



# Continuity/constancy of the Hamiltonian function in a Pontryagin maximum principle for optimal sampled-data control problems with free sampling times

Loïc Bourdin, Gaurav Dhar

## ► To cite this version:

Loïc Bourdin, Gaurav Dhar. Continuity/constancy of the Hamiltonian function in a Pontryagin maximum principle for optimal sampled-data control problems with free sampling times. *Mathematics of Control, Signals, and Systems*, 2019. hal-02160221

**HAL Id: hal-02160221**

**<https://hal.science/hal-02160221>**

Submitted on 19 Jun 2019

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

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

# Continuity/constancy of the Hamiltonian function in a Pontryagin maximum principle for optimal sampled-data control problems with free sampling times

Loïc Bourdin\*, Gaurav Dhar†

## Abstract

In a recent paper by Bourdin and Trélat, a version of the Pontryagin maximum principle (in short, PMP) has been stated for general nonlinear finite-dimensional optimal sampled-data control problems. Unfortunately their result is only concerned with fixed sampling times, and thus it does not take into account the possibility of free sampling times. The present paper aims to fill this gap in the literature. Precisely we establish a new version of the PMP that can handle free sampling times. As in the aforementioned work by Bourdin and Trélat, we obtain a first-order necessary optimality condition written as a nonpositive averaged Hamiltonian gradient condition. Furthermore, from the freedom of choosing sampling times, we get a new and additional necessary optimality condition which happens to coincide with the continuity of the Hamiltonian function. In an autonomous context, even the constancy of the Hamiltonian function can be derived. Our proof is based on the Ekeland variational principle. Finally a linear-quadratic example is numerically solved using shooting methods, illustrating the possible discontinuity of the Hamiltonian function in the case of fixed sampling times and highlighting its continuity in the instance of optimal sampling times.

**Keywords:** Sampled-data control, optimal control, optimal sampling times, Pontryagin maximum principle, Hamiltonian continuity, Hamiltonian constancy, Ekeland variational principle.

**AMS Classification:** 34K35, 34H05, 49J15, 49K15, 93C15, 93C57, 93C62, 93C83.

## 1 Introduction

In mathematics a *dynamical system* describes the evolution of a point (usually called the *state* of the system) in an appropriate set following an evolution rule (known as the *dynamics* of the system). Dynamical systems are of many different natures and they can be categorized in different classes such as: continuous systems versus discrete systems (see details hereafter), deterministic systems versus stochastic systems, etc. A *continuous system* is a dynamical system in which the state evolves in a continuous way in time (for instance, ordinary differential equations, evolution partial differential equations, etc.), while a *discrete system* is a dynamical system in which the state evolves in a discrete way in time (for instance, difference equations, quantum differential equations, etc.). A *control system* is a dynamical system in which a control parameter influences the evolution of the state. Finally an *optimal control problem* consists of determining a control which allows to steer the state of a control system from a specified configuration to some desired target while minimizing a given criterion. Established in [29] by Pontryagin et al. at the end of the 1950's, the *Pontryagin maximum principle* (in short, PMP) is the milestone of optimal control theory. It provides first-order necessary optimality conditions for continuous optimal control problems in

---

\*Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France. [loic.bourdin@unilim.fr](mailto:loic.bourdin@unilim.fr)

†Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France. [gaurav.dhar@unilim.fr](mailto:gaurav.dhar@unilim.fr)

which the dynamics is described by a general nonlinear ordinary differential equation. Roughly speaking, the classical PMP ensures the existence of an adjoint vector such that the optimal control satisfies the so-called *Hamiltonian maximization condition*. Soon afterwards and even nowadays, the PMP has been adapted to many situations, for control systems of different natures, with various constraints, etc. It is not the aim of the present paper to give a state of the art. Nevertheless we precise that several versions of the PMP were derived for discrete optimal control problems in which the dynamics is described by a difference equation (see, e.g., [9, 23, 24]). In these discrete versions of the PMP, the historical Hamiltonian maximization condition does not hold in general (see a counterexample in [9, Examples 10.1-10.4 p.59-62]) and has to be replaced by a weaker condition known as a *nonpositive Hamiltonian gradient condition* (see, e.g., [9, Theorem 42.1 p.330]). Note that some appropriate convexity conditions on the dynamics have been considered in order to recover the Hamiltonian maximization condition in the discrete case (see, e.g., [24]).

In this paper we are interested in *sampled-data control systems* in which the state evolves continuously in time while the control evolves discretely in time. More precisely the value of the control is authorized to be modified only a finite number of times. The times in which the control can be modified are usually called the *sampling times*. Note that sampled-data control systems have the peculiarity of presenting a mixed continuous/discrete structure. They have been considered as models mostly in Engineering implemented by *digital controllers* which have a finite precision (see, e.g., [30, 36]). Numerous texts and articles have developed control theory for sampled-data control systems (see, e.g., [1, 2, 3, 18, 25] and references therein). For instance, global controllability for sampled-data control systems has been investigated in [20]. Sampled-data control systems are used in Automation, notably in model predictive control algorithms in which the control value at each sampling time is chosen as the first value of a finite sequence of control values optimizing the given cost on a fixed finite horizon (see, e.g., [21]). Optimal sampled-data control problems have been investigated in the literature with different approaches. One approach has been to apply  $\mathcal{H}_2$ - $\mathcal{H}_\infty$  optimization theory (see [7, 15]) where the closed-loop transfer matrix under the  $\mathcal{H}_2$ - and  $\mathcal{H}_\infty$ -norms is taken as the criterion. Another approach involves the Karush-Kuhn-Tucker necessary conditions and dynamic programming (see [6]). However one should note that the aforementioned results are not formulated in terms of a PMP. Recently Bourdin and Trélat have obtained in [13] a version of the PMP for general nonlinear optimal sampled-data control problems. In that sampled-data control framework, as in the purely discrete case addressed in the previous paragraph, the usual Hamiltonian maximization condition does not hold in general and has to be replaced by a weaker condition known as a *nonpositive averaged Hamiltonian gradient condition* (see [13, Theorem 2.6 p.62]). Note that the PMP enunciated in [13, Theorem 2.6 p.62] is actually stated in the more general framework of *time scale calculus* and a version which does not take into account such a generality, and therefore closer to the considerations of the present paper, can be found in [12, Theorem 1 p.81] or [13, Theorem 1.1 p.55]. Unfortunately this PMP is only concerned with fixed sampling times, and thus it does not take into account the possibility of free sampling times that can be chosen from a given interval. The main objective of the present paper is to fill this gap in the literature by deriving a PMP for general nonlinear optimal sampled-data control problems with free sampling times. We mention that *optimal sampling times problems* have already been investigated in the literature but, to the best of our knowledge, never from a PMP point of view. For example many authors consider the related problem of finding the optimal fixed sampling interval (or time step) such as in [26, 28]. Nonuniform sampling partitions have also been studied but in specific cases such as for the linear-quadratic integrator in [31]. In [32] the optimal sampled-data control problem is transformed into a purely discrete one by integrating the state over the sampling intervals and then is treated as an usual optimization problem.

The main theoretical result of the present paper (Theorem 2.1 in Section 2.3) is a PMP for nonlinear optimal sampled-data control problems with free sampling times. Similarly to the PMP derived

in [12, Theorem 1 p.81] or [13, Theorem 1.1 p.55] for fixed sampling times, we obtain a first-order necessary optimality condition described by a nonpositive averaged Hamiltonian gradient condition (see Inequality (2)). Furthermore, from the freedom of choosing sampling times, we get a new and additional necessary optimality condition (see Equality (3)) which happens to coincide with the continuity of the Hamiltonian function. In an autonomous context, even the constancy of the Hamiltonian function can be derived. We refer to Section 2.4 for a detailed discussion on the continuity/constancy of the Hamiltonian function. Moreover, in case of additional constraints on the size of sampling intervals (in practice one can expect a minimum size for instance), the continuity of the Hamiltonian function is replaced by a weaker inequality (see Remarks 2.17 and 2.18 for details).

We must remark that in the classical case of purely continuous optimal control problems, the (absolute) continuity of the Hamiltonian function is a very well-known fact (see, e.g., [19, Theorem 2.6.3 p.73]). With the help of a simple linear-quadratic example, we show in Section 3 that this classical property does not hold in general for optimal sampled-data control problems with fixed sampling times (see Figure 1). On the other hand, the present work proves that this continuity property is recovered when considering optimal sampling times, which is illustrated with the same aforementioned linear-quadratic example (see Figure 2). Furthermore the linear-quadratic example developed in Section 3 allows us to prove the interest of our main result since it is numerically solved by using, in one hand, the Riccati theory developed in [14, Theorem 2 and Corollary 1 p.276] and, in the other hand, a shooting method based on the Hamiltonian continuity condition derived in Theorem 2.1. We conclude this paragraph by mentioning that, in the context of hybrid optimal control problems, a similar Hamiltonian continuity condition at crossing times (resp. at switching times) can be found in [22, Remark 1.3] by Haberkorn and Trélat (resp. in [34, Definition 13] by Sussmann under the name of *Hamiltonian value condition*). Nevertheless, due to the nature of the sampling times and of the sampled-data controls considered in the present paper, our main result (in particular the nonpositive averaged Hamiltonian gradient condition) cannot, to the best of our knowledge, be seen as a direct consequence of the works [22, 34].

In this paragraph our aim is to give some details about the strategy adopted in this paper and the major difficulties encountered. The proof of our main result is detailed in Appendix A and, similarly to [13, Theorem 2.6 p.62], it is based on the classical Ekeland variational principle [17, Theorem 1.1 p.324]. This leads us to consider a sequence of sampled-data controls converging in  $L^1$ -norm to the optimal one. A first difficulty emerges in the fact that the associated sampling times do not necessarily converge to the optimal sampling times. Indeed a degenerate situation can occur if the optimal control is constant over two consecutive sampling intervals. Moreover another obstacle is the possible phenomenon of accumulation of sampling times. These two difficulties are overcome by introducing a technical control set (see Equation (5) in Appendix A.1) which guarantees that the sampling times produced by the Ekeland variational principle, firstly, remain unchanged for the ones corresponding to the consecutive sampling intervals on which the optimal control is constant (avoiding thus the first difficulty) and, secondly, are contained in disjoint intervals for the others (avoiding thus the second difficulty). We refer to Proposition A.2 in Appendix A.1 for details. A final obstacle lies in the non-convexity of the set of  $N$ -piecewise constant functions (where  $N \in \mathbb{N}^*$  is fixed). Therefore the standard procedure of considering convex  $L^\infty$ -perturbations of the control (as in [13, Lemma 4.17 p.84]) has to be adapted by considering convex  $L^\infty$ -perturbations respecting the same  $N$ -partition. We refer to the proof of Lemma A.10 for details. We conclude this paragraph by mentioning that, thanks to the reviewing process of the present paper, an alternative proof of our main result has been brought to our attention. This different approach is based on a remarkable technique exposed in the paper [16] by Dmitruk and Kaganovich. We refer to Remark 2.12 for the details.

The paper is organized as follows. Section 2 is dedicated to our main result (Theorem 2.1). Precisely we introduce in Section 2.1 the notations and the functional framework required for our work. The optimal sampled-data control problem considered in this paper (with possibly free sampling times) is presented in details in Section 2.2 (see Problem (OSCP)). The corresponding Pontryagin maximum principle (Theorem 2.1) is stated in Section 2.3 and a list of general comments is in order. Finally we devote Section 2.4 to a discussion on the continuity/constancy of the Hamiltonian function. In Section 3 we numerically solve a simple linear-quadratic optimal sampled-data control problem and we compare the two following situations: fixed sampling times versus free sampling times. As expected from our main result, the Hamiltonian function admits discontinuities in the first case (see Figure 1), while it does not in the second case (see Figure 2). Appendix A is devoted to the detailed proof of Theorem 2.1.

## 2 Main result and comments

Section 2.1 is devoted to the notations and the functional framework required for our work. In Section 2.2, the general nonlinear optimal sampled-data control problem considered in this paper (with possibly free sampling times) is presented with terminology and assumptions. In Section 2.3, the corresponding Pontryagin maximum principle, which constitutes our main result, is stated. A list of general comments is in order. Finally Section 2.4 is dedicated to a discussion about the continuity/constancy of the Hamiltonian function.

### 2.1 Notations and functional framework

Let  $n \in \mathbb{N}^*$  and  $I \subset \mathbb{R}$  be a subinterval of  $\mathbb{R}$ . In this paper we denote by:

- $L^1(I, \mathbb{R}^n)$  the Lebesgue space of integrable functions defined on  $I$  with values in  $\mathbb{R}^n$ , endowed with its usual norm  $\|\cdot\|_{L^1(I, \mathbb{R}^n)}$ ;
- $L^\infty(I, \mathbb{R}^n)$  the Lebesgue space of essentially bounded functions defined on  $I$  with values in  $\mathbb{R}^n$ , endowed with its usual norm  $\|\cdot\|_{L^\infty(I, \mathbb{R}^n)}$ ;
- $C(I, \mathbb{R}^n)$  the space of continuous functions defined on  $I$  with values in  $\mathbb{R}^n$ , endowed with the standard uniform norm  $\|\cdot\|_{C(I, \mathbb{R}^n)}$ ;
- $AC(I, \mathbb{R}^n)$  the subspace of  $C(I, \mathbb{R}^n)$  of absolutely continuous functions.

