal to zero, and *normal* otherwise.) In particular, in [26] a generalization of Warga's criterion to a vast class of endpoint-constrained minimum problems' extensions has been recently achieved through the combined use of the notion of *abundance* (see [17, 36, 37]) and of a suitable set separation theorem.

However, the contrapositive statement that normality guarantees the absence of

*Received by the editors July 6, 2021; accepted for publication February 14, 2022; published electronically May 26, 2022.
https://doi.org/10.1137/21M1431692
Funding: This research was partially supported by the INdAM-GNAMPA Project 2020 "Extended Control Problems: Gap, Higher Order Conditions and Lyapunov Functions" and by the Padua University grant SID 2018 "Controllability, Stabilizability and Infimum Gaps for Control Systems," BIRD 187147.
†Department of Mathematics "Tullio Levi-Civita," Padua University, 35121 Padua, Italy (motta@math.unipd.it, rampazzo@math.unipd.it).
‡Department of Information Engineering, Computer Science and Mathematics (DISIM), University of L'Aquila, via Vetoio, 67100 L'Aquila, Italy (michele.palladino@univaq.it).

a gap is far from being a necessary condition. Let us point out two aspects of the question. On the one hand, in [26] the occurrence of a gap is shown to imply the linear separability of an approximating cone to the original reachable set from any approximating cone to the endpoint constraint (see subsection 2.3 for the concept of *approximating cone*). This linear separability translates into the abnormality of the minimizer. On the other hand, in the first order Maximum Principle the approximating cones to the reachable set are built as convex hulls of the so-called needle variations, while several higher order Maximum Principles are basically obtained by enlarging these first order approximating cones by means of new *higher order* (in the time-scale) variations involving iterated Lie brackets (see, e.g., [2, 3, 10, 19, 18, 30, 6, 7, 33] and references therein).

So, the question arises *whether we can prove a gap-abnormality relation involving iterated Lie brackets.* In this paper we provide a positive answer to this question in the case of an optimal unbounded control problem and its *impulsive* extension. Our main achievement consists of a new gap-abnormality relation, which, besides generalizing the known ones (see [22, 12]), involves a notion of *higher order abnormality.*

Let us introduce the minimum problem and its extension. The state space is a Riemannian differential manifold $(\mathscr{M}, \langle\cdot,\cdot\rangle)$, and the (space-time) endpoint constraint, or *target*, is a closed subset $\mathfrak{T} \subseteq \mathbb{R}_+ \times \mathscr{M}$. We consider a *cost function* $\Psi : \mathbb{R}_+ \times \mathscr{M} \to \mathbb{R}$ and an *energy upper bound* $K > 0$ (possibly $= +\infty$). The set $\mathscr{U}$ of (unbounded) *strict sense controls* is defined as

$$\mathscr{U} := \bigcup_{T>0} \left(\{T\} \times L^1\big([0,T], \mathscr{C} \times A\big)\right),$$

where $\mathscr{C} \subseteq \mathbb{R}^m$ is a closed cone and $A \subset \mathbb{R}^q$ is a compact subset. For any strict sense control $(T, u, a) \in \mathscr{U}$, we call (T, u, a, x, v) a *strict sense process* if (x, v) is the (unique) Carathéodory solution on $[0, T]$ to the Cauchy problem

$$\begin{cases} \dfrac{dx}{dt}(t) = f(x(t), a(t)) + \displaystyle\sum_{i=1}^{m} g_i(x(t))u^i(t), \\ \dfrac{dv}{dt}(t) = |u(t)|, \\ (x, v)(0) = (\check{x}, 0), \end{cases} \tag{1.1}$$

where, for every $a \in A$, $f(\cdot, a)$, $g_1, \ldots, g_m$ are given vector fields. Furthermore, a strict sense process (T, u, a, x, v) is said to be *feasible* if $(T, x(T)) \in \mathfrak{T}$ and $v(T) \leq K$. Incidentally, notice that for every $t \in [0, T]$, $v(t)$ coincides with the L^1 norm of the control u on $[0, t]$. The *original optimal control problem* is defined as

$$\text{(P)} \qquad \inf\Big\{\Psi(T, x(T)), \quad (T, u, a, x, v) \ \text{ feasible strict sense process}\Big\}.$$

Since the controls u are unbounded and no growth conditions avoid the occurrence of minimizing sequences of strict sense trajectories which converge to discontinuous paths, following the *graph completion* approach (see, e.g., [9, 21, 20, 8, 15, 16, 5, 38, 4, 23]), we embed problem (P) into the *extended optimal control problem*

$$\text{(P}_\text{e}\text{)} \quad \inf\Big\{\Psi(y^0(S), y(S)), \quad (S, w^0, w, \alpha, y^0, y, \beta) \ \text{ feasible extended sense process}\Big\},$$

where $(S, w^0, w, \alpha, y^0, y, \beta)$ is an *extended sense process* when (S, w^0, w, α) belongs to the set $\mathscr{W}$ of *extended sense controls*, defined as

$$\mathscr{W} := \bigcup_{S>0} \Big(\{S\} \times \Big\{ (w^0, w, \alpha) \in L^\infty([0,S], \mathbb{R}_+ \times \mathscr{C} \times A) : \, w^0 + |w| = 1 \text{ almost everywhere (a.e.)} \Big\} \Big),$$

and (y^0, y, β) is the unique solution on $[0, S]$ to

$$\begin{cases} \dfrac{dy^0}{ds}(s) = w^0(s), \\ \dfrac{dy}{ds}(s) = f(y(s), \alpha(s)) w^0(s) + \displaystyle\sum_{i=1}^m g_i(y(s)) w^i(s), \\ \dfrac{d\beta}{ds}(s) = |w(s)|, \\ (y^0, y, \beta)(0) = (0, \check{x}, 0). \end{cases} \tag{1.2}$$

As in the original problem, $(S, w^0, w, \alpha, y^0, y, \beta)$ is said to be *feasible* provided $(y^0, y, \beta)(S) \in \mathfrak{T} \times [0, K]$. Notice that by regarding y^0 as a time reparameterization, the class of strict sense processes can be identified with the subclass of extended sense processes with $w^0 > 0$ a.e. The elements of this subclass are referred to as *embedded strict sense processes*. Precisely, given a strict sense process (T, u, a, x, v), through the time change $y^0 := \sigma^{-1}$, where $[0, T] \ni t \mapsto \sigma(t) := t + v(t)$, we define the associated embedded strict sense process as follows:

$$(S, w^0, w, \alpha, y^0, y, \beta) := \left(\sigma(T), \frac{dy^0}{ds}, (u \circ y^0) \cdot \frac{dy^0}{ds}, a \circ y^0, y^0, x \circ y^0, v \circ y^0 \right). \tag{1.3}$$

Conversely, given any embedded strict sense process $(S, w^0, w, \alpha, y^0, y, \beta)$, we get the corresponding strict sense process (T, u, a, x, v) using the time change $\sigma := (y^0)^{-1}$. This is a one-to-one relation from the set of strict sense processes to the subset of embedded strict sense processes, in which clearly (T, u, a, x, v) is feasible if and only if the associated process $(S, w^0, w, \alpha, y^0, y, \beta)$ is feasible, and $\Psi(T, x(T)) = \Psi(y^0(S), y(S))$. The *impulsive extension* consists of allowing the time derivative w^0 to vanish on a set of positive measure. In particular, if $I \subseteq [0, S]$ is an interval on which w^0 vanishes, at time $t := y^0(I)$ the state y *evolves instantaneously*. (Alternatively, one can provide an equivalent t-based description of this extension using bounded variation trajectories, as done in [16, 5, 4, 23, 24].)

To state our main result, let us introduce the notion of a local infimum gap. We say that at a feasible extended sense process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ there is *a local infimum gap* if there exists some $\delta > 0$ such that

$$\Psi\big(\bar{y}^0(\bar{S}), \bar{y}(\bar{S})\big) < \inf \Psi(y^0(S), y(S)),$$

where the infimum is taken over the set of feasible embedded strict sense processes $(S, w^0, w, \alpha, y^0, y, \beta)$ such that $\mathbf{d}\big((S, y^0, y, \beta), (\bar{S}, \bar{y}^0, \bar{y}, \bar{\beta})\big) < \delta$, the distance $\mathbf{d}$ being defined in Definition 2.4 below.

We also need to specify the concept of *higher order abnormality*, in connection with a higher order Maximum Principle for the extend problem recently obtained in [3]. Let us define the Hamiltonian H by setting

$$H(x, p, p_0, \pi, w^0, w, a) := p_0 w^0 + p \cdot \left(f(x, a) w^0 + \sum_{i=1}^m g_i(x) w^i \right) + \pi |w|.$$

For simplicity, in this introduction we consider the special case where $\mathscr{M} = \mathbb{R}^n$, $\mathscr{C} = \mathbb{R}^m$, and the vector fields $f(\cdot, a)$, $g_1, \dots, g_m$ are of class C^∞ (see Definition 3.5 for the general case). We say that a feasible extended sense process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a (first order) *extremal* of the extended problem ($\mathrm{P_e}$) if there exists a nontrivial four-tuple (p_0, p, π, λ) (with $p_0 \in \mathbb{R}$, $p = p(\cdot)$ absolutely continuous, and $(\pi, \lambda) \in \mathbb{R}_- \times \mathbb{R}_+$ denoting the multiplier of the energy variable β and the cost multiplier, respectively) that verifies the usual conditions of the Maximum Principle, namely, the H-maximality, adjoint equation, and transversality (w.r.t. a given approximating cone $\mathscr{K}$ to the target $\mathfrak{T}$ at $(\bar{y}^0, \bar{y})(\bar{S})$). As is customary, a (first order) extremal is said to be *normal* if for all choices of the multipliers (p_0, p, π, λ), the cost multiplier λ is different from zero, while the extremal is said to be *abnormal* in the opposite situation. If, in addition, $\bar{\beta}(\bar{S}) < K$ and the *higher order condition*

$$\begin{cases} p(s) \cdot B(\bar{y}(s)) = 0 & \forall s \in [0, \bar{S}], \\ p(s) \cdot \left[f_{\bar{\alpha}(s)}, B\right](\bar{y}(s))\, \bar{w}^0(s) = 0 & \text{for a.e. } s \in [0, \bar{S}] \end{cases}$$

(where $f_a(\cdot) := f(\cdot, a)$) is satisfied for every iterated Lie bracket B of $g_1, \dots, g_m$, we say that $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a *higher order extremal*. Finally, a higher order extremal is said to be *normal* if for any choice of multipliers (p_0, p, π, λ), one has $\lambda \neq 0$, while it is said to be *abnormal* when it is not normal. Referring the reader to section 3 for details and a precise statement, we present here the main result of the paper in a simplified form.

THEOREM. *If $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a feasible extended sense process at which there is a local infimum gap and $\bar{\beta}(\bar{S}) < K$, then, for any approximating cone to the target at $(\bar{y}^0(\bar{S}), \bar{y}(\bar{S}))$, $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a higher order abnormal extremal* (see Theorem 3.1).

Let us point out that, as illustrated by the example in section 5, a higher order extremal might be an abnormal first order extremal while being a normal higher order extremal. Therefore, Theorem 3.1 actually provides a new weaker sufficient condition for the avoidance of an infimum gap.

Let us conclude this introduction with a couple of comments. First, in the present paper we adopt the concept of *Quasi Differential Quotient approximating cone* (QDQ *approximating cone*). A *Quasi Differential Quotient* is a notion of generalized differential for set-valued maps. It was introduced in [26] as a special case of Sussmann's Approximate Generalized Differential Quotient (AGDQ) (see [32]), by requiring some additional continuity properties for the involved selections. The main fact about QDQ is the validity of a standard, *nonpunctured*, Open Mapping result, which (is not valid for general AGDQs and) allows one to deduce a set separation criterion for QDQ approximating cones (see Corollary 2.1). In turn, the latter proves crucial in the proof of Theorem 3.1. The second comment concerns the possibility of establishing sufficient no-gap conditions through a notion of higher order normality for more general systems. As mentioned above, by means of the concept of *abundance*, such a program was pursued in [26] for the case of first order normality. Though utilizing abundance for higher order variations in a general system does not look straightforward, nevertheless the result presented in this paper might represent an initial step in that direction.

This paper is organized as follows. In section 2 we introduce our precise assumptions, define the QDQ approximating cone, and describe its key properties. Section 3 is devoted to stating the main result, whose proof is given in section 4. Finally, in section 5 we present an example.

2. Preliminaries.

2.1. Basic notation and main assumptions. We use the notation $\mathbb{R}_+ := [0, +\infty[$ and $\mathbb{R}_- :=]-\infty, 0]$. For any pair $a, b \in \mathbb{R}$, we set $a \wedge b := \min\{a, b\}$. If $N \geq 1$ is an integer and $\delta > 0$, for any $\check{x} \in \mathbb{R}^N$ we set $B_\delta^N(\check{x}) := \{x \in \mathbb{R}^N : |x - \check{x}| \leq \delta\}$ and $B_\delta^N := B_\delta^N(0)$. When the dimension is clear from the context, we omit the superscript N. Moreover, for every subset $E \subseteq \mathbb{R}^N$, we use $\mathbf{1}_E$ to denote the characteristic function of E, namely, $\mathbf{1}_E(x) = 1$ if $x \in E$ and $\mathbf{1}_E(x) = 0$ otherwise. For any integer $r \geq 0$, $(\mathscr{M}, \langle \cdot, \cdot \rangle)$ is a Riemannian differential manifold of class C^{r+1} if $\mathscr{M}$ is a C^{r+1} differential manifold and $\langle \cdot, \cdot \rangle$ is a C^r Riemannian metric on $\mathscr{M}$. For every $x \in \mathscr{M}$ and $e, f \in T_x\mathscr{M}$, $\langle e, f \rangle_x$ denotes the corresponding scalar product of e, f, and $|e|_x := \sqrt{\langle e, e \rangle_x}$ is called the *norm* of e. We often omit the subscript and write $\langle e, f \rangle$ and $|e|$ instead of $\langle e, f \rangle_x$ and $|e|_x$. Given an interval I and a subset $X \subseteq \mathscr{M}$, we write $AC(I, X)$ for the space of absolutely continuous functions, $C^0(I, X)$ for the space of continuous functions, $L^1(I, X)$ for the space of Lebesgue integrable functions, and $L^\infty(I, X)$ for the space of Lebesgue measurable, essentially bounded functions defined on I and with values in X. We use $\|\cdot\|_{L^\infty(I,X)}$ and $\|\cdot\|_{L^1(I,X)}$ to denote the essential supremum norm and the L^1 norm, respectively. When no confusion may arise, we simply write $\|\cdot\|_{L^\infty(I)}$, $\|\cdot\|_{L^1(I)}$, or also $\|\cdot\|_\infty$, $\|\cdot\|_1$.

Throughout this paper, $(\mathscr{M}, \langle \cdot, \cdot \rangle)$ is a C^∞ Riemannian differential manifold, the target $\mathfrak{T} \subseteq \mathbb{R}_+ \times \mathscr{M}$ is a closed set, the control set $A \subset \mathbb{R}^q$ is compact, while the unbounded control set $\mathscr{C} \subseteq \mathbb{R}^m$ is a closed cone of the form $\mathscr{C} = \mathscr{C}_1 \times \mathscr{C}_2$, where $m_1, m_2 \in \mathbb{N}$, $m_1 + m_2 = m$, and, if $m_1 \geq 1$, $\mathscr{C}_1 \subseteq \mathbb{R}^{m_1}$ is a closed cone that contains the lines $\{r\mathbf{e}_i : r \in \mathbb{R}\}$, for $i = 1, \dots, m_1$, and $\mathscr{C}_2 \subset \mathbb{R}^{m_2}$ is a closed cone which does not contain any straight line.

Furthermore, we assume the following regularity hypotheses:

(i) the *drift dynamics* f and the partial derivative $D_x f$ are continuous on $\mathscr{M} \times A$, the vector fields $g_1, \dots, g_m$ are of class C^1 on $\mathscr{M}$;

(ii) the *final cost* $\Psi : \mathbb{R} \times \mathscr{M} \to \mathbb{R}$ is of class C^1 on $\mathbb{R} \times \mathscr{M}$.

Remark 2.1. The hypotheses on the cone $\mathscr{C}$ are not at all restrictive. Indeed, they can be recovered by replacing the single vector fields g_i with suitable linear combinations of $\{g_1, \dots, g_m\}$ and by considering a corresponding linear transformation of coordinates in $\mathbb{R}^m$.

Remark 2.2. Through minor changes, the requests that the set A is a compact subset of $\mathbb{R}^q$ and that f and $D_x f$ are continuous w.r.t. $a \in A$ could be replaced by the assumption that A is just a set of parameters and the functions f, $D_x f$ are locally bounded in x, uniformly w.r.t. A.

