Mathematical Control and Related Fields
Vol. 14, No. 4, December 2024, pp. 1306-1333
doi:10.3934/mcrf.2024012

IMPULSIVE OPTIMAL CONTROL PROBLEMS WITH TIME DELAYS IN THE DRIFT TERM

Giovanni Fusco[1] and Monica Motta[*1]

[1]Department of Mathematics, University of Padova, Via Trieste, 63, Padova 35121, Italy

Abstract. We introduce a notion of bounded variation solution for a new class of nonlinear control systems with ordinary and impulsive controls, in which the drift function depends not only on the state, but also on its past history, through a finite number of time delays. After proving the well-posedness of such solutions and the continuity of the corresponding input-output map with respect to suitable topologies, we establish necessary optimality conditions for an associated optimal control problem. The approach, which involves approximating the problem by a non-impulsive optimal control problem with time delays and using Ekeland's principle combined with a recent, nonsmooth version of the Maximum Principle for conventional delayed systems, allows us to deal with mild regularity assumptions and a general endpoint constraint.

1. **Introduction.** We establish necessary optimality conditions, in the form of a nonsmooth Maximum Principle, for the following impulsive optimal control problem with time delays, labeled (P):

$$\text{Minimize } \Phi(x(0),x(T)) + \int_0^T l_0\left(t,\{x(t-h_k)\}_{k=0}^N,\alpha(t)\right)dt + \int_{[0,T]} l_1(t,\alpha(t))\mu(dt)$$

over the set of control pairs (μ,α) with $\mu \in C^*_{\mathcal{K}}([0,T])$ and $\alpha \in \mathcal{A}_\mu$, and trajectories $x : [-h,T] \to \mathbb{R}^n$, of bounded variation on $[0,T]$, satisfying the control system

$$\begin{cases} x(t) = x(0) + \int_0^t f\left(s,\{x(s-h_k)\}_{k=0}^N,\alpha(s)\right)ds + \int_{[0,t]} G(s,\alpha(s))\mu(ds) \ \forall t \in]0,T], \\ x(t) = \zeta(t), \quad \text{a.e. } t \in [-h,0[, \end{cases} \tag{1}$$

and the endpoint constraint

$$(x(0),x(T)) \in \mathcal{T}. \tag{2}$$

Problem (P) involves both measurable functions and vector-valued measures as controls, since, fixed a time horizon $T > 0$, $C^*_{\mathcal{K}}([0,T])$ denotes the set of regular measures μ on the Borel subsets of $[0,T]$ with range belonging to a closed convex cone $\mathcal{K} \subset \mathbb{R}^m$, and $\mathcal{A}_\mu := \{\alpha : [0,T] \to \mathbb{R}^q$ measurable s.t. $\alpha(t) \in A(t)$ a.e. and μ-a.e.$\}$, where $A : [0,T] \rightsquigarrow \mathbb{R}^q$ is a set-valued map. Furthermore, the data comprise real numbers $0 = h_0 < h_1 < \cdots < h_N =: h$, functions $\Phi : \mathbb{R}^{2n} \to \mathbb{R}$, $l_0 : [0,T] \times (\mathbb{R}^n)^{N+1} \times \mathbb{R}^q \to \mathbb{R}$, $l_1 : [0,T] \times \mathbb{R}^q \to \mathbb{R}^m$, $f : [0,T] \times (\mathbb{R}^n)^{N+1} \times \mathbb{R}^q \to \mathbb{R}^n$,

2020 *Mathematics Subject Classification.* Primary: 49N25, 34K35; Secondary: 93C43, 49K21.
Key words and phrases. Optimal control, impulse control, maximum principle, time delay systems, nonsmooth analysis.
The second author is supported by the INdAM-GNAMPA Project 2023, CUP E53C22001930001.
*Corresponding author: Monica Motta.

$G : [0,T] \times \mathbb{R}^q \to \mathbb{R}^{n\times m}$, $\zeta : [-h,0] \to \mathbb{R}^n$ (which describes the "history" of the state trajectories before time 0), and a closed subset $\mathcal{T} \subset \mathbb{R}^{2n}$ (the target). Precise assumptions and definitions will be given in Sec. 2.

This problem can be interpreted as an extension of the following non-impulsive optimal control problem with time delays, labeled (P′), in which dynamics and cost depend linearly on an unbounded control ω, in addition to an ordinary control α:

$$\text{Minimize } \Phi(x(0),x(T)) + \int_0^T \left[l_0\left(t,\{x(t-h_k)\}_{k=0}^N,\alpha(t)\right) + l_1(t,\alpha(t))\omega(t)\right] dt$$

over the set of controls $(\omega,\alpha) \in L^1([0,T],\mathbb{R}^m \times \mathbb{R}^q)$ such that $\omega(t) \in \mathcal{K}$ and $\alpha(t) \in A(t)$ for a.e. $t \in [0,T]$, and trajectories $x : [-h,T] \to \mathbb{R}^n$, absolutely continuous on $[0,T]$ and satisfying the conventional control system with time delays

$$\begin{cases} \dot{x}(t) = f\left(t,\{x(t-h_k)\}_{k=0}^N,\alpha(t)\right) + G(t,\alpha(t))\omega(t), & t \in]0,T], \\ x(t) = \zeta(t) \quad \text{a.e. } t \in [-h,0[, \end{cases} \tag{3}$$

and the endpoint constraint (2). Indeed, given a non-impulsive control pair (ω,α) and a corresponding solution x to (3), (ω,α) can be identified with the impulsive control $(\mu,\alpha) = (\omega\, dt,\alpha)$ and x clearly satisfies (1) for such (μ,α).

As first results, we prove that this dynamics' extension is proper and the notion of impulsive trajectory is robust. Specifically, under mild hypotheses we establish: *Well-posedness:* given an initial condition $\xi \in \mathbb{R}^n$ and a control (μ,α), there is one and only one corresponding bounded variation trajectory x to (1) such that $x(0) = \xi$; *Density:* for any (μ,α) and x as above, there exists a sequence of non-impulsive controls (ω_i,α_i) and corresponding trajectories x_i to (3), such that $dx_i \rightharpoonup^* dx$ and $x_i(t) \to x(t)$ in a full measure subset of the interval $[0,T]$ containing 0 and T. *Continuity of the input-output map:* the map $(\xi,\mu,\alpha) \mapsto x$, where x is the solution to (1) associated with (μ,α) and such that $x(0) = \xi$, is continuous in a suitable sense (see Thm. 2.4). Afterwards, the main result of the paper is expressed as a *Maximum Principle* for the impulsive optimal control problem with time delays, (P).

We emphasize that these results are obtained without invoking a well-known reparameterization technique, usually employed in impulsive control without time delays since early work [19, 22]. In particular, this technique leads to the introduction of an auxiliary free end-time optimal control problem with ordinary controls only, in which time is considered as a state variable (see also [5, 14, 16, 4]). In fact, this procedure seems hardly extendable to the case with time delays, since the auxiliary problem that one obtains is not standard, as it involves time delays depending on the control itself. Furthermore, this method would require Lipschitz continuity of the data in their t-dependence and a fixed control set A in place of the set-valued map $A(t)$, as t plays the role of a state. We follow instead a different approach proposed in [21], which involves approximating our impulsive optimal control problem with time delays by a conventional one, without any time change. We point out that, as in the case without delays, this approach is applicable because the Lagrangian cost function l_1 and the *fast dynamics* G do not depend on the state variable.

The Maximum Principle established in this paper extends, on the one hand, the impulsive Maximum Principle obtained in [21] to time delayed systems, and, on the other hand, the Maximum Principle for time delayed problems of [3] to the impulsive setting considered here. Furthermore, it might also be useful for the applications' relevant problem of parameter identification, whenever ordinary controls in the system are treated as unknown parameters. Our primary goal here is to provide the weakest hypotheses our methods permit under which necessary conditions can be formulated and proved for problem (P). In particular, we do not address the infimum gap phenomenon, that is, we do not look for conditions guaranteeing that the infimum of the extended problem (P) coincides with that of (P′). Results of this kind for very general impulsive problems without delays can be found e.g. in [2, 17, 8, 9, 10, 11, 18, 15] and references therein.

It is worth mentioning that in the literature there are several results on the stabilizability of delayed impulsive control systems and on the optimization of some specific related problems, but they all concern the so-called 'impulse model', where impulsive controls essentially reduce to a finite or countable number of jump instants, with a preassigned jump-function. Therefore, the line of research starting with this paper, based on a different notion of impulsive control system, could have interesting implications for applications (e.g. to fed-batch fermentation [23, 12] or to impulsive control of delayed neural networks [13]), as it would allow the development of new nonlinear models, impulsive and with time delays, with a significant freer allocation of the impulses.

The paper is organized as follows. In Sec. 2 we introduce in a rigorous way the concepts of control and trajectory for the delayed impulsive control system considered in problem (P) and prove some fundamental properties of the set of the impulsive trajectories. In Sec. 3 we establish the Maximum Principle, whose proof in given in Sec. 4. An Appendix with some technical proofs concludes the paper.

1.1. **Notations and preliminaries.** We write $\mathcal{B}$ and $\mathscr{L}$ for the sets of the Borel subsets and the Lebesgue subsets of $[0,T]$, respectively, and we denote the set of Borel subsets of $\mathbb{R}^l$ by $\mathcal{B}^l$. We use $\mathbb{R}^{n\times m}$ to denote the set of $n\times m$ real matrices. Given an interval $I\subseteq\mathbb{R}$ and a set $\Omega\subseteq\mathbb{R}^l$, we write $\mathcal{M}(I,\Omega)$, $C(I,\Omega)$, $W^{1,1}(I,\Omega)$, $BV(I,\Omega)$ for the space of measurable, continuous, absolutely continuous and bounded variation functions on I and with values in Ω, respectively. We will use $\|\cdot\|_{L^\infty(I)}$ and $\|\cdot\|_{L^1(I)}$ to denote the ess-sup norm on I and the L^1-norm on I, respectively. When the domain is clear, we will sometimes simply write $\|\cdot\|_{L^\infty}$ and $\|\cdot\|_{L^1}$. We denote by $C^*_\Omega([0,T])$ the set of signed and regular measures $\mu:\mathcal{B}\to\Omega$ (from now on we will refer to such μ simply as measures), and we set $C^\oplus:=C^*_{\mathbb{R}_{\geq0}}([0,T])$, where $\mathbb{R}_{\geq0}:=[0,+\infty[$. For all these classes of functions, we will not specify domain and codomain when the meaning is clear. Given $\mu\in C^*_{\mathbb{R}^l}([0,T])$, μ-*a.e.* means "almost everywhere w.r.t. μ", and when we do not specify μ we implicitly refer to the Lebesgue measure. It is known that there is a bijection between $C^*_{\mathbb{R}^l}([0,T])$ and the set of the equivalence classes of functions of bounded variation from $[0,T]$ to $\mathbb{R}^l$ which are right continuous on $]0,T[$ and differ for a constant. In particular, given $\mu\in C^*_{\mathbb{R}^l}([0,T])$, we can associate with it the discontinuous trajectory z such that $dz(t)=\mu(dt)$, given by

$$z(t)=\int_{[0,t]}dz(s)=\int_{[0,t]}\mu(ds)\qquad\forall t\in]0,T]. \tag{4}$$

Given $\mu\in C^*_{\mathbb{R}^l}([0,T])$, the *total variation measure* is the measure $|\mu|\in C^\oplus$ given by

$$|\mu|=\sum_{j=1}^{l}(\mu^j_++\mu^j_-),$$

where $\mu_+^j, \mu_-^j \in C^\oplus$ are the elements of the Jordan decomposition of the j-th component μ^j of μ. In $C^*_{\mathbb{R}^l}([0,T])$, the norm is the sum of the total variations of the component measures, i.e. $\|\mu\|_{C^*_{\mathbb{R}^l}([0,T])} := \sum_{j=1}^l |\mu^j|([0,T])$. Note that the components of $\mu \in C^*_{\mathbb{R}^l}([0,T])$ are absolutely continuous with respect to the total variation measure $|\mu|$ (we write $\mu^j \ll |\mu|$ for any j), hence there exists a function $\omega : [0,T] \to \mathbb{R}^l$, which we will refer to as the *Radon-Nikodym derivative of μ with respect to* $|\mu|$, such that $\mu(dt) = \omega(t)|\mu|(dt)$. Equivalently, we will write $\omega = \frac{d\mu}{d|\mu|}$. A vector-valued function $\Psi : [0,T] \to \mathbb{R}^l$ is said to be μ-integrable if $t \mapsto |\Psi(t) \cdot \omega(t)|$ is integrable with respect to $|\mu|$, and in this case we set [1]

$$\int_B \Psi(t) \cdot \mu(dt) = \int_B \Psi(t) \cdot \omega(t)|\mu|(dt), \qquad \forall B \in \mathcal{B}. \tag{5}$$

Given a sequence $(\mu_i)_i \subset C^*_{\mathbb{R}^l}([0,T])$ and $\mu \in C^*_{\mathbb{R}^l}([0,T])$, we write $\mu_i \rightharpoonup^* \mu$ if

$$\lim_i \int_{[0,T]} \varphi(t)\mu_i^j(dt) = \int_{[0,T]} \varphi(t)\mu^j(dt), \qquad \forall \varphi \in C([0,T],\mathbb{R}),\ \forall j = 1,\dots,l.$$

We denote by $\ell(\Omega)$, $\mathrm{co}(\Omega)$, $\overline{\Omega}$, and $\partial\Omega$ the Lebesgue measure, the convex hull, the closure, and the boundary of Ω, respectively. As customary, χ_Ω is the characteristic function of Ω, namely $\chi_\Omega(x) = 1$ if $x \in \Omega$ and $\chi_\Omega(x) = 0$ if $x \in \mathbb{R}^l \setminus \Omega$. For any $a, b \in \mathbb{R}$, we write $a \wedge b := \min\{a, b\}$. Given $r > 0$, we denote the closed ball of radius r in $\mathbb{R}^l$ by $r\mathbb{B}_l$, omitting the dimension when it is clear from the context. Given a closed set $C \subseteq \mathbb{R}^l$ and a point $z \in \mathbb{R}^l$, we define the distance of z from C as $d_C(z) := \min_{y\in C} |z - y|$, and we define the *support function* σ_C of the set C as $\sigma_C(y) := \sup\{z \cdot y : z \in C\}$ for any $y \in \mathbb{R}^l$.

Some standard constructs from nonsmooth analysis are employed in this paper. For background material we refer the reader for instance to [6, 7, 20]. A set $\mathcal{K} \subseteq \mathbb{R}^l$ is a *cone* if $\alpha x \in \mathcal{K}$ for any $\alpha > 0$, whenever $x \in \mathcal{K}$. Take a closed set $C \subseteq \mathbb{R}^l$ and a point $\bar{x} \in C$, the *limiting normal cone* $N_C(\bar{x})$ of C at $\bar{x}$ is defined by

$$N_C(\bar{x}) := \left\{ \xi : \exists x_i \xrightarrow{C} \bar{x},\, \xi_i \to \xi \ \text{ such that } \ \limsup_{x\to x_i} \frac{\xi_i \cdot (x - x_i)}{|x - x_i|} \le 0 \ \ \forall i \right\},$$

in which the notation $x_i \xrightarrow{C} \bar{x}$ is used to indicate that all points in the sequence $(x_i)_i$ lay in C. Take a lower semicontinuous function $H : \mathbb{R}^l \to \mathbb{R}$ and a point $\bar{x} \in \mathbb{R}^l$, the *limiting subdifferential* of H at $\bar{x}$ is

$$\partial H(\bar{x}) := \left\{ \xi : \exists \xi_i \to \xi,\, x_i \to \bar{x} \text{ s.t. } \limsup_{x\to x_i} \frac{\xi_i \cdot (x - x_i) - H(x) + H(x_i)}{|x - x_i|} \le 0 \ \forall i \right\}.$$

If $H : \mathbb{R}^l \times \mathbb{R}^h \to \mathbb{R}$ is a lower semicontinuous function and $(\bar{x}, \bar{y}) \in \mathbb{R}^l \times \mathbb{R}^h$, we write $\partial_x H(\bar{x}, \bar{y})$ [resp. $\partial_y H(\bar{x}, \bar{y})$] to denote the *partial limiting subdifferential of H at $(\bar{x}, \bar{y})$ w.r.t. x* [resp. w.r.t. y], and we write $\tilde{\partial}_x H(\bar{x}, \bar{y})$ [resp. $\tilde{\partial}_y H(\bar{x}, \bar{y})$] to denote the *projected limiting subdifferential w.r.t. x* [resp. w.r.t. y], i.e. the projection of the limiting subdifferential of $H(\cdot,\cdot)$ at $(\bar{x}, \bar{y})$ onto the x-coordinate [resp. y-coordinate]. Given a locally Lipschitz continuous function $G : \mathbb{R}^k \to \mathbb{R}^l$ and $\bar{x} \in \mathbb{R}^k$, we write $DG(\bar{x})$ to denote the *Clarke generalized Jacobian of G at $\bar{x}$*,

[1] This definition is more general than $\int_B \Psi(t) \cdot \mu(dt) = \sum_j \int_B \Psi_j(t)\mu^j(dt)$, since the latter requires that Ψ_j is μ^j-integrable for any j.

defined as

$$DG(\bar{x}) := \text{co}\left\{\xi\colon \exists\, x_i \overset{\text{diff}(G)\setminus\{\bar{x}\}}{\longrightarrow} \bar{x} \text{ and } \nabla G(x_i) \to \xi\right\},$$

where ∇G denotes the classical Jacobian matrix of G and diff(G) denotes the full measure set of differentiability points of G. We recall that the set-valued map $x \rightsquigarrow DG(x)$ has nonempty, compact, convex values and is upper semicontinuous.