Let  $m \in \mathbb{N}^*$  and  $\tau > 0$  be fixed in this whole section. For all  $N \in \mathbb{N}^*$ , the set of all  $N$ -partitions of the interval  $[0, \tau]$  is defined by

$$\mathcal{P}_N^\tau := \{\mathbb{T} = (t_i)_{i=0, \dots, N} \mid 0 = t_0 < t_1 < \dots < t_{N-1} < t_N = \tau\}.$$

Then, for all  $N \in \mathbb{N}^*$  and all  $\mathbb{T} = (t_i)_{i=0, \dots, N} \in \mathcal{P}_N^\tau$ , the set of all *piecewise constant functions over  $[0, \tau]$  respecting the  $N$ -partition  $\mathbb{T}$*  is defined by

$$PC_N^\mathbb{T}([0, \tau], \mathbb{R}^m) := \{u \in L^\infty([0, \tau], \mathbb{R}^m) \mid \forall i = 0, \dots, N-1, \exists u_i \in \mathbb{R}^m, u(t) = u_i \text{ a.e. } t \in [t_i, t_{i+1}]\}.$$

In this paper, as usual in the Lebesgue space  $L^\infty([0, \tau], \mathbb{R}^m)$ , two functions in  $PC_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$  which are equal almost everywhere on  $[0, \tau]$  will be identified. Precisely, if  $u \in PC_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$ , then  $u$  is identified to the function

$$u(t) = \begin{cases} u_i & \text{if } t \in [t_i, t_{i+1}), \quad i \in \{0, \dots, N-2\}, \\ u_{N-1} & \text{if } t \in [t_{N-1}, t_N], \end{cases}$$

for all  $t \in [0, \tau]$ . Note that  $PC_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$  is a linear subspace of  $L^\infty([0, \tau], \mathbb{R}^m)$ .

**Remark 2.1.** Note that the inclusion  $\text{PC}_M^{\mathbb{T}'}([0, \tau], \mathbb{R}^m) \subset \text{PC}_N^{\mathbb{T}}([0, \tau], \mathbb{R}^m)$  holds true for all  $M, N \in \mathbb{N}^*$  such that  $M \leq N$ , and all  $\mathbb{T}' \in \mathcal{P}_M^\tau, \mathbb{T} \in \mathcal{P}_N^\tau$  such that  $\mathbb{T}' \subset \mathbb{T}$ .

Finally, for all  $N \in \mathbb{N}^*$ , the set of all *piecewise constant functions over  $[0, \tau]$  respecting at least one  $N$ -partition* is defined by

$$\text{PC}_N([0, \tau], \mathbb{R}^m) := \bigcup_{\mathbb{T} \in \mathcal{P}_N^\tau} \text{PC}_N^{\mathbb{T}}([0, \tau], \mathbb{R}^m).$$

Note that  $\text{PC}_N([0, \tau], \mathbb{R}^m)$  is included in  $L^\infty([0, \tau], \mathbb{R}^m)$ , but it is not a linear subspace, neither a convex subset. As explained in Introduction, this lack of convexity constitutes a difficulty in the present work. Precisely it leads us to adapt in the proof of Lemma A.10 the standard procedure of convex  $L^\infty$ -perturbations of the control that can be found for example in [13, Lemma 4.17 p.84].

**Remark 2.2.** Similarly to Remark 2.1, note that  $\text{PC}_M([0, \tau], \mathbb{R}^m) \subset \text{PC}_N([0, \tau], \mathbb{R}^m)$  for all  $M, N \in \mathbb{N}^*$  such that  $M \leq N$ .

## 2.2 The optimal sampled-data control problem: terminology and assumptions

Let  $m, n, j, N \in \mathbb{N}^*$  be four positive integers fixed in the whole paper. In the present work we focus on the general optimal sampled-data control problem (OSCP) given by

$$\left. \begin{aligned} & \text{minimize} && \varphi(x(0), x(T), T) + \int_0^T L(x(t), u(t), t) dt, \\ & \text{subject to} && T > 0 \text{ fixed or free,} \\ & && \mathbb{T} = (t_i)_{i=0, \dots, N} \in \mathcal{P}_N^T \text{ fixed or free,} \\ & && x \in \text{AC}([0, T], \mathbb{R}^n), u \in \text{PC}_N^{\mathbb{T}}([0, T], \mathbb{R}^m), \\ & && \dot{x}(t) = f(x(t), u(t), t), \quad \text{a.e. } t \in [0, T], \\ & && g(x(0), x(T), T) \in S, \\ & && u_i \in \Omega, \quad \text{for all } i = 0, \dots, N-1. \end{aligned} \right\} \quad (\text{OSCP})$$

A solution to Problem (OSCP) is thus a quadruple  $(T, \mathbb{T}, x, u)$  which satisfies all above constraints and which minimizes the cost among all quadruples satisfying these constraints. Our aim in this section is to fix the terminology and the assumptions associated to Problem (OSCP).

In Problem (OSCP),  $x$  is the *state* function (also called *trajectory*) and  $u$  is the *control* function. In the classical literature about the Pontryagin maximum principle (see, e.g., [29, 33, 35] and references therein), the control  $u$  usually can be any function in  $L^\infty([0, T], \mathbb{R}^m)$ , satisfying moreover the constraint  $u(t) \in \Omega$  for almost every  $t \in [0, T]$ . In the present paper, the control  $u$  is constrained to be a piecewise constant function respecting at least one  $N$ -partition, where  $N \in \mathbb{N}^*$  is fixed. In other words, the value of the control is authorized to be modified at most  $N - 1$  times. In that situation the standard terminology adopted in the literature is to say that the control  $u$  in Problem (OSCP) is a *sampled-data control* (see, e.g., [6, 12, 13, 14] and references therein).

In Problem (OSCP), the *final time*  $T > 0$  can be fixed or not. In the case where the final time is free, it becomes a parameter to optimize. Similarly the  $N$ -partition  $\mathbb{T} = (t_i)_{i=0, \dots, N}$  can be fixed or not in Problem (OSCP). For  $i = 1, \dots, N - 1$ , the elements  $t_i$  of  $\mathbb{T}$  are called the *sampling times* because they correspond to the times in which the value of the sampled-data control  $u$  can be modified. We distinguish two situations:

- (i) If the  $N$ -partition is fixed in Problem (OSCP), we say that the sampling times  $t_i$  are fixed and Problem (OSCP) is an *optimal sampled-data control problem with fixed sampling times*;
- (ii) If the  $N$ -partition is free in Problem (OSCP), we say that the sampling times  $t_i$  are free and they become  $N - 1$  parameters to optimize. In that case, Problem (OSCP) is said to be an *optimal sampled-data control problem with free sampling times*.

In this paper we consider the following regularity and topology assumptions:

- the functions  $\varphi : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}$  and  $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$ , that describe respectively the *Mayer cost*  $\varphi(x(0), x(T), T)$  and the *Lagrange cost*  $\int_0^T L(x(t), u(t), t) dt$ , are of class  $C^1$ ;
- the set  $S \subset \mathbb{R}^j$  is a nonempty closed convex subset of  $\mathbb{R}^j$  and the function  $g : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}^j$ , that describes the *terminal state constraint*  $g(x(0), x(T), T) \in S$ , is of class  $C^1$ ;
- the set  $\Omega \subset \mathbb{R}^m$ , that describes the *control constraint*  $u_i \in \Omega$ , is a nonempty closed convex subset of  $\mathbb{R}^m$ ;
- the *dynamics*  $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+ \rightarrow \mathbb{R}^n$ , that drives the *state equation*  $\dot{x}(t) = f(x(t), u(t), t)$ , is of class  $C^1$ . In particular, for every compact subset  $K \subset \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+$ , there exists a nonnegative constant  $L \geq 0$  such that  $\|\partial_1 f(x, u, t)\|_{\mathbb{R}^n \times n} \leq L$ ,  $\|\partial_2 f(x, u, t)\|_{\mathbb{R}^n \times m} \leq L$  for all  $(x, u, t) \in K$ , and such that

$$\|f(x_2, u_2, t) - f(x_1, u_1, t)\|_{\mathbb{R}^n} \leq L(\|x_2 - x_1\|_{\mathbb{R}^n} + \|u_2 - u_1\|_{\mathbb{R}^m}), \quad (1)$$

for all  $(x_1, u_1, t), (x_2, u_2, t) \in K$ .

Since the total cost  $\varphi(x(0), x(T), T) + \int_0^T L(x(t), u(t), t) dt$  considered in Problem (OSCP) is written as the sum of a Mayer cost and a Lagrange cost, it is said to be of *Bolza form*.

**Remark 2.3.** In Problem (OSCP), since the  $N$ -partition depends on the final time, note that the problem has no sense if the final time is free while the  $N$ -partition is fixed. This case will not be treated in this paper. In order to deal with it, one should introduce a different framework that can handle partitions that are independent of the final time. This is not the aim of the present paper. Our objective here is to focus on the free sampling times case.

**Remark 2.4.** If the final time is fixed in Problem (OSCP), one can directly consider that the two functions  $\varphi$  and  $g$  are both independent of  $T$ , and we directly write the Mayer cost as  $\varphi(x(0), x(T))$  and the terminal state constraint as  $g(x(0), x(T)) \in S$ .

**Remark 2.5.** In the free sampling times case, Problem (OSCP) can be rewritten by removing the third line, and by replacing “ $u \in \text{PC}_N^{\mathbb{T}}([0, T], \mathbb{R}^m)$ ” in the fourth line by “ $u \in \text{PC}_N([0, T], \mathbb{R}^m)$ ”.

**Remark 2.6.** Let  $(T, \mathbb{T}, x, u)$  be a solution to Problem (OSCP). One can easily deduce, respectively from Remarks 2.1 and 2.2, that:

- if the sampling times are fixed in Problem (OSCP) and  $u \in \text{PC}_M^{\mathbb{T}'}([0, T], \mathbb{R}^m)$  for some  $M \leq N$  and some  $\mathbb{T}' \in \mathcal{P}_M^{\mathbb{T}}$  such that  $\mathbb{T}' \subset \mathbb{T}$ , then the quadruple  $(T, \mathbb{T}', x, u)$  is a solution to the same problem as Problem (OSCP) replacing  $N$  by  $M$  and  $\mathbb{T}$  by  $\mathbb{T}'$ . A similar remark was already done in [13, Remark 3 p.60];
- if the sampling times are free in Problem (OSCP) and  $u \in \text{PC}_M^{\mathbb{T}'}([0, T], \mathbb{R}^m)$  for some  $M \leq N$  and some  $\mathbb{T}' \in \mathcal{P}_M^{\mathbb{T}}$ , then the quadruple  $(T, \mathbb{T}', x, u)$  is a solution to the same problem as Problem (OSCP) replacing  $N$  by  $M$ .

**Remark 2.7.** Let  $(T, \mathbb{T}, x, u)$  be a quadruple satisfying all constraints of Problem (OSCP). From the state equation and since  $u$  is a piecewise constant function, it is clear that  $x$  is not only absolutely continuous but also piecewise smooth of class  $C^1$  over the interval  $[0, T]$ , in the sense that  $x$  is of class  $C^1$  over each interval  $[t_i, t_{i+1}]$ .

**Remark 2.8.** A Filippov-type theorem for the existence of a solution to Problem (OSCP) in case of fixed sampling times was derived in [13, Theorem 2.1 p.61]. The present work focuses only on necessary optimality conditions and thus it is not our aim to discuss the extension of the previously mentioned result to the case of free sampling times. Nevertheless we precise that, in the context of free sampling times, one would likely be faced with the same difficulty encountered in the proof of Theorem 2.1 developed in Appendix A. Precisely, considering a minimizing sequence of sampled-data controls would lead to a sequence of partitions and thus to the possibility of accumulation of sampling times. As a consequence, a cautious and rigorous mathematical treatment would be required in order to give a meaning to the limit of the sequence of sampled-data controls when accumulations of sampling times appear. Moreover, note that the standard Filippov's theorem is usually established in case of permanent controls, that is, with controls that belong to the infinite dimensional space  $L^\infty([0, T], \mathbb{R}^m)$ , while the sampled-data control framework considered here (with fixed or free sampling times) can be seen as a finite dimensional optimization problem. This fundamental difference could potentially lead to existence results in case of sampled-data controls with relaxed assumptions with respect to the case of permanent controls. As a conclusion, the issue of existence in case of sampled-data controls with free sampling times will be the central topic of a future and full fledged paper.

### 2.3 Pontryagin maximum principle and general comments

The main objective of the present paper is to state a Pontryagin maximum principle for Problem (OSCP). As mentioned in the previous section, one of the novelties of Problem (OSCP) with respect to the classical literature is to consider sampled-data controls. Note that this framework was already considered by Bourdin and Trélat in [12, 13] in which a Pontryagin maximum principle was already established. However, in contrary to the framework considered in [12, 13], the sampling times  $t_i$  in Problem (OSCP) are not necessarily fixed and can be free. Hence the major contribution of the present paper is to state a Pontryagin maximum principle that can handle, not only sampled-data controls, but also free sampling times. In that particular case, a new necessary optimality condition is derived (see Equality (3) in Theorem 2.1 below). This additional necessary optimality condition happens to coincide with the continuity of the Hamiltonian function. A discussion devoted to this phenomenon is provided in Section 2.4. Before stating our main result, we first need to recall the three following standard notions.

**Definition 2.1** (Hamiltonian). *The Hamiltonian  $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  associated to Problem (OSCP) is defined by*

$$H(x, u, p, p^0, t) := \langle p, f(x, u, t) \rangle_{\mathbb{R}^n} + p^0 L(x, u, t),$$

for all  $(x, u, p, p^0, t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+$ .

**Definition 2.2** (Normal cone). *The normal cone to  $S \subset \mathbb{R}^j$  at a point  $z \in S$  is defined as the set*

$$N_S[z] := \{z' \in \mathbb{R}^j \mid \forall z'' \in S, \langle z', z'' - z \rangle_{\mathbb{R}^j} \leq 0\}.$$

*It is a closed convex cone containing  $0_{\mathbb{R}^j}$ .*

**Definition 2.3** (Submersiveness). *We say that  $g$  is submersive at a point  $(x_1, x_2, \tau) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+$  if its differential at this point, that is*

$$Dg(x_1, x_2, \tau) = \begin{pmatrix} \partial_1 g(x_1, x_2, \tau) & \partial_2 g(x_1, x_2, \tau) & \partial_3 g(x_1, x_2, \tau) \end{pmatrix} \in \mathbb{R}^{j \times (2n+1)},$$

*is surjective.*

We are now in a position to state the main result of the present paper.