Remark 2.3. As is customary, adding a state variable equal to the integral of the Lagrangian, one might consider a more general cost

$$\Psi(T, x(T)) + \int_0^T \ell_0(x(t), a(t)) + \ell_1(x(t), |u(t)|)\, dt,$$

where ℓ_0, ℓ_1 are nonnegative and such that the *extended Lagrangian* $L(x, w^0, r, a) := \ell_0(x, a)w^0 + \lim_{\rho \to w^0} \ell_1\left(x, \rho^{-1} r\right) \rho$ and the partial derivative $D_x L$ are continuous in all variables, and $L(x, 0, r, a) \equiv 0$ (see [3]). Furthermore, by adding the new state variables x^0, $\hat{x}$ and by considering the trivial equations $\frac{dx^0}{dt} = 1$, $\frac{d\hat{x}}{dt} = u$, where $\hat{x} = (x^{n+1}, \dots, x^{n+m})$, and the initial conditions $x^0(0) = 0$, $\hat{x}(0) = 0$, one can even allow f, $g_1, \dots, g_m$ to depend on t and the function $U(t) := \int_0^t u(\tau) d\tau$ as well.

In order to specify the notion of local infimum gap anticipated in the introduction, let us introduce a concept of *distance* between extended trajectories.

DEFINITION 2.4. *Let $i = 1, 2$, and for all $(S_i, y_i^0, y_i, \beta_i)$ with $S_i > 0$ and continuous functions $(y_i^0, y_i, \beta_i) : [0, S_i] \to \mathbb{R} \times \mathscr{M} \times \mathbb{R}$, define the distance*

$$\begin{aligned}\mathbf{d}\big((S_1, y_1^0, y_1, \beta_1), (S_2, y_2^0, y_2, \beta_2)\big)& \\ := |S_1 - S_2| + \sup_{s \in \mathbb{R}_+} d\big((\tilde{y}_1^0, \tilde{y}_1, \tilde{\beta}_1)(s), (\tilde{y}_2^0, \tilde{y}_2, \tilde{\beta}_2)(s)\big),\end{aligned} \tag{2.1}$$

where d is the distance on $\mathbb{R} \times \mathscr{M} \times \mathbb{R}^1$ and $(\tilde{y}_i^0, \tilde{y}_i, \tilde{\beta}_i)$ denotes the extension to $\mathbb{R}_+$ of the function (y_i^0, y_i, β_i), such that $(\tilde{y}_i^0, \tilde{y}_i, \tilde{\beta}_i)(s) := (y_i^0, y_i, \beta_i)(S_i)$ for all $s > S_i$.

Remark 2.5. Given a feasible extended sense process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$, the main results of this paper, of local nature, are still valid if the regularity hypothesis (i) is assumed just on a **d**-neighborhood of the reference trajectory, and, instead of (ii), we suppose Ψ merely continuous on a **d**-neighborhood of $(\bar{S}, \bar{y}^0, \bar{y}, \bar{\beta})$ and differentiable at the final point $(\bar{y}^0, \bar{y})(\bar{S})$.

2.2. Cone transversality. Let us recall some elementary notions concerning cones of linear spaces (see, e.g., [31, 32]). Let E be a finite-dimensional, real, linear space, and let E^* be its dual space. A subset $\mathscr{K} \subset E$ is a *cone* if $\alpha k \in \mathscr{K}$ for all $(\alpha, k) \in [0, +\infty[\times \mathscr{K}$. If $D \subset E$ is any subset, let us set

$$\begin{aligned}&\mathrm{span}^+ D \doteq \Big\{\textstyle\sum_{i=1}^{\ell} \alpha_i v_i : \ \ell \in \mathbb{N}, \ \alpha_i \geq 0, \ \ v_i \in D \ \ \forall i = 1, \dots, \ell\Big\} \subset E, \\ &D^\perp \doteq \{p \in E^* : \ p \cdot w \leq 0 \ \ \forall w \in D\} \subset E^*.\end{aligned}$$

The convex cones $\mathrm{span}^+ D$ and $D^\perp$ are called the *conic hull* of D and the *polar cone* of D, respectively. Let $\mathscr{K}_1$, $\mathscr{K}_2 \subseteq E$ be convex cones. We say that $\mathscr{K}_1$ and $\mathscr{K}_2$ are *transversal* if $\mathscr{K}_1 - \mathscr{K}_2 := \big\{k_1 - k_2 : \ (k_1, k_2) \in \mathscr{K}_1 \times \mathscr{K}_2\big\} = E$. $\mathscr{K}_1$ and $\mathscr{K}_2$ are *strongly transversal* if they are transversal and $\mathscr{K}_1 \cap \mathscr{K}_2 \supsetneq \{0\}$.

PROPOSITION 2.1. *Two convex cones $\mathscr{K}_1$, $\mathscr{K}_2 \subseteq E$ are transversal if and only if they are either strongly transversal or are complementary linear subspaces, namely $\mathscr{K}_1 \oplus \mathscr{K}_2 = E$ (i.e., $\mathscr{K}_1 + \mathscr{K}_2 = E$ and $\mathscr{K}_1 \cap \mathscr{K}_2 = \{0\}$).*

Saying that two convex cones $\mathscr{K}_1$ and $\mathscr{K}_2$ are not transversal is equivalent to saying they are linearly separable.

PROPOSITION 2.2. *Two convex cones $\mathscr{K}_1, \mathscr{K}_2 \subseteq E$ are not transversal if and only if* $\mathscr{K}_1$ and $\mathscr{K}_2$ are linearly separable, by which we mean that $(-\mathscr{K}_1^\perp \cap \mathscr{K}_2^\perp) \backslash \{0\} \neq \emptyset$, namely, there exists a linear form $\lambda \in E^* \backslash \{0\}$ such that for all $(k_1, k_2) \in \mathscr{K}_1 \times \mathscr{K}_2$, one has $\lambda \cdot k_1 \geq 0$ and $\lambda \cdot k_2 \leq 0$.

2.3. Quasi-differential quotients and approximating multicones. We call a function $\rho : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{+\infty\}$ a *pseudo-modulus* if it is monotonically nondecreasing and $\lim_{s \to 0^+} \rho(s) = \rho(0) = 0$.

DEFINITION 2.6 (see [26, Def. 2.3]). *Let $G : \mathbb{R}^N \rightsquigarrow \mathbb{R}^n$ be a set-valued map, $(\bar{\varepsilon}, \bar{y}) \in \mathbb{R}^N \times \mathbb{R}^n$, let $\Lambda \subset Lin\{\mathbb{R}^N, \mathbb{R}^n\}$ be a compact set, and let $\Gamma \subset \mathbb{R}^N$ be any subset. We say that Λ is a* Quasi Differential Quotient (QDQ) of G at $(\bar{\varepsilon}, \bar{y})$ in the direction of Γ *if there exists a pseudo-modulus ρ enjoying the property that for*

[1]If $x_1, x_2 \in \mathscr{M}$, the distance $d_{\mathscr{M}}(x_1, x_2)$ is defined as the minimum among the $\langle \cdot, \cdot \rangle$-lengths of the absolutely continuous curves with x_1, x_2 as endpoints.

any $\delta > 0$ with $\rho(\delta) < +\infty$ there is a continuous map $(L_\delta, h_\delta) : (\bar\varepsilon + B_\delta) \cap \Gamma \to Lin\{\mathbb{R}^N, \mathbb{R}^n\} \times \mathbb{R}^n$ such that for all $\varepsilon \in (\bar\varepsilon + B_\delta) \cap \Gamma$,

$$\bar y + L_\delta(\varepsilon) \cdot (\varepsilon - \bar\varepsilon) + h_\delta(\varepsilon) \in G(\varepsilon), \quad \min_{L' \in \Lambda} |L_\delta(\varepsilon) - L'| \le \rho(\delta), \quad |h_\delta(\varepsilon)| \le \delta\rho(\delta).$$

The notion of QDQ extends to differential manifolds as follows.

DEFINITION 2.7 (see [26, Def. 2.4]). *Let $\mathscr{N}$, $\mathscr{M}$ be differential manifolds of class C^1. Assume that $\tilde G : \mathscr{N} \rightsquigarrow \mathscr{M}$ is a set-valued map, $(\bar\varepsilon, \bar y) \in \mathscr{N} \times \mathscr{M}$, $\tilde\Lambda \subset Lin\{T_{\bar\varepsilon}\mathscr{N}, T_{\bar y}\mathscr{M}\}$ is a compact set, and $\tilde\Gamma \subset \mathscr{N}$ is any subset. Moreover, let $\phi : U \to \mathbb{R}^N$ and $\psi : V \to \mathbb{R}^n$ be charts defined on neighborhoods U and V of $\bar\varepsilon$ and $\bar y$, respectively, and assume that $\phi(\bar\varepsilon) = 0$, $\psi(\bar y) = 0$. Let $G : \mathbb{R}^N \to \mathbb{R}^n$ any extension of the map $\psi \circ \tilde G \circ \phi^{-1} : \phi(U) \to \mathbb{R}^n$. We say that $\tilde\Lambda$ is a QDQ of $\tilde G$ at $(\bar\varepsilon, \bar y)$ in the direction of $\tilde\Gamma$ if $\Lambda := D\psi(\bar y) \cdot \Lambda \cdot D\phi^{-1}(0)$ is a QDQ of G at $(0, 0)$ in the direction of $\Gamma := \phi(\tilde\Gamma \cap U)$.*

In a linear space E, let us call any family of convex cones of E *convex multicone*.

DEFINITION 2.8 (see [26, Def. 2.5]). *Let $\mathscr{M}$ be a C^1 differential manifold, $\mathscr{E} \subset \mathscr{M}$ a set, and $z \in \mathscr{E}$. A* QDQ approximating multicone *to $\mathscr{E}$ at z is a convex multicone $\mathscr{K} \subseteq T_z\mathscr{M}$ such that there exist an integer $N \ge 0$, a set-valued map $G : \mathbb{R}^N \rightsquigarrow \mathscr{M}$, a convex cone $\Gamma \subset \mathbb{R}^N$, and a QDQ Λ of G at $(0, z)$ in the direction of Γ such that $G(\Gamma) \subset \mathscr{E}$ and $\mathscr{K} = \{L \cdot \Gamma \ : L \in \Lambda\}$.*[2] *We say that such a* triple (G, Γ, Λ) generates the multicone $\mathscr{K}$. *If the triple (G, Γ, Λ) defining a QDQ approximating cone $\mathscr{K}$ can be chosen so that $G(\Gamma) \subset \mathscr{E}\backslash\{z\}$, then we say that* the QDQ approximating multicone $\mathscr{K}$ is z-ignoring.

Remark 2.9. Because of the local character of the notion of QDQ for a set-valued map, one can equivalently say that *a QDQ approximating multicone $\mathscr{K}$ to $\mathscr{E}$ at z is z-ignoring if $G(B_\delta \cap \Gamma) \subset \mathscr{E}\backslash\{z\}$ for some $\delta > 0$.*

Remark 2.10. The classical Boltyanski approximating cone is a special case of a QDQ approximating cone.

DEFINITION 2.11. *Let $\mathscr{X}$ be a topological space, and let $\mathscr{D}_1, \mathscr{D}_2 \subset \mathscr{X}$, $y \in \mathscr{D}_1 \cap \mathscr{D}_2$. We say that $\mathscr{D}_1$ and $\mathscr{D}_2$ are* locally separated at y *provided there exists a neighborhood V of y such that $\mathscr{D}_1 \cap \mathscr{D}_2 \cap V = \{y\}$.*

The following open-mapping-based result, obtained in [26], characterizes set separation in terms of linear separation of QDQ approximating cones. It includes a crucial approximation property (see (ii) below) whenever the cones are complementary linear subspaces.

THEOREM 2.1 (see [26, Thm. 2.3]). *Let $\mathscr{E}_1, \mathscr{E}_2$ be subsets of a C^1 differential manifold $\mathscr{M}$ and let $z \in \mathscr{E}_1 \cap \mathscr{E}_2$. Assume that $\mathscr{K}_1$, $\mathscr{K}_2$ are QDQ approximating cones of $\mathscr{E}_1$ and $\mathscr{E}_2$, respectively, at z. Then the following assertions hold true:*

(i) *if $\mathscr{K}_1$, $\mathscr{K}_2$ are strongly transversal, the sets $\mathscr{E}_1$, $\mathscr{E}_2$ are not locally separated;*

(ii) *if $\mathscr{K}_1$, $\mathscr{K}_2$ are linear subspaces, $\mathscr{K}_1 \oplus \mathscr{K}_2 = T_z\mathscr{M}$,*[3] *and, for each $i = 1, 2$, $(G_i, \Gamma_i, \Lambda_i)$ is a triple that generates $\mathscr{K}_i$ with $\Lambda_i = \{L_i\}$, $L_i \in Lin\{\mathbb{R}^{N_i}, \mathbb{R}^n\}$, then there exists a sequence $(\gamma_{1_k}, \gamma_{2_k}) \in \Gamma_1 \times \Gamma_2$ such that $z_k \in G_1(\gamma_{1_k}) \cap G_2(\gamma_{2_k})$ and $z_k \to z$.*

[2]When a QDQ approximating multicone is a singleton, namely $\mathscr{K} = \{K\}$, we say that K is a QDQ approximating cone to $\mathscr{E}$ at z.

[3]In view of Proposition 2.1, only in this case nontransversality differs from nonstrong-transversality.

COROLLARY 2.1. *Let $\mathscr{E}_1, \mathscr{E}_2$ be subsets of a C^1 differential manifold $\mathscr{M}$, and let $z \in \mathscr{E}_1 \cap \mathscr{E}_2$. Assume that $\mathscr{K}_1$, $\mathscr{K}_2$ are QDQ approximating cones of $\mathscr{E}_1$ and $\mathscr{E}_2$, respectively, at z, and that $\mathscr{K}_1$ (or, equivalently, $\mathscr{K}_2$) is z-ignoring. If the cones $\mathscr{K}_1$ and $\mathscr{K}_2$ are transversal, then the sets $\mathscr{E}_1$ and $\mathscr{E}_2$ are not locally separated.*

Proof. In view of Theorem 2.1 we need only prove the result in the case when $\mathscr{K}_1$ and $\mathscr{K}_2$ are transversal but not strongly transversal. This means that $\mathscr{K}_1$ and $\mathscr{K}_2$ are linear subspaces and $\mathscr{K}_1 \oplus \mathscr{K}_2 = T_z\mathscr{M}$. By Theorem 2.1 and the fact that $\mathscr{K}_1$ is z-ignoring, we deduce that there are a nonnegative integer N, a convex cone $\Gamma \subset \mathbb{R}^N$, and a set-valued map $G : \mathbb{R}^N \rightsquigarrow \mathscr{M}$ such that $\Lambda = \{L\} \subset Lin(\mathbb{R}^N, T_z\mathscr{M})$ is a QDQ of G at $(0, z)$ in the direction of Γ, $\mathscr{K}_1 = L \cdot \Gamma$, $G(\Gamma) \subseteq \mathscr{E}_1 \backslash \{z\}$, and there is a sequence $(\gamma_k) \subset \Gamma$ such that $z_k \in G(\gamma_k) \cap \mathscr{E}_2$ and $z_k \to z$. Since $G(\gamma_k) \subseteq \mathscr{E}_1 \backslash \{z\}$ for every $k \in \mathbb{N}$, one has $z_k \in \mathscr{E}_1 \cap \mathscr{E}_2 \backslash \{z\}$, so that $\mathscr{E}_1$ and $\mathscr{E}_2$ are not locally separated. ◻

3. The main result.

3.1. Infimum gap. Let us use $\mathcal{S}_{\mathscr{W}}$ to denote the set of extended sense processes. Furthermore, let $\mathcal{S}_{\mathscr{W}_+} \subset \mathcal{S}_{\mathscr{W}}$ be the subset of *embedded strict sense processes*, by which we mean those processes with controls in

$$\mathscr{W}_+ := \{(S, w^0, w, \alpha) \in \mathscr{W} : \ w^0 > 0 \text{ a.e.}\}.$$

Given a feasible extended sense process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ and $r > 0$, the sets $\mathscr{R}'^r_{\mathscr{W}_+}$, $\mathscr{R}'^r_{\mathscr{W}} \subset \mathbb{R} \times \mathscr{M} \times \mathbb{R}$, defined as

$$\mathscr{R}'^r_{\mathscr{W}_+} := \Big\{(y^0, y, \beta)(S) : \ (S, w^0, w, \alpha, y^0, y, \beta) \in \mathcal{S}_{\mathscr{W}_+}, \ \mathbf{d}\big((S, y^0, y, \beta), (\bar{S}, \bar{y}^0, \bar{y}, \bar{\beta})\big) < r\Big\},$$
$$\mathscr{R}'^r_{\mathscr{W}} := \Big\{(y^0, y, \beta)(S) : \ (S, w^0, w, \alpha, y^0, y, \beta) \in \mathcal{S}_{\mathscr{W}}, \ \mathbf{d}\big((S, y^0, y, \beta), (\bar{S}, \bar{y}^0, \bar{y}, \bar{\beta})\big) < r\Big\},$$

will be called the *reachable set* and the *extended reachable set*, respectively. The occurrence of a local infimum gap is captured by the following definition.