2. **BV trajectories of delayed impulsive control systems.**

In this section we introduce in a rigorous way the concepts of control and trajectory for an impulsive control system with N time delays $0 < h_1 < \cdots < h_N$, of the form

$$\begin{cases} dx(t) = f(t, x(t), x(t-h_1), \dots, x(t-h_N), \alpha(t))\,dt + G(t,\alpha(t))\mu(dt), & t \in [0,T], \\ x(t) = \zeta(t) \ \text{ a.e. } t \in [-h, 0[\qquad (h = h_N), \end{cases} \tag{6}$$

and establish the main properties of the set of the corresponding trajectories.

2.1. **Statements and main results.**

We define the set $\mathcal{U}$ of *impulsive controls* as follows:

$$\mathcal{U} := \{(\mu,\alpha): \quad \mu \in C^*_{\mathcal{K}}([0,T]), \quad \alpha \in \mathcal{A}_\mu, \ t \mapsto G(t,\alpha(t)) \ \mu\text{-integrable}\}, \tag{7}$$

where

$$\mathcal{A}_\mu := \{\alpha : [0,T] \to \mathbb{R}^q : \ \alpha \text{ is Borel measurable and } \alpha(t) \in A(t) \text{ a.e. and } \mu\text{-a.e.}\}.$$

A solution to (6) is defined as a function $x : [-h,T] \to \mathbb{R}^n$, with bounded variation on $[0,T]$, satisfying

$$\begin{cases} x(t) = x(0) + \displaystyle\int_0^t f\left(s, \{x(s-h_k)\}, \alpha(s)\right) ds + \int_{[0,t]} G(s,\alpha(s))\mu(ds) \ \ \forall t \in]0,T], \\ x(t) = \zeta(t) \quad \text{a.e. } t \in [-h,0[, \end{cases} \tag{8}$$

where we write $\{x(t-h_k)\}$ in place of $\{x(t-h_k)\}_{k=0}^N := (x(t), x(t-h_1), \dots, x(t-h_N))$ and use the notion (5) of μ-integrability. We will refer to x as an *impulsive trajectory* (or simply as a trajectory) associated with (μ,α) and to the triple (μ,α,x) as an *impulsive process* (or simply as a process) for (6). When the measure μ is absolutely continuous w.r.t. the Lebesgue measure ℓ, so that there exists $\omega \in L^1([0,T],\mathcal{K})$ such that $\mu(dt) = \omega(t)\,dt$, the trajectory x is absolutely continuous on $[0,T]$ and the impulsive control system (6) becomes the following conventional control system with time delays,

$$\begin{cases} \dot{x}(t) = f(t, \{x(s-h_k)\}, \alpha(t)) + G(t,\alpha(t))\omega(t), & t \in [0,T], \\ x(t) = \zeta(t) \qquad \text{a.e. } t \in [-h,0[. \end{cases} \tag{9}$$

As in the case without delays, the impulsive control system (6) can be seen as an extension of (9). For this reason, a process (μ,α,x) with $\mu \ll \ell$ will be referred to as a *strict sense process* for (6).

We shall consider the following hypotheses:

(H1) *The set-valued map* $[0,T] \ni t \rightsquigarrow A(t) \subseteq \mathbb{R}^q$ *has graph* $\mathrm{Gr}(A) := \{(t,a) : a \in A(t)\}$ *which is* $\mathscr{L} \times \mathcal{B}^q$*-measurable.* [2]

[2] $\mathscr{L} \times \mathcal{B}^q$ denotes the product σ-algebra of $\mathscr{L}$ and $\mathcal{B}^q$.

(H2) *The function* ζ *belongs to* $L^\infty([-h,0],\mathbb{R}^n)$. *For every* $\{z_k\}\in(\mathbb{R}^n)^{N+1}$, *the function* $(t,a)\mapsto f(t,\{z_k\},a)$ *is* $\mathscr{L}\times\mathcal{B}^q$*-measurable and, for any* $M>0$, *there exists* $L_M\in L^1([0,T],\mathbb{R}_{\geq 0})$ *such that*

$$
\begin{aligned}
&|f(t,\{z_k\},a)-f(t,\{y_k\},a)|\leq L_M(t)|\{z_k-y_k\}|,\\
&\forall\{z_k\},\ \{y_k\}\in M\,\mathbb{B}_{n(N+1)},\ \forall a\in A(t),\quad a.e.\ t\in[0,T].
\end{aligned}
\tag{10}
$$

(H3) *There exists* $c\in L^1([0,T],\mathbb{R}_{\geq 0})$ *such that*

$$
\sup_{a\in A(t)}|f(t,\{z_k\},a)|\leq c(t)(1+|\{z_k\}|)\qquad\forall\{z_k\}\in(\mathbb{R}^n)^{N+1},\ \ a.e.\ t\in[0,T]. \tag{11}
$$

(H4) *The set valued function* $t\rightsquigarrow\overline{\mathrm{co}}\,\mathfrak{g}(t,A(t),\tilde{\mathcal{K}})$, $t\in[0,T]$, *is uniformly bounded and continuous with respect to the Hausdorff metric, where* $\mathfrak{g}:\mathbb{R}\times\mathbb{R}^q\times\mathbb{R}^m\to\mathbb{R}^n$ *is the function defined by*

$$
\mathfrak{g}(t,a,w):=\frac{G(t,a)\cdot w}{1+\sum_{i=1}^n|\sum_{j=1}^m g_{ij}(t,a)w^j|}, \tag{12}
$$

with the g_{ij}*'s being the components of* G *and* $\tilde{\mathcal{K}}:=\big\{w\in\mathcal{K}:\ \ |w^j|\leq 1,\ j=1,\dots,m\big\}$.

Remark 2.1. Hypothesis (H4) is a kind of continuity assumption on a compactification of the fast dynamics G. As we will see in Prop. 2.7 below, (H4) allows us to replace without loss of generality the original control system with an equivalent *auxiliary* control system with time delays, still impulsive but in which the vector-valued measure μ is replaced by a nonnegative and scalar measure ν and an ordinary control $\omega\in L^1([0,T],\tilde{\mathcal{K}})$, while the possibly highly irregular and unbounded term "$G(t,a)$" is replaced by "$\mathfrak{g}(t,a,w)$". In the approximation results, where essentially we want to deduce from the convergence of suitable scalar auxiliary controls $\nu_i\rightharpoonup^*\nu$, the convergence of the corresponding vector-valued measures $\mu_i\rightharpoonup^*\mu$ (see Prop. 2.3 and Lemma 2.10 below), we replace assumption (H4) with the following stronger condition (H4)*:

(H4)* *The set valued function* $t\rightsquigarrow\overline{\mathrm{co}}\,(\mathfrak{h},\mathfrak{g})(t,A(t),\tilde{\mathcal{K}})$, $t\in[0,T]$, *is uniformly bounded and continuous with respect to the Hausdorff metric, where* $\mathfrak{g}$ *is as in* (12) *and* $\mathfrak{h}:\mathbb{R}\times\mathbb{R}^q\times\mathbb{R}^m\to\mathbb{R}^m$ *is the function defined by*

$$
\mathfrak{h}(t,a,w):=\frac{w}{1+\sum_{i=1}^n|\sum_{j=1}^m g_{ij}(t,a)w^j|}. \tag{13}
$$

Both assumptions (H4) and (H4)* are automatically satisfied by autonomous systems and also by systems where the multifunction $t\rightsquigarrow G(t,A(t))$ is uniformly bounded on $[0,T]$ and continuous in the sense of Kuratowski.

Let us state some fundamental properties of impulsive trajectories.

Proposition 2.2 (Well-posedness)**.** *Assume hypotheses* (H1)–(H3). *Then, for any control* $(\mu,\alpha)\in\mathcal{U}$ *and any initial condition* $\xi\in\mathbb{R}^n$, *there exists one and only one impulsive trajectory* x *of* (6) *associated with* (μ,α) *and such that* $x(0)=\xi$.

Proof. Given a control $(\mu,\alpha)\in\mathcal{U}$ and an initial condition $\xi\in\mathbb{R}^n$, let us consider the following (non-impulsive) ODE with time delays

$$
\begin{cases}
\dot{\tilde{x}}(t)=\varphi(t,\{\tilde{x}(t-h_k)\})\qquad\text{a.e. } t\in[0,T],\\
\tilde{x}(t)=\zeta(t)\quad\text{a.e. } t\in[-h,0[,\quad\tilde{x}(0)=\xi,
\end{cases}
\tag{14}
$$

where $\varphi : [0,T] \times (\mathbb{R}^n)^{N+1} \to \mathbb{R}^n$ is given by

$$\varphi(t,\{z_k\}) := f(t,\{z_k + \tilde{x}_0(t-h_k)\},\alpha(t)) \qquad \forall t \in [0,T],\ \forall \{z_k\} \in (\mathbb{R}^n)^{N+1},$$

and $\tilde{x}_0 : [-h,T] \to \mathbb{R}^n$ is defined as

$$\tilde{x}_0(t) := \int_{[0,t]} G(s,\alpha(s)) \cdot \mu(ds), \quad t \in]0,T], \quad \tilde{x}_0(t) := 0 \ \ \forall t \in [-h,0].$$

In particular, $\tilde{x}_0$ is a BV function uniquely determined by the control pair (μ,α), and the function $t \mapsto \varphi(t,\{z_k\})$ is $\mathscr{L}$-measurable for any $\{z_k\} \in (\mathbb{R}^n)^{N+1}$.

The proof is complete as soon as we show that (14) admits a unique solution $\tilde{x}$, since $x := \tilde{x} + \tilde{x}_0$ turns out to be the unique solution of (6) associated with (μ,α) and such that $x(0) = \xi$. Indeed, if there are two solutions x, x', with $x' \neq x$, to (6) associated with (μ,α) and such that $x(0) = \xi$, then it follows immediately that (14) also has the distinct solutions $\tilde{x} := x - \tilde{x}_0$ and $\tilde{x}' := x' - \tilde{x}_0$.

To this aim, we define $\tilde{c} \in L^1([0,T],\mathbb{R}_{\geq 0})$, $M > 0$ and $\tilde{M} > 0$ as

$$\begin{aligned} &\tilde{c}(t) := \big(1 + (N+1)\|\tilde{x}_0\|_{L^\infty(0,T)}\big)\, c(t), \\ &M := \Big[|\xi| + \big(1 + (N+1)\|\zeta\|_{L^\infty(-h,0)}\big)\|\tilde{c}\|_{L^1(0,T)}\Big] e^{(N+1)\|\tilde{c}\|_{L^1(0,T)}}, \\ &\tilde{M} := 2M + (N+1)\|\tilde{x}_0\|_{L^\infty(0,T)}, \end{aligned} \tag{15}$$

where c is as in (H3). Hence, we introduce a function $\eta \in C^\infty((\mathbb{R}^n)^{N+1},[0,1])$ which is equal to 1 in $M\mathbb{B}_{n(N+1)}$, equal to 0 in $(\mathbb{R}^n)^{N+1} \setminus 2M\mathbb{B}_{n(N+1)}$, and such that $\|\nabla\eta(\{z_k\})\|_{L^\infty} \leq 1$. Using a standard truncation technique, in place of (14) we consider the following ODE with time delays

$$\begin{cases} \dot{\tilde{x}}(t) = \tilde{\varphi}(t,\{\tilde{x}(t-h_k)\}) \qquad \text{a.e. } t \in [0,T], \\ \tilde{x}(t) = \zeta(t) \ \ \text{a.e. } t \in [-h,0[, \ \ \tilde{x}(0) = \xi, \end{cases} \tag{16}$$

where, the function $\tilde{\varphi} : [0,T] \times (\mathbb{R}^n)^{N+1} \to \mathbb{R}^n$ is given by

$$\tilde{\varphi}(t,\{z_k\}) := \eta(\{z_k\})\varphi(t,\{z_k\}) \qquad \forall t \in [0,T],\ \forall \{z_k\} \in (\mathbb{R}^n)^{N+1}.$$

By (H3), for any $t \in [0,T]$ and any $\{z_k\} \in (\mathbb{R}^n)^{N+1}$, one has

$$\begin{aligned} |\tilde{\varphi}(t,\{z_k\})| \leq |\varphi(t,\{z_k\})| &= |f(t,\{z_k + \tilde{x}_0(t-h_k)\},\alpha(t))| \\ &\leq c(t)\big(1 + (N+1)\|\tilde{x}_0\|_{L^\infty(0,T)} + |\{z_k\}|\big) \\ &\leq \tilde{c}(t)(1 + |\{z_k\}|). \end{aligned} \tag{17}$$

Moreover, there exists a function $\tilde{L} \in L^1([0,T],\mathbb{R}_{\geq 0})$ such that

$$|\tilde{\varphi}(t,\{z_k\}) - \tilde{\varphi}(t,\{y_k\})| \leq \tilde{L}(t)|\{z_k - y_k\}| \qquad \text{a.e. } t,\ \forall \{z_k\},\ \{y_k\} \in (\mathbb{R}^n)^{N+1}. \tag{18}$$

Indeed, the inequality $|\tilde{\varphi}(t,\{z_k\}) - \tilde{\varphi}(t,\{y_k\})|$ is zero when both $\{z_k\}$, $\{y_k\} \in (\mathbb{R}^n)^{N+1} \setminus 2M\mathbb{B}_{n(N+1)}$, while, if, e.g. $\{y_k\} \in (\mathbb{R}^n)^{N+1} \setminus 2M\mathbb{B}_{n(N+1)}$, so that $\eta(\{y_k\}) = 0$, but $\{z_k\} \in 2M\mathbb{B}_{n(N+1)}$, we have

$$\begin{aligned} |\tilde{\varphi}(t,\{z_k\}) - \tilde{\varphi}(t,\{y_k\})| &= |\eta(\{z_k\})|\,|\varphi(t,\{z_k\})| \\ &= |\eta(\{z_k\}) - \eta(\{y_k\})|\,|f(t,\{z_k + \tilde{x}_0(t-h_k)\},\alpha(t))| \\ &\leq c(t)\big(1 + \tilde{M}\big)\,|\{z_k\} - \{y_k\}|. \end{aligned}$$

If instead both $\{z_k\}$, $\{y_k\} \in 2M\mathbb{B}_{n(N+1)}$, we get

$$\begin{aligned}
&|\tilde\varphi(t,\{z_k\}) - \tilde\varphi(t,\{y_k\})| \\
&\le |\eta(\{z_k\})|\,|f(t,\{z_k + \tilde x_0(t-h_k)\},\alpha(t)) - f(t,\{y_k + \tilde x_0(t-h_k)\},\alpha(t))| \\
&\quad + |\eta(\{z_k\}) - \eta(\{y_k\})|\,|f(t,\{y_k + \tilde x_0(t-h_k)\},\alpha(t))| \\
&\le L_{\tilde M}(t)\,|\{z_k\} - \{y_k\}| + c(t)\big(1+\tilde M\big)\,|\{z_k\} - \{y_k\}|,
\end{aligned}$$

where $L_{\tilde M}$ is as in (H2), for $\tilde M$ as in (15). Hence, (18) holds with $\tilde L(t) := L_{\tilde M}(t) + c(t)\big(1+\tilde M\big)$. At this point, (16) has a unique solution $\tilde x$ by [3, Thm. 4.1].

In order to conclude, it remains to show that $\|\tilde x\|_{L^\infty(0,T)} \le M$. From (17) it follows that

$$|\tilde x(t)| \le |\xi| + \int_0^t |\tilde\varphi(s,\{\tilde x(s-h_k)\})|\,ds \le |\xi| + \int_0^t \tilde c(s)(1+|\{\tilde x(s-h_k)\}|)\,ds$$

for all $t \in [0,T]$, so that we obtain

$$\|\tilde x\|_{L^\infty(0,t)} \le |\xi| + \int_0^t \tilde c(s)\big[1 + (N+1)\big(\|\tilde x\|_{L^\infty(0,s)} + \|\zeta\|_{L^\infty(-h,0)}\big)\big]\,ds.$$

A straightforward application of the Gronwall Lemma to the function $t \mapsto \|\tilde x\|_{L^\infty(0,t)}$, and the very definition of M in (15), imply that $\|\tilde x\|_{L^\infty(0,T)} \le M$. Hence, $\tilde x$ is the unique solution to (14) and $x = \tilde x + \tilde x_0$ is the unique solution to (6) associated with (μ,α) and such that $x(0) = \xi$. The proof is thus complete. □

Next proposition establishes that the set of strict sense trajectories is dense, in a suitable sense, in the set of impulsive, BV trajectories.