**Theorem 2.1** (Pontryagin maximum principle). *Let  $(T, \mathbb{T}, x, u)$  be a solution to Problem (OSCP). If  $g$  is submersive at  $(x(0), x(T), T)$ , then there exists a nontrivial couple  $(p, p^0) \in \text{AC}([0, T], \mathbb{R}^n) \times \mathbb{R}_-$  such that:*

(i) **Adjoint equation:**  $p$  satisfies

$$-\dot{p}(t) = \partial_1 f(x(t), u(t), t)^\top \times p(t) + p^0 \partial_1 L(x(t), u(t), t),$$

*for almost every  $t \in [0, T]$ ;*

(ii) **Transversality conditions on the adjoint vector:**  $p$  satisfies

$$\begin{aligned} -p(0) &= p^0 \partial_1 \varphi(x(0), x(T), T) + \partial_1 g(x(0), x(T), T)^\top \times \Psi, \\ p(T) &= p^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \Psi, \end{aligned}$$

*where  $\Psi \in \mathbb{R}^j$  is such that  $-\Psi \in \text{N}_S[g(x(0), x(T), T)]$ ;*

(iii) **Nonpositive averaged Hamiltonian gradient condition:** the condition

$$\left\langle \int_{t_i}^{t_{i+1}} \partial_2 H(x(t), u_i, p(t), p^0, t) dt, \omega - u_i \right\rangle_{\mathbb{R}^m} \leq 0, \quad (2)$$

*is satisfied for all  $\omega \in \Omega$  and all  $i = 0, \dots, N-1$ ;*

(iv) **If moreover the sampling times are free in Problem (OSCP):** the optimal sampling times  $t_i$  satisfy the Hamiltonian continuity condition

$$H(x(t_i), u_{i-1}, p(t_i), p^0, t_i) = H(x(t_i), u_i, p(t_i), p^0, t_i), \quad (3)$$

*for all  $i = 1, \dots, N-1$ ;*

(v) **If moreover the final time is free in Problem (OSCP):** the optimal final time  $T$  satisfies the transversality condition

$$-H(x(T), u_{N-1}, p(T), p^0, T) = p^0 \partial_3 \varphi(x(0), x(T), T) + \partial_3 g(x(0), x(T), T)^\top \times \Psi,$$

*where  $\Psi \in \mathbb{R}^j$  is introduced in the transversality conditions on the adjoint vector.*

Appendix A is dedicated to the proof of Theorem 2.1. A list of comments is in order. We just point out, as detailed in Remarks 2.14 and 2.15 below, that the submersion property considered in Theorem 2.1 is not restrictive. The reader who is interested in the continuity/constancy of the Hamiltonian function may jump directly to the specific Section 2.4.

**Remark 2.9.** *The nontrivial couple  $(p, p^0)$  in Theorem 2.1, which is a Lagrange multiplier, is defined up to a positive multiplicative scalar. In the normal case  $p^0 \neq 0$ , it is usual to normalize the Lagrange multiplier so that  $p^0 = -1$ .*

**Remark 2.10.** Let us consider the framework of Theorem 2.1. One can easily see that the couple  $(x, p)$  satisfies the Hamiltonian system

$$\dot{x}(t) = \partial_3 H(x(t), u(t), p(t), p^0, t), \quad -\dot{p}(t) = \partial_1 H(x(t), u(t), p(t), p^0, t),$$

for almost every  $t \in [0, T]$ .

**Remark 2.11.** Our strategy in Appendix A in order to prove Theorem 2.1 is based on the Ekeland variational principle [17, Theorem 1.1 p.324]. It requires the closedness of  $\Omega$  in order to define the corresponding penalized functional on a complete metric set (see details in Appendix A.3). The closure of  $\Omega$  is thus a crucial assumption in our strategy. On the other hand, the convexity of  $\Omega$  is also an essential hypothesis for our strategy in order to consider convex  $L^\infty$ -perturbation of the control (see the proof of Lemma A.10).

**Remark 2.12.** The authors of the present paper are very grateful to the anonymous reviewer for bringing to their attention an alternative proof of Theorem 2.1. By adapting a remarkable technique exposed in the paper [16] by Dmitruk and Kaganovich, optimal sampled-data control problems with free sampling times, considered in this paper, can indeed be reparameterized such that each sampling interval  $[t_i, t_{i+1}]$  maps to the interval  $[0, 1]$ . In that situation, the free sampling times  $t_i$  play the role of free terminal states which lead, through the application of the classical PMP, to transversality conditions which exactly coincide with the Hamiltonian continuity condition (3), while the values  $u_i$  of the sampled-data control play the role of parameters which lead, through the application of a “PMP with parameters” (see, e.g., [11, Remark 5 p.3790]), to a necessary optimality condition written in integral form which exactly coincides with the nonpositive averaged Hamiltonian gradient condition (2). This alternative approach should undoubtedly be the subject of a forthcoming work. On the other hand, the authors of the present paper are particularly interested in the extension of optimal control problems using the tools of time scale calculus (see, e.g., [11, 13]) and fractional calculus (see, e.g., [5]). It has to be noted that the method developed in [16] is strongly based on the standard chain rule which has no analogue neither in time scale calculus, nor in fractional calculus. As a consequence, the technique developed in [16] cannot be used in order to extend our main result (Theorem 2.1) in both of these contexts, while the proof based on Ekeland’s variational principle proposed in the present paper probably can. We take this occasion to mention that the continuity of the Hamiltonian function in the framework of fractional optimal control problems remains an open challenge (see details in [5, Section 5.1]).

**Remark 2.13.** If  $\Omega = \mathbb{R}^m$  (that is, if there is no control constraint in Problem (OSCP)), then the nonpositive averaged Hamiltonian gradient condition in Theorem 2.1 (see Inequality (2)) can be rewritten as

$$\int_{t_i}^{t_{i+1}} \partial_2 H(x(t), u_i, p(t), p^0, t) dt = 0_{\mathbb{R}^m},$$

for all  $i = 0, \dots, N - 1$ .

**Remark 2.14.** In this remark, for simplicity, we suppose that the final time is fixed in Problem (OSCP). Our aim here is to describe some typical terminal state constraint  $g(x(0), x(T)) \in S$  and the corresponding transversality conditions on the adjoint vector derived in Theorem 2.1:

- If the terminal points are fixed in Problem (OSCP) (that is,  $x(0) = x_0$  and  $x(T) = x_f$ ), one may consider  $j = 2n$ ,  $g$  as the identity function and  $S = \{x_0\} \times \{x_f\}$ . In that case, the normal cone to  $S$  is the entire space, and thus the transversality conditions on the adjoint vector in Theorem 2.1 do not provide any additional information.

- If the initial point is fixed (that is,  $x(0) = x_0$ ) and the final point  $x(T)$  is free in Problem (OSCP), one may consider  $j = 2n$ ,  $g$  as the identity function and  $S = \{x_0\} \times \mathbb{R}^n$ . In that case, the nontriviality of the couple  $(p, p^0)$  and the second transversality condition on the adjoint vector in Theorem 2.1 imply that  $p^0 \neq 0$  (which we normalize to  $p^0 = -1$ , see Remark 2.9) and  $p(T) = -\partial_2 \varphi(x(0), x(T))$ .
- If the initial point is fixed (that is,  $x(0) = x_0$ ) and the final point  $x(T)$  is subject to inequality constraints  $G_\ell(x(T)) \geq 0$  for  $\ell = 1, \dots, n_G$ , for some  $n_G \in \mathbb{N}^*$ , one may consider  $j = n + n_G$ ,  $g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n+n_G}$  given by  $g(x_1, x_2) := (x_1, G(x_2))$  where  $G = (G_1, \dots, G_{n_G}) : \mathbb{R}^n \rightarrow \mathbb{R}^{n_G}$  and  $S = \{x_0\} \times (\mathbb{R}_+)^{n_G}$ . If  $G$  is of class  $C^1$  and the differential  $DG(x_2) \in \mathbb{R}^{n_G \times n}$  is surjective at any point  $x_2 \in G^{-1}((\mathbb{R}_+)^{n_G})$ , then the second transversality condition on the adjoint vector in Theorem 2.1 can be written as  $p(T) = p^0 \partial_2 \varphi(x(0), x(T)) + \sum_{\ell=1}^{n_G} \lambda_\ell \nabla G_\ell(x(T))$ , for some  $\lambda_\ell \geq 0$  satisfying moreover the slackness condition  $\lambda_\ell G_\ell(x(T)) = 0$ , for all  $\ell = 1, \dots, n_G$ .
- If there is no Mayer cost (that is,  $\varphi = 0$ ) and the periodic condition  $x(0) = x(T)$  is considered in Problem (OSCP), one may consider  $j = n$ ,  $g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  given by  $g(x_1, x_2) := x_2 - x_1$  and  $S = \{0_{\mathbb{R}^n}\}$ . In that case the transversality conditions on the adjoint vector in Theorem 2.1 yield that  $p(0) = p(T)$ .

We point out that, in all examples above, the submersiveness condition is satisfied.

**Remark 2.15.** Let  $(T, \mathbb{T}, x, u)$  be a solution to Problem (OSCP). If the submersion property is not satisfied, one can easily go back to the submersive case by noting that  $(T, \mathbb{T}, x, u)$  is also a solution to the same problem as Problem (OSCP) replacing  $j$  by  $\tilde{j} := 2n + 1$ ,  $g$  by the identity function  $\tilde{g}$  and  $S$  by the singleton  $\tilde{S} := \{x(0)\} \times \{x(T)\} \times \{T\}$ . With this new problem the submersion property is obviously satisfied and Theorem 2.1 can be applied. However, with this new problem, the normal cone to  $\tilde{S}$  is the entire space, and thus the transversality conditions on the adjoint vector and on the final time do not provide any information. In other words, if the submersion property is not satisfied, then Theorem 2.1 is still valid by removing the two items (ii) and (v).

**Remark 2.16.** Following the proof of Theorem 2.1 in Appendix A, one can easily see that the theorem is still valid for a quadruple  $(T, \mathbb{T}, x, u)$  which is solution to Problem (OSCP) in (only) a local sense to precise.

**Remark 2.17.** In the case of free sampling times in Problem (OSCP), one may be interested by the additional constraint  $t_{i+1} - t_i \geq \delta_{\min}$  for all  $i = 0, \dots, N-1$ , for some  $\delta_{\min} > 0$  fixed. Following the proof of Theorem 2.1 in Appendix A, one can easily see that Equality (3) is preserved for all  $i \in \{1, \dots, N-1\}$  such that  $\min(t_i - t_{i-1}, t_{i+1} - t_i) > \delta_{\min}$ , but has to be replaced by the weaker condition

$$H(x(t_i), u_{i-1}, p(t_i), p^0, t_i) \leq H(x(t_i), u_i, p(t_i), p^0, t_i),$$

for all  $i \in \{1, \dots, N-1\}$  such that  $t_i - t_{i-1} = \delta_{\min}$  and  $t_{i+1} - t_i > \delta_{\min}$ , and by the weaker condition

$$H(x(t_i), u_{i-1}, p(t_i), p^0, t_i) \geq H(x(t_i), u_i, p(t_i), p^0, t_i),$$

for all  $i \in \{1, \dots, N-1\}$  such that  $t_i - t_{i-1} > \delta_{\min}$  and  $t_{i+1} - t_i = \delta_{\min}$ . However, if  $t_i - t_{i-1} = t_{i+1} - t_i = \delta_{\min}$ , then no necessary optimality condition on  $t_i$  can be derived from our strategy in Appendix A.

**Remark 2.18.** Remark 2.17 can be easily adapted to the case of the additional constraint  $t_{i+1} - t_i \leq \delta_{\max}$  for all  $i = 0, \dots, N-1$ , for some  $\delta_{\max} > 0$  fixed. One can also obtain a similar remark than Remark 2.17 by considering the additional constraint  $\delta_{\min} \leq t_{i+1} - t_i \leq \delta_{\max}$  for all  $i = 0, \dots, N-1$ , for some  $0 < \delta_{\min} < \delta_{\max}$  fixed.

**Remark 2.19.** *Thanks to the reviewing process of the present paper, several research perspectives have been brought to our attention. We take this occasion to thank the anonymous reviewers and to provide a (nonexhaustive) list of perspectives for future works:*

- (i) *In the context of linear-quadratic problems, the authors of [14] prove that the optimal sampled-data controls (with fixed sampling times) converge pointwisely to the optimal permanent control when the lengths of sampling intervals tend uniformly to zero. The convergence of the corresponding costs and the uniform convergence of the corresponding states and costates are also derived. An interesting research perspective would be to get similar convergence results in the context of the present work. Several directions can be investigated: nonlinear dynamics, terminal state constraints, free sampling times (whereas sampling times are fixed in [14]). In context of free sampling times, a wonderful challenge would be to study the asymptotic behavior when letting  $N$  tend to  $+\infty$  (which is a weaker condition than the uniform convergence to zero of the lengths of sampling intervals).*
- (ii) *In view of initializations of numerical algorithms, it would be relevant to get theoretical results about the distribution of optimal sampling times with respect to  $N$  and/or with respect to the data (cost, dynamics, constraints) of the considered problem.*
- (iii) *The present paper focuses on finite partitions. In view of handling chattering phenomenon, it would be relevant to extend the present framework to the case of infinite partitions. For example, one may consider the case of countably infinite partitions with exactly one accumulation point (as in the well-known Fuller's example).*
- (iv) *As evoked in the survey [27] by Margaliot, a sufficient condition for the stabilization of a switched system is obviously given by the stability of the trajectory associated to the worst-case switching law, itself implied by the stability of the trajectories associated to the convexified differential inclusion (in which a permanent control is introduced for playing the role of convexifying parameter). Having in mind this strategy, finding the most destabilizing permanent control can be solved using variational approaches, as well illustrated in [27]. In the other hand, in case of switched systems with a maximal number of switches, we note that the stability of the trajectories associated to the convexified differential inclusion by considering sampled-data controls with free sampling times is a weaker sufficient condition than considering permanent controls. As a consequence, the application of Theorem 2.1 in order to study the stabilization of switched systems in the previously mentioned context constitutes an attractive perspective for future works.*
- (v) *Last (but not least) a relevant research perspective would concern the extension of the present paper to the more general framework in which the values of the free sampling times  $t_i$  intervene explicitly in the cost to minimize and/or in the dynamics. Let us take this occasion to mention the paper [4] in which the authors derive Pontryagin-type conditions for a specific problem from medicine that can be written as an optimal sampled-data control problem in which the sampling times  $t_i$  are free and intervene explicitly in the expression of the dynamics. We precise that, even in this very particular context, giving an expression of the necessary optimality conditions in an Hamiltonian form still remains an open mathematical question.*

## 2.4 Continuity/constancy of the Hamiltonian function

Let us first recall the following standard notion.