DEFINITION 3.1. *Let $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ be a feasible extended sense process. We say that* at $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ there is a local infimum gap *if there exists $r > 0$ such that one has*

$$\Psi\big(\bar{y}^0(\bar{S}), \bar{y}(\bar{S})\big) < \inf_{(y^0, y, \beta) \in \mathscr{R}'^r_{\mathscr{W}_+} \cap (\mathfrak{T} \times [0, K])} \Psi(y^0, y). \tag{3.1}$$

Despite the name, the local infimum gap condition (3.1) is a fully dynamical property. Indeed, it reflects the fact that no feasible embedded strict sense trajectories do exist in a sufficiently small $\mathbf{d}$-neighborhood of the extended trajectory $(\bar{S}, \bar{y}^0, \bar{y}, \bar{\beta})$. To make this rigorous, let us introduce the notion of *isolated process*.

DEFINITION 3.2. *A feasible extended sense process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is called* isolated *if $\mathscr{R}'^r_{\mathscr{W}_+} \cap (\mathfrak{T} \times [0, K]) = \emptyset$ for some $r > 0$.*

The following result is straightforward.

LEMMA 3.1. *Let $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ be a feasible extended sense process. The following statements are equivalent:*

- (i) *there is a local infimum gap at $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$;*
- (ii) *the process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is isolated;*
- (iii) *there exists some $\bar{r} > 0$ such that $\inf_{(y^0, y, \beta) \in \mathscr{R}'^r_{\mathscr{W}_+} \cap (\mathfrak{T} \times [0, K])} \tilde{\Psi}(y^0, y) = +\infty$ for every continuous function $\tilde{\Psi}$ and every $r \in [0, \bar{r}]$.*

3.2. Iterated Lie brackets. Before giving the notion of higher order extremal, let us recall some basic facts concerning iterated Lie brackets.

If h_1, h_2 are C^1 vector fields on a differential manifold $\mathscr{M}$,[4] the *Lie bracket of* h_1 *and* h_2 is defined, on any chart, as

$$[h_1, h_2](x) := Dh_2(x) \cdot h_1(x) - Dh_1(x) \cdot h_2(x)(= -[h_2, h_1](x)).$$

As is well known, the map $[h_1, h_2]$ is a true vector field, i.e., it can be defined intrinsically. Therefore, if the vector fields are sufficiently regular, one can iterate the bracketing process: for instance, given a 4-tuple $\mathbf{h} := (h_1, h_2, h_3, h_4)$ of vector fields one can construct the brackets $[[h_1, h_2], h_3]$, $[[h_1, h_2], [h_3, h_4]]$, $[[[h_1, h_2], h_3], h_4]$, $[[h_2, h_3], h_4]$. Accordingly, one can consider the *(iterated) formal brackets* $B_1 := [[X_1, X_2], X_3]$, $B_2 := [[X_1, X_2], [X_3, X_4]]$, $B_3 := [[[X_1, X_2], X_3], X_4]$, $B_4 := [[X_2, X_3], X_4]$ (regarded as sequence of letters $X_1, \ldots, X_4$, commas, and left and right square parentheses), so that, with obvious meaning of the notation, one has $B_1(\mathbf{h}) = [[h_1, h_2], h_3]$, $B_2(\mathbf{h}) = [[h_1, h_2], [h_3, h_2]]$, $B_3(\mathbf{h}) = [[[h_1, h_2], h_3], h_4]$, $B_4(\mathbf{h}) = [[h_2, h_3], h_4]$.

The *length* of a formal bracket is the number of letters that are involved in it. For instance, the brackets B_1, B_2, B_3, B_4 have lengths equal to 3, 4, 4, and 3, respectively. By convention, we declare that a single variable X_i is a formal bracket of length 1.

The *switch-number* of a (formal) bracket B is the number $r_{_B}$ defined recursively on the nested structure of the bracket as

$$r_{_B} := 1 \text{ if } B \text{ has length } 1; \quad r_{_B} := 2\big(r_{_{B_1}} + r_{_{B_2}}\big) \text{ if } B = [B_1, B_2].$$

For instance, the switch-numbers of $[[X_3, X_4], [[X_5, X_6], X_7]]$ and $[[X_5, X_6], X_7]$ are 28 and 10, respectively. If there is no danger of confusion, we sometimes take the liberty of speaking of "length and switch-number of Lie brackets of vector fields."

The regularity of a (nonformal) iterated Lie bracket depends on both the nested structure of the underlying formal brackets and the involved vector fields. We will use the following notion of *bracket regularity* for a string of vector fields (for a more rigorous definition we refer to [11]).

DEFINITION 3.3 (bracket regularity). *Fix* $k \in \mathbb{N}$. *If* $\mu \geq 0$, $r \geq 1$, *and* $\nu \geq \mu + r$ *are integers,* $B = B(X_{\mu+1}, \ldots, X_{\mu+r})$ *is an iterated formal bracket, and* $\mathbf{h} = (h_1, \ldots, h_\nu)$ *is a string of vector fields, we say that* $\mathbf{h}$ is of class C^{B+k} *if there is a* ν*-tuple* $(j_1, \ldots, j_\nu) \in \mathbb{N}^\nu$ *such that* h_i *is of class* C^{j_i} *for any* $i = 1, \ldots, \nu$ *and* $B(\mathbf{h})$ *is a vector field of class* C^k. *In this case, we call* $(B, \mathbf{h})$ *an* admissible C^k bracket pair.

For instance, if $B = [[[X_3, X_4], [X_5, X_6]], X_7]$, for any $k \geq 0$, a string $\mathbf{h} = (h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8)$ is of class C^{B+k} provided the vector fields h_3, h_4, h_5, h_6 are of class C^{3+k} and h_7 is of class C^{1+k}.

3.3. Higher order extremals. Let us set

$$\mathbf{C} := \{(w^0, w) \in \mathbb{R}_+ \times \mathscr{C} : \, w^0 + |w| = 1\}, \tag{3.2}$$

and let us consider the *unmaximized Hamiltonian* $H : T^*\mathscr{M} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \times \mathscr{C} \times A \to \mathbb{R}$,

$$H(x, p, p_0, \pi, w^0, w, a) := p_0 w^0 + p \cdot \left(f(x, a) w^0 + \sum_{i=1}^{m} g_i(x) w^i \right) + \pi |w|.$$

To give the notion of higher order extremal, we need one more definition.

[4]The regularity of the differential manifold $\mathscr{M}$ will be always assumed such that all considered brackets can be classically defined. For simplicity, one can assume that $\mathscr{M}$ is of class C^∞.

DEFINITION 3.4. *For every integer $k \geq 0$, we will use $\mathfrak{Br}^k$ to denote the (possibly empty) set of admissible C^k bracket pairs $(B, \mathbf{h})$, such that $\mathbf{h} := (h_1, \ldots, h_\nu)$ is a ν-tuple of vector fields $h_j \in \{g_1, \ldots, g_{m_1}\}$ for every $j = 1, \ldots, \nu$.*

Let us consider, for instance, $m_1 \geq 10$ and the pair $B = [[X_3, [X_4, X_5]], X_6]$, $\mathbf{h} = (h_1, \ldots, h_6) := (g_8, g_{10}, g_1, g_4, g_3, g_1)$. Then, the vector field $B(\mathbf{h})$ coincides with the iterated Lie bracket $[[g_1, [g_4, g_3]], g_1]$. Moreover, the pair $(B, \mathbf{h}) \in \mathfrak{Br}^k$ (which implies that $B(\mathbf{h}) \in C^k$) provided $g_1 \in C^{2+k}$ and $g_3, g_4 \in C^{3+k}$.

In what follows, we will use the notation $f_a(\cdot) := f(\cdot, a)$ for all $a \in A$.

DEFINITION 3.5. *Let $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ be a feasible extended sense process. Let $\mathscr{K}$ be a QDQ approximating cone to the target $\mathfrak{T}$ at $(\bar{y}^0, \bar{y})(\bar{S})$. We say that the process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a* Ψ-higher order extremal *if there exist a lift $(\bar{y}, p) \in AC([0, \bar{S}], T^*\mathscr{M})$ and multipliers $(p_0, \pi, \lambda) \in \mathbb{R} \times \mathbb{R}_- \times \mathbb{R}_+$ such that conditions* (i)–(vi) *below are valid.*

(i) (nontriviality) *The triple (p_0, p, λ) is nontrivial, i.e.,*

$$(p_0, p, \lambda) \neq (0, 0, 0)\,. \tag{3.3}$$

Furthermore, if the trajectory $\bar{y}$ is not instantaneous, namely, if $\bar{y}^0(\bar{S}) > 0$, then (3.3) *can be strengthened to*

$$(p, \lambda) \neq (0, 0). \tag{3.4}$$

(ii) (nontranversality)

$$(p_0, p(\bar{S}), \pi) \in \left[-\lambda D\Psi\big((\bar{y}^0, \bar{y})(\bar{S})\big) - \mathscr{K}^\perp\right] \times J_K, \tag{3.5}$$

where $J_K := \{0\}$ if $\bar{\beta}(\bar{S}) < K$, and $J_K :=]0, +\infty[$ if $\bar{\beta}(\bar{S}) = K$. In particular,

$$\pi = 0 \quad \textit{provided} \quad \bar{\beta}(\bar{S}) < K. \tag{3.6}$$

(iii) (Hamiltonian equations) *The path $(\bar{y}, p)$ verifies, for a.e. $s \in [0, \bar{S}]$,*

$$\frac{d}{ds}(\bar{y}, p)(s) \;=\; \mathbf{X}_{\bar{H}}\left(s, \bar{y}(s), p(s)\right), \tag{3.7}$$

where $\bar{H} = \bar{H}(s, y, p) := H\big(y, p, p_0, \pi, \bar{w}^0(s), \bar{w}(s), \bar{\alpha}(s)\big)$, and $\mathbf{X}_{\bar{H}}$ denotes the (s-dependent) Hamiltonian vector field corresponding to $\bar{H}$.[5]

(iv) (first order maximization) *For a.e. $s \in [0, \bar{S}]$,*

$$\begin{aligned} &H\Big(\bar{y}(s), p(s), p_0, \pi, \bar{w}^0(s), \bar{w}(s), \bar{\alpha}(s)\Big) \\ &\qquad\qquad = \max_{(w^0, w, a) \in \mathbf{C} \times A} H(\bar{y}(s), p(s), p_0, \pi, w^0, w, a), \end{aligned} \tag{3.8}$$

and, as soon as $\bar{\beta}(S) < K$,

$$p(s) \cdot g_i(\bar{y}(s)) = 0 \qquad \forall\, s \in [0, \bar{S}],\ i = 1, \ldots, m_1, \tag{3.9}$$

[5]If $K = K(s, y, p)$ is a differentiable map on the cotangent bundle $T^*\mathscr{M}$, in any local system of canonical coordinates $(\mathfrak{y}, \mathfrak{p})$, the Hamiltonian vector field $\mathbf{X}_K$ corresponding to K is defined as $\mathbf{X}_K(s, \mathfrak{y}, \mathfrak{p}) := (D_{\mathfrak{p}} K, -D_{\mathfrak{y}} K)\,(s, \mathfrak{y}, \mathfrak{p})$, so that (3.7) coincides with the extended system coupled with the usual adjoint equation.

(v) (vanishing of the Hamiltonian)

$$\max_{(w^0,w,a)\in\mathbf{C}\times A} H(\bar y(s), p(s), p_0, \pi, w^0, w, a) = 0 \quad \forall\, s \in [0, \bar S]. \tag{3.10}$$

(vi) (higher order conditions) *If* $\bar\beta(S) < K$ *and* $(B, \mathbf{h}) \in \mathfrak{Br}^0$,

$$p(s) \cdot B(\mathbf{h})(\bar y(s)) = 0 \quad \forall\, s \in [0, \bar S]. \tag{3.11}$$

Furthermore, if $(B, \mathbf{h}) \in \mathfrak{Br}^1$ *for a.e.* $s \in [0, S]$, *one has*

$$p(s) \cdot \left(\big[f_{\bar\alpha(s)}, B(\mathbf{h})\big](\bar y(s))\, \bar w^0(s) + \sum_{j=m_1+1}^{m} \big[g_j, B(\mathbf{h})\big](\bar y(s))\, \bar w^j(s) \right) = 0. \tag{3.12}$$

Definition 3.6. *Let* $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ *be a feasible extended sense process, which, given a QDQ approximating cone* $\mathscr{K}$ *of the target* $\mathfrak{T}$ *at* $(\bar y^0, \bar y)(\bar S)$, *is a higher order* Ψ*-extremal. We say that* $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ *is a* normal higher order Ψ-extremal *if for any choice of the multipliers* (p_0, p, π, λ), *one has* $\lambda \neq 0$. *Otherwise, we say that it is an* abnormal higher order extremal.[6]

Remark 3.7. The notion of extremal depends on the approximating cone $\mathscr{K}$, so one might more properly speak of extremal w.r.t. $\mathscr{K}$. In this regard, the form of (3.5) derives from the fact that, given $\mathscr{K}$, as an approximating cone to the (y^0, y, β)-target $\mathfrak{T} \times [0, K]$ at $(\bar y^0, \bar y, \bar\beta)(\bar S)$, one chooses $\mathscr{K} \times \mathbb{R}$ if $\bar\beta(\bar S) < K$ and $\mathscr{K} \times]-\infty, 0]$ if $\bar\beta(\bar S) = K$. In particular, $(\mathscr{K} \times \mathbb{R})^\perp = \mathscr{K}^\perp \times \{0\}$ if $\bar\beta(\bar S) < K$, while one has $(\mathscr{K} \times]-\infty, 0])^\perp = \mathscr{K}^\perp \times \mathbb{R}_+$ when $\bar\beta(\bar S) = K$.

Remark 3.8. This notion of higher order extremal is slightly more general than the one utilized in the Higher Order Maximum Principle established in [3]. Indeed, besides considering a state y ranging on a Riemannian manifold (rather than on a mere Euclidean space), we use QDQ approximating cones here, which are more general than the Boltyanski approximating cones considered in [3].

3.4. Higher order normality and no-gap. Let us state our main result.

Theorem 3.1 (gap and higher order abnormality). *Let* $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ *be a feasible extended sense process at which there is a local infimum gap. Then, for every QDQ approximating cone* $\mathscr{K}$ *to the target* $\mathfrak{T}$ *at* $(\bar y^0, \bar y)(\bar S)$, *the process* $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ *is an abnormal higher order extremal.*

The proof of this result will be given in section 4.

As a straightforward consequence of Theorem 3.1, we deduce the following sufficient condition for the absence of gap.

Theorem 3.2 (higher order normality and no-gap). *Let* $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ *be a feasible extended sense process which satisfies, for some* $r > 0$,

$$\Psi((\bar y^0, \bar y)(\bar S)) \le \inf_{(y^0,y,\beta)\in \mathscr{R}'^{\,r}_{\mathscr{W}_+} \cap (\mathfrak{T}\times[0,K])} \Psi(y^0, y).$$

If $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ *is a normal higher order* Ψ*-extremal for some QDQ approximating cone* $\mathscr{K}$ *to* $\mathfrak{T}$ *at* $(\bar y^0, \bar y)(\bar S)$, *then at* $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ *there is no local infimum gap.*

[6] The fact that a higher order extremal, as well as a classical extremal, is abnormal does not depend on the cost function Ψ.

Remark 3.9. As illustrated in the example of section 5, if $\beta(S) < K$, it may happen that a higher order normal extremal is abnormal as a first order extremal. Actually, this is the main motivation for a result such as the one stated in Theorem 3.2.

4. Proof of Theorem 3.1. The proof of the theorem relies on a set separation argument, whose application is made possible by Theorem 4.1 below, which states that the reachable set can be approximated by suitable "higher order" QDQ approximating cones of the extended reachable set. This is far from being obvious, as the extended reachable set may be quite larger than the reachable set (see Remark 4.9 below).

Let $(\bar S, \bar w^0, \bar w, \bar\alpha, \bar y^0, \bar y, \bar\beta)$ be a feasible extended sense process at which there is a local infimum gap. In case $\bar\beta(\bar S) = K$, the statement of Theorem 3.1 reduces to the first order conditions (3.3), (3.4), (3.5), (3.7), (3.8), and (3.10), which follow from [26, Theorem 5.2]. Hence, from now on we suppose $\bar\beta(\bar S) < K$.

We will assume the simplifying hypothesis

$$\mathscr{M} \textit{ is an open subset of } \mathbb{R}^n\textit{, so that we can identify } T_{\bar y(\bar S)}\mathscr{M} \textit{ with } \mathbb{R}^n,$$

and the stronger regularity hypothesis

(H)$_b$ *All the general regularity assumptions are verified and, moreover, f and the partial derivatives $D_{x^1}f, \dots, D_{x^n}f$ are uniformly continuous and bounded on $\mathscr{M} \times A$; the vector fields $g_1, \dots, g_m$, their derivatives, and, for all bracket pairs $(B, \mathbf{h}) \in \mathfrak{Br}^0$, the Lie brackets $B(\mathbf{h})$, are uniformly continuous and bounded.*

Both hypotheses are not restrictive, because of the local character of the result.