Proposition 2.3 (Density). *Let hypotheses* (H1)–(H3) *and* (H4)* *be satisfied. Then, for any impulsive process $(\bar\mu,\bar\alpha,\bar x)$, there exists a sequence of strict sense processes $(\mu_i,\alpha_i,x_i)_i$, such that*

$$\mu_i \rightharpoonup^* \bar\mu, \quad \ell\big(\{t\in[0,T]:\ \alpha_i(t) = \bar\alpha(t)\}\big) \to T, \quad dx_i \rightharpoonup^* d\bar x, \quad x_i(t) \to \bar x(t) \ \ \forall t \in \mathcal{E},$$

where $\mathcal{E}$ is a full measure subset of $[0,T]$ which contains 0 and T.

Finally, under an additional hypothesis, we obtain the following continuity property of the input-output map.

Theorem 2.4 (Continuity of the input-output map). *Assume hypotheses* (H1)–(H3)*, let G be independent of a, i.e. $G(t,a) = \hat G(t)$, and let $\hat G$ be continuous on $[0,T]$. Then, the input-output map $\mathcal{I} : \mathbb{R}^n \times \mathcal{U} \mapsto L^\infty([-h,T],\mathbb{R}^n) \cap BV([0,T],\mathbb{R}^n)$, given by*

$$(\bar\xi,\bar\mu,\bar\alpha) \mapsto \mathcal{I}(\bar\xi,\bar\mu,\bar\alpha) = \bar x, \tag{19}$$

where $\bar x$ is the impulsive trajectory associated with the control $(\bar\mu,\bar\alpha)$ and such that $\bar x(0) = \bar\xi$, is (well-defined and) continuous, in the sense that all sequences $(\xi_i,\mu_i,\alpha_i)_i \subset \mathbb{R}^n \times \mathcal{U}$ that satisfy

$$\xi_i \to \bar\xi, \quad (\mu_i,|\mu_i|) \rightharpoonup^* (\bar\mu,\lambda) \text{ for some } \lambda \in C^\oplus,\ \ell(\{t\in[0,T]:\ \alpha_i(t) = \bar\alpha(t)\}) \to T, \tag{20}$$

have corresponding trajectories $x_i := \mathcal{I}(\xi_i,\mu_i,\alpha_i)$ that satisfy

$$dx_i \rightharpoonup^* d\bar x, \qquad x_i(t) \to \bar x(t) \quad \forall t \in \mathcal{E} \tag{21}$$

where $\mathcal{E}$ is a full measure subset of $[0,T]$ which contains 0 and T.

The proofs of Prop. 2.3 and Thm. 2.4, which require some preliminary results, will be given in the next subsection.

The additional assumptions that G does not depend on a and is continuous are crucial for the validity of Thm. 2.4, as shown by the following simple, delay-free examples.

Example 2.5. Consider $T = 1$, $G(t,a) = a$, $A(t) \equiv [0,1]$, $\mathcal{K} = \mathbb{R}_{\geq 0}$, and $f \equiv 0$. Let $\xi_i \equiv \xi_0 = 0$, $\mu_i \equiv \mu = \delta_{\{0\}}$ be the Dirac unit measure concentrated at 0, $\alpha(t) \equiv 0$, and let $(\alpha_i)_i \subset \mathcal{A}$ be defined by

$$\alpha_i(t) := \begin{cases} 1 & t \in \left[0, \frac{1}{i}\right], \\ 0 & \text{otherwise.} \end{cases}$$

Of course, $\mu_i \rightharpoonup^* \mu$ and $\ell(\{t : \alpha_i(t) \neq \alpha(t)\}) = \frac{1}{i} \to 0$, but $dx_i(t) = \alpha_i(t)\mu_i(dt)$ does not weakly* converge to $dx(t) = 0$ in C^*. Indeed, take any $\varphi \in C([0,1],\mathbb{R})$ with $\varphi(0) \neq 0$, then $\int_{[0,1]} \varphi(t)\alpha_i(t)\mu_i(dt) = \varphi(0)\alpha_i(0) = \varphi(0)$ for any i, while $\int_{[0,1]} \varphi(t)dx(t) = 0$. Furthermore, for any $t > 0$, we have $x(t) = 0$, while $x_i(t) = 1$ for any $i > 1/t$, so that $x_i(t) \to 1 \neq x(t)$.

Example 2.6. Let $T = 1$, $G(t) = \chi_{]0,1]}(t)$, $\mu = \delta_{\{0\}}$, $f = 0$, and $\mu_i(dt) = m_i(t)dt$, where $m_i(t)$ is given by

$$m_i(t) := \begin{cases} i & \forall t \in \left[0, \frac{1}{i}\right], \\ 0 & \text{otherwise.} \end{cases}$$

Take $\varphi \in C([0,1])$. By the mean value theorem, for any i there exists $t_i \in \left[0, \frac{1}{i}\right]$ such that

$$\lim_i \int_{[0,1]} \varphi(t)\mu_i(dt) = \lim_i \; i\int_0^{\frac{1}{i}} \varphi(t) = \lim_i \varphi(t_i) = \varphi(0) = \int_{[0,1]} \varphi(t)\mu(dt).$$

Thus, $\mu_i \rightharpoonup^* \mu$. However, we have that $\int_{[0,1]} G(t)\mu(dt) = 0$ but $\int_{[0,1]} G(t)\mu_i(dt) = \int_0^1 m_i(t)dt = 1$ for all i. Therefore, choosing $\varphi \equiv 1$, we see that $dx_i(t) = G(t)\mu_i(dt)$ does not weakly* converge in C^* to $dx(t) = G(t)\mu(dt)$. Similarly to the previous example, for any $t > 0$, $x(t) = 0$ and $x_i(t) = 1$ for any $i > 1/t$, so that we still have $x_i(t) \to 1 \neq x(t)$.

2.2. **Proofs of Prop. 2.3 and Thm. 2.4.** In these proofs, equivalence between the given impulsive system and an auxiliary impulsive system, and some approximation results for measures play a key role. So, let us start with these preliminary results. Consider the set of *auxiliary controls* $\tilde{\mathcal{U}}$, given by

$$\tilde{\mathcal{U}} := \left\{(\nu,\alpha,\omega) : \; \nu \in C^{\oplus}, \; \alpha \in \mathcal{A}_\nu, \;\; \omega \in \mathcal{W}_\nu, \;\; t \mapsto \mathfrak{g}(t,\alpha(t),\omega(t)) \; \nu\text{-integrable}\right\}, \tag{22}$$

where $\mathfrak{g}$ is as in (12) and, for $\tilde{\mathcal{K}}$ as in (H4), $\mathcal{W}_\nu$ is given by

$$\mathcal{W}_\nu := \{\omega : [0,T] \to \mathbb{R}^m, \; \omega \text{ is Borel measurable}, \; \omega(t) \in \tilde{\mathcal{K}} \text{ a.e. and } \nu\text{-a.e.}\}.$$

We introduce the following *auxiliary impulsive control system* with time delays,

$$\begin{cases} dx(t) = f\left(t, \{x(t-h_k)\}, \alpha(t)\right) dt + \mathfrak{g}(t,\alpha(t),\omega(t))\nu(dt), & t \in [0,T], \\ x(t) = \zeta(t) \;\; \text{a.e. } t \in [-h, 0[, \end{cases} \tag{23}$$

Precisely, for any auxiliary control (ν, α, ω), we call *auxiliary trajectory* any bounded variation function x, satisfying

$$\begin{cases} x(t) = x(0) + \displaystyle\int_0^t f\left(s, \{x(s-h_k)\}, \alpha(s)\right) ds + \int_{[0,t]} \mathfrak{g}(s, \alpha(s), \omega(s))\nu(ds) \ \forall t \in]0, T], \\ x(t) = \zeta(t) \quad \text{a.e. } t \in [-h, 0[, \end{cases} \tag{24}$$

and refer to (ν, α, ω, x) as an *auxiliary process*. When $\nu \ll \ell$, we say that (ν, α, ω, x) is a *strict sense auxiliary process*, and we call x a *strict sense auxiliary trajectory* and (ν, α, ω) a *strict sense auxiliary control*.

Proposition 2.7. *There is a one-to-one correspondence between processes for* (6) *and auxiliary processes. Precisely, we have that*

(i) given a process (μ, α, x) for (6), the four-tuple (ν, α, ω, x), where

$$\omega = \frac{d\mu}{d|\mu|}, \qquad \nu(dt) := \Big(1 + \sum\nolimits_{i=1}^n \Big|\sum\nolimits_{j=1}^m g_{ij}(t, \alpha(t))\omega^j(t)\Big|\Big)|\mu|(dt), \tag{25}$$

is an auxiliary process;

(ii) given an auxiliary process (ν, α, ω, x), then the triple (μ, α, x), where μ is given by

$$\mu(dt) := \frac{\omega(t)\nu(dt)}{1 + \sum_{i=1}^n \big|\sum_{j=1}^m g_{ij}(t, \alpha(t))\omega^j(t)\big|}, \tag{26}$$

is a process for (6).

In addition, a process (μ, α, x) is strict sense if and only if the corresponding auxiliary process (ν, α, ω, x) is strict sense.

Proof. The proof follows the same lines as the proof of [21, Lemma 6.1], so we illustrate only the key points, omitting details. Concerning statement (i), given a process (μ, α, x) we consider the control triple (ν, α, ω), where ω and ν are defined as above. Note that $\big(1 + \sum_{i=1}^n \big|\sum_{j=1}^m g_{ij}(t, \alpha(t))\omega^j(t)\big|\big)$ is $|\mu|$ integrable in view of the definition of integrability that we have adopted; consequently ν is well defined. From (5) it follows that the function $t \mapsto \mathfrak{g}(t, \alpha(t), \omega(t))$ is ν-integrable and

$$\int_B G(t, \alpha(t))\mu(dt) = \int_B \mathfrak{g}(t, \alpha(t), \omega(t))\nu(dt) \qquad \forall B \in \mathcal{B}. \tag{27}$$

Since clearly $\nu \ll |\mu|$ (in fact, $\mathcal{A}_\mu \subseteq \mathcal{A}_\nu$), this implies that $(\nu, \alpha, \omega) \in \tilde{\mathcal{U}}$. As a direct consequence of (27) we deduce that (ν, α, ω, x) is an auxiliary process. Furthermore, when (μ, α, x) is a strict sense process, namely $\mu \ll \ell$, one also has $|\mu| \ll \ell$, so that $\nu \ll |\mu|$ implies $\nu \ll \ell$ and (ν, α, ω, x) is strict sense too.

Conversely, take an auxiliary process (ν, α, ω, x) and consider the control pair (μ, α), in which μ is defined by (26). It is straightforward that $\mu \ll \nu$. Moreover, the function $t \mapsto G(t, \alpha(t))$ is μ-integrable and (27) still holds. Therefore, $(\mu, \alpha) \in \mathcal{U}$ and (μ, α, x) is a process for (6). Finally, the relation $\mu \ll \nu$ implies that (μ, α, x) is a strict sense process as soon as (ν, α, ω, x) is. ☐

We will also use the following three lemmas.

Lemma 2.8. *Consider some sequences $(\gamma_i)_i \subset C^*_{\mathbb{R}^n}([0, T])$, $(\xi_i)_i \subset \mathbb{R}^n$, $(\varphi_i)_i \subset \mathcal{M}([0, T] \times (\mathbb{R}^n)^{N+1}, \mathbb{R}^n)$, and some $\gamma_0 \in C^*_{\mathbb{R}^n}([0, T])$, $\xi_0 \in \mathbb{R}^n$, $\varphi_0 \in \mathcal{M}([0, T] \times (\mathbb{R}^n)^{N+1}, \mathbb{R}^n)$, and $\eta \in L^\infty([-h, 0], \mathbb{R}^n)$ satisfying the following conditions* (i)–(iv)*:*

(i) *For any $M > 0$, there exists $\tilde{L}_M \in L^1([0,T], \mathbb{R}_{\geq 0})$ such that, for each $i \geq 1$,*

$$|\varphi_i(t, \{z_k\}) - \varphi_i(t, \{y_k\})| \leq \tilde{L}_M(t) |\{z_k\} - \{y_k\}|, \tag{28}$$

for all $\{z_k\}$, $\{y_k\} \in M\,\mathbb{B}_{n(N+1)}$ and for a.e. $t \in [0,T]$. Moreover, one has

$$\ell(\{t \in [0,T] : \varphi_i(t, \{z_k\}) = \varphi_0(t, \{z_k\}) \quad \forall \{z_k\} \in (\mathbb{R}^n)^{N+1}\}) \to T. \tag{29}$$

(ii) *Given a function $z_0 \in BV([-h,T], \mathbb{R}^n)$ satisfying*

$$\begin{cases} z_0(t) = \xi_0 + \int_0^t \varphi_0(s, \{z_0(s - h_k)\}) ds + \int_{[0,t]} \gamma_0(ds) & \forall t \in]0,T], \\ z_0(t) = \eta(t) \quad a.e.\ t \in [-h, 0[, \quad z_0(0) = \xi_0, \end{cases}$$

there exists $\tilde{c} \in L^1([0,T], \mathbb{R}_{\geq 0})$ such that, for each $i \geq 1$,

$$|\varphi_i(t, \{z_0(t - h_k)\})| \leq \tilde{c}(t) \qquad a.e.\ t \in [0,T]. \tag{30}$$

(iii) *The sequence $(\gamma_i)_i$ is uniformly bounded in total variation and there exists a Borel subset $\mathcal{E} \subset [0,T]$ with $\ell(\mathcal{E}) = T$ such that*

$$\lim_i \int_{[0,t]} \gamma_i(ds) = \int_{[0,t]} \gamma_0(ds) \qquad \forall t \in \mathcal{E}.$$

(iv) *The sequence $\xi_i \to \xi_0$.*

Then, for any integer i large enough there exists $z_i \in BV([-h,T], \mathbb{R}^n)$, that satisfies

$$\begin{cases} z_i(t) = \xi_i + \displaystyle\int_0^t \varphi_i(s, \{z_i(s - h_k)\}) ds + \int_{[0,t]} \gamma_i(ds) & \forall t \in]0,T], \\ z_i(t) = \eta(t) \quad a.e.\ t \in [-h, 0[, \quad z_i(0) = \xi_i. \end{cases} \tag{31}$$

Moreover, $dz_i \rightharpoonup^* dz_0$ and $z_i(t) - \int_{[0,t]} \gamma_i(ds) \to z_0(t) - \int_{[0,t]} \gamma_0(ds)$ uniformly in $[0,T]$, so that $z_i(t) \to z(t)$ for all $t \in \mathcal{E}$.

The proof of Lemma 2.8, extending to time delayed systems the results of [21, Prop. 5.1], will be given in the Appendix.

Lemma 2.9. *Let $\gamma \in C^*_{\mathbb{R}^k}([0,T])$ and $(\gamma_i)_i \subset C^*_{\mathbb{R}^k}([0,T])$. Then, the following properties hold true.*

(i) If $(\gamma_i)_i$ is a uniformly bounded sequence in total variation and there exists a Borel subset $\mathcal{E} \subset [0,T]$ with $\ell(\mathcal{E}) = T$, $T \in \mathcal{E}$, and such that

$$\lim_i \int_{[0,t]} \gamma_i(ds) = \int_{[0,t]} \gamma(ds) \qquad \forall t \in \mathcal{E},$$

then $\gamma_i \rightharpoonup^* \gamma$.

(ii) If $(\gamma_i, |\gamma_i|) \rightharpoonup^* (\gamma, \lambda)$ for some $\lambda \in C^\oplus$, then there exists a Borel subset $\mathcal{E} \subset [0,T]$ containing T, such that $[0,T] \setminus \mathcal{E}$ is at most countable and

$$\lim_i \int_{[0,t]} \Psi(s) \cdot \gamma_i(ds) = \int_{[0,t]} \Psi(s) \cdot \gamma(ds) \qquad \forall t \in \mathcal{E},$$

for any integer $n \geq 1$ and every continuous function $\Psi : [0,T] \to \mathbb{R}^{n \times k}$.

(iii) If $\gamma_i \rightharpoonup^* \gamma$, there exists a subsequence $(\gamma_{i_l})_l$ of $(\gamma_i)_i$ such that $(\gamma_{i_l}, |\gamma_{i_l}|) \rightharpoonup^* (\gamma, \lambda)$ for some $\lambda \in C^\oplus$.

Statement (i) coincides with [21, Prop. 5.2 (a)], while (ii) and (iii) slightly extend the results of [21, Prop. 5.2 (b)]. We will then provide a concise proof of (ii), (iii) in the Appendix.