**Definition 2.4** (Hamiltonian function). *With the framework of Theorem 2.1, the corresponding Hamiltonian function  $\mathcal{H} : [0, T] \rightarrow \mathbb{R}$  is defined by*

$$\mathcal{H}(t) := H(x(t), u(t), p(t), p^0, t),$$

for all  $t \in [0, T]$ .

**Remark 2.20.** If the final time is free in Problem (OSCP), note that the transversality condition on the optimal final time  $T$  in Theorem 2.1 can be rewritten as

$$-\mathcal{H}(T) = p^0 \partial_3 \varphi(x(0), x(T), T) + \partial_3 g(x(0), x(T), T)^\top \times \Psi,$$

since  $u(T) = u_{N-1}$ .

**Remark 2.21.** In the classical case of purely continuous optimal control problems, we recall that the (absolute) continuity of the Hamiltonian function  $\mathcal{H}$  is a very well-known fact (see, e.g., [19, Theorem 2.6.3 p.73]). Moreover it holds that

$$\dot{\mathcal{H}}(t) = \partial_5 H(x(t), u(t), p(t), p^0, t),$$

for almost all  $t \in [0, T]$ .

Let us consider the framework of Theorem 2.1. Similarly to the trajectory  $x$  (see Remark 2.7), it can easily be seen from the adjoint equation that the adjoint vector  $p$  is not only absolutely continuous but also piecewise smooth of class  $C^1$  over the interval  $[0, T]$ , in the sense that  $p$  is of class  $C^1$  over each interval  $[t_i, t_{i+1}]$ . Since moreover  $u$  is piecewise constant, it is clear that the Hamiltonian function  $\mathcal{H}$  is piecewise smooth of class  $C^1$  over  $[0, T]$ , in the sense that  $\mathcal{H}$  is of class  $C^1$  over each semi-open interval  $[t_i, t_{i+1})$  for  $i = 0, \dots, N-2$  and over the closed interval  $[t_{N-1}, t_N]$ . Moreover, since the couple  $(x, p)$  satisfies the Hamiltonian system (see Remark 2.10), it clearly holds that

$$\dot{\mathcal{H}}(t) = \partial_5 H(x(t), u(t), p(t), p^0, t),$$

over each semi-open interval  $[t_i, t_{i+1})$  for  $i = 0, \dots, N-2$  and over the closed interval  $[t_{N-1}, t_N]$ .

However, in contrary to the couple  $(x, p)$ , the Hamiltonian function  $\mathcal{H}$  is not continuous over  $[0, T]$  in general. It may admit a discontinuity at each sampling times  $t_i$ . We provide an example of this phenomenon in Section 3 (see Figure 1 in which the sampling times are fixed). Nevertheless, if the sampling times are free in Problem (OSCP), Equality (3) in Theorem 2.1 implies that the optimal sampling times  $t_i$  satisfy

$$\lim_{\substack{t \rightarrow t_i \\ t < t_i}} \mathcal{H}(t) = H(x(t_i), u_{i-1}, p(t_i), p^0, t_i) = H(x(t_i), u_i, p(t_i), p^0, t_i) = \mathcal{H}(t_i),$$

for all  $i = 1, \dots, N-1$ , which correspond exactly to the continuity of  $\mathcal{H}$  at each optimal sampling time  $t_i$ . In that situation we conclude that the Hamiltonian function  $\mathcal{H}$  is continuous over the whole interval  $[0, T]$ . The following result summarizes the previous remarks.

**Proposition 2.1.** *Let us consider the framework of Theorem 2.1. Then, the Hamiltonian function  $\mathcal{H}$  is piecewise smooth of class  $C^1$  over the interval  $[0, T]$ , in the sense that  $\mathcal{H}$  is of class  $C^1$  over each semi-open interval  $[t_i, t_{i+1})$  for  $i = 0, \dots, N-2$  and over the closed interval  $[t_{N-1}, t_N]$ , with the derivative*

$$\dot{\mathcal{H}}(t) = \partial_5 H(x(t), u(t), p(t), p^0, t). \quad (4)$$

Moreover:

- (i) If the sampling times are fixed in Problem (OSCP), then  $\mathcal{H}$  may admit a discontinuity at each sampling time  $t_i$ .
- (ii) If the sampling times are free in Problem (OSCP), then the Hamiltonian function  $\mathcal{H}$  is continuous at each optimal sampling time  $t_i$ . In that case,  $\mathcal{H}$  is continuous over the whole interval  $[0, T]$ .

Proposition 2.1 is illustrated with a simple linear-quadratic example numerically solved in the next Section 3.

We conclude this section by discussing the case where Problem (OSCP) is *autonomous*, in the sense that the dynamics  $f$  and the Lagrange cost function  $L$  are independent of the variable  $t$ . In that case, the Hamiltonian  $H$  is also independent of the variable  $t$ . In that situation, from Equality (4), we deduce that the Hamiltonian function  $\mathcal{H}$  is constant over each semi-open interval  $[t_i, t_{i+1})$  for  $i = 0, \dots, N-2$  and over the closed interval  $[t_{N-1}, t_N]$ . If moreover the sampling times are free in Problem (OSCP), we deduce from Proposition 2.1 that the Hamiltonian function  $\mathcal{H}$  is constant over the whole interval  $[0, T]$ .

### 3 Numerical illustration with a simple linear-quadratic example

Bourdin and Trélat have recently extended in [14, Section 3 p.275] the classical Riccati theory to the sampled-data control framework. In particular they provide in [14, Theorem 2 and Corollary 1 p.276] a numerical way to compute the optimal sampled-data control for linear-quadratic problems in the case of fixed sampling times. Our aim in this section is to adapt this method in order to numerically solve a simple linear-quadratic example with free sampling times, and to illustrate our discussion in Section 2.4 about the continuity of the Hamiltonian function. Precisely we will focus in this section on the following unidimensional linear-quadratic optimal sampled-data control problem (LQOSCP) given by

$$\left. \begin{array}{ll} \text{minimize} & x(1)^2 + \int_0^1 3x(t)^2 + u(t)^2 dt, \\ \text{subject to} & \mathbb{T} = (t_i)_{i=0,\dots,N} \in \mathcal{P}_N^1 \text{ free,} \\ & x \in \text{AC}([0, 1], \mathbb{R}), \ u \in \text{PC}_N^{\mathbb{T}}([0, 1], \mathbb{R}), \\ & \dot{x}(t) = x(t) - u(t) + t, \quad \text{a.e. } t \in [0, 1], \\ & x(0) = -4, \end{array} \right\} \quad (\text{LQOSCP})$$

with different values of  $N \in \mathbb{N}^*$ . Note that Problem (LQOSCP) satisfies all the assumptions of Section 2.2 and of Theorem 2.1, with the final time  $T = 1$  being fixed.

For the needs of this section, for all  $N$ -partitions  $\mathbb{T} \in \mathcal{P}_N^1$ , we will denote by  $(\text{LQOSCP}_{\mathbb{T}})$  the same problem as Problem (LQOSCP) replacing “free” by “fixed”, that is, Problem  $(\text{LQOSCP}_{\mathbb{T}})$  corresponds to Problem (LQOSCP) but with the fixed partition  $\mathbb{T}$ . As recalled in the beginning of the section, [14, Theorem 2 and Corollary 1 p.276] allows us to numerically compute, for all  $N$ -partitions  $\mathbb{T} \in \mathcal{P}_N^1$ , the optimal cost (denoted by  $C_{\mathbb{T}}$ ) and the Hamiltonian function (denoted by  $\mathcal{H}_{\mathbb{T}}$ ) corresponding to Problem  $(\text{LQOSCP}_{\mathbb{T}})$ . Hence, in order to numerically solve Problem (LQOSCP) (with free sampling times), we can follow two different methods:

- (i) Firstly we directly minimize the optimal cost mapping  $\mathbb{T} \mapsto C_{\mathbb{T}}$  (using the MATLAB function *fmincon*).
- (ii) Secondly, following the Hamiltonian continuity conditions (3) in Theorem 2.1, we apply a shooting method (based on the MATLAB function *fsolve*) on the Hamiltonian discontinuities mapping given by

$$\mathbb{T} = (t_i)_{i=0,\dots,N} \mapsto \left( \mathcal{H}_{\mathbb{T}}(t_i) - \lim_{\substack{t \rightarrow t_i \\ t < t_i}} \mathcal{H}_{\mathbb{T}}(t) \right)_{i=1,\dots,N-1}.$$

Both methods yield the same optimal sampling times. Hence Problem (LQSCP) is numerically solved and we present hereafter some numerical simulations for different values of  $N$ . In particular we compare the results with the fixed uniform partition case (see Table 1). As expected, one can clearly observe that the optimal cost  $C_T$  is lower for the optimal sampling times than for the fixed uniform partition.

$N$	Fixed uniform partition	$C_T$	Optimal sampling times	$C_T$
2	$\mathbb{T} = \{0, 0.5, 1\}$	46.6828	$\mathbb{T} = \{0, 0.3592, 1\}$	46.0285
4	$\mathbb{T} = \{0, 0.25, 0.5, 0.75, 1\}$	44.5131	$\mathbb{T} = \{0, 0.1574, 0.3544, 0.6163, 1\}$	44.3159
8	$\mathbb{T} = \{0, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 1\}$	43.9704	$\mathbb{T} = \{0, 0.0744, 0.1567, 0.2487, 0.3529, 0.4729, 0.6140, 0.7847, 1\}$	43.9191

Table 1: Comparison of optimal costs  $C_T$  (fixed uniform partition versus optimal sampling times).

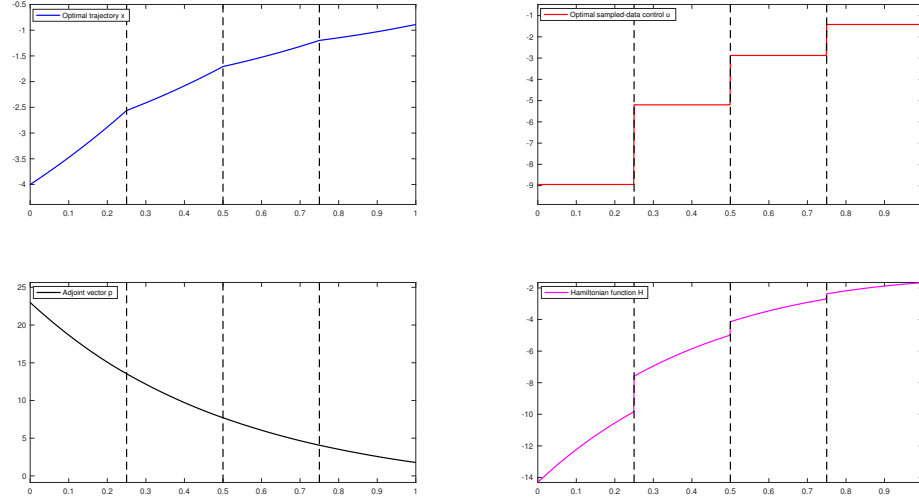


Figure 1: The case  $N = 4$  with fixed uniform partition.

In Figure 1 (with  $N = 4$  and fixed uniform partition), as expected from Section 2.4, we observe that the Hamiltonian function  $\mathcal{H}$  is continuous over each semi-open interval  $[t_i, t_{i+1})$  for  $i = 0, 1, 2$  and over the closed interval  $[t_3, t_4]$ . However, since the uniform partition is not optimal in that situation, the Hamiltonian function  $\mathcal{H}$  has discontinuities at each  $t_i$ . On the contrary, in Figure 2 (with  $N = 4$  and optimal sampling times), we observe that the Hamiltonian function  $\mathcal{H}$  is continuous over the whole interval  $[0, 1]$ . These numerical results are coherent with the discussion addressed in Section 2.4.

To conclude this section, note that the above numerical results emphasize the effectiveness of our two methods in order to compute the optimal sampling times of a simple linear-quadratic example. Numerous perspectives can be investigated by using other methods than the Riccati theory from [14, Section 3 p.275] and by considering more sophisticated problems such as nonlinear problems, multidimensional problems and by handling final state and/or control constraints, with or without free final time, etc.

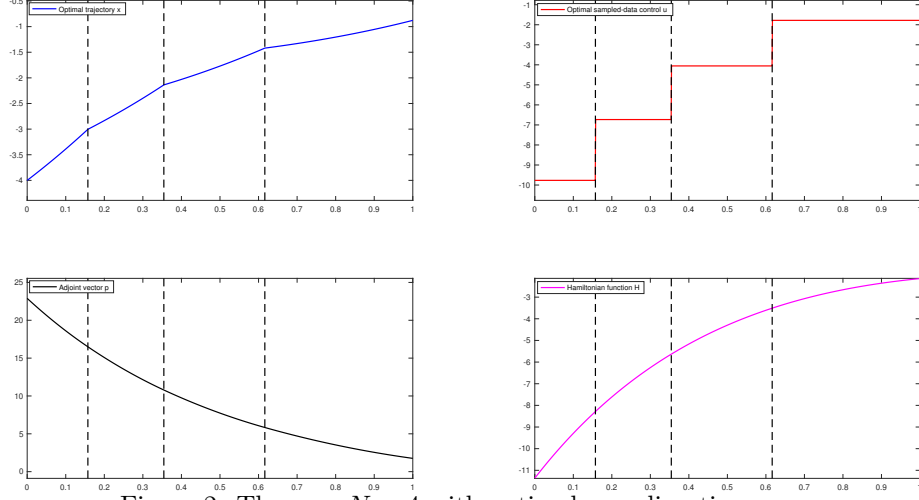


Figure 2: The case  $N = 4$  with optimal sampling times.

## A Proof of Theorem 2.1

This appendix is devoted to the detailed proof of Theorem 2.1 in the case  $L = 0$  (without Lagrange cost). Indeed, reducing a Bolza problem (that is, with  $L \neq 0$ ) to a Mayer problem (that is, with  $L = 0$ ) is very standard in the literature (see, e.g., [8, Section 2.1.4 p.12]).