4.1. Rate-independence of the extended control system. Let us enlarge the set of extended sense processes considered up to now by introducing the larger set of extended sense controls $\tilde{\mathscr{W}} \supset \mathscr{W}$, defined as

$$\tilde{\mathscr{W}} := \bigcup_{S>0} \left(\{S\} \times \left\{ (w^0, w, \alpha) \in L^\infty([0,S], \mathbb{R}_+ \times \mathscr{C} \times A) : \operatorname{ess\,inf}(w^0 + |w|) > 0 \right\} \right),$$

and the subset $\tilde{\mathscr{W}}_+ := \{(S, w^0, w, \alpha) \in \tilde{\mathscr{W}} : w^0 > 0 \text{ a.e.}\} \subset \tilde{\mathscr{W}}$. Let $\mathcal{S}_{\tilde{\mathscr{W}}}$, $\mathcal{S}_{\tilde{\mathscr{W}}_+}$ denote the set of processes $(S, w^0, w, \alpha, y^0, y, \beta)$, where $(S, w^0, w, \alpha) \in \tilde{\mathscr{W}}$ and $(S, w^0, w, \alpha) \in \tilde{\mathscr{W}}_+$, respectively, while (y^0, y, β) is the corresponding solution on $[0, S]$ of (1.2).

In section 4 we will refer to the elements of $\mathcal{S}_{\tilde{\mathscr{W}}}$ as *extended sense processes*, while any element of the subset $\mathcal{S}_{\tilde{\mathscr{W}}_+}$ will be called an *embedded strict sense process*. Finally, we will say that $(S, w^0, w, \alpha, y^0, y, \beta) \in \mathcal{S}_{\tilde{\mathscr{W}}}$ is *canonical* when $(S, w^0, w, \alpha) \in \mathscr{W}$, i.e., $w^0 + |w| = 1$ a.e. With this convention, all of the extended sense processes considered so far were canonical.

By *rate-independence* of the extended control system (1.2) we mean that, given any strictly increasing, surjective, absolutely continuous function $\sigma : [0, S] \to [0, \tilde S]$ with absolutely continuous inverse, $(\tilde S, \tilde w^0, \tilde w, \tilde\alpha, \tilde y^0, \tilde y, \tilde\beta)$ is an extended sense process if and only if the process $(S, w^0, w, \alpha, y^0, y, \beta)$ given by

$$(w^0, w) := \Big((\tilde w^0, \tilde w) \circ \sigma \Big) \frac{d\sigma}{ds}, \quad (\alpha, y^0, y, \beta) := \Big(\tilde\alpha, \tilde y^0, \tilde y, \tilde\beta \Big) \circ \sigma$$

is an extended sense process (see [21, sect. 3]). If we call any two extended sense processes as above *equivalent* and write $(\tilde S, \tilde w^0, \tilde w, \tilde\alpha, \tilde y^0, \tilde y, \tilde\beta) \sim (S, w^0, w, \alpha, y^0, y, \beta)$,[7] we can single out a special representative in any $\sim$equivalence class.

[7] Actually, $\sim$ is an equivalence relation on the set $\mathcal{S}_{\tilde{\mathscr{W}}}$.

DEFINITION 4.1. *Given* $(S, w^0, w, \alpha, y^0, y, \beta) \in \mathcal{S}_{\tilde{\mathscr{W}}}$*, we set* $[0, S] \ni s \mapsto \sigma(s) := y^0(s) + \beta(s)$*, and define the* canonical parameterization *of* $(S, w^0, w, \alpha, y^0, y, \beta)$ *as*

$$(S_c, w_c^0, w_c, \alpha_c, y_c^0, y_c, \beta_c) := \left(\sigma(S), ((w^0, w) \circ \sigma^{-1}) \cdot \frac{d\sigma^{-1}}{ds}, (\alpha, y^0, y, \beta) \circ \sigma^{-1}\right).$$

Note that $w_c^0(s) + |w_c(s)| = 1$ for a.e. $s \in [0, S_c]$, so that $(S_c, w_c^0, w_c, \alpha_c, y_c^0, y_c, \beta_c)$ is a canonical extended sense process. One can easily verify that an extended sense process is canonical if and only if it coincides with its canonical parameterization.

By considering the enlarged set $\tilde{\mathscr{W}}$ of extended sense controls, the original control system (1.1) can be embedded into the extended system (1.2) when the latter is thought of as defined on the $\sim$quotient space, and the set of strict sense processes can be identified with $\mathcal{S}_{\tilde{\mathscr{W}}_+}$. Precisely, any strict sense process (T, u, a, x, v) is in one-to-one correspondence with the $\sim$equivalence class $[(S, w^0, w, \alpha, y^0, y, \beta)]$, where $(S, w^0, w, \alpha, y^0, y, \beta)$ is the canonical embedded strict sense process defined in (1.3). Notice that every $(\tilde{S}, \tilde{w}^0, \tilde{w}, \tilde{\alpha}, \tilde{y}^0, \tilde{y}, \tilde{\beta})$ which is $\sim$equivalent to $(S, w^0, w, \alpha, y^0, y, \beta)$ verifies $\tilde{w}^0 > 0$ a.e.; hence it belongs to $\mathcal{S}_{\tilde{\mathscr{W}}_+}$.

Since an extended sense process is feasible if and only if any equivalent process is feasible (and the costs of equivalent processes do coincide), in the extended optimization problem and the definition of local infimum gap we have the freedom to consider indifferently one of the following classes of extended sense processes: (i) the set of *all extended sense processes*, or (ii) the subclass of *canonical extended sense processes* (as we have done in the previous sections), or even (iii) *any subclass of extended sense processes* $(S, w^0, w, \alpha, y^0, y, \beta)$ *such that* $R_1 \le w^0 + |w| \le R_2$ *a.e.* $0 < R_1 < R_2$.

Precisely, from the rate-independence of the extended system it easily follows that

(a) *at a feasible process* $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta}) \in \mathcal{S}_{\tilde{\mathscr{W}}}$ *there is a local infimum gap (w.r.t. feasible processes in* $\mathcal{S}_{\tilde{\mathscr{W}}_+}$*) if and only if there is a local infimum gap at every process* $(S, w^0, w, \alpha, y^0, y, \beta) \sim (\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$;

(b) *at a feasible canonical extended sense process* $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ *there is a local infimum gap w.r.t. feasible processes in* $\mathcal{S}_{\tilde{\mathscr{W}}_+}$ *if and only there is a local infimum gap w.r.t. feasible canonical processes in* $\mathcal{S}_{\mathscr{W}_+}$ *only.*

On the one hand, as a consequence of (a) it is by no means restrictive to investigate the gap issue just for a canonical process. On the other hand, in view of Lemma 3.1 the property (b) implies that the reference process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is isolated among *all* feasible processes in $\mathcal{S}_{\tilde{\mathscr{W}}_+}$, namely, there exists some $r > 0$ such that

$$\mathscr{R}'^{r}_{\tilde{\mathscr{W}}_+} \cap (\mathfrak{T} \times [0, K]) = \emptyset. \tag{4.1}$$

Here, for any $\bar{r} > 0$ we define the sets $\mathscr{R}'^{\bar{r}}_{\tilde{\mathscr{W}}_+}$ and $\mathscr{R}'^{\bar{r}}_{\tilde{\mathscr{W}}}$ as the reachable set and the extended reachable set of subsection 3.1, except that the sets of canonical controls $\mathscr{W}_+$ and $\mathscr{W}$ are replaced with $\tilde{\mathscr{W}}_+$ and $\tilde{\mathscr{W}}$, respectively. In any case, in this section we refer to $\mathscr{R}'^{\bar{r}}_{\tilde{\mathscr{W}}_+}$ and $\mathscr{R}'^{\bar{r}}_{\tilde{\mathscr{W}}}$ as the *reachable set* and the *extended reachable set,* respectively.

4.2. Higher order QDQ approximating cones.

4.2.1. Some technical preliminaries. Hypothesis $\mathbf{(H)}_b$ guarantees that if we introduce the compact set

$$\mathbf{W}' := \left\{(w^0, w, a) \in \mathbb{R}_+ \times \mathscr{C} \times A : \quad \frac{1}{2} \le w^0 + |w| \le 4\right\}$$

for any $(w^0, w, \alpha) \in L^1([0, \bar{S}], \mathbf{W}')$ there exists a unique solution $(y^0, y, \beta)[w^0, w, \alpha]$ to (1.2) defined on the whole interval $[0, \bar{S}]$. Moreover, it is straightforward to check that the input-output map

$$\Phi : L^1([0, \bar{S}], \mathbf{W}') \to C^0([0, \bar{S}], \mathbb{R} \times \mathscr{M} \times \mathbb{R})$$
$$(w^0, w, \alpha) \mapsto (y^0, y, \beta)[w^0, w, \alpha]$$

is continuous when $L^1([0, \bar{S}], \mathbf{W}')$ is endowed with the L^1 norm. In particular, for some constant $C > 0$ the map Φ enjoys the Lipschitz continuity properties

$$\|(y^0, y, \beta)[w^0, w, \alpha] - (y^0, y, \beta)[\tilde{w}^0, \tilde{w}, \tilde{\alpha}]\|_\infty \leq C \operatorname{meas}\Big\{s \in [0, \bar{S}] : \ (w^0, w, \alpha)(s) \neq (\tilde{w}^0, \tilde{w}, \tilde{\alpha})(s)\Big\}, \tag{4.2}$$

$$\|(y^0, y, \beta)[w^0, w, \alpha] - (y^0, y, \beta)[\tilde{w}^0, \tilde{w}, \alpha]\|_\infty \leq C \,\|(w^0, w) - (\tilde{w}^0, \tilde{w})\|_\infty \tag{4.3}$$

for every (w^0, w, α), $(\tilde{w}^0, \tilde{w}, \alpha)$, $(\tilde{w}^0, \tilde{w}, \tilde{\alpha}) \in L^1([0, \bar{S}], \mathbf{W}')$.[8] Actually, a stronger Lipschitz condition involving the distance between the primitives of the controls (w^0, w) holds true trivially. Moreover, at least in the case when the drift is a-independent, an even stronger Lipschitz condition involving the *Fréchet distance* on both sides is valid (see, e.g., [9, 21, 15].)

Consider now the subset $\mathbf{W} \subset \mathbf{W}'$ of extended control values given by

$$\mathbf{W} := \left\{(w^0, w, a) \in \mathbb{R}_+ \times \mathscr{C} \times A : \ \ \frac{1}{2} \leq w^0 + |w| \leq 2\right\},$$

the subclass of extended sense controls *with fixed endtime*

$$\tilde{\mathscr{W}}^{\bar{S}} := \Big\{(w^0, w, \alpha) : \ (\bar{S}, w^0, w, \alpha) \in \tilde{\mathscr{W}}, \ (w^0, w, \alpha)(s) \in \mathbf{W} \text{ for a.e. } s \in [0, \bar{S}]\Big\},$$

and the subset

$$\tilde{\mathscr{W}}^{\bar{S}}_+ := \Big\{(w^0, w, \alpha) : \ \ (w^0, w, \alpha) \in \tilde{\mathscr{W}}^{\bar{S}}, \ \ w^0(s) > 0 \text{ for a.e. } s \in [0, \bar{S}]\Big\} \subset \tilde{\mathscr{W}}^{\bar{S}}.$$

Since $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a canonical process, clearly $(\bar{w}^0, \bar{w}, \bar{\alpha}) \in \tilde{\mathscr{W}}^{\bar{S}}$. Furthermore, it is an isolated process even when one considers only embedded strict sense processes defined on the fixed interval $[0, \bar{S}]$ and with controls assuming values in the set $\mathbf{W}$. More precisely, if we introduce the *reachable set* and *the extended reachable set at fixed time* $\bar{S}$, defined as

$$\begin{aligned} {\mathscr{R}'}^{r}_{\tilde{\mathscr{W}}^{\bar{S}}_+} &:= \Big\{(y^0, y, \beta)[w^0, w, \alpha](\bar{S}) \in {\mathscr{R}'}^{r}_{\tilde{\mathscr{W}}_+} : (w^0, w, \alpha) \in \tilde{\mathscr{W}}^{\bar{S}}\Big\} \subseteq {\mathscr{R}'}^{r}_{\tilde{\mathscr{W}}_+}, \\ {\mathscr{R}'}^{r}_{\tilde{\mathscr{W}}^{\bar{S}}} &:= \Big\{(y^0, y, \beta)[w^0, w, \alpha](\bar{S}) \in {\mathscr{R}'}^{r}_{\tilde{\mathscr{W}}} : (w^0, w, \alpha) \in \tilde{\mathscr{W}}^{\bar{S}}\Big\} \subseteq {\mathscr{R}'}^{r}_{\tilde{\mathscr{W}}}, \end{aligned} \tag{4.4}$$

respectively, we have that ${\mathscr{R}'}^{r}_{\tilde{\mathscr{W}}^{\bar{S}}_+} \cap (\mathfrak{T} \times [0, K]) = \emptyset$ for the same r as in (4.1).

We will need the following technical result, which, for a given extended sense control (w^0, w, α), establishes some regularity properties of a certain operator $\theta_{(w^0, w, \alpha)}(\cdot)$, mapping any $\delta \in]0, 1]$ into an embedded strict sense control.

[8]Unlike inequality (4.2), which is valid for general control systems, condition (4.3) is an easy consequence of the fact that the original dynamics function is affine in u.

LEMMA 4.1. *There exists some positive constant $D > 0$ such that for every extended sense control $(w^0, w, \alpha) \in \tilde{\mathscr{W}}^{\bar{S}}$, the map $\theta_{(w^0,w,\alpha)} : [0,1] \to \tilde{\mathscr{W}}^{\bar{S}}$, defined by setting $\theta_{(w^0,w,\alpha)}(\delta) := \big(w^0 + \frac{\delta}{1+\delta}|w|,\ \frac{1}{1+\delta}\, w, \alpha\big)$ for all $\delta \in [0,1]$, verifies*

$$\theta_{(w^0,w,\alpha)}(]0,1]) \subseteq \tilde{\mathscr{W}}_+^{\bar{S}},$$

$$\|(y^0,y,\beta)[w^0,w,\alpha] - (y^0,y,\beta)\big[\theta_{(w^0,w,\alpha)}(\delta)\big]\|_\infty \le D\,\delta, \tag{4.5}$$

$$\|(y^0,y,\beta)\big[\theta_{(w^0,w,\alpha)}(\delta_1)\big] - (y^0,y,\beta)\big[\theta_{(w^0,w,\alpha)}(\delta_2)\big]\|_\infty \le D|\delta_1 - \delta_2| \tag{4.6}$$

for all δ, δ_1, $\delta_2 \in [0,1]$.

Proof. Observe that $(\theta_{(w^0,w,\alpha)}(0) = (w^0, w, \alpha))$ and for any $\delta \in]0,1]$ one has $\theta_{(w^0,w,\alpha)}(\delta) \in \tilde{\mathscr{W}}_+^{\bar{S}} \subset \tilde{\mathscr{W}}^{\bar{S}}$. Indeed, for almost every $s \in [0,\bar{S}]$ one has

$$\begin{aligned} &w^0(s) + \frac{\delta}{1+\delta}|w(s)| \ge \frac{\delta}{2(1+\delta)} > 0,\\ &w^0(s) + \frac{\delta}{1+\delta}|w(s)| + \frac{1}{1+\delta}\,|w(s)| = w^0(s) + |w(s)| \in \left[\frac{1}{2}\,,\,2\right]. \end{aligned}$$

Moreover, by the trivial estimates

$$\left\| w^0 + \frac{\delta_1}{1+\delta_1}|w| - \left(w^0 + \frac{\delta_2}{1+\delta_2}|w|\right)\right\|_\infty \le \frac{2|\delta_1-\delta_2|}{(1+\delta_1)(1+\delta_2)},$$

$$\left\| \frac{1}{1+\delta_1}\,|w| - \frac{1}{1+\delta_2}\,|w| \right\|_\infty \le \frac{2|\delta_1-\delta_2|}{(1+\delta_1)(1+\delta_2)},$$

valid for all $\delta_1,\ \delta_2 \in [0,1]$, and by the Lipschitz property (4.3), we get (4.5), (4.6). ☐

Let $[0,\bar{S}]^2 \ni (s,s_1) \mapsto \tilde{M}(s,s_1) \in Lin(\mathbb{R}\times\mathbb{R}^n\times\mathbb{R};\mathbb{R}\times\mathbb{R}^n\times\mathbb{R})$ be the *fundamental matrix solution* of the *variational equation*

$$\frac{d\tilde{v}}{ds}(s) = D_{(y^0,y,\beta)}\begin{pmatrix} \bar{w}^0(s) \\ f(\bar{y}(s),\bar{\alpha}(s))\bar{w}^0(s) + \sum_{i=1}^m g_i(\bar{y}(s))\ \bar{w}^i(s) \\ |\bar{w}(s)| \end{pmatrix} \cdot \tilde{v}(s) \tag{4.7}$$

corresponding to the control system (1.2) (and to the reference process).[9] Let us use $\tilde{M}_{i,j}$ to denote the entry (i,j) of $\tilde{M}$, where i and j range from 0 to $n+1$, so that, for all $s \in [0,\bar{S}]$, one has

$$\begin{aligned} &\tilde{M}_{0,j}(s,s_1) = \tilde{M}_{j,0}(s,s_1) = \delta_{0,j} \quad \text{for } j = 0,\dots,n+1,\\ &\tilde{M}_{n+1,j}(s,s_1) = \tilde{M}_{j,n+1}(s,s_1) = \delta_{n+1,j} \quad \text{for } j = 0,\dots,n+1,\\ &\tilde{M}_{i,r}(s,s_1) = M_{i,r}(s,s_1) \text{ for } i,r = 1,\dots,n, \end{aligned}$$

where M denotes the fundamental matrix solution of the state-variational equation

$$\frac{dv}{ds}(s) = \left(D_x f(\bar{y}(s),\bar{\alpha}(s))w^0(s) + \sum_{i=1}^m Dg_i(\bar{y}(s))\ \bar{w}^i(s)\right)\cdot v(s). \tag{4.8}$$

[9] Namely, for each vector $\tilde{v} :\in \mathbb{R}^{1+n+1}$ and each $s_1 \in [0,\bar{S}]$, the function $\tilde{v}(\cdot) := \tilde{M}(\cdot,s_1)\tilde{v}$ is the solution of (4.7) with initial condition $\tilde{v}(s_1) = \tilde{v}$.