To state the third lemma, let us define the sets $\mathcal{A}$, $\mathcal{W}$, as

$$\mathcal{A} := \{\alpha \in \mathcal{M}([0,T], \mathbb{R}^q) : \alpha(t) \in A(t) \text{ a.e.}\}, \quad \mathcal{W} := \mathcal{M}([0,T], \tilde{\mathcal{K}}).$$

Lemma 2.10. [21, Prop. 5.3] *Assume* (H1), (H4)*, and let* (ν, α, ω) *be an auxiliary control. Then, there exist sequences* $(m_i)_i \subset L^1([0,T], \mathbb{R}_{\geq 0})$ *and* $(\alpha_i, \omega_i)_i \subset \mathcal{A} \times \mathcal{W}$ *such that*

$$m_i(t)dt \rightharpoonup^* \nu(dt), \quad \ell(\{t \in [0,T] : (\alpha_i, \omega_i)(t) \neq (\alpha, \omega)(t)\}) \to 0.$$

Moreover, $t \mapsto \mathfrak{g}(t, \alpha_i(t), \omega_i(t))m_i(t)$ *is Lebesgue integrable for any* i *and*

$$\mathfrak{g}(t, \alpha_i(t), \omega_i(t))m_i(t)dt \rightharpoonup^* \mathfrak{g}(t, \alpha(t), \omega(t))\nu(dt).$$

If (H4) *is replaced with the stronger assumption* (H4)**, then the function* $t \mapsto (\mathfrak{g}, \mathfrak{h})(t, \alpha_i(t), \omega_i(t))m_i(t)$ *is Lebesgue integrable for any* i *and*

$$(\mathfrak{g}, \mathfrak{h})(t, \alpha_i(t), \omega_i(t))m_i(t)dt \rightharpoonup^* (\mathfrak{g}, \mathfrak{h})(t, \alpha(t), \omega(t))\nu(dt).$$

Proof of Prop. 2.3. Given an impulsive process $(\bar{\mu}, \bar{\alpha}, \bar{x})$ as in the statement, let us consider the corresponding auxiliary process $(\bar{\nu}, \bar{\alpha}, \bar{\omega}, \bar{x})$, as defined in Thm. 2.7. From Lemma 2.10 applied to the auxiliary control $(\bar{\nu}, \bar{\alpha}, \bar{\omega})$, it follows that there exists a sequence $(\nu_i, \alpha_i, \omega_i)_i \subset C^{\oplus} \times \mathcal{A} \times \mathcal{W}$ such that, for any i, $\nu_i(dt) = m_i(t)dt$ for some $m_i \in L^1([0,T], \mathbb{R}_{\geq 0})$, $m_i(t)dt \rightharpoonup^* \bar{\nu}(dt)$, $\ell(\{t \in [0,T] : (\alpha_i, \omega_i)(t) \neq (\bar{\alpha}, \bar{\omega})(t)\}) \to 0$, $t \mapsto (\mathfrak{g}, \mathfrak{h})(t, \alpha_i(t), \omega_i(t))m_i(t)$ is Lebesgue integrable for any i, and $(\mathfrak{g}, \mathfrak{h})(t, \alpha_i(t), \omega_i(t))m_i(t)dt \rightharpoonup^* (\mathfrak{g}, \mathfrak{h})(t, \bar{\alpha}(t), \bar{\omega}(t))\bar{\nu}(dt)$. Notice that, for any i, $(\nu_i, \alpha_i, \omega_i)$ is a strict sense auxiliary control, since $\nu_i \ll \ell$. Recalling that $\bar{\mu}(dt) = \mathfrak{h}(t, \bar{\alpha}(t), \bar{\omega}(t))\bar{\nu}(dt)$, if we set, for any i,

$$\mu_i(dt) := \mathfrak{h}(t, \alpha_i(t), \omega_i(t))m_i(t)\, dt, \quad \gamma_i(dt) := \mathfrak{g}(t, \alpha_i(t), \omega_i(t))m_i(t)\, dt$$

then $(\mu_i, \alpha_i)_i \subset \mathcal{U}$ and we have

$$\mu_i \rightharpoonup^* \bar{\mu}, \qquad \gamma_i \rightharpoonup^* \gamma_0 := \mathfrak{g}(t, \bar{\alpha}(t), \bar{\omega}(t))\bar{\nu}(dt) = G(t, \bar{\alpha}(t)) \cdot \bar{\mu}(dt).$$

From Lemma 2.9,(ii)-(iii), with Ψ identity matrix of $\mathbb{R}^{m\times m}$, passing possibly to a subsequence (we do not relabel), it follows that there exists a Borel subset $\mathcal{E} \subset [0,T]$ containing T, such that $[0,T] \setminus \mathcal{E}$ is at most countable, and

$$\lim_i \int_{[0,t]} \gamma_i(ds) = \int_{[0,t]} \gamma_0(ds) \qquad \forall t \in \mathcal{E}.$$

Then, a straightforward application of Lemma 2.8 in which, in particular, $z_0 \equiv \bar{x}$ and $\varphi_i(t, \{z_k\}) := f(t, \{z_k\}, \alpha_i(t))$, so that (H3) implies $|\varphi_i(t, \{\bar{x}(t-h_k)\})| \leq c(t)(1 + \|\bar{x}\|_{L^\infty}) =: \tilde{c}(t) \in L^1([0,T], \mathbb{R}_{\geq 0})$, leads us to deduce that, for any i sufficiently large, there exists a strict sense trajectory x_i associated with the control (μ_i, α_i), for which $x_i(0) = \bar{x}(0)$, $dx_i \rightharpoonup^* d\bar{x}$, and $x_i(t) \to \bar{x}(t)$ on $\mathcal{E}$. □

Proof of Thm. 2.4. Let $\bar{\xi} \in \mathbb{R}^n$ and $(\bar{\mu}, \bar{\alpha}) \in \mathcal{U}$, let $\bar{x} = \mathcal{I}(\bar{\xi}, \bar{\mu}, \bar{\alpha})$ (of course $\mathcal{I}$ is well defined in view of Prop. 2.2) and let $(\xi_i, \alpha_i, \mu_i)_i \subset \mathbb{R}^n \times \mathcal{U}$ satisfy the convergence conditions in (20). Since the function $t \mapsto \hat{G}(t)$ is continuous and $(\mu_i, |\mu_i|) \rightharpoonup^* (\bar{\mu}, \lambda)$, from Lemma 2.9, (ii) it follows that there exists a Borel subset $\mathcal{E} \subset [0,T]$ containing T such that $[0,T] \setminus \mathcal{E}$ is at most countable and

$$\lim_i \int_{[0,t]} \hat{G}(s) \cdot \mu_i(ds) = \int_{[0,t]} \hat{G}(s) \cdot \bar{\mu}(ds) \qquad \forall t \in \mathcal{E}.$$

At this point, taking $\varphi_i(t,\{z_k\}) = f(t,\{z_k\},\alpha_i(t))$, $\varphi_0(t,\{z_k\}) = f(t,\{z_k\},\bar\alpha(t))$, $\gamma_i(dt) = \hat G(t)\mu_i(dt)$, $z_0 \equiv \bar x$, and $\gamma_0(dt) = \hat G(t)\bar\mu(dt)$ (again, $|f(t,\{\bar x(t-h_k)\},\alpha_i(t))| \le c(t)(1+\|\bar x\|_{L^\infty}) \in L^1([0,T],\mathbb{R}_{\ge 0})$), we can apply Lemma 2.8 and conclude that $dx_i \rightharpoonup^* d\bar x$ and $\lim_i x_i(t) = \bar x(t)$ for any $t\in\mathcal{E}$. □

Remark 2.11. Actually, the set $\mathcal{E}$ of Thm. 2.4 is exactly the set of continuity points of λ, weak limit of $(|\mu_i|)_i$. In particular, when the cone $\mathcal{K}$ coincides the first orthant, $\mathcal{E}$ is the set of continuity points of $|\mu|$, as in this case $\lambda \equiv |\mu|$.

3. **Necessary conditions of optimality.** This section is devoted to introducing necessary optimality conditions for the following impulsive optimal control problem with time delays, labeled (P):

$$\text{Minimize}\quad \Phi(x(0),x(T)) + \int_0^T l_0\left(t,\{x(t-h_k)\},\alpha(t)\right)dt + \int_{[0,T]} l_1(t,\alpha(t))\mu(dt)$$

over the set of $(\mu,\alpha,x) \in \mathcal{U}\times(L^\infty([-h,T],\mathbb{R}^n)\cap BV([0,T],\mathbb{R}^n))$, satisfying

$$\begin{cases} x(t) = x(0) + \displaystyle\int_0^t f\left(s,\{x(s-h_k)\},\alpha(s)\right)ds + \int_{[0,t]} G(s,\alpha(s))\mu(ds) \quad \forall t\in]0,T], \\ x(t) = \zeta(t) \quad \text{a.e. } t\in[-h,0[, \end{cases}$$

the endpoint constraint

$$(x(0),x(T)) \in \mathcal{T},$$

and such that the function $t\mapsto l_1(t,\alpha(t))$ is μ-integrable.

The functions f, G and ζ, as well as the control set $\mathcal{U}$, are as in the previous section. The data now also comprise the cost functions $\Phi:\mathbb{R}^{2n}\to\mathbb{R}$, $l_0:[0,T]\times\mathbb{R}^{n(N+1)}\times\mathbb{R}^q\to\mathbb{R}$ and $l_1:[0,T]\times\mathbb{R}^q\to\mathbb{R}^m$, and the target set $\mathcal{T}\subset\mathbb{R}^{2n}$.
Let us introduce the subset $\mathcal{U}_1\subset\mathcal{U}$ of controls, given by

$$\mathcal{U}_1 := \{(\mu,\alpha)\in\mathcal{U}: \quad [0,T]\ni t\mapsto l_1(t,\alpha(t)) \ \text{ is } \mu\text{-integrable}\}. \tag{32}$$

We say that a triple (μ,α,x) is a *feasible process* if it is an impulsive process as defined in Sec. 2, such that $(\mu,\alpha)\in\mathcal{U}_1$ and $(x(0),x(T))\in\mathcal{T}$. Thus, the optimization problem (P) can be reformulated as the minimization of the functional

$$\mathcal{J}(\mu,\alpha,x) := \Phi(x(0),x(T)) + \int_0^T l_0\left(t,\{x(t-h_k)\},\alpha(t)\right)dt + \int_{[0,T]} l_1(t,\alpha(t))\mu(dt) \tag{33}$$

over feasible processes.

Definition 3.1. We say that a feasible process $(\bar\mu,\bar\alpha,\bar x)$ is *optimal* for (P) if

$$\mathcal{J}(\bar\mu,\bar\alpha,\bar x) \le \mathcal{J}(\mu,\alpha,x)$$

for any feasible process (μ,α,x).

Given a feasible process $(\bar\mu,\bar\alpha,\bar x)$, which we will call the *reference process*, in addition to hypotheses (H1)-(H4), we shall invoke also the following assumptions.

(H5) *The function l_0 satisfies hypotheses* (H2)-(H3) *and the function l_1 satisfies assumption* (H4)*, when they are inserted in place of f and G, respectively.*

(H6) *Φ is Lipschitz continuous on a neighborhood of $(\bar x(0),\bar x(T))$ and $\mathcal{T}$ is a closed set.*

Furthermore, for any $t \in \mathbb{R}$, we define $\bar{l}_0[t]$ and $\bar{f}[t]$ as

$$\bar{l}_0[t] := l_0(t, \{\bar{x}(t-h_k)\}, \bar{\alpha}(t)), \qquad \bar{f}[t] := f(t, \{\bar{x}(t-h_k)\}, \bar{\alpha}(t)). \tag{34}$$

Let us state the main result of this section.

Theorem 3.2 (Maximum Principle)**.** *Let $(\bar{\mu}, \bar{\alpha}, \bar{x})$ be an optimal process for problem* (P) *and assume that the data satisfy* (H1)–(H4) *and* (H5)–(H6)*. Then, there exist $\lambda \geq 0$ and $p_k \in W^{1,1}([-h_k, T], \mathbb{R}^n)$, $k = 0, \dots, N$, such that*

$$\dot{p}_k(t) = 0 \quad a.e.\ t \in [-h_k, 0], \quad p_k(t) = 0 \quad \forall t \in [(T-h_k) \vee 0, T], \tag{35}$$

for any $k = 1, \dots, N$, and satisfying conditions (36)–(41) *below:*

$$\|p\|_{L^\infty} + \lambda \neq 0; \tag{36}$$

$$\begin{aligned}\big(-\dot{p}_0((t-h_0)\vee 0), \dots, -\dot{p}_N((t-h_N)\vee 0)\big)& \\ \in \operatorname{co}\partial_{x_0,\dots,x_N}\big(p(t)\cdot \bar{f}[t] - \lambda \bar{l}_0[t]\big)& \quad a.e.\ t \in [0,T];\end{aligned} \tag{37}$$

$$(p(0), -p(T)) \in N_{\mathcal{T}}(\bar{x}(0), \bar{x}(T)) + \lambda\, \partial\Phi(\bar{x}(0), \bar{x}(T)); \tag{38}$$

$$\begin{aligned}\max_{a\in A(t)} \{p(t)\cdot f(t, \{\bar{x}(t-h_k)\}, a) - \lambda l_0(t, \{\bar{x}(t-h_k)\}, a)\}& \\ = p(t)\cdot \bar{f}[t] - \lambda \bar{l}_0[t]& \quad a.e.\ t \in [0,T];\end{aligned} \tag{39}$$

$$\sup_{a\in A(t)} \{\sigma_{\mathcal{K}}(p(t)\cdot G(t,a) - \lambda l_1(t,a))\} \leq 0 \quad \forall t \in [0,T]; \tag{40}$$

$$\sigma_{\mathcal{K}}\big(p(t)\cdot G(t,\bar{\alpha}(t)) - \lambda l_1(t,\bar{\alpha}(t))\big) = 0 \quad \bar{\mu}\text{-}a.e.\ t \in [0,T], \tag{41}$$

where $p \in W^{1,1}([0,T], \mathbb{R}^n)$ is given by

$$p(t) := \sum_{k=0}^{N} p_k(t) \qquad \forall t \in [0,T]. \tag{42}$$

It follows from (37) *and* (42) *that the adjoint arc p satisfies the following "advance" functional differential inclusion*

$$-\dot{p}(t) \in \sum_{j=0}^{N} \operatorname{co}\tilde{\partial}_{x_j}\Big(p(t+h_j)\cdot \bar{f}[t+h_j] - \lambda \bar{l}_0[t+h_j]\Big)\,\chi_{[0,T]}(t+h_j) \quad a.e.\ t \in [0,T]. \tag{43}$$

(Here $\tilde{\partial}_{x_j}$ denotes the projected limiting subdifferential onto the j-th delayed state coordinate). Furthermore, when μ is scalar valued and $\mathcal{K} = \mathbb{R}_{\geq 0}$, conditions (40) *and* (41) *can be replaced as follows:*

$$\sup_{a\in A(t)} \{p(t)\cdot G(t,a) - \lambda l_1(t,a)\} \leq 0 \quad \forall t \in [0,T], \tag{44}$$

$$p(t)\cdot G(t,\bar{\alpha}(t)) - \lambda l_1(t,\bar{\alpha}(t)) = 0 \quad \bar{\mu}\text{-}a.e.\ t \in [0,T]. \tag{45}$$

Remark 3.3. Assumptions (H1)–(H3) are slightly different from those used in the maximum principle with time delays in [3, Thm. 2.1], of which Thm. 3.2 can be seen as an extension to our impulsive problem. In particular, condition (H2) is stronger than the usual local Lipschitz continuity condition in the state variable, while (H3) is usually replaced by an integrably boundedness assumption in some ε-neighborhood of the reference trajectory $\{\bar{x}(t-h_k)\}$ in the L^∞-norm. Actually, the proof of Thm. 3.2 relies on Ekeland's variational principle, but, because of the presence of impulses, our approximating trajectories converge in general to the reference trajectory only almost everywhere, not in the L^∞-norm (see Sec. 4 below).

Hypothesis (H3) is then used to guarantee the existence of solutions to (6) laying in some compact subset of $\mathbb{R}^n$ (as well as to ensure the continuity of the input-output map (19)). Anyway, as it is easy to see, with regard to the drift term f, (H1)–(H3) imply the assumptions of [3, Thm. 2.1].

Remark 3.4. In a nonsmooth setting, condition (37) is a more precise condition than (43), as the subdifferential $\partial_{\{x_k\}}(p\cdot f)$ is a subset, and in some cases a strict subset, of the product of projected partial subdifferentials $\tilde\partial_{x_0}(p\cdot f)\times\cdots\times\tilde\partial_{x_N}(p\cdot f)$.