We start with some required preliminaries in Section A.1. Then, the proof is based on the sensitivity analysis of the state equation in Section A.2 and on the application of the Ekeland variational principle in Section A.3.

### A.1 Preliminaries

The first part of this section is devoted to some basics of convex analysis. Let  $d_S : \mathbb{R}^j \rightarrow \mathbb{R}_+$  denote the standard distance function to  $S$  defined by  $d_S(z) := \inf_{z' \in S} \|z - z'\|_{\mathbb{R}^j}$  for all  $z \in \mathbb{R}^j$ . We recall that, for all  $z \in \mathbb{R}^j$ , there exists a unique element  $P_S(z) \in S$  (called the *projection* of  $z$  onto  $S$ ) such that  $d_S(z) = \|z - P_S(z)\|_{\mathbb{R}^j}$ . It can easily be shown that the map  $P_S : \mathbb{R}^j \rightarrow S$  is 1-Lipschitz continuous. Moreover it holds that  $\langle z - P_S(z), z' - P_S(z) \rangle_{\mathbb{R}^j} \leq 0$  for all  $z \in \mathbb{R}^j$  and all  $z' \in S$ . Let us recall the three following useful lemmas.

**Lemma A.1.** *It holds that  $z - P_S(z) \in N_S[P_S(z)]$  for all  $z \in \mathbb{R}^j$ .*

**Lemma A.2.** *Let  $(z_k)_{k \in \mathbb{N}}$  be a sequence in  $\mathbb{R}^j$  converging to some point  $z \in S$  and let  $(\zeta_k)_{k \in \mathbb{N}}$  be a sequence in  $\mathbb{R}_+$ . If  $\zeta_k(z_k - P_S(z_k))$  converges to some  $\bar{z} \in \mathbb{R}^j$ , then  $\bar{z} \in N_S[z]$ .*

**Lemma A.3.** *The map*

$$\begin{aligned} d_S^2 : \mathbb{R}^j &\longrightarrow \mathbb{R}_+ \\ z &\longmapsto d_S^2(z) := d_S(z)^2, \end{aligned}$$

*is differentiable on  $\mathbb{R}^j$ , and its differential  $Dd_S^2(z)$  at every  $z \in \mathbb{R}^j$  can be expressed as*

$$Dd_S^2(z)(z') = 2\langle z - P_S(z), z' \rangle_{\mathbb{R}^j},$$

*for all  $z' \in \mathbb{R}^j$ .*

The second part of this section is devoted to additional notions about piecewise constant functions and to several technical results useful for the proof of Theorem 2.1. Precisely we will introduce a technical control set (see Equation (5)) which allows to avoid two degenerate situations in the behavior of the sequence of sampling times produced by the Ekeland variational principle in Section A.3. We also refer to Introduction and to Proposition A.2 for details. Let  $\tau > 0$  and  $N \in \mathbb{N}^*$  be fixed. For all  $\mathbb{T} = (t_i)_{i=0,\dots,N} \in \mathcal{P}_N^\tau$  and  $u \in \text{PC}_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$ , we denote by

$$\|\mathbb{T}\| := \min\{t_{i+1} - t_i \mid i = 0, \dots, N-1\} > 0,$$

and we define the set

$$\mathcal{P}_{N,(u,\mathbb{T})}^\tau := \left\{ \mathbb{T}' = (t'_i)_{i=0,\dots,N} \in \mathcal{P}_N^\tau \mid \forall i = 1, \dots, N-1, |t'_i - t_i| \leq \delta_{\{u_{i-1} \neq u_i\}} \frac{\|\mathbb{T}\|}{4} \right\},$$

where  $\delta_{\{u_{i-1} \neq u_i\}} = 1$  if  $u_{i-1} \neq u_i$ , and  $\delta_{\{u_{i-1} \neq u_i\}} = 0$  otherwise. In particular, if  $\mathbb{T}' = (t'_i)_{i=0,\dots,N} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$ , it holds that

$$\begin{aligned} 0 = t'_0 < t_1 - \frac{\|\mathbb{T}\|}{4} \leq t'_1 \leq t_1 + \frac{\|\mathbb{T}\|}{4} < t_2 - \frac{\|\mathbb{T}\|}{4} \leq t'_2 \leq t_2 + \frac{\|\mathbb{T}\|}{4} < \dots \\ \dots < t_{N-2} - \frac{\|\mathbb{T}\|}{4} \leq t'_{N-2} \leq t_{N-2} + \frac{\|\mathbb{T}\|}{4} < t_{N-1} - \frac{\|\mathbb{T}\|}{4} \leq t'_{N-1} \leq t_{N-1} + \frac{\|\mathbb{T}\|}{4} < t'_N = \tau, \end{aligned}$$

with  $t'_i = t_i$  for all  $i \in \{1, \dots, N-1\}$  such that  $u_{i-1} = u_i$ . Hence, for all  $\mathbb{T}' = (t'_i)_{i=0,\dots,N} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$ , the elements  $t'_i = t_i$  remain unchanged when  $u$  is constant over two consecutive sampling intervals  $[t_{i-1}, t_i]$  and  $[t_i, t_{i+1}]$  and all the elements  $t'_i$  live in intervals which are (strictly) disjoint. Finally we introduce the following technical control set

$$\text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m) := \bigcup_{\mathbb{T}' \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau} \text{PC}_N^{\mathbb{T}'}([0, \tau], \mathbb{R}^m). \quad (5)$$

Of course note that  $\mathbb{T} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$  and thus  $u \in \text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$ . Also note that the inclusion  $\mathcal{P}_{N,(u,\mathbb{T})}^\tau \subset \mathcal{P}_N^\tau$  holds and thus  $\text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$  is included in  $\text{PC}_N([0, \tau], \mathbb{R}^m) \subset L^\infty([0, \tau], \mathbb{R}^m)$ , but is not a linear subspace, neither a convex subset.

**Lemma A.4.** *Let  $\tau > 0$  and  $N \in \mathbb{N}^*$ . Let  $\mathbb{T} = (t_i)_{i=0,\dots,N} \in \mathcal{P}_N^\tau$  and  $u \in \text{PC}_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$ . Then  $\mathbb{T}$  is the unique element  $\mathbb{T}' \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$  such that  $u \in \text{PC}_N^{\mathbb{T}'}([0, \tau], \mathbb{R}^m)$ .*

*Proof.* Let  $\mathbb{T}' = (t'_i)_{i=0,\dots,N} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$  be such that  $u \in \text{PC}_N^{\mathbb{T}'}([0, \tau], \mathbb{R}^m)$ . Let us assume by contradiction that  $\mathbb{T}' \neq \mathbb{T}$ . Let  $i \in \{1, \dots, N-1\}$  such that  $t_i \notin \mathbb{T}'$ . Necessarily it holds that  $u_{i-1} \neq u_i$  and there exists  $j \in \{i-1, i\}$  such that  $t'_j < t_i < t'_{j+1}$ . Since  $u \in \text{PC}_N^{\mathbb{T}'}([0, \tau], \mathbb{R}^m)$ , there exists  $c \in \mathbb{R}^m$  such that  $u(t) = c$  for almost every  $t \in [t'_j, t'_{j+1}]$ . Since  $u(t) = u_{i-1}$  for almost every  $t \in [t_{i-1}, t_i]$  and  $u(t) = u_i$  for almost every  $t \in [t_i, t_{i+1}]$ , we deduce that  $c = u_{i-1}$  and  $c = u_i$  which raises a contradiction since  $u_{i-1} \neq u_i$ . The proof is complete.  $\square$

**Proposition A.1.** *Let  $\tau > 0$  and  $N \in \mathbb{N}^*$ . Let  $\mathbb{T} = (t_i)_{i=0,\dots,N} \in \mathcal{P}_N^\tau$  and  $u \in \text{PC}_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$ . The set  $\text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$  is a closed subset of  $L^1([0, \tau], \mathbb{R}^m)$ .*

*Proof.* Let  $(u_k)_{k \in \mathbb{N}}$  be a sequence in  $\text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$  converging in  $L^1([0, \tau], \mathbb{R}^m)$  to some  $u' \in L^1([0, \tau], \mathbb{R}^m)$ . Our aim is to prove that  $u' \in \text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$ . The proof is divided in three steps.

*First step:* Let  $\mathbb{T}_k = (t_{i,k})_{i=0,\dots,N} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$  be a partition associated to  $u_k$  for all  $k \in \mathbb{N}$ . It holds for all  $k \in \mathbb{N}$  that

$$\begin{aligned} 0 = t_{0,k} &< t_1 - \frac{\|\mathbb{T}\|}{4} \leq t_{1,k} \leq t_1 + \frac{\|\mathbb{T}\|}{4} < t_2 - \frac{\|\mathbb{T}\|}{4} \leq t_{2,k} \leq t_2 + \frac{\|\mathbb{T}\|}{4} < \dots \\ \dots &< t_{N-2} - \frac{\|\mathbb{T}\|}{4} \leq t_{N-2,k} \leq t_{N-2} + \frac{\|\mathbb{T}\|}{4} < t_{N-1} - \frac{\|\mathbb{T}\|}{4} \leq t_{N-1,k} \leq t_{N-1} + \frac{\|\mathbb{T}\|}{4} < t_{N,k} = \tau, \end{aligned}$$

and  $t_{i,k} = t_i$  for all  $i \in \{1, \dots, N-1\}$  such that  $u_{i-1} = u_i$ . Extracting a finite number of subsequences (that we do not relabel), we know that, for all  $i \in \{0, \dots, N\}$ ,  $t_{i,k}$  converges to some  $t'_i$  satisfying

$$\begin{aligned} 0 = t'_0 &< t_1 - \frac{\|\mathbb{T}\|}{4} \leq t'_1 \leq t_1 + \frac{\|\mathbb{T}\|}{4} < t_2 - \frac{\|\mathbb{T}\|}{4} \leq t'_2 \leq t_2 + \frac{\|\mathbb{T}\|}{4} < \dots \\ \dots &< t_{N-2} - \frac{\|\mathbb{T}\|}{4} \leq t'_{N-2} \leq t_{N-2} + \frac{\|\mathbb{T}\|}{4} < t_{N-1} - \frac{\|\mathbb{T}\|}{4} \leq t'_{N-1} \leq t_{N-1} + \frac{\|\mathbb{T}\|}{4} < t'_N = \tau, \end{aligned}$$

and  $t'_i = t_i$  for all  $i \in \{1, \dots, N-1\}$  such that  $u_{i-1} = u_i$ . Hence we have obtained a partition  $\mathbb{T}' := (t'_i)_{i=0,\dots,N} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$ .

*Second step:* Extracting a subsequence (that we do not relabel) from the partial converse of the Lebesgue dominated convergence theorem, we know that  $u_k(t)$  converges to  $u'(t)$  for almost every  $t \in [0, \tau]$ . We introduce the subset  $A$  of  $[0, \tau]$  of full measure defined by

$$A := \{t \in [0, \tau] \mid u_k(t) \text{ converges to } u'(t)\},$$

and the subset  $B$  of  $[0, \tau]$  of full measure defined by  $B := \cap_{k \in \mathbb{N}} B_k$  where

$$B_k := \bigcup_{i=0}^{N-1} \{t \in [t_{i,k}, t_{i+1,k}) \mid u_k(t) = u_{i,k}\},$$

for all  $k \in \mathbb{N}$ .

*Third step:* Let  $i \in \{0, \dots, N-1\}$  and let  $t \in (t'_i, t'_{i+1}) \cap (A \cap B)$ . For  $k \in \mathbb{N}$  sufficiently large, it holds that  $t \in (t_{i,k}, t_{i+1,k})$ . Since  $t \in A \cap B$ , we know that  $u_k(t) = u_{i,k}$  which converges to  $u'(t)$ . Since the convergence of  $u_{i,k}$  to  $u'(t)$  is independent of the choice of  $t \in (t'_i, t'_{i+1}) \cap (A \cap B)$ , we deduce that  $u$  is equal almost everywhere over  $[t'_i, t'_{i+1}]$  to a constant. Since the last sentence is true for every  $i \in \{0, \dots, N-1\}$ , we conclude that  $u' \in \text{PC}_N^{\mathbb{T}'}([0, \tau], \mathbb{R}^m) \subset \text{PC}_{N,(u,\mathbb{T})}^{\mathbb{T}'}([0, \tau], \mathbb{R}^m)$ . The proof is complete.  $\square$

**Proposition A.2.** *Let  $\tau > 0$  and  $N \in \mathbb{N}^*$ . Let  $\mathbb{T} = (t_i)_{i=0,\dots,N} \in \mathcal{P}_N^\tau$  and  $u \in \text{PC}_N^{\mathbb{T}}([0, \tau], \mathbb{R}^m)$ . Let  $(u_k)_{k \in \mathbb{N}}$  be a sequence in  $\text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$  converging in  $L^1([0, \tau], \mathbb{R}^m)$  to  $u$ . Let  $\mathbb{T}_k = (t_{i,k})_{i=0,\dots,N} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$  be a partition associated to  $u_k$  for all  $k \in \mathbb{N}$ . Then there exists a subsequence of  $(u_k)_{k \in \mathbb{N}}$  (that we do not relabel) such that:*

- (i)  $u_k(t)$  converges to  $u(t)$  for almost every  $t \in [0, \tau]$ ;
- (ii)  $t_{i,k}$  converges to  $t_i$  for all  $i = 0, \dots, N$ ;
- (iii)  $u_{i,k}$  converges to  $u_i$  for all  $i = 0, \dots, N-1$ .