4.2.2. Needle variations. Let us recall the basic notion of needle variation, which is used to produce a first order tangent approximation of the reachable set $\mathscr{R}'^{r}_{\tilde{\mathscr{W}}}$ at $(\bar{y}^0, \bar{y}, \bar{\beta})(\bar{S})$.

DEFINITION 4.2. *Any* $\xi = (w^0, w, a) \in \mathbf{W}$ *will be called a* needle variation generator *or a* variation generator of length 1.

DEFINITION 4.3. *Let* $\xi = (w^0, w, a)$ *be a needle variation generator, and let* $\bar{s} \in]0, \bar{S}]$. *For any control* $(\hat{w}^0, \hat{w}, \hat{\alpha}) \in \tilde{\mathscr{W}}^{\bar{S}}$, *the family*

$$\left\{ (\hat{w}^0, \hat{w}, \hat{\alpha})^{\varepsilon}_{\xi,\bar{s}}(s) := \begin{cases} (w^0, w, \alpha) & \text{if } s \in [\bar{s}-\varepsilon, \bar{s}], \\ (\hat{w}^0, \hat{w}, \hat{\alpha})(s) & \text{if } s \in [0, \bar{s}-\varepsilon[\cup]\bar{s}, \bar{S}], \end{cases} \quad \varepsilon \in]0, \bar{s}[\right\}$$

is called a needle control approximation of $(\hat{w}^0, \hat{w}, \hat{\alpha})$ at $\bar{s}$ associated to ξ.

At a given time $\bar{s}$ and to every variation generator ξ, let us associate an infinitesimal variation of the reference extended sense trajectory $(\bar{y}^0, \bar{y}, \bar{\beta})$, whose y-component coincides with a standard needle variation.

DEFINITION 4.4. *Let us consider the (full measure) subset* $]0, \bar{S}]_{\mathrm{Leb}} \subset]0, \bar{S}]$ *of Lebesgue points*[10] *of* $s \mapsto \big(\bar{w}^0(s), f(\bar{y}(s), \bar{\alpha}(s))\bar{w}^0(s) + \sum_{i=1}^m g_i(\bar{y}(s))\bar{w}^i(s), |\bar{w}|(s)\big)$. *For every* $\bar{s} \in]0, \bar{S}]_{\mathrm{Leb}}$ *and every needle variation generator* $\xi = (w^0, w, a)$, *we call* needle variation *at* $\bar{s}$ *the vector* $(\mathbf{v}^0_{\xi,\bar{s}}, \mathbf{v}_{\xi,\bar{s}}, \mathbf{v}^v_{\xi,\bar{s}})$ *given by*

$$\begin{aligned} &\mathbf{v}^0_{\xi,\bar{s}} := w^0 - \bar{w}^0(\bar{s}), \\ &\mathbf{v}_{\xi,\bar{s}} := f(\bar{y}(\bar{s}), a)w^0 - f(\bar{y}(\bar{s}), \bar{\alpha}(\bar{s}))\bar{w}^0(s) + \sum_{i=1}^m g_i(\bar{y}(\bar{s}))\big(w^i - \bar{w}^i(\bar{s})\big), \\ &\mathbf{v}^v_{\xi,\bar{s}} := |w| - |\bar{w}(\bar{s})|. \end{aligned} \tag{4.9}$$

Standard continuity estimates imply the following fact.

LEMMA 4.2. *Assume that* $\bar{s} \in]0, \bar{S}]_{\mathrm{Leb}}$. *For every needle variation generator* $\xi = (w^0, w, a)$ *we get*

$$\begin{pmatrix} y^{0\varepsilon}(\bar{S}) - \bar{y}^0(\bar{S}) \\ y^{\varepsilon}(\bar{S}) - \bar{y}(\bar{S}) \\ \beta^{\varepsilon}(\bar{S}) - \bar{\beta}(\bar{S}) \end{pmatrix} = \varepsilon \tilde{M}(\bar{S}, \bar{s}) \cdot \begin{pmatrix} \mathbf{v}^0_{\xi,\bar{s}} \\ \mathbf{v}_{\xi,\bar{s}} \\ \mathbf{v}^v_{\xi,\bar{s}} \end{pmatrix} + h(\varepsilon) = \varepsilon \begin{pmatrix} \mathbf{v}^0_{\xi,\bar{s}} \\ M(\bar{S}, \bar{s}) \cdot \mathbf{v}_{\xi,\bar{s}} \\ \mathbf{v}^v_{\xi,\bar{s}} \end{pmatrix} + h(\varepsilon), \tag{4.10}$$

where $(y^{0\varepsilon}, y^{\varepsilon}, \beta^{\varepsilon}) := (y^0, y, \beta)[(\bar{w}^0, \bar{w}, \bar{\alpha})^{\varepsilon}_{\xi,\bar{s}}]$, $(\bar{w}^0, \bar{w}, \bar{\alpha})^{\varepsilon}_{\xi,\bar{s}}$ *is the needle control approximation of* $(\bar{w}^0, \bar{w}, \bar{\alpha})$ *at* $\bar{s}$ *associated to* ξ, *and* h *is continuous and verifies* $|h(\varepsilon)| \leq \varepsilon\rho(\varepsilon)$, *where* ρ *is a pseudo-modulus.*[11]

4.2.3. Bracket-like variations. To every admissible bracket pair $(B, \mathbf{h})$, with B of length ≥ 2, one can link a control which generates a trajectory locally approximating the iterated Lie bracket $B(\mathbf{h})$. Such a control can be defined using one of the approaches outlined in [1, 10, 11, 18, 19] and references therein.

PROPOSITION 4.1 (see [11]). *Assume hypothesis* $\mathbf{(H)}_b$. *Let* $(B, \mathbf{h})$ *be an admissible bracket pair in* $\mathfrak{Br}^0$ *of length* $l \geq 2$. *Then, for every point* $\tilde{x} \in \mathbb{R}^n$, *there exists*

[10] Given $G \in L^1([a,b], \mathbb{R}^N)$, $s \in (a,b]$ is a Lebesgue point if $\lim_{\delta\to 0^+} \frac{1}{\delta}\int_{s-\delta}^{(s+\delta)\wedge\bar{S}} |G(\sigma) - G(s)| d\sigma = 0$. By the Lebesgue differentiation theorem, the set of Lebesgue points has measure $b - a$.

[11] See subsection 2.3 for the definition of "pseudo-modulus."

$\bar\varepsilon > 0$ such that for any $\sigma \in]0, \bar\varepsilon^{1/l}]$, one can construct a piecewise constant control $w_{B,\sigma}$ such that

$$y_\sigma(\sigma) = \tilde x + \left(\frac{\sigma}{r_B}\right)^l B(\mathbf{h})(\tilde x) + o(\sigma^l), \tag{4.11}$$

where y_σ denotes the solution to the Cauchy problem $\frac{dy}{ds} = \sum_{i=1}^n g_i(y) w^i_{B,\sigma}$, $y(0) = \tilde x$, and r_B is the switch-number of the formal bracket B.

DEFINITION 4.5. *A bracket pair $\xi = (B, \mathbf{h}) \in \mathfrak{Br}^0$ of length l (≥ 2) will also be called a* bracket-like variation generator of length l.

Bracket-like approximation is a classical issue in geometric control theory (see, e.g., [1, 6, 7, 10, 18]). Here, we fully exploit the unboundedness of the controls u, which translates into the possibility of choosing $w^0 = 0$ *during the variation*: namely, the variation is implemented during an interval in which the (original) time is constant.

DEFINITION 4.6 (bracket-like variation). *For every $\bar s \in]0, \bar S]$ and every bracket-like variation generator $\xi = (B, \mathbf{h}) \in \mathfrak{Br}^0$ of length l, $l \geq 2$, we call* bracket-like variation *at $\bar s$, the vector*

$$(\mathbf{v}^0_{\xi,\bar s}, \mathbf{v}_{\xi,\bar s}) := \left(0, \frac{B(\mathbf{h})(\bar y(\bar s))}{(r_B)^l}\right).[12] \tag{4.12}$$

DEFINITION 4.7 (bracket-like approximation). *Fix $\bar s \in]0, \bar S]$, and let $\xi = (B, \mathbf{h})$ in $\mathfrak{Br}^0$ be a bracket-like variation generator of length l. Set $\check\varepsilon := \min\left\{\bar\varepsilon, \left(\frac{\bar s}{2}\right)^l\right\}$,[13] and for each $\varepsilon \in]0, \check\varepsilon]$, consider the dilation $\varphi^\varepsilon : [\bar s - 2\varepsilon^{1/l}, \bar s - \varepsilon^{1/l}] \to [\bar s - 2\varepsilon^{1/l}, \bar s]$ defined by setting, for all $\sigma \in [\bar s - 2\varepsilon^{1/l}, \bar s - \varepsilon^{1/l}]$,*

$$\varphi^\varepsilon(\sigma) := (\bar s - 2\varepsilon^{1/l}) + 2\big(\sigma - (\bar s - 2\varepsilon^{1/l})\big). \tag{4.13}$$

For any control $(\hat w^0, \hat w, \hat\alpha) \in \tilde{\mathscr{W}}^{\bar S}$, let us define

$$(\hat w^0, \hat w, \hat\alpha)^\varepsilon_{\xi,\bar s}(s) := \begin{cases} \left(2\,\hat w^0, 2\,\hat w, \hat\alpha\right) \circ \varphi^\varepsilon(s) & \text{if } s \in [\bar s - 2\varepsilon^{1/l}, \bar s - \varepsilon^{1/l}[, \\ \left(0, w_{\xi,\varepsilon^{1/l}}(s - (\bar s - \varepsilon^{1/l})), a\right) & \text{if } s \in [\bar s - \varepsilon^{1/l}, \bar s], \\ \big(\hat w^0, \hat w, \hat\alpha\big)(s) & \text{if } s \in [0, \bar s - 2\varepsilon^{1/l}[\cup]\bar s, S], \end{cases} \tag{4.14}$$

where $a \in A$ is arbitrary and $w_{\xi,\varepsilon^{1/l}}$ is defined as in Proposition 4.1. We refer to the family of controls $\left\{(\hat w^0, \hat w, \hat\alpha)^\varepsilon_{\xi,\bar s} : \varepsilon \in]0, \check\varepsilon]\right\}$ as a bracket-like control approximation of $(\hat w^0, \hat w, \hat\alpha)$ at $\bar s$ associated to $\xi = (B, \mathbf{h})$.[14]

LEMMA 4.3 (asymptotics of bracket-like variations). *Let us consider a bracket-like variation generator $\xi = (B, \mathbf{h}) \in \mathfrak{Br}^0$ of length l. For every $(\bar s, \varepsilon) \in]0, \bar S] \times]0, \check\varepsilon]$,*[15]

[12] We do not define a β-component of the bracket-like variation since when we implement a control $w_{B,\sigma}$, the increment in the (y^0, y)-direction is of order ε^l, while in the β-direction it is of order ε.

[13] $\bar\varepsilon$ is defined in Proposition 4.1.

[14] Observe that $(\hat w^0, \hat w, \hat\alpha)^\varepsilon_{\xi,\bar s} \in L^1([0, \bar S], \mathbf{W}')$ but it may not belong to the set $\tilde{\mathscr{W}}^{\bar S}$. However, it belongs to $\tilde{\mathscr{W}}^{\bar S}$ when $(\hat w^0, \hat w, \hat\alpha)$ is a canonical control.

[15] $\check\varepsilon$ is as in Definition 4.7.

consider the bracket-like control approximation $(\bar w^0,\bar w,\bar\alpha)^\varepsilon_{\xi,\bar s}$ *of* $(\bar w^0,\bar w,\bar\alpha)$ *at* $\bar s$ *associated to* ξ. *Then, setting* $(y^{0\varepsilon},y^\varepsilon,\beta^\varepsilon):=(y^0,y,\beta)[(\bar w^0,\bar w,\bar\alpha)^\varepsilon_{\xi,\bar s}]$, *one has*

$$\begin{aligned}
&y^{0\varepsilon}(\bar S)-\bar y^0(\bar S)=\varepsilon\mathbf{v}^0_{\xi,\bar s}=0,\\
&y^\varepsilon(\bar S)-\bar y(\bar S)=\varepsilon M(\bar S,\bar s)\cdot\mathbf{v}_{\xi,\bar s}+h(\varepsilon)=\varepsilon M(\bar S,\bar s)\cdot\frac{B(\mathbf{h})(\bar y(\bar s))}{(r_B)^l}+h(\varepsilon),\\
&\beta^\varepsilon(\bar S)-\bar\beta(\bar S)=\varepsilon^{\frac1l},
\end{aligned}$$

with $h:[0,\check\varepsilon]\mapsto\mathbb{R}^n$ *continuous and verifying* $|h(\varepsilon)|\le\varepsilon\rho(\varepsilon)$ $\forall\varepsilon\in[0,\check\varepsilon]$ *for some pseudo-modulus* ρ.

Proof. The results concerning $y^{0\varepsilon}$ and β^ε are trivial. As for y^ε, it is not difficult to prove that (see [3, Lemma 6.9])

$$y^\varepsilon(\bar S)-\bar y(\bar S)=\varepsilon\, M(\bar S,\bar s)\cdot\frac{B(\mathbf{h})(\bar y(\bar s))}{(r_B)^l}+h(\varepsilon),\qquad h(\varepsilon)=o(\varepsilon). \tag{4.15}$$

Hence, if we set $\rho(\varepsilon):=\frac1\varepsilon\max_{\eta\in[0,\varepsilon]}|h(\eta)|$ for any $\varepsilon\in]0,\check\varepsilon]$, $\rho(0):=\lim_{\varepsilon\to0^+}\rho(\varepsilon)=0$, and $\rho(\varepsilon):=+\infty$ for any $\varepsilon>\check\varepsilon$, ρ turns out to be a pseudo-modulus and $h:[0,\check\varepsilon]\to\mathbb{R}^n$ verifies $|h(\varepsilon)|\le\varepsilon\rho(\varepsilon)$ for any $\varepsilon\in[0,\check\varepsilon]$. To prove the continuity of $\varepsilon\mapsto h(\varepsilon)$, let us begin observing that it is equivalent to the continuity of the map $\varepsilon\mapsto y^\varepsilon(\bar S)$.