Remark 3.5. Conditions (40)-(41) locate the support of $\bar\mu$. This might appear to provide rather sparse information, but the fact that the optimal trajectory must satisfy the terminal constraints and the conditions on the conventional control, embodies additional implicit information about $\bar\mu$. In fact, at least in the case without delays, it is a simple matter to show that for linear convex problems, under a normality hypothesis, the conditions of Thm. 3.2 are also sufficient for optimality of $(\bar x,\bar\alpha,\bar\mu)$ (see e.g. [21]). This shows the strength of the necessary conditions.

In order to prove the Maximum Principle, we introduce the following notion.

Definition 3.6. Given a function $\Psi:\mathbb{R}^n\to\mathbb{R}^l$ and a subset $\mathcal{C}\subset\mathbb{R}^n$, we say that an impulsive process $(\bar\mu,\bar\alpha,\bar x)$ of (6) is a $(\Psi,\mathcal{C})$-*boundary process* if $\bar x(0)\in\mathcal{C}$ and $\Psi(\bar x(T))$ belongs to the boundary of the $(\Psi,\mathcal{C})$-*reachable set* $\mathcal{R}^{\mathcal{C}}_\Psi$ *associated with* (6), given by

$$\mathcal{R}^{\mathcal{C}}_\Psi := \{\Psi(x(T)) : (\mu,\alpha,x) \text{ is a process for (6) with } x(0)\in\mathcal{C}\}.$$

We will first establish necessary conditions for $(\Psi,\mathcal{C})$-boundary processes. These conditions are of interest in their own right. But they also can be used simply to derive the necessary conditions of Thm. 3.2.

Theorem 3.7 (Necessary conditions for boundary processes)**.** *Let $(\bar\mu,\bar\alpha,\bar x)$ be a $(\Psi,\mathcal{C})$-boundary process for some $\Psi:\mathbb{R}^n\to\mathbb{R}^l$, which is Lipschitz continuous on a neighborhood of $\bar x(T)$, and for some closed subset $\mathcal{C}\subset\mathbb{R}^n$. Let the data satisfy hypotheses* (H1)–(H4)*. Then, there exist $\eta\in\mathbb{R}^l\setminus\{0\}$ and $p_k\in W^{1,1}([-h_k,T],\mathbb{R}^n)$, $k=0,\dots,N$, such that (35) is satisfied for any $k=1,\dots,N$ and conditions (46)–(51) below hold:*

$$\begin{aligned}\big(-\dot p_0((t-h_0)\vee 0),\dots,-\dot p_N((t-h_N)\vee 0)\big)&\\ \in \operatorname{co}\partial_{x_0,\dots,x_N}\big(p(t)\cdot\bar f[t]\big)&\quad a.e.\ t\in[0,T];\end{aligned}\tag{46}$$

$$p(0)\in N_{\mathcal{C}}(\bar x(0));\tag{47}$$

$$-p(T)\in\eta\cdot\partial\Psi(\bar x(1));\tag{48}$$

$$p(t)\cdot\bar f[t]=\max_{a\in A(t)}\{p(t)\cdot f(t,\{\bar x(t-h_k)\},a)\}\qquad a.e.\ t\in[0,T];\tag{49}$$

$$\sup_{a\in A(t)}\{\sigma_{\mathcal{K}}(p(t)\cdot G(t,a))\}\le 0\quad\forall t\in[0,T];\tag{50}$$

$$\sigma_{\mathcal{K}}\big(p(t)\cdot G(t,\bar\alpha(t))\big)=0\quad\bar\mu\text{-}a.e.\ t\in[0,T],\tag{51}$$

where p is as in (42). It follows from (42) and (46) that p satisfies the following "advance" functional differential inclusion

$$-\dot p(t)\in\sum_{j=0}^N\operatorname{co}\tilde\partial_{x_j}\Big(p(t+h_j)\cdot\bar f[t+h_j]\Big)\,\chi_{[0,T]}(t+h_j)\quad a.e.\ t\in[0,T].\tag{52}$$

Furthermore, when μ is scalar valued and $\mathcal{K} = \mathbb{R}_{\geq 0}$, conditions (50) *and* (51) *can be replaced, respectively, by:*

$$\sup_{a \in A(t)} \{p(t) \cdot G(t,a)\} \leq 0 \quad \forall t \in [0,T], \tag{53}$$

$$p(t) \cdot G(t, \bar{\alpha}(t)) = 0 \quad \bar{\mu}\text{-a.e. } t \in [0,T]. \tag{54}$$

The proof of Thm. 3.7 will be given in the next section. We now show that, as anticipated, Thm. 3.2 can be deduced as a corollary of Thm. 3.7.

Proof of Thm. 3.2. Let $(\bar{\mu}, \bar{\alpha}, \bar{x})$ be an optimal process for problem (P), as in the statement. We introduce a new control system, where we consider as processes the 7-uples $(\mu, \alpha, y_1, \dots, y_5)$, in which $(\mu, \alpha) \in \mathcal{U}_1$, $y_1 \in L^\infty([-h,T], \mathbb{R}^n) \cap BV([0,T], \mathbb{R}^n)$, $y_2 \in BV([0,T], \mathbb{R})$, $(y_3, y_4, y_5) \in W^{1,1}([0,T], \mathbb{R}^n \times \mathbb{R} \times \mathbb{R})$, and that satisfy

$$\begin{cases} y_1(t) = y_1(0) + \displaystyle\int_0^t f(s, \{y_1(s-h_k)\}, \alpha(s))ds + \int_{[0,t]} G(s, \alpha(s))\mu(ds) \quad \forall t \in]0,T], \\ y_2(t) = y_2(0) + \displaystyle\int_0^t l_0(s, \{y_1(s-h_k)\}, \alpha(s))ds + \int_{[0,t]} l_1(s, \alpha(s))\mu(ds) \quad \forall t \in]0,T], \\ \dot{y}_3(t) = 0, \quad \dot{y}_4(t) = 0, \quad \dot{y}_5(t) = 0, \qquad \text{a.e. } t \in [0,T], \\ y_1(t) = \zeta(t) \qquad \text{a.e. } t \in [-h, 0[. \end{cases} \tag{55}$$

Define the Lipschitz function $\Psi : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$, by

$$\Psi(z_1, z_2, z_3, z_4, z_5) = (z_1 - z_3, z_2 - z_4, z_5),$$

the closed set $\mathcal{C} \subset \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$, by

$$\mathcal{C} := \{(z_1, z_2, z_3, z_4, z_5) : (z_1, z_3) \in \mathcal{T}, \; z_2 = 0, \; z_5 \geq \Phi(z_1, z_3) + z_4\},$$

and let $\tilde{\mathcal{R}}^{\mathcal{C}}_{\Psi}$ denote the $(\Psi, \mathcal{C})$-reachable set associated with this new (impulsive) control system. Observe that the process $(\tilde{\mu}, \tilde{\alpha}, \tilde{y}_1, \dots, \tilde{y}_5)$, in which $(\tilde{\mu}, \tilde{\alpha}, \tilde{y}_1) := (\bar{\mu}, \bar{\alpha}, \bar{x})$, and, for every $t \in [0,T]$,

$$\begin{aligned} &\tilde{y}_2(t) := \int_0^t l_0(s, \{\bar{x}(s-h_k)\}, \bar{\alpha}(s))ds + \int_{[0,t]} l_1(s, \bar{\alpha}(s))\bar{\mu}(ds), \\ &\tilde{y}_3(t) \equiv \bar{x}(T), \\ &\tilde{y}_4(t) \equiv \int_0^T l_0(t, \{\bar{x}(t-h_k)\}, \bar{\alpha}(t))dt + \int_{[0,T]} l_1(t, \bar{\alpha}(t))\bar{\mu}(dt) \; (= \tilde{y}_2(T)), \\ &\tilde{y}_5(t) \equiv \mathcal{J}(\bar{\mu}, \bar{\alpha}, \bar{x}). \end{aligned}$$

satisfies (55) with $(\tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \tilde{y}_4, \tilde{y}_5)(0) \in \mathcal{C}$. Moreover, $\Psi((\tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \tilde{y}_4, \tilde{y}_5)(T)) \in \partial \tilde{\mathcal{R}}^{\mathcal{C}}_{\Psi}$, otherwise $(0, 0, \mathcal{J}(\bar{\mu}, \bar{\alpha}, \bar{x}) - \rho) \in \tilde{\mathcal{R}}^{\mathcal{C}}_{\Psi}$ for some $\rho > 0$ sufficiently small, but this clearly contradicts the optimality of $(\bar{\mu}, \bar{\alpha}, \bar{x})$. Thus, by Thm. 3.7 there exist $\mathbf{p}_k := (p_{1_k}, p_{2_k}, p_{3_k}, p_{4_k}, p_{5_k}) \in W^{1,1}([-h_k, T], \mathbb{R}^{n+1+n+1+1})$ for $k = 0, \dots, N$, and $(\eta_1, \eta_2, \eta_3) \in \mathbb{R}^{n+1+1} \setminus \{0\}$ such that (35) holds with $\mathbf{p}_k$ replacing p_k for any $k =$

$1,\dots,N$, and such that conditions (i)–(vii) below are met for some $\lambda \geq 0$:

(i) $\begin{aligned}&(-\dot p_{1_0}(t),\dots,-\dot p_{1_N}((t-h_N)\vee 0))\\ &\qquad\in \operatorname{co}\partial_{x_0,\dots,x_N}\big\{p_1(t)\cdot \bar f[t] + p_2(t)\bar l_0[t]\big\} \quad \text{a.e. } t\in[0,T];\end{aligned}$

(ii) $\dot p_2(t)=0,\;\; \dot p_3(t)=0,\;\; \dot p_4(t)=0,\;\; \dot p_5(t)=0, \qquad \text{a.e. } t\in[0,T];$

(iii) $(p_1,p_3)(0)\in N_{\mathcal T}(\bar x(0),\bar x(T)) + \lambda\partial\Phi(\bar x(0),\bar x(T)),\;\; p_4(0)=\lambda,\;\; p_5(0)=-\lambda;$

(iv) $(p_1,p_2,p_3,p_4,p_5)(T)\in -(\eta_1,\eta_2,\eta_3)\cdot\partial\Psi = (-\eta_1,-\eta_2,\eta_1,\eta_2,-\eta_3);$

(v) $\begin{aligned}&\max_{a\in A(t)}\{p_1(t)\cdot f(t,\{\bar x(t-h_k)\},a) + p_2(t) l_0(t,\{\bar x(t-h_k)\},a)\}\\ &\qquad\qquad = p_1(t)\cdot \bar f[t] + p_2(t)\bar l_0[t] \quad \text{a.e. } t\in[0,T];\end{aligned}$

(vi) $\sup_{a\in A(t)}\{\sigma_{\mathcal K}(p_1(t)\cdot G(t,a) + p_2(t) l_1(t,a))\}\leq 0 \quad \forall t\in[0,T];$

(vii) $\sigma_{\mathcal K}\big(p_1(t)\cdot G(t,\bar\alpha(t)) + p_2(t) l_1(t,\bar\alpha(t))\big) = 0 \quad \bar\mu\text{-a.e. } t\in[0,T],$

where $\bar l_0$ and $\bar f$ are as in (34) and $\mathbf p := (p_1,p_2,p_3,p_4,p_5)$ is given by

$$\mathbf p(t) := \sum_{k=0}^N \mathbf p_k(t) \qquad \forall t\in[0,T].$$

In particular, (iii) follows by [7, Ex. 1.11.26 and Ex. 2.9.11]. Clearly, in view of (i) and the very definition of $\mathbf p$, for a.e. $t\in[0,T]$, p_1 satisfies [3]

$$-\dot p_1(t)\in\sum_{j=0}^N \operatorname{co}\tilde\partial_{x_j}\big\{p_1(t+h_j)\cdot\bar f[t+h_j] + p_2(t+h_j)\bar l_0[t+h_j]\big\}\,\chi_{[0,T]}(t+h_j).$$

By (ii) we deduce that p_2, p_3, p_4, and p_5 are constants and (iv) implies that $p_3 = -p_1(T)$. Considering also (iii), we notice that $p_2 = p_5 = -p_4 = -\lambda$ and $\eta_3 = \eta_2$. We point out that $\|p_1\|_{L^\infty} + \lambda \neq 0$, otherwise $\eta_1 = -p_1(T) = 0$ and $\eta_2 = \eta_3 = p_4 = -\lambda = 0$, which contradicts $(\eta_1,\eta_2,\eta_3)\neq 0$. Therefore, λ and the paths $p_k := p_{1_k}$, $k=0,\dots,N$, satisfy the requirements of Thm. 3.2. $\square$

4. **Proof of Thm. 3.7.** The proof consists of several steps. First of all, we show that it is sufficient to prove the theorem for the auxiliary process associated with the given impulsive boundary process. Then, we construct a sequence of optimization problems having as admissible controls only strict sense controls, and with costs measuring how much a trajectory is distant from the reference one in the L^2-norm. Using the Ekeland variational principle, minimizers are constructed for these problems that converge to the initial boundary process. Moreover, by applying the Maximum Principle in [3, Thm. 2.1] to these problems with reference to the above mentioned minimizers, we obtain in the limit a set of multipliers that satisfy conditions (35)–(41).

Step 1. In view of Prop. 2.7, we can associate with the $(\Psi,\mathcal C)$-boundary process $(\bar\mu,\bar\alpha,\bar x)$ the auxiliary process $(\bar\nu,\bar\alpha,\bar\omega,\bar x)$ in which $\bar\omega = \frac{d\bar\mu}{d|\bar\mu|}$ and $\bar\nu$ is the scalar

[3]With $\tilde\partial_{x_j}$ we mean here the projected limiting subdifferential w.r.t. the j-th component of the variable y_1.

nonnegative measure defined as in (25), so that $(\bar\nu, \bar\alpha, \bar\omega) \in \tilde{\mathcal{U}}$. In particular, the trajectory $\bar x$ can be also interpreted as a solution of the auxiliary control system,

$$\begin{cases} dx(t) = f\left(t, \{x(t-h_k)\}, \alpha(t)\right) dt + \mathfrak{g}(t, \alpha(t), \omega(t))\nu(dt), \quad t \in [0,T], \\ x(t) = \zeta(t) \quad \text{a.e. } t \in [-h, 0[, \end{cases} \tag{56}$$

where $\mathfrak{g}$ is as in (12). Clearly, $(\bar\nu, \bar\alpha, \bar\omega, \bar x)$ turns out to be a $(\Psi, \mathcal{C})$-boundary process for the $(\Psi, \mathcal{C})$-reachable set associated with control system (56).

Suppose we have proved Thm. 3.7 for $(\bar\nu, \bar\alpha, \bar\omega, \bar x)$. Then, since the drift term f is the same for the original and the auxiliary control systems, there exist $\eta \in \mathbb{R}^l \setminus \{0\}$ and $p_k \in W^{1,1}([-h_k, T], \mathbb{R}^n)$, $k = 0, \dots, N$, such that (35) is satisfied for any $k = 1, \dots, N$, and such that conditions (46), (47), (48), (49), and (52) are met. Moreover, conditions (53) and (54) for the auxiliary control system reduce respectively to

$$\sup_{a \in A(t),\ w \in \tilde{\mathcal{K}}} \left\{ \frac{p(t) \cdot G(t,a)w}{1 + \sum_{i=1}^m \left| \sum_{j=1}^m g_{ij}(t,a) w^j \right|} \right\} \le 0 \qquad \forall t \in [0,T] \tag{57}$$

and

$$\frac{p(t) \cdot G(t, \bar\alpha(t))\bar\omega(t)}{1 + \sum_{i=1}^m \left| \sum_{j=1}^m g_{ij}(t, \bar\alpha(t))\bar\omega^j(t) \right|} = 0 \qquad \bar\nu\text{-a.e. } t \in [0,T]. \tag{58}$$

Let us observe that for any $w \in \mathcal{K}$ there exist $\tilde w \in \tilde{\mathcal{K}}$ and $r \ge 0$ such that $w = r\tilde w$. Moreover, from (25) it follows that for any $B \in \mathcal{B}$ one has $\bar\nu(B) > 0$ provided $\bar\mu(B) > 0$. Hence, it is immediate to see that (57) and (58) imply (50) and (51) for the original process $(\bar\mu, \bar\alpha, \bar x)$, respectively. Furthermore, in case $\mathcal{K} = \mathbb{R}_{\ge 0}$ then $\bar\mu \equiv |\bar\mu|$, so that $\bar\omega \equiv 1$ in view of (25). Accordingly, from conditions (57) and (58) we can easily deduce the strengthened relations (53) and (54), respectively. Therefore, Thm. 3.7 is proved for $(\bar\mu, \bar\alpha, \bar x)$ as soon as it is proved for the auxiliary process $(\bar\nu, \bar\alpha, \bar\omega, \bar x)$, and this will be our goal from now on.