*Proof.* Following exactly the same steps as in the proof of Proposition A.1 (replacing  $u'$  by  $u$ ), we construct a partition  $\mathbb{T}' = (t'_i)_{i=0,\dots,N} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$  such that  $u \in \text{PC}_N^{\mathbb{T}'}([0, \tau], \mathbb{R}^m)$ . From Lemma A.4,

it implies that  $\mathbb{T}' = \mathbb{T}$ . From the construction of  $\mathbb{T}'$ , we conclude that, up to subsequences (which we do not relabel),  $t_{i,k}$  converges to  $t_i$  for all  $i = 0, \dots, N$ . Let us prove the last statement. Let us consider the sets  $A$  and  $B$  defined in the proof of Proposition A.1 and let us introduce the subset  $B'$  of  $[0, \tau]$  of full measure defined by

$$B' := \bigcup_{i=0}^{N-1} \{t \in [t_i, t_{i+1}) \mid u(t) = u_i\}.$$

Let  $i = 0, \dots, N-1$  and let  $t \in (t_i, t_{i+1}) \cap (A \cap B \cap B')$ . For  $k \in \mathbb{N}$  sufficiently large, it holds that  $t \in (t_{i,k}, t_{i+1,k})$ . Moreover, since  $t \in A \cap B \cap B'$ , we know that  $u_k(t) = u_{i,k}$  converges to  $u(t) = u_i$ . Since the last statement is true for all  $i = 0, \dots, N-1$ , the proof is complete.  $\square$

## A.2 Sensitivity analysis of the state equation

In this section we focus on the Cauchy problem given by

$$\begin{cases} \dot{x}(t) = f(x(t), u(t), t), & \text{a.e. } t \geq 0, \\ x(0) = x_0, \end{cases} \quad (\text{CP})$$

for any  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . Before proceeding to the sensitivity analysis of the Cauchy problem (CP) (with respect to the control  $u$  and the initial condition  $x_0$ ), we first recall some definitions and results from the classical Cauchy-Lipschitz (or Picard-Lindelöf) theory.

**Definition A.1.** Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . A (local) solution to the Cauchy problem (CP) is a couple  $(x, I)$  such that:

- (i)  $I$  is an interval such that  $\{0\} \subsetneq I \subset \mathbb{R}_+$ ;
- (ii)  $x \in \text{AC}([0, \tau], \mathbb{R}^n)$ , with  $\dot{x}(t) = f(x(t), u(t), t)$  for almost every  $t \in [0, \tau]$ , for all  $\tau \in I$ ;
- (iii)  $x(0) = x_0$ .

Let  $(x_1, I_1)$  and  $(x_2, I_2)$  be two (local) solutions to the Cauchy problem (CP). We say that  $(x_2, I_2)$  is an extension (resp. strict extension) to  $(x_1, I_1)$  if  $I_1 \subset I_2$  (resp.  $I_1 \subsetneq I_2$ ) and  $x_2(t) = x_1(t)$  for all  $t \in I_1$ . A maximal solution to the Cauchy problem (CP) is a (local) solution that does not admit any strict extension. Finally a global solution to the Cauchy problem (CP) is a solution  $(x, I)$  such that  $I = \mathbb{R}_+$ . In particular a global solution is necessarily a maximal solution.

**Lemma A.5.** Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . Any (local) solution to the Cauchy problem (CP) can be extended into a maximal solution.

**Lemma A.6.** Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . A couple  $(x, I)$  is a (local) solution to the Cauchy problem (CP) if and only if:

- (i)  $I$  is an interval such that  $\{0\} \subsetneq I \subset \mathbb{R}_+$ ;
- (ii)  $x \in C(I, \mathbb{R}^n)$ ;
- (iii)  $x$  satisfies the integral representation  $x(t) = x_0 + \int_0^t f(x(s), u(s), s) ds$  for all  $t \in I$ .

**Proposition A.3.** For all  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ , the Cauchy problem (CP) admits a unique maximal solution denoted by  $(x(\cdot, u, x_0), I(u, x_0))$ . Moreover the maximal interval  $I(u, x_0)$  is semi-open and we write  $I(u, x_0) = [0, \tau(u, x_0))$  where  $\tau(u, x_0) \in (0, +\infty]$ . Furthermore, if  $\tau(u, x_0) < +\infty$ , that is, if the maximal solution  $(x(\cdot, u, x_0), I(u, x_0))$  is not global, then  $x(\cdot, u, x_0)$  is not bounded over  $I(u, x_0) = [0, \tau(u, x_0))$ .

**Remark A.1.** Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . The maximal solution  $(x(\cdot, u, x_0), I(u, x_0))$  to the Cauchy problem (CP) coincides with the maximal extension (see Lemma A.5) of any other local solution.

Our aim in the next subsections is to study the behaviour of  $x(\cdot, u, x_0)$  with respect to perturbations on the control  $u$  and on the initial condition  $x_0$ .

### A.2.1 A continuity result

Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . In the sequel, for the ease of notations, we denote by  $\|\cdot\|_{L^\infty} := \|\cdot\|_{L^\infty(\mathbb{R}_+, \mathbb{R}^m)}$  and we introduce two sets:

- (i) For all  $R \geq \|u\|_{L^\infty}$  and all  $0 < \tau < \tau(u, x_0)$ , we denote by

$$K((u, x_0), (R, \tau)) := \{(y, v, t) \in \mathbb{R}^n \times \overline{B}_{\mathbb{R}^m}(0_{\mathbb{R}^m}, R) \times [0, \tau] \mid \|y - x(t, u, x_0)\|_{\mathbb{R}^n} \leq 1\}.$$

Firstly note that  $K((u, x_0), (R, \tau))$  is convex with respect to its first two variables. Secondly, since  $x(\cdot, u, x_0)$  is continuous over  $[0, \tau]$ , then  $K((u, x_0), (R, \tau))$  is a compact subset of  $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+$ . Thus we denote by  $L((u, x_0), (R, \tau)) \geq 0$  the Lipschitz constant of  $f$  over the compact subset  $K((u, x_0), (R, \tau))$  (see Inequality (1) in Section 2.2).

- (ii) For all  $R \geq \|u\|_{L^\infty}$  and all  $0 < \tau < \tau(u, x_0)$ , we denote by

$$\mathcal{N}((u, x_0), (R, \tau), \varepsilon)$$

$$:= \left\{ (u', x'_0) \in \left( \overline{B}_{L^1([0, \tau], \mathbb{R}^m)}(u, \varepsilon) \cap \overline{B}_{L^\infty}(0_{L^\infty}, R) \right) \times \overline{B}_{\mathbb{R}^n}(x_0, \varepsilon) \mid u' = u \text{ over } [\tau, +\infty) \right\},$$

for all  $\varepsilon > 0$ , which can be seen as a neighborhood of the couple  $(u, x_0)$  in the  $L^1([0, \tau], \mathbb{R}^m) \times \mathbb{R}^n$ -space. The second part of the above definition, imposing that  $u' = u$  over  $[\tau, +\infty)$ , allows us in the sequel to endow the above set with the  $L^1([0, \tau], \mathbb{R}^m) \times \mathbb{R}^n$ -distance.

In the next proposition we state a continuous dependence result for the trajectory  $x(\cdot, u, x_0)$  with respect to the couple  $(u, x_0)$ .

**Proposition A.4.** Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . For all  $R \geq \|u\|_{L^\infty}$  and all  $0 < \tau < \tau(u, x_0)$ , there exists  $\varepsilon > 0$  such that

$$\forall (u', x'_0) \in \mathcal{N}((u, x_0), (R, \tau), \varepsilon), \quad \tau(u', x'_0) > \tau.$$

Moreover, considering the  $L^1([0, \tau], \mathbb{R}^m) \times \mathbb{R}^n$ -distance over the set  $\mathcal{N}((u, x_0), (R, \tau), \varepsilon)$ , the map

$$(u', x'_0) \in \mathcal{N}((u, x_0), (R, \tau), \varepsilon) \longmapsto x(\cdot, u', x'_0) \in C([0, \tau], \mathbb{R}^n),$$

is Lipschitz continuous and

$$(x(t, u', x'_0), u'(t), t) \in K((u, x_0), (R, \tau)),$$

for almost every  $t \in [0, \tau]$  and for all  $(u', x'_0) \in \mathcal{N}((u, x_0), (R, \tau), \varepsilon)$ .

*Proof.* The proof is standard and left to the reader. For similar statements with detailed proofs, we refer to [11, Lemmas 1 and 3 p.3795-3797], [13, Lemmas 4.3 and 4.5 p.73-74] (in the general framework of time scale calculus) or to [10, Propositions 1 and 2 p.4-5] (in a more classical framework, closer to the present considerations).  $\square$

**Remark A.2.** Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . Let  $R \geq \|u\|_{L^\infty}$  and  $0 < \tau < \tau(u, x_0)$ . Let  $\varepsilon > 0$  given in Proposition A.4. Let  $(u_k, x_{0,k})_{k \in \mathbb{N}}$  be a sequence in  $\mathcal{N}((u, x_0), (R, \tau), \varepsilon)$  and let  $(u', x'_0) \in \mathcal{N}((u, x_0), (R, \tau), \varepsilon)$ . From Proposition A.4, if  $(u_k, x_{0,k})$  converges to  $(u', x'_0)$  in  $L^1([0, \tau], \mathbb{R}^m) \times \mathbb{R}^n$ , then the sequence  $(x(\cdot, u_k, x_{0,k}))_{k \in \mathbb{N}}$  uniformly converges to  $x(\cdot, u', x'_0)$  over  $[0, \tau]$ .

### A.2.2 Perturbation of the control

In the next proposition we state a differentiability result for the trajectory  $x(\cdot, u, x_0)$  with respect to a convex  $L^\infty$ -perturbation of the control  $u$ .

**Proposition A.5.** *Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$  and  $0 < \tau < \tau(u, x_0)$ . Let  $v \in L^\infty(\mathbb{R}_+, \mathbb{R}^m)$  be fixed. We consider the convex  $L^\infty$ -perturbation given by*

$$u_v(\cdot, \alpha) := \begin{cases} u + \alpha(v - u) & \text{over } [0, \tau), \\ u & \text{over } [\tau, +\infty), \end{cases}$$

for all  $0 \leq \alpha \leq 1$ . Then:

- (i) *there exists  $0 < \alpha_0 \leq 1$  such that  $\tau(u_v(\cdot, \alpha), x_0) > \tau$  for all  $0 \leq \alpha \leq \alpha_0$ ;*
- (ii) *the map*

$$\alpha \in [0, \alpha_0] \mapsto x(\cdot, u_v(\cdot, \alpha), x_0) \in C([0, \tau], \mathbb{R}^n),$$

*is differentiable at  $\alpha = 0$  and its derivative is equal to  $w_v$  being the unique solution (that is global) to the linear Cauchy problem given by*

$$\begin{cases} \dot{w}(t) = \partial_1 f(x(t, u, x_0), u(t), t) \times w(t) \\ \quad + \partial_2 f(x(t, u, x_0), u(t), t) \times (v(t) - u(t)), & \text{a.e. } t \in [0, \tau], \\ w(0) = 0_{\mathbb{R}^n}. \end{cases}$$

*Proof.* The proof is standard and left to the reader. For a similar statement with detailed proof, we refer to [11, Lemma 4 and Proposition 1 p.3797-3798].  $\square$

We conclude this section by a technical lemma on the convergence of the variation vectors. This result is needed in the proof of our main result (see Section A.3.2).

**Lemma A.7.** *Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . Let  $R \geq \|u\|_{L^\infty}$  and  $0 < \tau < \tau(u, x_0)$ . We take  $\varepsilon > 0$  as in Proposition A.4. Let  $(u_k, x_{0,k})_{k \in \mathbb{N}}$  be a sequence of elements in  $\mathcal{N}((u, x_0), (R, \tau), \varepsilon)$  such that  $x_{0,k}$  converges to  $x_0$  and  $u_k(t)$  converges to  $u(t)$  for almost every  $t \in [0, \tau]$ . Let  $(v_k)_{k \in \mathbb{N}}$  be a sequence in  $L^\infty([0, \tau], \mathbb{R}^m)$  converging in  $L^1([0, \tau], \mathbb{R}^m)$  to some  $v \in L^\infty([0, \tau], \mathbb{R}^m)$ . Finally let  $w_{v_k}^k$  be the unique solution (that is global) to the linear Cauchy problem given by*

$$\begin{cases} \dot{w}(t) = \partial_1 f(x(t, u_k, x_{0,k}), u_k(t), t) \times w(t) \\ \quad + \partial_2 f(x(t, u_k, x_{0,k}), u_k(t), t) \times (v_k(t) - u_k(t)), & \text{a.e. } t \in [0, \tau], \\ w(0) = 0_{\mathbb{R}^n}, \end{cases}$$

*for all  $k \in \mathbb{N}$ . Then the sequence  $(w_{v_k}^k)_{k \in \mathbb{N}}$  uniformly converges to  $w_v$  over  $[0, \tau]$  where  $w_v$  is defined as in Proposition A.5.*

*Proof.* The proof is standard and left to the reader. For a similar statement with detailed proof, we refer to [11, Lemmas 4.8 and 4.9 p.77-78].  $\square$

### A.2.3 Perturbation of the initial condition

In the next proposition we state a differentiability result for the trajectory  $x(\cdot, u, x_0)$  with respect to a simple perturbation of the initial condition  $x_0$ .

**Proposition A.6.** *Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$  and  $0 < \tau < \tau(u, x_0)$ . Let  $y \in \mathbb{R}^n$  be fixed. Then:*

- (i) *there exists  $\alpha_0 > 0$  such that  $\tau(u, x_0 + \alpha y) > \tau$  for all  $0 \leq \alpha \leq \alpha_0$ ;*
- (ii) *the map*

$$\alpha \in [0, \alpha_0] \longmapsto x(\cdot, u, x_0 + \alpha y) \in C([0, \tau], \mathbb{R}^n),$$

*is differentiable at  $\alpha = 0$  and its derivative is equal to  $w_y$  being the unique solution (that is global) to the linear homogeneous Cauchy problem given by*

$$\begin{cases} \dot{w}(t) = \partial_1 f(x(t, u, x_0), u(t), t) \times w(t), & \text{a.e. } t \in [0, \tau], \\ w(0) = y. \end{cases}$$

*Proof.* The proof is standard and left to the reader. For similar statements with detailed proofs, we refer to [11, Lemma 10 and Proposition 3 p.3802-3803] and to [13, Lemma 4.13 and Proposition 5 p.81-83].  $\square$

We conclude this section by a technical lemma on the convergence of the variation vectors. This result is needed in the proof of our main result (see Section A.3.2).