The right continuity at $\varepsilon_0=0$ is straightforward. Let us show that $\varepsilon\mapsto y^\varepsilon(\bar S)$ is right continuous at any $\varepsilon_0\in]0,\check\varepsilon[$. By the continuity of the input-output map, in order to show that $\lim_{\varepsilon\to\varepsilon_0^+}y^\varepsilon(\bar S)=y^{\varepsilon_0}(\bar S)$ it is enough to estimate, for every $\varepsilon\in]\varepsilon_0,\check\varepsilon]$, the L^1-distance $\int_0^{\bar S}|(\bar w^0,\bar w,\bar\alpha)^\varepsilon_{\xi,\bar s}(s)-(\bar w^0,\bar w,\bar\alpha)^{\varepsilon_0}_{\xi,\bar s}(s)|\,ds$. Let us observe that

$$\begin{aligned}
&\int_0^{\bar S}|(\bar w^0,\bar w,\bar\alpha)^\varepsilon_{\xi,\bar s}(s)-(\bar w^0,\bar w,\bar\alpha)^{\varepsilon_0}_{\xi,\bar s}(s)|\,ds=\int_{\bar s-2\varepsilon^{1/l}}^{\bar s}|\bar w^\varepsilon_{\xi,\bar s}(s)-\bar w^{\varepsilon_0}_{\xi,\bar s}(s)|\,ds\\
&=\int_{\bar s-2\varepsilon^{1/l}}^{\bar s-2\varepsilon_0^{1/l}}|2\bar w\circ\varphi^\varepsilon(s)-\bar w(s)|\,ds+\int_{\bar s-2\varepsilon_0^{1/l}}^{\bar s-\varepsilon^{1/l}}2|\bar w\circ\varphi^\varepsilon(s)-\bar w\circ\varphi^{\varepsilon_0}(s)|\,ds\\
&\quad+\int_{\bar s-\varepsilon^{1/l}}^{\bar s-\varepsilon_0^{1/l}}|w_\varepsilon(s)-2\bar w\circ\varphi^{\varepsilon_0}(s)|\,ds+\int_{\bar s-\varepsilon_0^{1/l}}^{\bar s}|w_\varepsilon(s)-w_{\varepsilon_0}(s)|\,ds,
\end{aligned}\tag{4.16}$$

where we have set $w_\varepsilon(s):=w_{\xi,\varepsilon^{1/l}}(s-(\bar s-\varepsilon^{1/l}))$ and $w_{\varepsilon_0}(s):=w_{\xi,\varepsilon_0^{1/l}}(s-(\bar s-\varepsilon_0^{1/l}))$ for a.e. s. Each of the first and the third terms in the right-hand side of (4.16) is clearly bounded above by $6|\varepsilon-\varepsilon_0|$. As for the remaining terms, let us choose a continuous L^1-approximation $\bar w^c$ of $\bar w$ such that $\int_0^{\bar s}|\bar w^c(s)-\bar w(s)|ds<|\varepsilon_0^{1/l}-\varepsilon^{1/l}|$. Let ρ_c be the modulus of continuity of $\bar w^c$. Then, the second term on the right-hand side of (4.16) can be estimated as

$$\begin{aligned}
&\int_{\bar s-2\varepsilon_0^{1/l}}^{\bar s-\varepsilon^{1/l}}2|\bar w\circ\varphi^\varepsilon(s)-\bar w\circ\varphi^{\varepsilon_0}(s)|\,ds\le\int_{\bar s-2\varepsilon_0^{1/l}}^{\bar s-\varepsilon^{1/l}}2|\bar w\circ\varphi^\varepsilon(s)-\bar w^c\circ\varphi^\varepsilon(s)|\,ds\\
&+\int_{\bar s-2\varepsilon_0^{1/l}}^{\bar s-\varepsilon^{1/l}}2|\bar w^c\circ\varphi^\varepsilon(s)-\bar w^c\circ\varphi^{\varepsilon_0}(s)|\,ds+\int_{\bar s-2\varepsilon_0^{1/l}}^{\bar s-\varepsilon^{1/l}}2|\bar w^c\circ\varphi^{\varepsilon_0}(s)-\bar w\circ\varphi^{\varepsilon_0}(s)|\,ds\\
&=\int_{\bar s+2\varepsilon^{1/l}-4\varepsilon_0^{1/l}}^{\bar s}|\bar w(s)-\bar w^c(s)|\,ds+\int_{\bar s-2\varepsilon_0^{1/l}}^{\bar s-\varepsilon^{1/l}}2|\bar w^c\circ\varphi^\varepsilon(s)-\bar w^c\circ\varphi^{\varepsilon_0}(s)|\,ds\\
&+\int_{\bar s-2\varepsilon_0^{1/l}}^{\bar s-2\left(\varepsilon^{1/l}-\varepsilon_0^{1/l}\right)}|\bar w(s)-\bar w^c(s)|\,ds\le2|\varepsilon_0^{1/l}-\varepsilon^{1/l}|+\rho_c(|\varepsilon_0-\varepsilon|).
\end{aligned}$$

Let us examine the fourth term on the right-hand side of (4.16). One has

$$\int_{\bar s-\varepsilon_0^{1/l}}^{\bar s} |w_\varepsilon(s)-w_{\varepsilon_0}(s)|\,ds = \int_0^{\varepsilon_0^{1/l}} \left|w_{\xi,\varepsilon^{1/l}}\left(\sigma+\varepsilon^{1/l}-\varepsilon_0^{1/l}\right)-w_{\xi,\varepsilon_0^{1/l}}(\sigma)\right|d\sigma$$
$$= \int_0^{\varepsilon_0^{1/l}} \left|w_{\xi,\varepsilon_0^{1/l}}\left(\frac{\varepsilon_0}{\varepsilon}\left(\sigma+\varepsilon^{1/l}-\varepsilon_0^{1/l}\right)\right)-w_{\xi,\varepsilon_0^{1/l}}(\sigma)\right|d\sigma,$$

where, in turn, we have changed the integral variable $\sigma = s-(\bar s-\varepsilon_0^{1/l})$ and used the two (equivalent) relations

$$(4.17)\qquad w_\varepsilon(s) = w_{\varepsilon_0}\left(s\frac{\varepsilon_0}{\varepsilon}\right)\quad \forall\, s\in[0,\varepsilon],\qquad w_{\varepsilon_0}(s) = w_\varepsilon\left(s\frac{\varepsilon}{\varepsilon_0}\right)\quad \forall\, s\in[0,\varepsilon_0].$$

Let $w^c_{\varepsilon_0}$ be a continuous map satisfying the relation $\int_0^{\bar s}|w^c_{\varepsilon_0}(\sigma)-w_{\xi,\varepsilon_0^{1/l}}(\sigma)|ds < |\varepsilon_0^{1/l}-\varepsilon^{1/l}|$, and let ρ_{c,ε_0} be a modulus of continuity for $w^c_{\varepsilon_0}$. Then, we get

$$\int_{\bar s-\varepsilon_0^{1/l}}^{\bar s} |w_\varepsilon(s)-w_{\varepsilon_0}(s)|\,ds$$
$$\le 2\left(\frac{\varepsilon_0}{\varepsilon}\right)^{1/l}|\varepsilon_0^{1/l}-\varepsilon^{1/l}| + \rho_{c,\varepsilon_0}\left(\left(\frac{\varepsilon_0}{\varepsilon}\right)^{1/l}|\varepsilon_0^{1/l}-\varepsilon^{1/l}|+|\varepsilon_0^{1/l}-\varepsilon^{1/l}|\right),$$

where $\frac{\varepsilon_0}{\varepsilon}<1$. From the previous relations we deduce that the map $\varepsilon\mapsto y^\varepsilon(\bar s)$ is right continuous at any $\varepsilon_0>0$.

It remains to show that the function $\varepsilon\mapsto y^\varepsilon(\bar S)$ is left continuous at any $\varepsilon_0\in]0,\check\varepsilon]$. To this aim, one can proceed similarly as above by writing for any $\varepsilon\in\left]\frac{\varepsilon_0}{2},\varepsilon_0\right[$ a relation similar to (4.16), in which the roles of ε and ε_0 are interchanged. The resulting estimates are very similar to the previous ones, with exception of the term

$$(4.18)\qquad \begin{aligned}\int_{\bar s-\varepsilon^{1/l}}^{\bar s} |w_\varepsilon(s)-w_{\varepsilon_0}(s)|\,ds &= \int_0^{\varepsilon^{1/l}} \left|w_{\xi,\varepsilon^{1/l}}(\sigma)-w_{\xi,\varepsilon_0^{1/l}}(\sigma+\varepsilon_0^{1/l}-\varepsilon^{1/l})\right|d\sigma\\ &= \int_0^{\varepsilon^{1/l}} \left|w_{\xi,\varepsilon_0^{1/l}}\left(\frac{\varepsilon_0}{\varepsilon}\sigma\right)-w_{\xi,\varepsilon_0^{1/l}}(\sigma+\varepsilon_0^{1/l}-\varepsilon^{1/l})\right|d\sigma,\end{aligned}$$

where, in turn, we have used the change of variable $\sigma = s-(\bar s-\varepsilon^{1/l})$ and relation (4.17). If $w^c_{\varepsilon_0}$ is a continuous L^1-approximation of $w_{\xi,\varepsilon_0^{1/l}}$ satisfying the relation $\int_0^{\bar s}|w^c_{\varepsilon_0}(\sigma)-w_{\xi,\varepsilon_0^{1/l}}(\sigma)|ds < |\varepsilon_0^{1/l}-\varepsilon^{1/l}|$ and $\tilde\rho_{c,\varepsilon_0}$ is a modulus of continuity for $w^c_{\varepsilon_0}$, one can obtain an estimate akin to (4.18), where now $\frac{\varepsilon_0}{\varepsilon}<2$. This concludes the proof of the continuity of the map $\varepsilon\mapsto y^\varepsilon(\bar S)$ (and hence of $\varepsilon\mapsto h(\varepsilon)$) for any $\varepsilon\in[0,\check\varepsilon]$. □

4.2.4. Composition of several variations. Let us define *the family of variation generators* as the set

$$\Xi := \mathbf{W}\cup\mathfrak{Br}^0.$$

Let us fix $(\xi,\bar s)\in\Xi\times]0,\bar S]$ and, for an $\tilde\varepsilon\in]0,1]$ small enough and any $\varepsilon\in]0,\tilde\varepsilon]$, let us introduce the operator

$$\begin{aligned}\mathscr{A}^\varepsilon_{\xi,\bar s}: L^\infty\left([0,\bar S],\mathbb{R}_+\times\mathscr{C}\times A\right) &\to L^\infty\left([0,\bar S],\mathbb{R}_+\times\mathscr{C}\times A\right)\\ (w^0,w,\alpha) &\mapsto \mathscr{A}^\varepsilon_{\xi,\bar s}(w^0,w,\alpha) := (w^0,w,\alpha)^\varepsilon_{\xi,\bar s}.\end{aligned}$$

LEMMA 4.4. *Let N be a positive integer, and let us consider an N-tuple of variation generators $\vec{\xi} := (\xi_1, \dots, \xi_N) \in \Xi^N$ of lengths $(l_1, \dots, l_N)$. Fix $(\bar{s}_1, \dots, \bar{s}_N) \in]0, \bar{S}]^N$, where $0 =: \bar{s}_0 < \bar{s}_1 < \cdots < \bar{s}_N \le \bar{S}$ and $\bar{s}_j \in]0, \bar{S}]_{\rm Leb}$ as soon as $l_j = 1$. For each $\vec{\varepsilon} := (\varepsilon_1, \dots, \varepsilon_N) \in]0, \tilde{\varepsilon}[^N$ for some $\tilde{\varepsilon} \in]0, 1[$ small enough,*[16] *set*

$$(w^{0\vec{\varepsilon}}, w^{\vec{\varepsilon}}, \alpha^{\vec{\varepsilon}}) := \mathscr{A}^{\varepsilon_N}_{\xi_N, \bar{s}_N} \circ \cdots \circ \mathscr{A}^{\varepsilon_j}_{\xi_j, \bar{s}_j} \circ \cdots \circ \mathscr{A}^{\varepsilon_1}_{\xi_1, \bar{s}_1} (\bar{w}^0, \bar{w}, \bar{\alpha}),^{17}$$

and let $\big(\bar{S}, w^{0\vec{\varepsilon}}, w^{\vec{\varepsilon}}, \alpha^{\vec{\varepsilon}}, y^{0\vec{\varepsilon}}, y^{\vec{\varepsilon}}, \beta^{\vec{\varepsilon}}\big)$ denote the corresponding process of the extended system (1.2). *Then, there exist a pseudo-modulus ρ and a continuous map $h : [0, \tilde{\varepsilon}]^N \to \mathbb{R}^{1+n}$ such that for every $s \in]\bar{s}_N, \bar{S}]$ and every $\vec{\varepsilon} \in]0, \tilde{\varepsilon}[^N$, one has*

$$\begin{pmatrix} y^{0\vec{\varepsilon}}(s) - \bar{y}^0(s) \\ y^{\vec{\varepsilon}}(s) - \bar{y}(s) \end{pmatrix} = \sum_{j=1}^{N} \varepsilon_j \begin{pmatrix} \mathbf{v}^0_{\xi_j, \bar{s}_j} \\ M(s, \bar{s}_j)\mathbf{v}_{\xi_j, \bar{s}_j} \end{pmatrix} + h(\vec{\varepsilon}), \quad |h(\vec{\varepsilon})| \le |\vec{\varepsilon}|\rho(|\vec{\varepsilon}|), \tag{4.19}$$

and

$$\beta^{\vec{\varepsilon}}(s) - \bar{\beta}(s) = \sum_{j \in I_1} \varepsilon_j \left(|w_j| - |\bar{w}(\bar{s}_j)|\right) + |\vec{\varepsilon}|\rho(|\vec{\varepsilon}|) + \sum_{j \in \{1, \dots, N\} \setminus I_1} (\varepsilon_j)^{\frac{1}{l_j}}, \tag{4.20}$$

where $I_1 := \{j = 1, \dots, N : \; l_j = 1\}$. In particular, if all ξ_j are needle variations, i.e. $\xi_j := (w^0_j, w_j, \alpha_j)$ for every $j = 1, \dots, N$, then

$$\beta^{\vec{\varepsilon}}(s) - \bar{\beta}(s) = \sum_{j=1}^{N} \varepsilon_j \big(|w_j| - |\bar{w}(\bar{s}_j)|\big) + |\vec{\varepsilon}|\rho(|\vec{\varepsilon}|).$$

Proof. Apart from the specification on the continuity of the function h, this result is proved in [3, Lemma 6.10]. Instead, the continuity of h is a straightforward consequence of the continuity properties established in Lemmas 4.2 and 4.3. ∎

4.2.5. Approximating the original reachable set by higher order extended cones. Let $\mathbf{Pr} : \mathbb{R} \times \mathscr{M} \times \mathbb{R} \to \mathbb{R} \times \mathscr{M}$ denote the *projection* defined by setting $\mathbf{Pr}(y^0, y, \beta) := (y^0, y)$ for all $(y^0, y, \beta) \in \mathbb{R} \times \mathscr{M} \times \mathbb{R}$.

For any $r > 0$, let us introduce the projections $\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}_+}$ and $\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}}$ of the reachable set and the extended reachable set at fixed time $\bar{S}$, respectively,

$$\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}_+} := \mathbf{Pr}({\mathscr{R}'}^r_{\tilde{\mathscr{W}}^{\bar{S}}_+}), \quad \mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}} := \mathbf{Pr}({\mathscr{R}'}^r_{\tilde{\mathscr{W}}^{\bar{S}}}).$$

DEFINITION 4.8. *Let $N > 0$ be an integer, let $\vec{\xi} := (\xi_1, \dots, \xi_N) \in \Xi^N$ be an N-tuple of variation generators, and fix $(\bar{s}_1, \dots, \bar{s}_N) \in]0, \bar{S}]^N$ as in Lemma* 4.4. *The convex cone in $\mathbb{R}^{1+n}$,*

$$\mathbf{R}_{\vec{\xi}} := \text{span}^+ \left\{ \begin{pmatrix} \mathbf{v}^0_{\xi_j, \bar{s}_j} \\ M(\bar{S}, \bar{s}_j) \cdot \mathbf{v}_{\xi_j, \bar{s}_j} \end{pmatrix} : \; j = 1, \dots, N \right\} \subseteq \mathbb{R}^{1+n}, \tag{4.21}$$

will be called a higher order extended variational cone *corresponding to the feasible extended sense process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$. In the case when all N variations are*

[16] Precisely, we require that $\tilde{\varepsilon} \le \big(\frac{\bar{s}_1 - \bar{s}_0}{2^{(l_1 - 1) \wedge 1}} \wedge \bar{\varepsilon}^{1/l_1}\big) \wedge \cdots \wedge \big(\frac{\bar{s}_N - \bar{s}_{N-1}}{2^{(l_N - 1) \wedge 1}} \wedge \bar{\varepsilon}^{1/l_N}\big)$, where $\bar{\varepsilon}$ is as in Proposition 4.1.