Step 2. Let $(\xi_i)_i \subset \mathbb{R}^l \setminus \mathcal{R}^{\mathcal{C}}_\Psi$ be a sequence such that $\xi_i \to \Psi(\bar x(T))$, which exists as $\Psi(\bar x(T)) \in \partial \mathcal{R}^{\mathcal{C}}_\Psi$. From Lemma 2.10 it follows that there exist sequences $(\tilde m_i)_i \subset L^1([0,T], \mathbb{R}_{\ge 0})$ and $(\tilde\alpha_i, \tilde\omega_i)_i \subset \mathcal{A} \times \mathcal{W}$ such that

$$\tilde m_i(t) dt \rightharpoonup^* \bar\nu(dt), \qquad \ell(\{t \in [0,T] : (\tilde\alpha_i, \tilde\omega_i)(t) \ne (\bar\alpha, \bar\omega)(t)\}) \to 0, \tag{59}$$

and $\mathfrak{g}(t, \tilde\alpha_i(t), \tilde\omega_i(t)) \tilde m_i(t) dt \rightharpoonup^* \mathfrak{g}(t, \bar\alpha(t), \bar\omega(t))\bar\nu(dt)$. Hence, by Lemma 2.9, (ii)-(iii), possibly passing to a subsequence (we do not relabel here and in the rest of the proof), we have that there exists a subset $\mathcal{E} \subset [0,T]$ with $T \in \mathcal{E}$ and $[0,T] \setminus \mathcal{E}$ at most countable, such that

$$\lim_i \int_{[0,t]} \mathfrak{g}(t, \tilde\alpha_i(t), \tilde\omega_i(t)) \tilde m_i(t)\, dt = \int_{[0,t]} \mathfrak{g}(t, \bar\alpha(t), \bar\omega(t))\bar\nu(dt) \quad \forall t \in \mathcal{E}.$$

Since the sequence $\left(\mathfrak{g}(t, \tilde\alpha_i(t), \tilde\omega_i(t)) \tilde m_i(t) dt\right)_i$ is uniformly bounded in total variation, by (H4) and (59), and since (H3) implies $|f(t, \{\bar x(t-h_k)\}, \tilde\alpha_i(t))| \le c(t)(1 + (N+1)\|\bar x\|_{L^\infty}) \in L^1([0,T], \mathbb{R}_{\ge 0})$, from Lemma 2.8 it now follows that there exists a sequence $(\tilde x_i)_i$ of functions from $[-h, T]$ to $\mathbb{R}^n$ whose restriction to $[0,T]$ is absolutely continuous, such that, for any i, $\tilde x_i$ is a solution to

$$\begin{cases} x(t) = \bar x(0) + \int_0^t f(s, \{x(s-h_k)\}, \tilde\alpha_i(s)) ds \\ \qquad\qquad + \int_{[0,t]} \mathfrak{g}(s, \tilde\alpha_i(s), \tilde\omega_i(s)) \tilde m_i(s) ds \quad \forall t \in]0,T] \\ x(t) = \zeta(t) \quad \text{a.e. } t \in [-h, 0[, \quad x(0) = \bar x(0). \end{cases}$$

Moreover, $d\tilde{x}_i \rightharpoonup^* d\bar{x}$, so that $(d\tilde{x}_i)_i$ is uniformly bounded in total variation and $(\tilde{x}_i)_i$ is uniformly bounded in L^∞ by (4). Furthermore, $\tilde{x}_i(t) \to \bar{x}(t)$ for all $t \in \mathcal{E}$. Hence, by the dominated convergence theorem we deduce that $\tilde{x}_i \to \bar{x}$ in $L^2([0,T],\mathbb{R}^n)$, so that $\varepsilon_i \to 0$, being $(\varepsilon_i)_i \subset \mathbb{R}_{\geq 0}$ the sequence defined by

$$\varepsilon_i^2 := \int_0^T |\tilde{x}_i(t) - \bar{x}(t)|^2 dt + |\tilde{x}_i(T) - \bar{x}(T)|^2 + |\xi_i - \Psi(\tilde{x}_i(T))|.$$

Possibly passing to a subsequence, we can suppose that $\varepsilon_i \leq 1$ for all i. Let $(C_i)_i$ be the sequence of L^1 functions defined by

$$C_i(t) := i + \max_{j \leq i}\{\tilde{m}_j(t)\}. \tag{60}$$

Clearly, this sequence is monotone, $\tilde{m}_i(t) \leq C_i(t)$, and $C_i(t) \to +\infty$ for all $t \in [0,T]$. For each i, consider the following (non-impulsive) optimization problem with time delays,

$$(\tilde{\mathrm{P}}_\mathrm{i}) \begin{cases} \text{Minimize } \int_0^T |x(t) - \bar{x}(t)|^2 dt + |x(T) - \bar{x}(T)|^2 + |\xi_i - \Psi(x(T))| \\ \text{over controls } (m,\alpha,\omega) \in L^1([0,T],\mathbb{R}_{\geq 0}) \times \mathcal{M}([0,T],\mathbb{R}^q \times \mathbb{R}^m) \\ \text{and arcs } x \in L^\infty([-h,T],\mathbb{R}^n) \cap W^{1,1}([0,T],\mathbb{R}^n), \text{ satisfying} \\ \dot{x}(t) = f(t,\{x(t-h_k)\},\alpha(t)) + \mathfrak{g}(t,\alpha(t),\omega(t))m(t) \quad \text{a.e. } t \in [0,T] \\ x(t) = \zeta(t) \quad \text{a.e. } t \in [-h,0[, \quad x(0) \in \mathcal{C}, \\ (m(t),\alpha(t),\omega(t)) \in [0,C_i(t)] \times A(t) \times \tilde{\mathcal{K}} \quad \text{a.e. } t \in [0,T]. \end{cases}$$

For each i, let Γ_i be the set of elements (m,α,ω,x_0) for which $x_0 \in \mathcal{C}$ and (m,α,ω) is a measurable control such that $(m(t),\alpha(t),\omega(t)) \in [0,C_i(t)] \times A(t) \times \tilde{\mathcal{K}}$ a.e.. Under our hypotheses, for every $(m,\alpha,\omega,x_0) \in \Gamma_i$ there exists exactly one solution

$$x := x[m,\alpha,\omega,x_0] \in L^\infty([-h,T],\mathbb{R}^n) \cap W^{1,1}([0,T],\mathbb{R}^n)$$

to the delayed control system in $(\tilde{P}_i)$ with initial condition $x(0) = x_0$, so that problem $(\tilde{P}_i)$ can be reformulated as

$$\begin{cases} \text{Minimize } \mathcal{J}_i(m,\alpha,\omega,x_0) := \int_0^T |x(t) - \bar{x}(t)|^2 dt + |x(T) - \bar{x}(T)|^2 + |\xi_i - \Psi(x(T))| \\ \text{over } (m,\alpha,\omega,x_0) \in \Gamma_i. \end{cases}$$

The set Γ_i is a complete metric space, when equipped with the metric $\mathbf{d}$ defined by

$$\begin{aligned} \mathbf{d}((m,\alpha,\omega,x_0),(\tilde{m},\tilde{\alpha},\tilde{\omega},\tilde{x}_0)) :=& \|m - \tilde{m}\|_{L^1} + |x_0 - \tilde{x}_0| \\ &+ \ell(\{t \in [0,T] : (\alpha,\omega)(t) \neq (\tilde{\alpha},\tilde{\omega})(t)\}). \end{aligned}$$

(see [6, Lemma 1, p. 202]). Moreover, by the continuity of the input-output map associated with the (conventional) delayed control system in $(\tilde{P}_i)$, for each i there exists a function $\rho_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with $\lim_{d \to 0^+} \rho_i(d) = 0$, such that, for any pair (m,α,ω,x_0), $(\tilde{m},\tilde{\alpha},\tilde{\omega},\tilde{x}_0) \in \Gamma_i$,

$$\|x[m,\alpha,\omega,x_0] - x[\tilde{m},\tilde{\alpha},\tilde{\omega},\tilde{x}_0]\|_{L^\infty} \leq \rho_i\big(\mathbf{d}((m,\alpha,\omega,x_0),(\tilde{m},\tilde{\alpha},\tilde{\omega},\tilde{x}_0))\big).$$

As a consequence, the map $\mathcal{J}_i$ is continuous on Γ_i with respect to $\mathbf{d}$.

Step 3. By the previous arguments, $(\tilde{m}_i,\tilde{\alpha}_i,\tilde{\omega}_i,\tilde{x}_i(0))$ is an ε_i^2-minimizer of problem $(\tilde{\mathrm{P}}_i)$, thus Ekeland's variational principle yields the existence of a sequence

$(m_i, \alpha_i, \omega_i, x_{0_i})$ which is optimal for the following optimization problem

$$(\mathrm{P}_i) \begin{cases} \text{Minimize } \ \mathcal{J}_i(m, \alpha, \omega, x_0) \\ +\varepsilon_i \left(|x_0 - x_{0_i}| + \int_0^T [|m(t) - m_i(t)| + \vartheta_i(t, \alpha(t), \omega(t))] dt \right) \\ \text{over } \ (m, \alpha, \omega, x_0) \in \Gamma_i, \end{cases}$$

where $\vartheta_i : [0, T] \times \mathbb{R}^q \times \mathbb{R}^m \to \{0, 1\}$ is defined by

$$\vartheta_i(t, a, w) := \begin{cases} 0 & \text{if } (a, w) = (\alpha_i(t), \omega_i(t)), \\ 1 & \text{otherwise.} \end{cases}$$

Moreover, it also holds

$$\mathbf{d}((m_i, \alpha_i, \omega_i, x_{0_i}), (\tilde{m}_i, \tilde{\alpha}_i, \tilde{\omega}_i, \tilde{x}_i(0))) \le \varepsilon_i. \tag{61}$$

By (59), (61), and the fact that $\tilde{x}_i(0) \equiv \bar{x}(0)$ we get

$$\ell(\{t \in [0, T] : (\alpha_i(t), \omega_i(t)) \neq (\bar{\alpha}(t), \bar{\omega}(t))\}) \to 0, \tag{62}$$

$$m_i(t)dt \rightharpoonup^* \bar{\nu}(dt), \tag{63}$$

$$x_{0_i} \to \bar{x}(0). \tag{64}$$

By (61), the fact that $(\tilde{m}_i, \tilde{\alpha}_i, \tilde{\omega}_i, \tilde{x}_{0_i})$ is an ε_i^2-minimizer for problem $(\tilde{\mathrm{P}}_i)$, and the fact that $(m_i, \alpha_i, \omega_i, x_{0_i})$ is optimal for (P_i), we deduce that $\mathcal{J}_i(m_i, \alpha_i, x_{0_i}) \to 0$, so that the trajectories $x_i := x[m_i, \alpha_i, \omega_i, x_{0_i}]$ satisfy $x_i \to \bar{x}$ in $L^2([0, T], \mathbb{R}^n)$ and $x_i(T) \to \bar{x}(T)$. Hence, possibly up to a subsequence, we have

$$x_i \to \bar{x} \quad \text{in } L^2, \qquad x_i(t) \to \bar{x}(t) \quad \text{a.e. } t \in [0, T], \qquad x_i(T) \to \bar{x}(T). \tag{65}$$

By (H4), (63), and (64) there exists $\tilde{C} > 0$ such that $(x_{0_i})_i \subset \tilde{C}\mathbb{B}_n$ and the sequence $\big(\mathfrak{g}(t, \alpha_i(t), \omega_i(t)) m_i(t) dt\big)_i$ is bounded in total variation by $\tilde{C}$. Accordingly, using (H3) we get

$$\begin{aligned} |x_i(t)| &\le |x_{0_i}| + \int_0^t |f(s, \{x_i(s - h_k)\}, \alpha_i(s))| \, ds + \int_0^t |\mathfrak{g}(s, \alpha_i(s), \omega_i(s))| \, |m_i(s)| ds \\ &\le 2\tilde{C} + \int_0^t c(s)(1 + |\{x_i(s - h_k)\}|) \, ds, \end{aligned}$$

from which we immediately deduce

$$\|x_i\|_{L^\infty(0,t)} \le 2\tilde{C} + \int_0^t c(s) \Big(1 + (N + 1) \big(\|x_i\|_{L^\infty(0,s)} + \|\zeta\|_{L^\infty(-h,0)} \big) \Big) \, ds.$$

A straightforward application of the Gronwall's Lemma implies that $(x_i)_i$ is a uniformly bounded sequence in L^∞, so that there exists $M > 0$ such that

$$\|\bar{x}\|_{L^\infty([-h,T])} \le M, \qquad \|x_i\|_{L^\infty([-h,T])} \le M \quad \text{for any } i. \tag{66}$$

As a consequence, we obtain

$$\begin{aligned} \int_0^T |\dot{x}_i(t)| dt &\le \int_0^T |f(t, \{x_i(t - h_k)\}, \alpha_i(t))| \, ds + \int_0^T |\mathfrak{g}(t, \alpha_i(t), \omega_i(t))| \, |m_i(t)| dt \\ &\le \int_0^T c(t) \Big(1 + (N + 1) \big(M + \|\zeta\|_{L^\infty(-h,0)} \big) \Big) \, ds + \tilde{C}. \end{aligned}$$

Therefore, the sequence of measures $(dx_i)_i$ associated with the x_i's is uniformly bounded in total variation and, in view of (64)-(65), satisfies

$$\int_{[0,t]} dx_i(s) \to \int_{[0,t]} d\bar{x}(s) \quad \forall t \in \tilde{\mathcal{E}},$$

where $\tilde{\mathcal{E}} \subset [0,T]$ has full Lebesgue measure and contains T. Hence, Lemma 2.9, (i), yields that $dx_i \rightharpoonup^* d\bar{x}$. Thanks to (62), (65), hypothesis (H3) and (66), we can apply the dominated convergence theorem to deduce that

$$f(t,\{x_i(t-h_k)\},\alpha_i(t))dt \rightharpoonup^* f(t,\{\bar{x}(t-h_k)\},\bar{\alpha}(t))dt. \tag{67}$$

Since $\mathfrak{g}(t,\alpha_i(t),\omega_i(t))m_i(t)dt = dx_i(t) - f(t,\{x_i(t-h_k)\},\alpha_i(t))dt$, we obtain that

$$\mathfrak{g}(t,\alpha_i(t),\omega_i(t))m_i(t)dt \rightharpoonup^* \mathfrak{g}(t,\bar{\alpha}(t),\bar{\omega}(t))\bar{\nu}(dt). \tag{68}$$

Step 4. Problem (P_i) is a non-impulsive optimization problem with time delays in the dynamics, for which a Maximum Principle is available. In particular, by applying [3, Thm. 2.1] with reference to the minimizer $(m_i,\alpha_i,\omega_i,x_i)$ we deduce the existence of arcs $p_{k_i} \in W^{1,1}([-h_k,T],\mathbb{R}^n)$, $k=0,\dots,N$, such that

$$\dot{p}_{k_i}(t) = 0 \ \text{ for a.e. } t \in [-h_k,0], \qquad p_{k_i}(t) = 0 \ \ \forall t \in [(T-h_k)\vee 0, T], \tag{69}$$

for $k=1,\dots,N$, and satisfying the following conditions: [4]

$$\begin{aligned}&(-\dot{p}_{0_i}((t-h_0)\vee 0),\dots,-\dot{p}_{N_i}((t-h_N)\vee 0))\\ &\qquad\in \mathrm{co}\,\partial_{x_0,\dots,x_N}\big(p_i(t)\cdot f_i[t]\big) - 2(x_i(t)-\bar{x}(t)) \qquad \text{a.e. } t\in[0,T];\end{aligned} \tag{70}$$

$$p_i(0) \in N_{\mathcal{C}}(x_i(0)) + \varepsilon_i\mathbb{B}_n; \tag{71}$$

$$-p_i(T) \in \partial_x|\xi_i - \Psi(x_i(T))| + 2(x_i(T)-\bar{x}(T)); \tag{72}$$

$$\begin{aligned}&p_i(t)\cdot\big(f_i[t] + \mathfrak{g}(t,\alpha_i(t),\omega_i(t))m_i(t)\big)\\ &\quad= \max_{a\in A(t),\ w\in\tilde{\mathcal{K}},\ m\in[0,C_i(t)]}\big\{p_i(t)\cdot\big[f(t,\{x_i(t-h_k)\},a) + \mathfrak{g}(t,a,w)m\big]\\ &\qquad\qquad -\varepsilon_i(\vartheta_i(t,a,w) + |m-m_i(t)|)\big\} \quad \text{a.e. } t\in[0,T],\end{aligned} \tag{73}$$

where, for any $t\in[0,T]$, $p_i(t)$ and $f_i[t]$ are given by

$$p_i(t) := \sum_{k=0}^{N} p_{k_i}(t), \qquad\qquad f_i[t] := f(t,\{x_i(t-h_k)\},\alpha_i(t)). \tag{74}$$