**Lemma A.8.** *Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . Let  $R \geq \|u\|_{L^\infty}$  and  $0 < \tau < \tau(u, x_0)$ . We take  $\varepsilon > 0$  as in Proposition A.4. Let  $(u_k, x_{0,k})_{k \in \mathbb{N}}$  be a sequence of elements in  $\mathcal{N}((u, x_0), (R, \tau), \varepsilon)$  such that  $x_{0,k}$  converges to  $x_0$  and  $u_k(t)$  converges to  $u(t)$  for almost every  $t \in [0, \tau]$ . Let  $y \in \mathbb{R}^n$  be fixed. Finally let  $w_y^k$  be the unique solution (that is global) to the linear homogeneous Cauchy problem given by*

$$\begin{cases} \dot{w}(t) = \partial_1 f(x(t, u_k, x_{0,k}), u_k(t), t) \times w(t), & \text{a.e. } t \in [0, \tau], \\ w(0) = y, \end{cases}$$

*for all  $k \in \mathbb{N}$ . Then the sequence  $(w_y^k)_{k \in \mathbb{N}}$  uniformly converges to  $w_y$  over  $[0, \tau]$  where  $w_y$  is defined as in Proposition A.6.*

*Proof.* The proof is standard and left to the reader. It is similar to the proof of Lemma A.7.  $\square$

#### A.2.4 Perturbation of a switching time

Let us introduce the following notion of *switching time* for a control  $u \in L^\infty(\mathbb{R}_+, \mathbb{R}^m)$ .

**Definition A.2** (Switching time). *Let  $u \in L^\infty(\mathbb{R}_+, \mathbb{R}^m)$ . We say that  $r > 0$  is a switching time of  $u$  if there exist  $0 < \eta_r \leq r$  and  $u(r^-), u(r^+) \in \mathbb{R}^m$  such that:*

- (i)  *$u = u(r^-)$  almost everywhere over  $[r - \eta_r, r]$ ;*
- (ii)  *$u = u(r^+)$  almost everywhere over  $[r, r + \eta_r]$ .*

This notion is particularly relevant when dealing with piecewise constant controls as in Problem (OSCP). Indeed, let  $u \in L^\infty(\mathbb{R}_+, \mathbb{R}^m)$  such that  $u \in \text{PC}_N^T([0, \tau], \mathbb{R}^m)$  for some  $\tau > 0$ ,  $N \in \mathbb{N}^*$  and  $\mathbb{T} = (t_i)_{i=0, \dots, N} \in \mathcal{P}_N^\tau$ . Then  $t_i$  is a switching time of  $u$  with  $u(t_i^-) = u_{i-1}$ ,  $u(t_i^+) = u_i$  and  $\eta_{t_i} = \min(t_i - t_{i-1}, t_{i+1} - t_i) > 0$  for every  $i \in \{1, \dots, N-1\}$ .

In the next proposition we prove a differentiability result for the trajectory  $x(\cdot, u, x_0)$  with respect to a perturbation of a switching time of the control  $u$ .

**Proposition A.7.** *Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$  and  $0 < \tau < \tau(u, x_0)$ . Let  $0 < r < \tau$  be a switching time of  $u$  and let  $\mu \in \{-1, 1\}$ . We consider the perturbation*

$$u_r^\mu(\cdot, \alpha) := \begin{cases} u(r^-) & \text{over } [r - \eta_r, r + \mu\alpha), \\ u(r^+) & \text{over } [r + \mu\alpha, r + \eta_r), \\ u & \text{otherwise,} \end{cases}$$

for all  $0 \leq \alpha \leq \frac{\eta_r}{2}$ . Then:

- (i) *there exists  $0 < \alpha_0 \leq \frac{\eta_r}{2}$  such that  $\tau(u_r^\mu(\cdot, \alpha), x_0) > \tau$  for all  $0 \leq \alpha \leq \alpha_0$ ;*
- (ii) *for any  $0 < \lambda \leq \tau - r$  fixed, the map*

$$\alpha \in [0, \alpha_0] \mapsto x(\cdot, u_r^\mu(\cdot, \alpha), x_0) \in C([r + \lambda, \tau], \mathbb{R}^n),$$

*is differentiable at  $\alpha = 0$  and its derivative is equal to  $w_r^\mu$  being the unique solution (that is global) to the linear homogeneous Cauchy problem given by*

$$\begin{cases} \dot{w}(t) = \partial_1 f(x(t, u, x_0), u(t), t) \times w(t), & \text{a.e. } t \in [r, \tau], \\ w(r) = \mu \left( f(x(r, u, x_0), u(r^-), r) - f(x(r, u, x_0), u(r^+), r) \right). \end{cases}$$

*Proof.* We only prove the case  $\mu = 1$  (the proof for the case  $\mu = -1$  is similar). First of all note that the variation vector  $w_r^\mu$  is global (in the sense that it is defined over the whole interval  $[r, \tau]$ ) since the corresponding Cauchy problem is linear. Let  $R := \|u\|_{L^\infty}$ . For the ease of notations we denote by  $K := K((u, x_0), (R, \tau))$  and by  $L := L((u, x_0), (R, \tau))$  (see the beginning of Section A.2.1 for these two notations). From Proposition A.4, there exists  $\varepsilon > 0$  such that  $\tau(u', x'_0) > \tau$  for all  $(u', x'_0) \in \mathcal{N}((u, x_0), (R, \tau), \varepsilon)$ . Let us take  $0 < \alpha_0 \leq \frac{\eta_r}{2}$  small enough such that  $r + \alpha_0 < \tau$  and  $2R\alpha_0 \leq \varepsilon$ . Then it holds that  $u_r^\mu(\cdot, \alpha) = u$  over  $[\tau, +\infty)$ ,  $\|u_r^\mu(\cdot, \alpha) - u\|_{L^1([0, \tau], \mathbb{R}^m)} \leq 2R\alpha \leq \varepsilon$  and  $\|u_r^\mu(\cdot, \alpha)\|_{L^\infty} \leq R$  for all  $0 \leq \alpha \leq \alpha_0$ . It follows that  $(u_r^\mu(\cdot, \alpha), x_0) \in \mathcal{N}((u, x_0), (R, \tau), \varepsilon)$  and  $\tau(u_r^\mu(\cdot, \alpha), x_0) > \tau$  for all  $0 \leq \alpha \leq \alpha_0$ . The first item of Proposition A.6 is proved. Since  $(u_r^\mu(\cdot, \alpha), x_0) \in \mathcal{N}((u, x_0), (R, \tau), \varepsilon)$  for all  $0 \leq \alpha \leq \alpha_0$  and  $u_r^\mu(\cdot, \alpha)$  converges to  $u$  in  $L^1([0, \tau], \mathbb{R}^m)$  as  $\alpha$  tends to zero, we know from Proposition A.4 that  $x(\cdot, u_r^\mu(\cdot, \alpha), x_0)$  converges uniformly to  $x(\cdot, u, x_0)$  over  $[0, \tau]$  as  $\alpha$  tends to zero and that  $(x(t, u_r^\mu(\cdot, \alpha), x_0), u_r^\mu(t, \alpha), t) \in K$  for almost every  $t \in [0, \tau]$  and for all  $0 \leq \alpha \leq \alpha_0$ . Now let us define the function

$$\varepsilon(t, \alpha) := \frac{x(t, u_r^\mu(\cdot, \alpha), x_0) - x(t, u, x_0)}{\alpha} - w_r^\mu(t),$$

for all  $t \in [r, \tau]$  and all  $\alpha \in (0, \alpha_0]$ . Let  $0 < \lambda \leq \tau - r$  be fixed. We will prove that  $\varepsilon(\cdot, \alpha)$  uniformly converges to the zero function on  $[r + \lambda, \tau]$  as  $\alpha$  tends to 0. From the integral representation of  $\varepsilon(\cdot, \alpha)$ , it holds that

$$\begin{aligned} \varepsilon(t, \alpha) = \varepsilon(r + \alpha, \alpha) + \int_{r+\alpha}^t & \left[ \frac{f(x(s, u_r^\mu(\cdot, \alpha), x_0), u(s), s) - f(x(s, u, x_0), u(s), s)}{\alpha} \right. \\ & \left. - \partial_1 f(x(s, u, x_0), u(s), s) \times w_r^\mu(s) \right] ds, \end{aligned}$$

for all  $t \in [r + \alpha, \tau]$  and all  $\alpha \in (0, \alpha_0]$ . Expanding this expression using Taylor's theorem with integral remainder, we obtain

$$\begin{aligned} \varepsilon(t, \alpha) = & \varepsilon(r + \alpha, \alpha) + \int_{r+\alpha}^t \int_0^1 \partial_1 f(\star_{\alpha\theta s}) d\theta \times \varepsilon(s, \alpha) ds \\ & + \int_{r+\alpha}^t \left( \int_0^1 \partial_1 f(\star_{\alpha\theta s}) - \partial_1 f(x(s, u, x_0), u(s), s) d\theta \right) \times w_r^\mu(s) ds, \end{aligned}$$

for all  $t \in [r + \alpha, \tau]$  and all  $\alpha \in (0, \alpha_0]$ , where

$$\star_{\alpha\theta s} := (x(s, u, x_0) + \theta(x(s, u_r^\mu(\cdot, \alpha), x_0) - x(s, u, x_0)), u(s), s) \in K,$$

since  $K$  is convex with respect to its first two variables. From the Triangle inequality it holds that

$$\|\varepsilon(t, \alpha)\|_{\mathbb{R}^n} \leq \|\varepsilon(r + \alpha, \alpha)\|_{\mathbb{R}^n} + \Phi(\alpha) + L \int_{r+\alpha}^t \|\varepsilon(s, \alpha)\|_{\mathbb{R}^n} ds,$$

for all  $t \in [r + \alpha, \tau]$  and all  $\alpha \in (0, \alpha_0]$ , where the term  $\Phi(\alpha)$  is defined to be:

$$\Phi(\alpha) := \int_r^\tau \int_0^1 \|\partial_1 f(\star_{\alpha\theta s}) - \partial_1 f(x(s, u, x_0), u(s), s)\|_{\mathbb{R}^n \times n} d\theta \|w_r^\mu(s)\|_{\mathbb{R}^n} ds.$$

From the Gronwall lemma, it holds that

$$\|\varepsilon(t, \alpha)\|_{\mathbb{R}^n} \leq (\|\varepsilon(r + \alpha, \alpha)\|_{\mathbb{R}^n} + \Phi(\alpha))e^{L\tau},$$

for all  $t \in [r + \alpha, \tau]$  and all  $\alpha \in (0, \alpha_0]$ . Since we only want to prove the uniform convergence of  $\varepsilon(\cdot, \alpha)$  to the zero function on  $[r + \lambda, \tau]$  as  $\alpha$  tends to 0 and since the estimate on the right-hand side is independent of  $t$ , we only need to prove that  $\varepsilon(r + \alpha, \alpha)$  tends to  $0_{\mathbb{R}^n}$  and  $\Phi(\alpha)$  tends to 0 as  $\alpha$  tends to zero. The convergence of  $\Phi(\alpha)$  can be obtained with the Lebesgue dominated convergence theorem. Now let us prove that  $\varepsilon(r + \alpha, \alpha)$  tends to  $0_{\mathbb{R}^n}$  as  $\alpha$  tends to zero. Since  $x(r, u_r^\mu(\cdot, \alpha), x_0) = x(r, u, x_0)$  and from the integral representations of  $x(\cdot, u_r^\mu(\cdot, \alpha), x_0)$  and  $x(\cdot, u, x_0)$ , it holds that

$$\begin{aligned} \varepsilon(r + \alpha, \alpha) = & \int_r^{r+\alpha} \frac{f(x(s, u_r^\mu(\cdot, \alpha), x_0), u_r^\mu(s, \alpha), s) - f(x(s, u, x_0), u(s), s)}{\alpha} ds - w_r^\mu(r + \alpha) \\ = & \int_r^{r+\alpha} \frac{f(x(s, u_r^\mu(\cdot, \alpha), x_0), u(r^-), s) - f(x(s, u, x_0), u(r^+), s)}{\alpha} ds - w_r^\mu(r + \alpha) \\ = & \int_r^{r+\alpha} \frac{f(x(s, u, x_0), u(r^-), s) - f(x(s, u, x_0), u(r^+), s)}{\alpha} ds - w_r^\mu(r + \alpha) \\ & + \int_r^{r+\alpha} \frac{f(x(s, u_r^\mu(\cdot, \alpha), x_0), u(r^-), s) - f(x(s, u, x_0), u(r^-), s)}{\alpha} ds, \end{aligned}$$

for all  $\alpha \in (0, \alpha_0]$ . Let us deal with the three terms above. Since the first above integrand is continuous, it is clear that  $r$  is a Lebesgue point and we get that

$$\begin{aligned} \lim_{\alpha \rightarrow 0} \int_r^{r+\alpha} \frac{f(x(s, u, x_0), u(r^-), s) - f(x(s, u, x_0), u(r^+), s)}{\alpha} ds \\ = f(x(r, u, x_0), u(r^-), r) - f(x(r, u, x_0), u(r^+), r) = w_r^\mu(r). \end{aligned}$$

Secondly, from the continuity of  $w_r^\mu$ , we know that  $w_r^\mu(r + \alpha)$  tends to  $w_r^\mu(r)$  as  $\alpha$  tends to 0. Finally, using the Lipschitz continuity of  $f$  over  $K$ , we get that

$$\begin{aligned} & \left\| \int_r^{r+\alpha} \frac{f(x(s, u_r^\mu(\cdot, \alpha), x_0), u(r^-), s) - f(x(s, u, x_0), u(r^-), s)}{\alpha} ds \right\|_{\mathbb{R}^n} \\ & \leq \frac{L}{\alpha} \int_r^{r+\alpha} \|x(s, u_r^\mu(\cdot, \alpha), x_0) - x(s, u, x_0)\|_{\mathbb{R}^n} ds \leq L \|x(\cdot, u_r^\mu(\cdot, \alpha), x_0) - x(\cdot, u, x_0)\|_{C([0, \tau], \mathbb{R}^n)}. \end{aligned}$$

Since  $x(\cdot, u_r^\mu(\cdot, \alpha), x_0)$  converges uniformly to  $x(\cdot, u, x_0)$  over  $[0, \tau]$  as  $\alpha$  tends to 0, the proof is complete.  $\square$

We conclude this section by a technical lemma on the convergence of the variation vectors. This result is needed in the proof of our main result (see Section A.3.2).