[17] Since $(\bar{w}^0, \bar{w}) \in L^1([0, \bar{S}], \mathbf{C})$ and the intervals $[\bar{s}_j - 2^{(l_j - 1) \wedge 1} \varepsilon_j^{1/l_j}, \bar{s}_j]$ are disjoint by the choice of $\tilde{\varepsilon}$, the control $(w^{0\vec{\varepsilon}}, w^{\vec{\varepsilon}}, \alpha^{\vec{\varepsilon}})$ turns out to belong to $\tilde{\mathscr{W}}_{\bar{S}}$.

needle variations, i.e., $\xi_j = (w_j^0, w_j, a_j)$, $j = 1, \dots N$, *we can also define the standard (first order) extended variational cone in* $\mathbb{R}^{1+n+1}$,

$$\mathbf{R}'_{\vec{\xi}} := \mathrm{span}^+ \left\{ \begin{pmatrix} \mathbf{v}^0_{\xi_j,\bar{s}_j} \\ M(\bar{S},\bar{s}_j)\cdot \mathbf{v}_{\xi_j,\bar{s}_j} \\ \mathbf{v}^v_{\xi_j,\bar{s}_j} \end{pmatrix} : \ j = 1,\dots N \right\} \subseteq \mathbb{R}^{1+n+1}. \tag{4.22}$$ [18]

In Theorem 4.1 below we establish that $\mathbf{R}_{\vec{\xi}}$ is a QDQ approximating cone to the union of $\{(\bar{y}^0,\bar{y})(\bar{S})\}$ with the reachable set $\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}_+}$. Though we already know that $\mathbf{R}_{\vec{\xi}}$ is a QDQ (actually, a Boltyanski) approximating cone to the extended reachable set $\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}}$ (see [3]), the fact that $\mathbf{R}_{\vec{\xi}}$ *approximates* $\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}_+} \cup \{(\bar{y}^0,\bar{y})(\bar{S})\}$ as well is trivial because of the strict inclusion $\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}_+} \cup \{(\bar{y}^0,\bar{y})(\bar{S})\} \subsetneq \mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}}$.

THEOREM 4.1. *Let* $(\bar{S},\bar{w}^0,\bar{w},\bar{\alpha},\bar{y}^0,\bar{y},\bar{\beta})$ *be the reference canonical, feasible extended sense process at which there is a local infimum gap such that* $\bar{\beta}(\bar{S}) < K$. *Then, for some* $r > 0$ *the higher order extended variational cone* $\mathbf{R}_{\vec{\xi}}$ *is a* $(\bar{y}^0,\bar{y})(\bar{S})$*-ignoring QDQ approximating cone to* $\mathscr{R}^r_{\tilde{\mathscr{W}}^{\bar{S}}_+} \cup \{(\bar{y}^0,\bar{y})(\bar{S})\}$ *at* $(\bar{y}^0,\bar{y})(\bar{S})$.

Remark 4.9. The fact that the QDQ approximating cone $\mathbf{R}_{\vec{\xi}}$ is $(\bar{y}^0,\bar{y})(\bar{S})$-ignoring is crucial in order to apply the set separation result of Corollary 2.1 in the proof of Theorem 3.1.

Remark 4.10. Whenever no variation ξ_j is bracket-like, then one can establish a result like Theorem 4.1 for the whole variable (y^0, y, β) as well. Indeed, if all ξ_j are needle variation generators, the (first order) extended variational cone $\mathbf{R}'_{\vec{\xi}} \subset \mathbb{R}^{1+n+1}$ turns out to be a $(\bar{y}^0,\bar{y},\bar{\beta})(\bar{S})$-ignoring, QDQ approximating cone to $\mathscr{R}'^r_{\tilde{\mathscr{W}}^{\bar{S}}_+} \cup \{(\bar{y}^0,\bar{y},\bar{\beta})(\bar{S})\}$ at $(\bar{y}^0,\bar{y},\bar{\beta})(\bar{S})$. This can be straightforwardly deduced from a general result established in [26] by means of the notion of "abundance."

Proof of Theorem 4.1. For some positive integer $N > 1$, let $\vec{\xi} := (\xi_1,\dots,\xi_N) \in \Xi^N$ be an N-tuple of variation generators and fix $(\bar{s}_1,\dots,\bar{s}_N) \in]0,\bar{S}]^N$ as in Definition 4.4. Set $\bar{\delta} := \tilde{\varepsilon}$, $\Gamma := [0,+\infty[^N$, and for any $\vec{\varepsilon} := (\varepsilon_1,\dots,\varepsilon_N) \in B_{\bar{\delta}} \cap \Gamma$, set

$$(\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}} := (\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}}_{\vec{\xi},\bar{s}}, \qquad (y^{0^{\vec{\varepsilon}}}, y^{\vec{\varepsilon}}) := (y^0,y)[(\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}}].$$

By Lemma 4.4, we have

$$\begin{pmatrix} y^{0^{\vec{\varepsilon}}}(\bar{S}) - \bar{y}^0(\bar{S}) \\ y^{\vec{\varepsilon}}(\bar{S}) - \bar{y}(\bar{S}) \end{pmatrix} = L\cdot\vec{\varepsilon} + h(\vec{\varepsilon}) \qquad \forall \vec{\varepsilon} := (\varepsilon_1,\dots,\varepsilon_N) \in B_{\bar{\delta}} \cap \Gamma, \tag{4.23}$$

where the linear operator $L \in Lin(\mathbb{R}^N,\mathbb{R}^n)$ is defined as

$$L := \begin{pmatrix} \mathbf{v}^0_{\xi_1,\bar{s}_1} & \cdots\cdots & \mathbf{v}^0_{\xi_N,\bar{s}_N} \\ M(\bar{S},\bar{s}_1)\mathbf{v}_{\xi_1,\bar{s}_1} & \cdots\cdots & M(\bar{S},\bar{s}_N)\mathbf{v}_{\xi_N,\bar{s}_N} \end{pmatrix},$$

and h is a continuous function which, for some pseudo-modulus ρ, verifies $|h(\vec{\varepsilon})| \le |\vec{\varepsilon}|\rho(|\vec{\varepsilon}|)$ for all $\vec{\varepsilon} \in B_{\bar{\delta}} \cap \Gamma$. Let us now define the set-valued map $G : B_{\bar{\delta}} \cap \Gamma \rightsquigarrow \mathbb{R}^{1+n}$ by setting

$$G(\vec{\varepsilon}) := \left\{ \begin{pmatrix} y^0\left[\theta_{(\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}}}(\delta^2)\right](\bar{S}) \\ y\left[\theta_{(\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}}}(\delta^2)\right](\bar{S}) \end{pmatrix} : \ \delta \in]0,\bar{\delta}] \right\} \qquad \forall \vec{\varepsilon} \in B_{\bar{\delta}} \cap \Gamma,$$

[18] In this case, one clearly has $\mathbf{R}_{\vec{\xi}} = \mathbf{Pr}(\mathbf{R}'_{\vec{\xi}})$.

where the function $r \mapsto \theta_{(\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}}}(r)$ is defined as in Lemma 4.1. It follows from the latter that the map

$$\hat{h}_\delta(\vec{\varepsilon}) := \begin{pmatrix} y^0\left[\theta_{(\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}}}(\delta^2)\right](\bar{S}) - y^{0^{\vec{\varepsilon}}}(\bar{S}) \\ y\left[\theta_{(\bar{w}^0,\bar{w},\bar{\alpha})^{\vec{\varepsilon}}}(\delta^2)\right](\bar{S}) - y^{\vec{\varepsilon}}(\bar{S}) \end{pmatrix}$$

satisfies the relation

$$\left|\hat{h}_\delta(\vec{\varepsilon})\right| \le E\delta^2 \qquad \forall \delta \in]0,\bar{\delta}], \qquad \forall \vec{\varepsilon} \in B_{\bar{\delta}} \cap \Gamma$$

for a suitable constant $E \ge 0$. Furthermore, using Lemmas 4.2 and 4.3 and standard ODE results about the continuous dependence of solutions on initial data, one deduces that the mapping $\vec{\varepsilon} \mapsto \hat{h}_\delta(\vec{\varepsilon})$ is continuous at each $\vec{\varepsilon} \in B_{\bar{\delta}} \cap \Gamma$. Then, Lemma 4.1 and (4.23) imply that for any $\delta \in]0,\bar{\delta}]$ and for every $\vec{\varepsilon} \in B_\delta \cap \Gamma$,

$$\begin{pmatrix} \bar{y}^0(\bar{S}) \\ \bar{y}(\bar{S}) \end{pmatrix} + L \cdot \vec{\varepsilon} + h_\delta(\vec{\varepsilon}) \in G(\vec{\varepsilon}),$$

where $h_\delta(\vec{\varepsilon}) := h(\vec{\varepsilon}) + \hat{h}_\delta(\vec{\varepsilon})$ is continuous and verifies

$$|h_\delta(\vec{\varepsilon})| \le |\vec{\varepsilon}|\rho(|\vec{\varepsilon}|) + E\delta^2 \le \delta\left(\rho(\delta) + E\delta\right) \quad \forall \vec{\varepsilon} \in B_\delta \cap \Gamma.$$

Setting $\Lambda = \{L\}$ and extending arbitrarily G to a set-valued map $G : \mathbb{R}^N \rightsquigarrow \mathbb{R}^{1+n}$, the previous arguments show that for every $r > 0$, $\mathbf{R}_{\vec{\xi}}$ is a QDQ approximating cone to $\mathscr{R}^r_{\tilde{\mathscr{W}}_+^{\bar{S}}} \cup \{(\bar{y}^0,\bar{y})(\bar{S})\}$ at $(\bar{y}^0,\bar{y})(\bar{S})$, generated by the triple (G,Γ,Λ). Furthermore, by the fact that $(\bar{S},\bar{w}^0,\bar{w},\bar{\alpha},\bar{y}^0,\bar{y},\bar{\beta})$ is an isolated process, there exists some $r > 0$ such that $\mathscr{R}'^r_{\tilde{\mathscr{W}}_+^{\bar{S}}} \cap (\mathfrak{T} \times [0,K]) = \emptyset$ (see (4.1)). In addition, since we are assuming $\bar{\beta}(\bar{S}) < K$, we can choose this r so that the projection $\mathscr{R}^r_{\tilde{\mathscr{W}}_+^{\bar{S}}}$ satisfies

$$\mathscr{R}^r_{\tilde{\mathscr{W}}_+^{\bar{S}}} \cap \mathfrak{T} = \emptyset. \tag{4.24}$$

In view of Lemma 4.1, for such r (by reducing $\bar{\delta}$, if necessary) one has $G(B_{\bar{\delta}} \cap \Gamma) \subset \mathscr{R}^r_{\tilde{\mathscr{W}}_+^{\bar{S}}}$. Hence, $(\bar{y}^0,\bar{y})(\bar{S})$, which is in $\mathfrak{T}$ by definition, cannot belong to $G(B_{\bar{\delta}} \cap \Gamma)$ and the QDQ approximating cone $\mathbf{R}_{\vec{\xi}}$ is $(\bar{y}^0,\bar{y})(\bar{S})$-ignoring (see also Remark 2.9). ◻

4.3. Conclusion of the proof of Theorem 3.1. Let us recall that the extended sense process $(\bar{S},\bar{w}^0,\bar{w},\bar{\alpha},\bar{y}^0,\bar{y},\bar{\beta})$ is canonical, isolated, feasible, and such that $\bar{\beta}(\bar{S}) < K$. Let $N > 0$ be an integer, let $\vec{\xi} := (\xi_1,\dots,\xi_N) \in \Xi^N$ be an N-tuple of variation generators and fix $(\bar{s}_1,\dots,\bar{s}_N) \in]0,\bar{S}]^N$ as in Definition 4.4. As observed in the proof of Theorem 4.1, for some $r > 0$ the condition (4.24) holds true, namely, the sets $\mathfrak{T}$ and $\mathscr{R}^r_{\tilde{\mathscr{W}}_+^{\bar{S}}} \cup \{(\bar{y}^0,\bar{y})(\bar{S})\}$ are locally separated at $(\bar{y}^0,\bar{y})(\bar{S})$. Let $\mathscr{K}$ be a QDQ approximating cone to the target set $\mathfrak{T}$ at $(\bar{y}^0,\bar{y})(\bar{S})$, as in the statement of Theorem 3.1. Since, by Theorem 4.1, the higher order extended variational cone $\mathbf{R}_{\vec{\xi}}$ is a $(\bar{y}^0,\bar{y})(\bar{S})$-ignoring QDQ approximating cone to $\mathscr{R}^r_{\tilde{\mathscr{W}}_+^{\bar{S}}} \cup \{(\bar{y}^0,\bar{y})(\bar{S})\}$ at $(\bar{y}^0,\bar{y})(\bar{S})$, from Corollary 2.1 it follows that the convex cones $\mathscr{K}$ and $\mathbf{R}_{\vec{\xi}}$ are not transversal. As a consequence, in view of Proposition 2.2 they are linearly separated, namely, there exists $(\zeta_0,\zeta) \in (T_{(\bar{y}^0,\bar{y})(\bar{S})}(\mathbb{R} \times \mathscr{M}))^* \backslash \{0\}$ such that $(\zeta_0,\zeta) \in -\mathscr{K}^\perp \cap \mathbf{R}^\perp_{\vec{\xi}}$. Setting

$p_0 := \zeta_0$, $p(s) := \zeta \cdot M(\bar{S}, s)$ for every $s \in [0, \bar{S}]$, where $M(\bar{S}, s)$ is the fundamental matrix of the state-variational equation defined in (4.8), we get

$$(p_0, p) \neq 0, \qquad (p_0, p(\bar{S})) \in -\mathscr{K}^{\perp},$$

and

$$\frac{dp}{ds}(s) = -p(s) \cdot \left(D_x f(\bar{y}(s), \bar{\alpha}(s)) \bar{w}^0(s) + \sum_{i=1}^{m} D g_i(\bar{y}(s)) \bar{w}^i(s) \right).$$

When coupled with the dynamics, the last equation can be expressed as the Hamiltonian system $\frac{d}{ds}(\bar{y}, p)(s) = \mathbf{X}_{\bar{H}}(s, (\bar{y}, p)(s))$ for almost every $s \in [0, \bar{S}]$. Moreover, from $(\zeta_0, \zeta) \in \left(\mathbf{R}_{\vec{\xi}}\right)^{\perp}$ it follows that, for every $j = 1, \ldots, N$,

$$0 \geq \zeta_0 \, \mathbf{v}^0_{\xi_j, \bar{s}_j} + \zeta \cdot \left(M(\bar{S}, \bar{s}_j) \cdot \mathbf{v}_{\xi_j, \bar{s}_j} \right) = p_0 \, \mathbf{v}^0_{\xi_j, \bar{s}_j} + p(\bar{s}_j) \cdot \mathbf{v}_{\xi_j, \bar{s}_j}. \tag{4.25}$$

Therefore, the lift $(\bar{y}, p)$ and the multipliers $p_0 = \zeta_0$, $\pi = 0$, and $\lambda = 0$ satisfy the nontriviality condition (3.3), the nontransversality condition (3.5), and the adjoint equation (3.7) of Definition 3.5. Moreover, for a needle variation generator $\xi_j = (w^0_j, w_j, a_j)$, by (4.25) we get

$$H\Big(\bar{y}(\bar{s}_j), p(\bar{s}_j), p_0, 0, w^0_j, w_j, a_j\Big) - H\Big(\bar{y}(\bar{s}_j), p(\bar{s}_j), p_0, 0, \bar{w}^0(\bar{s}_j), \bar{w}(\bar{s}_j), \bar{\alpha}(\bar{s}_j)\Big) \leq 0, \tag{4.26}$$

while for a bracket-like variation generator $\xi_j = (B_j, \mathbf{h}_j)$, we obtain

$$p(\bar{s}_j) \cdot B_j(\mathbf{h}_j)(\bar{y}(\bar{s}_j)) \leq 0. \tag{4.27}$$

So far, we have obtained the maximality condition (3.8), the condition (3.9), and the higher order conditions (3.11) for every finite set of variation generators $\vec{\xi} := (\xi_1, \ldots, \xi_N) \in \Xi^N$ of lengths $(l_1, \ldots, l_N)$ and of times $(\bar{s}_1, \ldots, \bar{s}_N) \in]0, \bar{S}]^N$ verifying $0 =: \bar{s}_0 < \bar{s}_1 < \cdots < \bar{s}_N \leq \bar{S}$, with $\bar{s}_j \in]0, \bar{S}]_{\mathrm{Leb}}$ as soon as $l_j = 1$. It remains to prove (i) the validity of (4.26) for every control $(w^0_j, w_j, a_j) \in \mathbf{W}$ and almost every time, and (ii) the validity of (4.27) for every admissible bracket pair $(B, \mathbf{h}) \in \mathfrak{Br}^0$ and for all times s. Actually, this can be obtained through Cantor's intersection theorem, as done in, e.g., [31, 26, 3]. The vanishing of the Hamiltonian (3.10) follows from the fact that we have proven the maximality condition on the larger set of control values $\mathbf{W}$, which contains the set of canonical values $\mathbf{C} \times A$ in its interior. Indeed, since the process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is canonical, by (4.26) one has

$$\begin{aligned} H\Big(\bar{y}(s), p(s), p_0, \pi, \bar{w}^0(s), \bar{w}(s), \bar{\alpha}(s)\Big) &= \max_{(w^0, w, a) \in \mathbf{W}} H(\bar{y}(s), p(s), p_0, \pi, w^0, w, a) \\ &= \max_{(w^0, w, a) \in \mathbf{C} \times A} H(\bar{y}(s), p(s), p_0, \pi, w^0, w, a) \end{aligned}$$

for a.e. $s \in [0, \bar{S}]$. Now, as one can easily verify, the two maxima cannot coincide if at some s $\max_{(w^0, w, a) \in \mathbf{C} \times A} H(\bar{y}(s), p(s), p_0, \pi, w^0, w, a) \neq 0$, so (3.10) is proved. Finally, the drift-involving higher order conditions (3.12) can be achieved by differentiation (see [3, Cor. 4.5]). ∎

5. Example: Higher order normality. In the following example the absence of an infimum gap, while being undetectable by means of first order conditions, is ensured by proving that the extended sense minimizer is a normal higher order extremal and applying Theorem 3.2.