We now deduce conditions (46)–(51) by passing to the limit in relations (70)–(73). To begin with, observe that, from (70) and the very definition of p_i in (74), using (66) together with (H2), we get

$$\begin{aligned}|p_i(t)| &\le |p_i(T)| + \textstyle\int_{[t,T]}\sum_{k=0}^N|\dot{p}_{k_i}(s)|\,ds\\ &\le |p_i(T)| + \textstyle\sum_{k=0}^N\int_{[t,T]}|p_i(s+h_k)|L_M(s+h_k)\chi_{[0,T]}(s+h_k)\,ds\\ &\quad +2\textstyle\int_{[0,T]}|x_i(s)-\bar{x}(s)|\,ds.\end{aligned}$$

Since $(m_i,\alpha_i,\omega_i,x_i)$ is a strict sense process and $\xi_i\in\mathbb{R}^l\setminus\mathcal{R}^{\mathcal{C}}_{\Psi}$ then $|\xi_i-\Psi(x_i(T))|\neq 0$, hence (72) and the Jacobian chain rule imply that there exists $\eta_i\in\partial\mathbb{B}_l$ such that

$$-p_i(T) \in \eta_i\cdot\partial\Psi(x_i(T)) + 2(x_i(T)-\bar{x}(T)). \tag{75}$$

Thus, by (65) and the fact that Ψ is Lipschitz continuous in a neighborhood of $\bar{x}(T)$ it follows that $(p_i(T))_i\subset\mathbb{R}^n$ and $\big(\int_{[0,T]}|x_i(s)-\bar{x}(s)|\,ds\big)_i\subset\mathbb{R}$ are uniformly

[4] Since problem (P_i) has free terminal point, the cost multiplier can be taken equal to 1.

bounded sequences. Thereby, a standard application of the Gronwall's Lemma to the map $t \mapsto \|p_i\|_{L^\infty([t,T])}$ allows us to deduce that $(p_i)_i \subset W^{1,1}([0,T])$ is a uniformly bounded sequence in L^∞, namely, there exists $P > 0$ such that

$$\|p_i\|_{L^\infty} \le P \qquad \forall i \in \mathbb{N}. \tag{76}$$

Accordingly, $(p_{k_i})_i$ is a sequence with uniformly integrably bounded derivatives for any $k = 0, \dots, N$, such that $p_{k_i}(T) = 0$ for $k = 1, \dots, N$ in view of (69), and the sequence $(p_{0_i}(T))_i \equiv (p_i(T))_i$ is uniformly bounded by the previous arguments. In view of the Ascoli-Arzelá's Theorem there exist functions $\tilde{p}_k \in W^{1,1}([-h_k, T], \mathbb{R}^n)$, $k = 0, \dots, N$, and a subsequence of $(p_{k_i})_i$ such that $p_{k_i} \to \tilde{p}_k$ in L^∞. As a consequence, for p_i as in (74), one has

$$p_i \to \tilde{p} := \sum_{k=0}^{N} \tilde{p}_k \quad \text{in } L^\infty. \tag{77}$$

Furthermore, the upper semicontinuity of the Clarke generalized Jacobian and (65) yield the existence of a sequence $(\tilde{r}_i)_i$ of measurable functions from $[0,T]$ into $\mathbb{R}_{\ge 0}$ such that $\tilde{r}_i(t) \to 0$ for a.e. $t \in [0,T]$, and for which we have

$$D_{x_0,\dots,x_N} f(t, \{x_i(t-h_k)\}, \bar{\alpha}(t)) \subset D_{x_0,\dots,x_N} \bar{f}[t] + \tilde{r}_i(t)\mathbb{B},$$

where $\bar{f}[\cdot]$ is as in (34). Observe that, in view of (H2), the sequence $(\tilde{r}_i)_i$ is uniformly integrably bounded. In particular, we have that $\tilde{r}_i(t) \le 2L_M(t)$ a.e, where M is as in (66). Therefore, from (70) and (76), setting

$$\Omega_i := \{t \in [0,T] : \alpha_i(t) = \bar{\alpha}(t)\},$$

for any $t \in \Omega_i$, we get [5]

$$\begin{aligned}
&\big(-\dot{p}_{0_i}((t-h_0)\vee 0), \dots, -\dot{p}_{N_i}((t-h_N)\vee 0)\big) \\
&\in p_i(t) \cdot D_{x_0,\dots,x_N} \bar{f}[t] + \Big(|p_i(t)|\tilde{r}_i(t) + 2|x_i(t) - \bar{x}(t)|\Big)\mathbb{B} \\
&\subset \operatorname{co} \partial_{x_0,\dots,x_N}\big(\tilde{p}(t)\cdot \bar{f}[t]\big) + \Big(|p_i(t) - \tilde{p}(t)|L_M(t) + P\tilde{r}_i(t) + 2|x_i(t) - \bar{x}(t)|\Big)\mathbb{B}.
\end{aligned}$$

In particular, we have shown that

$$\big(-\dot{p}_{0_i}((t-h_0)\vee 0), \dots, -\dot{p}_{N_i}((t-h_N)\vee 0))\big) \in \operatorname{co} \partial_{x_0,\dots,x_N}\Big(\tilde{p}(t)\cdot \bar{f}[t]\Big) + r_i(t)$$

for a.e. $t \in \Omega_i$, where $r_i : [0,T] \to \mathbb{R}_{\ge 0}$ is given by

$$r_i(t) := P\,\tilde{r}_i(t) + 2|x_i(t) - \bar{x}(t)| + \|p_i - \tilde{p}\|_{L^\infty} L_M(t).$$

By (65), (77), and the boundedness property of $\tilde{r}_i$, it follows that the sequence r_i converges to 0 a.e. on $[0,T]$ and is uniformly integrably bounded, hence the dominated convergence theorem implies $r_i \to 0$ in L^1. Therefore, from the compactness of trajectories theorem (see [20, Thm. 2.5.3]) it follows that there exist functions $p_k \in W^{1,1}([-h_k, T], \mathbb{R}^n)$, $k = 0, \dots, N$, and a subsequence of $(p_{k_i})_i$ such that

$$p_{k_i} \to p_k \quad \text{in } L^\infty, \tag{78}$$

and

$$\big(-\dot{p}_0((t-h_0)\vee 0), \dots, -\dot{p}_N((t-h_N)\vee 0)\big) \in \operatorname{co} \partial_{x_0,\dots,x_N}\Big(\tilde{p}(t)\cdot \bar{f}[t]\Big) \quad \text{a.e. } t \in [0,T].$$

---

[5] We recall that $p \cdot DG(x) = \operatorname{co} \partial(p \cdot G)(x)$ for any $p \in \mathbb{R}^l$ and for all $x \in \mathbb{R}^k$, see [20, Def. 6.2.2].

By the uniqueness of the uniform limit we deduce $p_k \equiv \tilde{p}_k$ for every $k = 0, \dots, N$, so that $p \equiv \tilde{p}$ and the adjoint equation (46) is confirmed. Moreover, in view of (69), also (35) holds for $k = 1, \dots, N$.

Passing to the limit of a proper subsequence in (75), using (65), the upper semicontinuity of the Clarke's generalized Jacobian, and the fact that $(\eta_i)_i$ is a bounded sequence, we deduce that there exists $\eta \in \partial\mathbb{B}_l$ (hence, $\eta \neq 0$) for which the transversality condition (48) at the final point holds. Furthermore, from (71), the properties of the limiting normal cone, and (64) we easily deduce the transversality condition (47) at the initial point.

Now we prove the maximality conditions (49) and (53), where the last one, for the auxiliary optimization problem, reduces to

$$\sup_{a \in A(t),\, w \in \tilde{\mathcal{K}}} \{p(t) \cdot \mathfrak{g}(t, a, w)\} \leq 0 \qquad \forall t \in [0, T]. \tag{79}$$

In the following, we will use the fact that, given a sequence $(\mathcal{N}_i)_i$ of subsets of $[0, T]$ with $\ell(\mathcal{N}_i) \to 0$, then there is a subsequence (we do not relabel) such that [6]

$$\lim_i \ \ell(\{t : t \in [0, T] \setminus \mathcal{N}_j \ \text{ for any } j \geq i\}) = T.$$

By (63) we have that $(m_i)_i$ is uniformly bounded in L^1, hence

$$\ell(\{t \in [0, T] : m_i(t) > 1/\sqrt{\varepsilon_i}\}) \to 0 \qquad \text{as } i \to \infty.$$

Moreover, using (76), (11), and (66), we get that the sequence $(\varphi_i)_i$ given by

$$\varphi_i(t) := \sup_{a \in A(t)} |p_i(t) \cdot f(t, \{x_i(t - h_k)\}, a)|, \quad t \in [0, T],$$

is uniformly pointwise bounded by $Pc(t)\big(1 + (N+1)(M + \|\zeta\|_{L^\infty(-h,0)})\big)$, so that

$$\ell(\{t \in [0, T] : \varphi_i(t) > \sqrt{C_i(t)}\}) \to 0 \qquad \text{as } i \to \infty,$$

since $C_i(t) \to +\infty$ as $i \to \infty$ for any $t \in [0, T]$. Furthermore, since (73) implies that $m_i(t) = C_i(t)$ as soon as $p_i(t) \cdot \mathfrak{g}(t, \alpha_i(t), \omega_i(t)) > \varepsilon_i$, again from (63) it holds

$$\ell(\{t \in [0, T] : p_i(t) \cdot \mathfrak{g}(t, \alpha_i(t), \omega_i(t)) > \varepsilon_i\}) \to 0 \qquad \text{as } i \to \infty.$$

Let $\mathcal{S}_i$ bet is the subset of $[0, T]$ of points that satisfy the following conditions:

$$m_j(t) \leq \frac{1}{\sqrt{\varepsilon_i}} \qquad \forall j \geq i \tag{80}$$

$$\sup_{a \in A(t)} |p_j(t) \cdot f(t, \{x_j(t - h_k)\}, a)| \leq \sqrt{C_j(t)} \qquad \forall j \geq i \tag{81}$$

$$p_j(t) \cdot \mathfrak{g}(t, \alpha_j(t), \omega_j(t)) \leq \varepsilon_j \qquad \forall j \geq i \tag{82}$$

$$\alpha_j(t) = \bar{\alpha}(t) \qquad \forall j \geq i \tag{83}$$

$$\begin{aligned} \max_{\substack{a \in A(t),\, w \in \tilde{\mathcal{K}} \\ m \in [0, C_j(t)]}} &\big\{p_j(t) \cdot \big(f(t, \{x_j(t - h_k)\}, a) + \mathfrak{g}(t, a, w)m\big) - \varepsilon_j(\vartheta_j(t, a, w) \\ &+ |m - m_j(t)|)\big\} = p_j(t) \cdot \big(f_j[t] + \mathfrak{g}(t, \alpha_j(t), \omega_j(t))m_j(t)\big) \quad \forall j \geq i \end{aligned} \tag{84}$$

$$x_j(t - h_k) \to \bar{x}(t - h_k) \qquad \forall k = 0, \dots, N, \ \text{ as } j \to \infty. \tag{85}$$

[6] Indeed, it is enough to consider a subsequence such that $\sum_i \ell(\mathcal{N}_i) < \infty$. Then, $\ell(\{t \in [0, T] \setminus \mathcal{N}_j \ \text{ for } j \geq i\}) \geq \ell(\{t \in [0, T] \setminus \cup_{j \geq i} \mathcal{N}_j\}) \geq T - \sum_{j \geq i} \ell(\mathcal{N}_j) \to T$ as $i \to \infty$.

By the previous arguments, (62), (65), and (73), it turns out that, up to a subsequence, $\ell(\mathcal{S}_i) \to T$. Clearly, $\mathcal{S} := \cup_i \mathcal{S}_i$ is a dense subset of $[0,T]$.

Given $t \in \mathcal{S}$, let $i \in \mathbb{N}$ be such that $t \in \mathcal{S}_i$. Fix arbitrary $\bar{w} \in \tilde{\mathcal{K}}$ and $\bar{a} \in A(t)$. Using (80), (82), (83) and taking $a = \bar{a}$ and $m = 0$, from (84) we obtain

$$p_j(t) \cdot f(t, \{x_j(t-h_k)\}, \bar{\alpha}(t)) + \sqrt{\varepsilon_j} \geq p_j(t) \cdot f(t, \{x_j(t-h_k)\}, \bar{a}) - \varepsilon_j \Big(1 + \frac{1}{\sqrt{\varepsilon_j}}\Big)$$

for any $j \geq i$. By (78) and (85) we can let $j \to \infty$ in the above relation. Therefore, we deduce that the drift-maximality condition (49) holds in the full measure subset $\mathcal{S}$. Now, for any $j \geq i$, choose $\bar{w}_j \in \tilde{\mathcal{K}}$, $\bar{a}_j \in A(t)$ satisfying

$$p_j(t) \cdot \mathfrak{g}(t, \bar{a}_j, \bar{w}_j) C_j(t) \geq \sup_{a \in A(t),\, w \in \tilde{\mathcal{K}}} \{p_j(t) \cdot \mathfrak{g}(t, a, w) C_j(t)\} - \varepsilon_j. \tag{86}$$

By (80)–(82), (86), and taking $a = \bar{a}_j$ and $m = C_j(t)$, from (84) we obtain

$$\sqrt{C_j(t)} + \sqrt{\varepsilon_j} \geq -\sqrt{C_j(t)} + \sup_{a \in A(t),\ w \in \tilde{\mathcal{K}}} \{p_j(t) \mathfrak{g}(t, a, w)\} C_j(t) - \varepsilon_j - \varepsilon_j - \varepsilon_j C_j(t)$$

for any $j \geq i$. Since $C_j(t) \to \infty$ as $j \to \infty$, if we divide the above inequality for $C_j(t)$ and we let $j \to \infty$, by the continuity of $p \mapsto \sup_{a \in A(t),\ w \in \tilde{\mathcal{K}}} \{p \cdot \mathfrak{g}(t, a, w)\}$ and (78) we get that

$$\sup_{a \in A(t),\ w \in \tilde{\mathcal{K}}} \{p(t) \cdot \mathfrak{g}(t, a, w)\} \leq 0 \qquad \forall t \in \mathcal{S}.$$

To obtain the maximality condition (79) on the whole interval $[0,T]$, it suffices to notice that $\mathcal{S}$ is a dense set in $[0,T]$, the function $t \mapsto \sup_{a \in A(t),\ w \in \tilde{\mathcal{K}}} \{p(t) \cdot \mathfrak{g}(t, a, w)\}$ is continuous thanks to (H4), and that $p(\cdot)$ is an absolutely continuous function.

In order to conclude the proof, it remains to prove (54), that for the auxiliary control system reduces to

$$p(t) \cdot \mathfrak{g}(t, \bar{\alpha}(t), \bar{\omega}(t)) = 0 \qquad \bar{\nu}\text{-a.e. } t \in [0,T]. \tag{87}$$

Taking $a = \alpha_i(t)$, $w = \omega_i(t)$, and $m = 0$, from (73) we obtain

$$p_i(t) \cdot \mathfrak{g}(t, \alpha_i(t), \omega_i(t)) m_i(t) \geq -\varepsilon_i m_i(t) \qquad \text{a.e. } t \in [0,T],$$

from which it follows that

$$p_i(t) \cdot \mathfrak{g}(t, \alpha_i(t), \omega_i(t)) \geq -\varepsilon_i \qquad \text{for a.e. } t \in \{\tau \in [0,T] : m_i(\tau) > 0\}. \tag{88}$$

Moreover, by (68) and (78) we deduce that

$$p_i(t) \cdot \mathfrak{g}(t, \alpha_i(t), \omega_i(t)) m_i(t) dt \rightharpoonup^* p(t) \cdot \mathfrak{g}(t, \bar{\alpha}(t), \bar{\omega}(t)) \bar{\nu}(dt).$$

Therefore, Lemma 2.9, (iii), implies that, restricting attention to a suitable subsequence, there exists a subset $\hat{\mathcal{E}} \subset [0,T]$ of full measure, containing T, such that

$$\lim_i \int_{[0,t]} p_i(s) \cdot \mathfrak{g}(s, \alpha_i(s), \omega_i(s)) m_i(s) ds = \int_{[0,t]} p(s) \cdot \mathfrak{g}(s, \bar{\alpha}(s), \bar{w}(s)) \bar{\nu}(ds) \quad \forall t \in \hat{\mathcal{E}}. \tag{89}$$

By (88) and the fact that m_i takes values in $\mathbb{R}_{\geq 0}$, for any $t \in \hat{\mathcal{E}}$ we have

$$\begin{aligned}
&\int_{[0,t]} p_i(s) \cdot \mathfrak{g}(s, \alpha_i(s), \omega_i(s)) m_i(s) ds \\
&= \int_{[0,t] \cap \{\tau \,:\, m_i(\tau) > 0\}} p_i(s) \cdot \mathfrak{g}(s, \alpha_i(s), \omega_i(s)) m_i(s) ds \\
&\geq -\varepsilon_i \int_{[0,t]} m_i(s) ds.
\end{aligned}$$

Recalling that $(m_i)_i$ is uniformly bounded in L^1 by (63), from the above relation and (89) we deduce $\int_{[0,t]} p(s) \cdot \mathfrak{g}(s, \bar{\alpha}(s), \bar{\omega}(s)) \bar{\nu}(ds) \geq 0$ for any $t \in \hat{\mathcal{E}}$. But this implies

$$\int_B p(t) \cdot \mathfrak{g}(t, \bar{\alpha}(t), \bar{\omega}(t)) \bar{\nu}(dt) \geq 0 \qquad \forall B \in \mathcal{B},$$

as the family of subsets $([0,t])_{t \in \hat{\mathcal{E}}}$ generates $\mathcal{B}$. From this last relation it follows that

$$p(t) \cdot \mathfrak{g}(t, \bar{\alpha}(t), \bar{\omega}(t)) \geq 0 \qquad \bar{\nu}\text{-a.e. } t \in [0,T].$$

The previous relation, condition (79) and the fact that $\bar{\alpha}(t) \in A(t)$ $\bar{\nu}$-a.e. $t \in [0,T]$ and $\bar{\omega}(t) \in \tilde{\mathcal{K}}$ $\bar{\nu}$-a.e. $t \in [0,T]$, imply (87). □

Appendix A.

Proof of Lemma 2.8. For any integer $i \geq 0$ consider the measure $\tilde{\gamma}_i$ on the Borel subsets of $[-h,T]$ defined by $\tilde{\gamma}_i(B) = 0$ for any Borel subset $B \subseteq [-h,0]$ and $\tilde{\gamma}_i \equiv \gamma_i$ on the Borel subsets of $]0,T]$. Hence, consider the function $\tilde{z}_0 : [-h,T] \to \mathbb{R}^n$, absolutely continuous on $[0,T]$, defined by $\tilde{z}_0(t) := z_0(t) - \int_{[0,t]} \tilde{\gamma}_0(ds)$ for any $t \in [-h,T]$, that turns out to satisfy the following delayed differential equation

$$\begin{cases} \dot{\tilde{z}}_0(t) = \varphi_0\left(t, \left\{\tilde{z}_0(t-h_k) + \int_{[0,t-h_k]} \tilde{\gamma}_0(ds)\right\}\right), & t \in]0,T] \\ \tilde{z}_0(t) = \eta(t) \quad \text{a.e. } t \in [-h,0[, \qquad \tilde{z}_0(0) = \xi_0. \end{cases}$$

Hence, let $M > 0$ be such that [7]

$$\|\tilde{z}_0(t)\|_{L^\infty([-h,T])} + \|\gamma_i\|_{C^*_{\mathbb{R}^k}} \leq \frac{M}{2(N+1)}$$

for any integer $i \geq 0$. Let $\tilde{L}_M \in L^1([0,T], \mathbb{R}_{\geq 0})$ be as in (28) and notice that (28) and (30) are valid for $i = 0$ as well, due to (29). Now, for any integer $i \geq 1$, consider the function $\beta_i : [0,T] \to \mathbb{R}_{\geq 0}$ defined as

$$\beta_i(t) := \left| \dot{\tilde{z}}_0(t) - \varphi_i\left(t, \left\{\tilde{z}_0(t-h_k) + \int_{[0,t-h_k]} \tilde{\gamma}_i(ds)\right\}\right)\right|$$

and the set $\mathcal{E}_i$ given by

$$\mathcal{E}_i := \{t \in [0,T] \,:\, \varphi_i(t, \{x_k\}) = \varphi_0(t, \{z_k\}) \ \forall \{z_k\} \in (\mathbb{R}^n)^{N+1}\}.$$

[7]Here we use the fact that $\eta(\cdot)$ is bounded.

Thus, for a.e. $t \in \mathcal{E}_i$, we get

$$\begin{aligned}\beta_i(t) = \Big|&\varphi_0\big(t, \big\{\tilde z_0(t-h_k) + \textstyle\int_{[0,t-h_k]} \tilde\gamma_0(ds)\big\}\big)\\ &-\varphi_0\big(t, \big\{\tilde z_0(t-h_k) + \textstyle\int_{[0,t-h_k]} \tilde\gamma_i(ds)\big\}\big)\Big|\\ \le{}& \tilde L_M(t)\big|\big\{ \textstyle\int_{[0,t-h_k]} \big[\gamma_i(ds) - \gamma_0(ds)\big]\big\}\big|.\end{aligned}$$

Instead, for a.e. $t \in [0,T] \setminus \mathcal{E}_i$ we have

$$\begin{aligned}\beta_i(t) \le{}& |\dot{\tilde z}_0(t)| + \big|\varphi_i\big(t, \big\{\tilde z_0(t-h_k) + \textstyle\int_{[0,t-h_k]} \tilde\gamma_0(ds)\big\}\big)\big|\\ &+\big|\varphi_i\big(t, \big\{\tilde z_0(t-h_k) + \textstyle\int_{[0,t-h_k]} \tilde\gamma_i(ds)\big\}\big)\\ &-\varphi_i\big(t, \big\{\tilde z_0(t-h_k) + \textstyle\int_{[0,t-h_k]} \tilde\gamma_0(ds)\big\}\big)\big|\\ \le{}& 2\tilde c(t) + \tilde L_M(t)\big|\big\{ \textstyle\int_{[0,t-h_k]} \big[\gamma_i(ds) - \gamma_0(ds)\big]\big\}\big|.\end{aligned}$$

Since $\left|\int_{[0,t]} [\gamma_i(ds) - \gamma_0(ds)]\right|$ is uniformly bounded in $[0,T]$ and converges to 0 for all t in the subset $\mathcal{E} \subseteq [0,T]$ of full measure and containing $t = T$, by applying the dominated convergence theorem and by (29) we obtain

$$\lim_i \int_0^T \beta_i(t)\,dt = 0. \tag{90}$$

Thanks to (90) we can apply the Filippov Theorem for delayed systems [3, Thm. 4.1] in order to deduce that for any integer i sufficiently large there exists a function $\tilde z_i \in W^{1,1}([0,T],\mathbb{R}^n)$ that solves [8]

$$\begin{cases}\dot{\tilde z}_i(t) = \varphi_i\left(t, \left\{\tilde z_i(t-h_k) + \int_{[0,t-h_k]} \tilde\gamma_i(ds)\right\}\right), & t \in]0,T]\\ \tilde z_i(t) = \eta(t) \quad \text{a.e. } t \in [-h,0[, & \tilde z_i(0) = \xi_i.\end{cases} \tag{91}$$

and that satisfies

$$\begin{aligned}\|\tilde z_i - \tilde z_0\|_{L^\infty([0,T])} &\le |\xi_i - \xi_0| + \int_0^T |\dot{\tilde z}_i(t) - \dot{\tilde z}_0(t)|\,dt\\ &\le e^{(N+1)\int_0^T \tilde L_M(t)dt}\left(|\xi_i - \xi_0| + \int_0^T \beta_i(t)dt\right).\end{aligned} \tag{92}$$

For any integer $i \ge 1$ we define the function $z_i : [-h,T] \to \mathbb{R}^n$ to be $z_i(t) = \tilde z_i(t) + \int_{[0,t]} \gamma_i(ds)$ for any $t \in]0,T]$, $z_i(0) = \xi_i$, and $z_i(t) = \eta(t)$ for $t \in [-h,0[$. In view of (91) we deduce that z_i belongs to $BV([-h,T],\mathbb{R}^n)$ and is a solution to (31). Furthermore, by (92), the fact that $\xi_i \to \xi_0$, and the hypothesis on the convergence of the γ_i, we immediately deduce that $dz_i \rightharpoonup^* dz_0$ and $z_i(t) - \int_{[0,t]} \gamma_i(ds) \to z_0(t) - \int_{[0,t]} \gamma_0(ds)$ uniformly in $[0,T]$, so that $z_i(t) \to z(t)$ for all $t \in \mathcal{E}$. $\square$

Proof of Lemma 2.9, (ii), (iii). Let $(\gamma_i, |\gamma_i|) \rightharpoonup^* (\gamma, \lambda)$ for some $\lambda \in C^\oplus$ and let $\Psi \in C([0,T],\mathbb{R}^{n\times k})$ be as in (ii). Let $\tilde{\mathcal{E}} \subseteq [0,T]$ be the subset of points t such that

[8] The function $(t,\{z_k\}) \mapsto \tilde\varphi_i(t,\{z_k\}) := \varphi_i\big(t, \{z_k + \int_{[0,t-h_k]} \tilde\gamma_i(ds)\}\big)$ is measurable in the t variable. Moreover, $\{z_k\} \in \{\tilde z_0(t-h_k)\} + \frac{M}{2}\mathbb{B}_{n(N+1)}$ implies that $\{z_k + \int_{[0,t-h_k]} \tilde\gamma_i(ds)\} \in M\mathbb{B}_{n(N+1)}$, so that $\tilde\varphi_i$ is $\tilde L_M(\cdot)$-Lipschitz continuous in the $\{z_k\}$ variables in the $\frac{M}{2}$-tube around $\{\tilde z_0(t-h_k)\}$.

$\lambda(\{t\}) = 0$. Clearly, $[0,T] \setminus \tilde{\mathcal{E}}$ is a countable set. For every $t \in \mathcal{E} := \tilde{\mathcal{E}} \cup \{T\}$ and any integer $i \geq 1$, consider

$$\int_{[0,t]} \Psi(s) \cdot \gamma_i(ds) = \left(\int_{[0,t]} \Psi_1(s) \cdot \gamma_i(ds), \dots, \int_{[0,t]} \Psi_n(s) \cdot \gamma_i(ds) \right),$$

where, for every $l = 1, \dots, n$, $\Psi_l(s)$ is the row vector $(\Psi_{l1}(s), \dots, \Psi_{lk}(s))$. Since this integral converges if and only if each component converges, it is sufficient to prove that

$$\int_{[0,t]} \Psi_l(s) \cdot \gamma_i(ds) \to \int_{[0,t]} \Psi_l(s) \cdot \gamma(ds) \quad \text{as } i \to \infty$$

for each $l = 1, \dots, n$. This is equivalent to show that [9]

$$\sum_{j=1}^{k} \int_{[0,T]} \Psi_{lj}(s) \chi_{[0,t]}(s) \gamma_i^j(ds) \to \sum_{j=1}^{k} \int_{[0,T]} \Psi_{lj}(s) \chi_{[0,t]}(s) \gamma^j(ds) \quad \text{as } i \to \infty,$$

which is certainly true if, for every $j = 1, \dots, k$, we have

$$\int_{[0,T]} \Psi_{lj}(s) \chi_{[0,t]}(s) \gamma_i^j(ds) \to \int_{[0,T]} \Psi_{lj}(s) \chi_{[0,t]}(s) \gamma^j(ds) \quad \text{as } i \to \infty. \tag{93}$$

Note that, for $t = T$, the thesis follows from the definition of weak convergence, in view of the continuity of Ψ. Thus, let $t \in \tilde{\mathcal{E}}$. At this point, to complete the proof of (ii) we can appeal to [1, Prop. 1.62], observing that, for each pair l, j, the scalar function $s \mapsto \Psi_{lj}(s) \chi_{[0,t]}(s)$ is a bounded, Borel measurable function such that its unique discontinuity point, t, has $\lambda(\{t\}) = 0$.

In order to prove statement (iii), notice that, since $\gamma_i \rightharpoonup^* \gamma$, if for each $j = 1, \dots, k$ we consider the Jordan decomposition $\gamma_i^j = (\gamma_i^j)^+ - (\gamma_i^j)^-$, by uniform boundedness, possibly extracting a subsequence, we obtain that $(\gamma_{i_l}^j)^+ \rightharpoonup^* \lambda^{j+}$ and $(\gamma_{i_l}^j)^- \rightharpoonup^* \lambda^{j-}$, for some positive, finite measures λ^{j+}, λ^{j-}. Then,

$$|\gamma_{i_l}^j| = (\gamma_{i_l}^j)^+ + (\gamma_{i_l}^j)^- \rightharpoonup^* \lambda^j := \lambda^{j+} + \lambda^{j-},$$

which implies that $|\gamma_{i_l}| = \sum_{j=1}^k |\gamma_{i_l}^j| \rightharpoonup^* \lambda := \sum_{j=1}^k \lambda^j \in C^\oplus$, so that, along this subsequence, we have $(\gamma_{i_l}, |\gamma_{i_l}|) \rightharpoonup^* (\gamma, \lambda)$. □

Acknowledgments. We would like to thank Prof. Richard Vinter for suggesting the problem and for helpful discussions.

REFERENCES

[1] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, the Clarendon Press, Oxford University Press, New York, 2000.

[2] M. S. Aronna, M. Motta and F. Rampazzo, Infimum gaps for limit solutions, *Set-Valued Var. Anal.*, **23** (2015), 3-22.

[3] A. Boccia and R. B. Vinter, The maximum principle for optimal control problems with time delays, *SIAM Journal of Control*, **55** (2017), 2905-2935.

[4] A. Bressan and B. Piccoli, *Introduction to The Mathematical Theory of Control*, AIMS Series on Applied Mathematics, 2, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

---

[9] Actually, according to our definition (5), for each i, $\int_{[0,t]} \Psi_l(s) \cdot \gamma_i(ds) = \int_{[0,t]} \Psi_l(s) \cdot \omega_i(s) |\gamma_i|(ds) = \int_{[0,t]} \sum_{j=1}^k \Psi_{lj}(s) \omega_i^j(s) |\gamma_i|(ds)$, where ω_i is the Radon-Nicodym derivative of the measure γ_i w.r.t. $|\gamma_i|$. Since, however, $[0,T] \ni t \mapsto \Psi_l(t)$ is continuous (and bounded), all components $\Psi_{lj}(s)$ are γ_i^j integrable and $\int_{[0,t]} \sum_{j=1}^k \Psi_{lj}(s) \omega_i^j(s) |\gamma_i|(ds) = \sum_{j=1}^k \int_{[0,t]} \Psi_{lj}(s) \gamma_i^j(ds)$.

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

[6] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.

[7] F. H. Clarke, Y. S. Ledyaev, R. J. Stern and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Graduate Texts in Mathematics, 178, Springer-Verlag, New York, 1998.

[8] G. Fusco and M. Motta, No infimum gap and normality in optimal impulsive control under state constraints, *Set-Valued Var. Anal.*, **29** (2021), 519-550.

[9] 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.

[10] G. Fusco and M. Motta, Gap phenomena and controllability in free end-time problems with active state constraints, *J. Math. Anal. Appl.*, **510** (2022), Paper No. 126021, 25 pp.

[11] G. Fusco and M. Motta, Strict sense minimizers which are relaxed extended minimizers in general optimal control problems, *Proceedings of the 60th IEEE Conference on Decision and Control, Austin, Texas*, (2021), 6000-6005.

[12] C. X. Gao, K. Z. Li, E. M. Feng and Z. L. Xiu, Nonlinear impulsive system of fed-batch culture in fermentative production and its properties, *Chaos. Soliton. Fract.*, **28** (2006), 271-277.

[13] X. D. Li, J. D. Cao and D. W. C. Ho, Impulsive control of nonlinear systems with time-varying delay and applications, *IEEE Transactions on Cybernetics*, **50** (2020), 2661-2673.

[14] B. M. Miller, The method of discontinuous time substitution in problems of the optimal control of impulse and discrete-continuous systems, *Automat. Remote Control*, **54** (1993), 1727-1750.

[15] M. Motta, M. Palladino and F. Rampazzo, Unbounded control, infimum Gaps, and higher order normality, *SIAM J. Control Optim.*, **60** (2022), 1436-1462.

[16] M. Motta and F. Rampazzo, Space-time trajectories of non linear systems driven by ordinary and impulsive controls, *Differ. Int. Eq.*, **8** (1995), 269-288.

[17] M. Motta, F. Rampazzo and R. Vinter, Normality and gap phenomena in optimal unbounded control, *ESAIM Control, Optimisation and Calculus of Variations*, **24** (2018), 1645-1673.

[18] M. Palladino and F. Rampazzo, A geometrically based criterion to avoid infimum gaps in optimal control, *J. Differ. Equ.*, **269** (2020), 10107-10142.

[19] R. W. Rishel, An extended Pontryagin principle for control systems whose control laws contain measures, *SIAM Journal of Control*, **3** (1965), 191-205.

[20] R. Vinter, *Optimal Control*, Birkhäuser, Boston, 2000.

[21] R. B. Vinter and F. M. F. L. Pereira, A Maximum principle for optimal processes with discontinuous trajectories, *SIAM Journal of Control*, 26 (1988), 205-229.

[22] J. Warga, Variational problems with unbounded controls, *SIAM Journal of Control*, **3** (1965), 424-438, .

[23] Z. L. Xiu, B. H. Song, L. H. Sun and A. P. Zeng, Theoretical analysis of effects of metabolic overflow and time delay on the performance and dynamic behavior of a two-stage fermentation process, *Biochem. Eng. J.*, **11** (2002), 101-109.

Received for publication September 2023; early access April 2024.