Let $\mathscr{U} := L^1([0,1],\mathbb{R}^3)$ and consider the optimal control problem

$$\text{(5.1)}\quad \begin{cases} \qquad\text{minimize} \quad \Psi(x(1)) \\ \text{over the processes } (u,x)\in\mathscr{U}\times AC([0,1],\mathbb{R}^4) \text{ such that} \\ \dfrac{dx}{dt}(t) = f(x(t)) + \displaystyle\sum_{i=1}^{3} g_i(x(t))u^i(t), \\ \dfrac{dv}{dt}(t) = |u(t)|, \\ (x,v)(0) = (-1,0,0,0,0), \qquad (x(1),v(1))\in\mathfrak{T}\times[0,3], \end{cases}$$

where $x = (x^1,x^2,x^3,x^4)$, $f(x) = (0,x^2,0,(x^3)^2)$, $g_1(x) = (1,0,-x^2,0)$, $g_2(x) = (0,-1,x^1,0)$, $g_3(x) = (1,0,0,0)$, $\mathfrak{T} := \{0\}\times\{0\}\times\{0\}\times\mathbb{R}$, and $\Psi(x) = x^4$. Set $\mathscr{W} := \bigcup_{S>0}\big(\{S\}\times\big\{(w^0,w)\in L^\infty([0,S],\mathbb{R}_+\times\mathbb{R}^3) :\ w^0+|w| = 1 \text{ a.e.}\big\}\big)$.

The corresponding extended optimal control problem reads

$$\text{(5.2)}\quad \begin{cases} \qquad\text{minimize} \quad \Psi(y(S)) \\ \text{over the extended sense controls } (S,w^0,w)\in\mathscr{W} \text{ and the functions} \\ \quad (y^0,y,\beta)\in AC([0,S],\mathbb{R}_+\times\mathbb{R}^4\times\mathbb{R}_+) \text{ such that} \\ \dfrac{dy^0}{ds}(s) = w^0(s), \\ \dfrac{dy}{ds}(s) = f(y(s))w^0(s) + \displaystyle\sum_{i=1}^{3} g_i(y(s))w^i(s), \\ \dfrac{d\beta}{ds}(s) = |w(s)|, \\ (y^0,y,\beta)(0) = (0,-1,0,0,0,0), \qquad (y^0,y,\beta)(S)\in\{1\}\times\mathfrak{T}\times[0,3], \end{cases}$$

where $y = (y^1,y^2,y^3,y^4)$ and $w = (w^1,w^2,w^3)$. The extended sense control

$$\big(\bar S,\bar w^0,\bar w^1,\bar w^2,\bar w^3\big) := \big(2,\mathbf{1}_{[1,2]},\mathbf{1}_{[0,1]},0,0\big)\in\mathscr{W}$$

generates the extended sense trajectory

$$\big(\bar y^0,\bar y^1,\bar y^2,\bar y^3,\bar y^4,\bar\beta\big)(s) := \big((s-1)\mathbf{1}_{[1,2]},(s-1)\mathbf{1}_{[0,1]},0,0,0,s\big), \quad s\in[0,2].$$

The extended sense process $(\bar S,\bar w^0,\bar w,\bar y^0,\bar y,\bar\beta)$ is clearly a minimizer for the extended problem (5.2), as the cost is increasing along all trajectories of the system. The unmaximized Hamiltonian of the problem is

$$\begin{aligned} &H(x,p,p_0,\pi,w^0,w) \\ &:= \Big(p_0 + p_2x^2 + p_4\left(x^3\right)^2\Big)w^0 + \left(p_1 - p_3x^2\right)w^1 + \left(-p_2 + p_3x^1\right)w^2 + p_1w^3 + \pi|w|, \end{aligned}$$

where $p = (p_1,p_2,p_3,p_4)$. By the first order maximum principle in [3, Thm. 3.1] there exist $\lambda\geq 0$, an adjoint arc $(p_0,p)\in AC([0,2],\mathbb{R}\times\mathbb{R}^4)$, and $\pi\leq 0$ satisfying the first order necessary conditions (i)–(v) in Definition 3.5 of higher order Ψ-extremal. In particular, $\pi = 0$ since $\bar\beta(2) = 2 < 3$, and λ, p_0, and p satisfy

(a) the nontriviality condition

$$(p_0(s),p(s),\lambda)\neq(0,0,0) \qquad \forall s\in[0,2];$$

(b) the adjoint end-time problem[19]

(5.3)
$$\begin{cases} \left(\dfrac{dp_0}{ds}, \dfrac{dp_1}{ds}, \dfrac{dp_2}{ds}, \dfrac{dp_3}{ds}, \dfrac{dp_4}{ds}\right)(s) = \Big(0, 0, p_3(s)\bar{w}^1(s) - \bar{w}^0(s)p_2(s), 0, 0\Big), \quad s \in [0,2], \\ \Big(p_0, p_1, p_2, p_3, p_4\Big)(2) = \Big(C_0, C_1, C_2, C_3, -\lambda\Big) \end{cases}$$

for some constants $C_0, C_1, C_2, C_3 \in \mathbb{R}$;

(c) the first order maximality and the vanishing condition, for a.e. $s \in [0,2]$,

$$0 = H(\bar{y}(s), p(s), p_0(s), \pi, \bar{w}^0(s), \bar{w}(s)) = \max_{(w^0,w)\in \mathbf{C}} H(\bar{y}(s), p(s), p_0(s), \pi, w^0, w);$$

(d) the first order relations for all $s \in [0,2]$,

$$p(s) \cdot g_i(\bar{y}(s)) = 0, \qquad i = 1, 2, 3.$$

Integrating (5.3), one gets $p_0 \equiv C_0$, $p_1 \equiv C_1$, $p_3 \equiv C_3$, $p_4 \equiv -\lambda$, and

(5.4)
$$p_2(s) := \begin{cases} C_2 e^{2-s}, & 1 \le s \le 2, \\ C_2 e + C_3(s-1), & 0 \le s < 1. \end{cases}$$

By condition (d) one easily obtains $C_1 = C_2 = 0$, while the maximality condition (c) on the interval $[1,2]$ implies that $C_0 = 0$. The maximality condition (c) on the interval $[0,1]$ now reads as

$$\max_{(w^0,w^1,w^2)} \left\{(-C_3(s-1) + C_3(s-1))w^2\right\} = 0.$$

Since this condition is verified for every $C_3 \in \mathbb{R}$, the previous analysis shows that the extended sense process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a (first order) abnormal extremal in that we can choose the multipliers $\lambda = 0$, $p_0 \equiv 0$, $p(s) = (0, C_3(s-1)\mathbf{1}_{[0,1]}, C_3, 0)$ on $[0,2]$ for some $C_3 \neq 0$. Hence, a first order normality criterion for no-gap, like the one obtained in [22], does not allow one to exclude the presence of a gap between the infimum cost of the extended problem and the original problem. However, the higher order maximum principle in [3, Thm. 4.1] establishes that, with reference to the process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{y}^0, \bar{y}, \bar{\beta})$, there are multipliers (p_0, p, π, λ) which, in addition to complying with the previous conditions (a)–(d), also satisfy, in particular, the following second order condition:

(5.5)
$$p(s) \cdot [g_3, g_2](\bar{y}(s)) = C_3 = 0.$$

Hence, $C_3 = 0$, so that $(p_0, p) \equiv 0$, and the nontriviality condition (a) implies that $\lambda \neq 0$ necessarily. Therefore, $(\bar{S}, \bar{w}^0, \bar{w}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a normal second order Ψ-extremal and the higher order normality criterion for no-gap in Theorem 3.2 guarantees the absence of an infimum gap.

REFERENCES

[1] A. Agrachev and Y. Sachkov, *Control Theory from the Geometric Viewpoint*, Encyclopedia Math. Sci. 87, Springer-Verlag, Berlin, 2013.

[19] We use the obvious approximating cone $\mathscr{K} := \{0\}^3 \times \mathbb{R}$ to the target $\mathfrak{T} = \{0\}^3 \times \mathbb{R}$ at $\bar{y}(2)$.

[2] M. S. Aronna, M. Motta, and F. Rampazzo, *Necessary conditions involving Lie brackets for impulsive optimal control problems*, in Proceedings of the IEEE 58th Annual Conference on Decision and Control (CDC), 2019, pp. 1474–1479, https://ieeexplore.ieee.org/abstract/document/9029273.

[3] M. S. Aronna, M. Motta, and F. Rampazzo, *A higher-order maximum principle for impulsive optimal control problems*, SIAM J. Control Optim., 58 (2020), pp. 814–844, https://doi.org/10.1137/19M1273785.

[4] M. S. Aronna and F. Rampazzo, *L^1 limit solutions for control systems*, J. Differential Equations, 258 (2015), pp. 954–979, https://doi.org/10.1016/j.jde.2014.10.013.

[5] A. Arutyunov, D. Karamzin, and F. L. Pereira, *Pontryagin's maximum principle for constrained impulsive control problems*, Nonlinear Anal., 75 (2012), pp. 1045–1057, https://doi.org/10.1016/j.na.2011.04.047.

[6] R. Bianchini and M. Kawski, *Needle variations that cannot be summed*, SIAM J. Control Optim., 42 (2003), pp. 218–238, https://doi.org/10.1137/S0363012902402876.

[7] R. Bianchini and G. Stefani, *Good needle-like variations*, in Differential Geometry and Control, Proc. Sympos. Pure Math. 64, American Mathematical Society, Providence, RI, 1999, pp. 91–102.

[8] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, AIMS Ser. Appl. Math. 2, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

[9] A. Bressan and F. Rampazzo, *On differential systems with vector-valued impulsive controls*, Boll. Un. Mat. Ital. B (7), 2 (1988), pp. 641–656.

[10] F. Chittaro and G. Stefani, *Minimum-time strong optimality of a singular arc: The multi-input non involutive case*, ESAIM Control Optim. Calc. Var., 22 (2016), pp. 786–810, https://doi.org/10.1051/cocv/2015026.

[11] E. Feleqi and F. Rampazzo, *Iterated Lie brackets for nonsmooth vector fields*, NoDEA Nonlinear Differential Equations Appl., 24 (2017), 61, https://doi.org/10.1007/s00030-017-0484-4.

[12] G. Fusco and M. Motta, *No infimum gap and normality in optimal impulsive control under state constraints*, Set-Valued Var. Anal., 29 (2021), pp. 519–550, https://doi.org/10.1007/s11228-021-00576-2.

[13] G. Fusco and M. Motta, *Nondegenerate abnormality, controllability, and gap phenomena in optimal control with state constraints*, SIAM J. Control Optim., 60 (2022), 280–309, https://doi.org/10.1137/20M1382465.

[14] G. Fusco and M. Motta, *Gap phenomena and controllability in free end-time problems with active state constraints*, J. Math. Anal. Appl., 510 (2022), 126021, https://doi.org/10.1016/j.jmaa.2022.126021.

[15] M. Guerra and A. Sarychev, *Fréchet generalized trajectories and minimizers for variational problems of low coercivity*, J. Dyn. Control Syst., 21 (2015), pp. 351–377, https://doi.org/10.1007/s10883-014-9231-x.

[16] D. Y. Karamzin, V. A. de Oliveira, F. L. Pereira, and G. N. Silva, *On the properness of an impulsive control extension of dynamic optimization problems*, ESAIM Control Optim. Calc. Var., 21 (2015), pp. 857–875, https://doi.org/10.1051/cocv/2014053.

[17] B. Kaskosz, *Extremality, controllability and abundant subsets of generalized control systems*, J. Optim. Theory Appl., 101 (1999), pp. 73–108, https://doi.org/10.1023/A:1021719027140.

[18] H. Knobloch, *Higher order necessary conditions in optimal control theory*, in Ordinary and Partial Differential Equations, W. Everitt ed., Lecture Notes in Math. 827, Springer, Berlin, Heidelberg, 1980, pp. 151–164, https://doi.org/10.1007/BFb0091379.

[19] A. Krener, *The high order maximal principle and its application to singular extremals*, SIAM J. Control Optim., 15 (1977), pp. 256–293, https://doi.org/10.1137/0315019.

[20] M. Miller and E. Y. Rubinovich, *Impulsive Control in Continuous and Discrete-Continuous Systems*, Kluwer Academic/Plenum Publishers, New York, 2003, https://doi.org/10.1007/978-1-4615-0095-7.

[21] M. Motta and F. Rampazzo, *Space-time trajectories of nonlinear systems driven by ordinary and impulsive controls*, Differential Integral Equations, 8 (1995), pp. 269–288.

[22] M. Motta, F. Rampazzo, and R. B. Vinter, *Normality and gap phenomena in optimal unbounded control*, ESAIM Control Optim. Calc. Var., 24 (2018), pp. 1645–1673, https://doi.org/10.1051/cocv/2018069.

[23] M. Motta and C. Sartori, *On $\mathscr{L}^1$ limit solutions in impulsive control*, Discrete Contin. Dyn. Syst. Ser. S, 11 (2018), pp. 1201–1218, https://doi.org/10.3934/dcdss.2018068.

[24] M. Motta and C. Sartori, *Normality and nondegeneracy of the maximum principle in optimal impulsive control under state constraints*, J. Optim. Theory Appl., 185 (2020), pp.

44–71, https://doi.org/10.1007/s10957-020-01641-w.

[25] M. Palladino and F. Rampazzo, *A no infimum-gap criterion*, in Proceedings of the IEEE 58th Annual Conference on Decision and Control (CDC), IEEE, 2019, pp. 535–540, https://doi.org/10.1109/CDC40024.2019.9029870.

[26] M. Palladino and F. Rampazzo, *A geometrically based criterion to avoid infimum gaps in optimal control*, J. Differential Equations, 269 (2020), pp. 10107–10142, https://doi.org/10.1016/j.jde.2020.06.066.

[27] M. Palladino and R. B. Vinter, *When does relaxation reduce the minimum cost of an optimal control problem?*, in Proceedings of the IEEE 52nd Annual Conference on Decision and Control (CDC), 2013, pp. 526–531, https://doi.org/10.1109/CDC.2013.6759935.

[28] M. Palladino and R. B. Vinter, *Minimizers that are not also relaxed minimizers*, SIAM J. Control Optim., 52 (2014), pp. 2164–2179, https://doi.org/10.1137/130909627.

[29] M. Palladino and R. B. Vinter, *When are minimizing controls also minimizing relaxed controls?*, Discrete Contin. Dyn. Syst. Ser. A, 35 (2015), pp. 4573–4592, https://doi.org/10.3934/dcds.2015.35.4573.

[30] H. Schättler and U. Ledzewicz, *Geometric Optimal Control: Theory, Methods and Examples*, Springer, New York, 2012, https://doi.org/10.1007/978-1-4614-3834-2.

[31] H. J. Sussmann, *Geometry and optimal control*, in Mathematical Control Theory, J. Baillieul and J. C. Willems, eds., Springer, New York, 1999, pp. 140–198, https://doi.org/10.1007/978-1-4612-1416-8_5.

[32] H. J. Sussmann, *Generalized differentials, variational generators, and the maximum principle with state constraints*, in Nonlinear and Optimal Control Theory, P. Nistri and G. Stefani, eds., Lecture Notes in Math. 1932, Springer, Berlin, Heidelberg, 2008, pp. 221–287, https://doi.org/10.1007/978-3-540-77653-6_4.

[33] H. J. Sussmann, *High-order point variations and generalized differentials*, in Geometric Control and Nonsmooth Analysis, Ser. Adv. Math. Appl. Sci. 76, 2008, pp. 327–357, https://doi.org/10.1142/9789812776075_0019.

[34] J. Warga, *Normal control problems have no minimizing strictly original solutions*, Bull. Amer. Math. Soc., 77 (1971), pp. 625–628.

[35] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972, https://doi.org/10.1016/C2013-0-11669-8.

[36] J. Warga, *Controllability, extremality, and abnormality in nonsmooth optimal control*, J. Optim. Theory Appl., 41 (1983), pp. 239–260, https://doi.org/10.1007/BF00934445.

[37] J. Warga, *Optimization and controllability without differentiability assumptions*, SIAM J. Control Optim., 21 (1983), pp. 837–855, https://doi.org/10.1137/0321051.

[38] P. Wolenski and S. Žabić, *A differential solution concept for impulsive systems*, Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal., 13B (2006), pp. 199–210.