**Lemma A.9.** *Let  $(u, x_0) \in L^\infty(\mathbb{R}_+, \mathbb{R}^m) \times \mathbb{R}^n$ . Let  $R \geq \|u\|_{L^\infty}$  and let  $0 < \tau < \tau(u, x_0)$ . We take  $\varepsilon > 0$  as in Proposition A.4. Let  $(u_k, x_{0,k})_{k \in \mathbb{N}}$  be a sequence of elements in  $\mathcal{N}((u, x_0), (R, \tau), \varepsilon)$  such that  $x_{0,k}$  converges to  $x_0$  and  $u_k(t)$  converges to  $u(t)$  for almost every  $t \in [0, \tau]$ . Let  $0 < r < \tau$  be a switching time of  $u$  and  $r_k$  be a switching time of  $u_k$  for all  $k \in \mathbb{N}$ . Let us assume that  $r_k$  converges to  $r$  and that  $u_k(r_k^-)$  and  $u_k(r_k^+)$  converge respectively to  $u(r^-)$  and  $u(r^+)$ . Finally let  $\mu \in \{-1, 1\}$  and let  $w_{r_k}^{\mu, k}$  be the unique solution (that is global) to the linear homogeneous Cauchy problem given by*

$$\begin{cases} \dot{w}(t) = \partial_1 f(x(t, u_k, x_{0,k}), u_k(t), t) \times w(t), & \text{a.e. } t \in [r_k, \tau], \\ w(r_k) = \mu \left( f(x(r_k, u_k, x_{0,k}), u_k(r_k^-), r_k) - f(x(r_k, u_k, x_{0,k}), u_k(r_k^+), r_k) \right), \end{cases}$$

for all  $k \in \mathbb{N}$ . Then, for any  $0 < \lambda \leq \tau - r$  fixed, the sequence  $(w_{r_k}^{\mu, k})_{k \in \mathbb{N}}$  uniformly converges to  $w_r^\mu$  over  $[r + \lambda, \tau]$ , where the variation vector  $w_r^\mu$  is defined as in Proposition A.7.

*Proof.* First of all, for all  $k \in \mathbb{N}$ , note that the variation vector  $w_{r_k}^{\mu, k}$  is global (in the sense that it is defined over the whole interval  $[r_k, \tau]$ ) since the corresponding Cauchy problem is linear. In this proof we denote by  $K := K((u, x_0), (R, \tau))$  and by  $L := L((u, x_0), (R, \tau))$  (see the beginning of Section A.2.1 for these two notations). From Proposition A.4, it is clear that  $\|\partial_1 f(x(t, u_k, x_{0,k}), u_k(t), t)\|_{\mathbb{R}^n \times \mathbb{R}^n} \leq L$  for almost every  $t \in [0, \tau]$  and all  $k \in \mathbb{N}$ . From the integral representation of  $w_{r_k}^{\mu, k}$ , it holds that

$$w_{r_k}^{\mu, k}(t) = w_{r_k}^{\mu, k}(r_k) + \int_{r_k}^t \partial_1 f(x(s, u_k, x_{0,k}), u_k(s), s) \times w_{r_k}^{\mu, k}(s) ds,$$

for all  $t \in [r_k, \tau]$  and all  $k \in \mathbb{N}$ . We deduce that

$$\|w_{r_k}^{\mu, k}(t)\|_{\mathbb{R}^n} \leq \|w_{r_k}^{\mu, k}(r_k)\|_{\mathbb{R}^n} + L \int_{r_k}^t \|w_{r_k}^{\mu, k}(s)\|_{\mathbb{R}^n} ds,$$

and, from the Gronwall lemma, that  $\|w_{r_k}^{\mu, k}(t)\|_{\mathbb{R}^n} \leq \|w_{r_k}^{\mu, k}(r_k)\|_{\mathbb{R}^n} e^{L\tau}$  for all  $t \in [r_k, \tau]$  and all  $k \in \mathbb{N}$ . From Proposition A.4, we know that  $x(r_k, u_k, x_{0,k})$  converges to  $x(r, u, x_0)$  and, from the continuity of  $f$  and the hypotheses, it is clear that  $w_{r_k}^{\mu, k}(r_k)$  tends to  $w_r^\mu(r)$ . We deduce that there exists a constant  $C \geq 0$  such that  $\|w_{r_k}^{\mu, k}(t)\|_{\mathbb{R}^n} \leq C$  for all  $t \in [r_k, \tau]$  and all  $k \in \mathbb{N}$ . Now we define  $\bar{r}_k := \max(r_k, r)$  for all  $k \in \mathbb{N}$ . Note that  $\bar{r}_k$  tends to  $r$ . From the integral representations of  $w_{r_k}^{\mu, k}$  and  $w_r^\mu$ , it holds that

$$\begin{aligned} w_{r_k}^{\mu, k}(t) - w_r^\mu(t) &= w_{r_k}^{\mu, k}(\bar{r}_k) - w_r^\mu(\bar{r}_k) + \int_{\bar{r}_k}^t \partial_1 f(x(s, u_k, x_{0,k}), u_k(s), s) \times (w_{r_k}^{\mu, k}(s) - w_r^\mu(s)) ds \\ &\quad + \int_{\bar{r}_k}^t (\partial_1 f(x(s, u_k, x_{0,k}), u_k(s), s) - \partial_1 f(x(s, u, x_0), u(s), s)) \times w_r^\mu(s) ds, \end{aligned}$$

for all  $t \in [\bar{r}_k, \tau]$  and all  $k \in \mathbb{N}$ . From the Triangle inequality, it holds that

$$\|w_{r_k}^{\mu,k}(t) - w_r^\mu(t)\|_{\mathbb{R}^n} \leq \|w_{r_k}^{\mu,k}(\bar{r}_k) - w_r^\mu(\bar{r}_k)\|_{\mathbb{R}^n} + \Gamma_k + L \int_{\bar{r}_k}^t \|w_{r_k}^{\mu,k}(s) - w_r^\mu(s)\|_{\mathbb{R}^n} ds,$$

for all  $t \in [\bar{r}_k, \tau]$  and all  $k \in \mathbb{N}$ , where the term  $\Gamma_k$  is defined to be:

$$\Gamma_k := \int_{\bar{r}_k}^{\tau} \|\partial_1 f(x(s, u_k, x_{0,k}), u_k(s), s) - \partial_1 f(x(s, u, x_0), u(s), s)\|_{\mathbb{R}^n \times n} \|w_r^\mu(s)\|_{\mathbb{R}^n} ds.$$

From the Gronwall lemma, we obtain

$$\|w_{r_k}^{\mu,k}(t) - w_r^\mu(t)\|_{\mathbb{R}^n} \leq (\|w_{r_k}^{\mu,k}(\bar{r}_k) - w_r^\mu(\bar{r}_k)\|_{\mathbb{R}^n} + \Gamma_k) e^{L\tau},$$

for all  $t \in [\bar{r}_k, \tau]$  and all  $k \in \mathbb{N}$ . Since we only want to prove the uniform convergence of  $w_{r_k}^{\mu,k}$  to  $w_r^\mu$  on  $[r + \lambda, \tau]$  (and since  $\bar{r}_k$  converges to  $r$ ) and since the estimate on the right-hand side is independent of  $t$ , we only need to prove that  $\|w_{r_k}^{\mu,k}(\bar{r}_k) - w_r^\mu(\bar{r}_k)\|_{\mathbb{R}^n}$  and  $\Gamma_k$  converge to 0 as  $k$  tends to  $+\infty$ . The convergence of  $\Gamma_k$  can be obtained with the Lebesgue dominated convergence theorem. Now let us prove that  $\|w_{r_k}^{\mu,k}(\bar{r}_k) - w_r^\mu(\bar{r}_k)\|_{\mathbb{R}^n}$  tends to 0 as  $k$  tends to  $+\infty$ . It holds that

$$\begin{aligned} & \|w_{r_k}^{\mu,k}(\bar{r}_k) - w_r^\mu(\bar{r}_k)\|_{\mathbb{R}^n} \\ & \leq \|w_{r_k}^{\mu,k}(\bar{r}_k) - w_{r_k}^{\mu,k}(r_k)\|_{\mathbb{R}^n} + \|w_{r_k}^{\mu,k}(r_k) - w_r^\mu(r_k)\|_{\mathbb{R}^n} + \|w_r^\mu(r_k) - w_r^\mu(\bar{r}_k)\|_{\mathbb{R}^n}, \end{aligned}$$

for all  $k \in \mathbb{N}$ . Let us deal with the three terms above. Firstly, from the integral representation of  $w_{r_k}^{\mu,k}$ , it holds that

$$\begin{aligned} \|w_{r_k}^{\mu,k}(\bar{r}_k) - w_{r_k}^{\mu,k}(r_k)\|_{\mathbb{R}^n} & \leq \int_{r_k}^{\bar{r}_k} \|\partial_1 f(x(s, u_k, x_{0,k}), u_k(s), s)\|_{\mathbb{R}^n \times n} \|w_{r_k}^{\mu,k}(s)\|_{\mathbb{R}^n} ds \\ & \leq LC(\bar{r}_k - r_k), \end{aligned}$$

for all  $k \in \mathbb{N}$ . Secondly we have already mentioned that  $w_{r_k}^{\mu,k}(r_k)$  tends to  $w_r^\mu(r)$  as  $k$  tends to  $+\infty$ . Thirdly we use the continuity of  $w_r^\mu$  to conclude the proof.  $\square$

### A.3 Application of the Ekeland variational principle in the case $L = 0$

From the sensitivity analysis of the state equation given in Section A.2, we are now in a position to give a proof of Theorem 2.1 based on the following simplified version of the Ekeland variational principle (see [17, Theorem 1.1 p.324]).

**Proposition A.8** (Ekeland variational principle). *Let  $(E, d_E)$  be a complete metric set. Let  $\mathcal{J} : E \rightarrow \mathbb{R}^+$  be a continuous nonnegative map. Let  $\varepsilon > 0$  and  $\lambda^* \in E$  such that  $\mathcal{J}(\lambda^*) \leq \varepsilon$ . Then there exists  $\lambda_\varepsilon \in E$  such that  $d_E(\lambda_\varepsilon, \lambda^*) \leq \sqrt{\varepsilon}$ , and  $-\sqrt{\varepsilon} d_E(\lambda, \lambda_\varepsilon) \leq \mathcal{J}(\lambda) - \mathcal{J}(\lambda_\varepsilon)$  for all  $\lambda \in E$ .*

Without loss of generality (see details at the beginning of Appendix A), we will assume that  $L = 0$  in Problem (OSCP). We will also assume that the final time and the  $N$ -partition are free in Problem (OSCP) (the two simpler cases where only the final time is fixed, and where both of them are fixed can both be treated in very similar ways).

Let  $(T, \mathbb{T}, x, u)$  be a solution to Problem (OSCP). In the sequel we will consider that  $u \in L^\infty(\mathbb{R}_+, \mathbb{R}^m)$  by considering the extension

$$\begin{cases} u & \text{over } [0, T), \\ u_{N-1} & \text{over } [T, +\infty). \end{cases}$$

In particular, using the notations of Section A.2, note that  $x = x(\cdot, u, x(0))$  and that  $\tau(u, x(0)) > T$ . In the rest of the proof we fix some  $\tau_0, \tau$  such that

$$\tau_0 := T - \frac{T - t_{N-1}}{3} \quad \text{and} \quad T < \tau < \min \left( T + \frac{T - t_{N-1}}{3}, \tau(u, x(0)) \right).$$

In particular it holds that  $t_{N-1} < \tau_0 < T < \tau < \tau(u, x(0))$ . Replacing  $t_N = T$  by  $t_N = \tau$ , it holds that  $\mathbb{T} \in \mathcal{P}_N^\tau$  and, with the above extension of  $u$ , it holds that  $u \in \text{PC}_N^{\mathbb{T}}([0, \tau], \mathbb{R}^m)$ . We conclude by noting that, with the new value of  $\|\mathbb{T}\|$ , it holds that  $t_{N-1} + \frac{\|\mathbb{T}\|}{4} < \tau_0$ .

### A.3.1 Fix $R \in \mathbb{N}$ such that $R \geq \|u\|_{L^\infty}$

In this section we fix  $R \in \mathbb{N}$  such that  $R \geq \|u\|_{L^\infty}$  and we denote by

$$\mathcal{N}_\varepsilon^R := \left\{ (u', x'_0) \in \mathcal{N}((u, x(0)), (R, \tau), \varepsilon) \mid u' \in \text{PC}_{N, (u, \mathbb{T})}([0, \tau], \mathbb{R}^m) \right. \\ \left. \text{with } u'(t) \in \Omega \text{ for a.e. } t \in [0, \tau] \right\},$$

where  $\varepsilon > 0$  is given in Proposition A.4. We endow the set  $\mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  with the  $L^1([0, \tau], \mathbb{R}^m) \times \mathbb{R}^n \times \mathbb{R}$ -distance. Endowed with this distance, it can be seen from Proposition A.1, from the closedness assumption on  $\Omega$  and from the partial converse of the Lebesgue dominated convergence theorem that  $\mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  is a complete metric set.

Let us consider a sequence  $(\varepsilon_k)_{k \in \mathbb{N}}$  converging to zero such that  $0 < \sqrt{\varepsilon_k} < \varepsilon$  for all  $k \in \mathbb{N}$ . Then we define the penalized functional  $\mathcal{J}_k^R : \mathcal{N}_\varepsilon^R \times [\tau_0, \tau] \rightarrow \mathbb{R}_+$  by

$$\mathcal{J}_k^R(u', x'_0, T') \\ := \sqrt{\left( \varphi(x'_0, x(T'), u', x'_0, T') - \varphi(x(0), x(T), T) + \varepsilon_k \right)^2 + d_S^2(g(x'_0, x(T'), u', x'_0, T'))},$$

for all  $(u', x'_0, T') \in \mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  and all  $k \in \mathbb{N}$ .

Since  $\varphi, g$  and  $d_S^2$  are continuous and from Proposition A.4, it follows that  $\mathcal{J}_k^R$  is a continuous nonnegative map over  $\mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  for all  $k \in \mathbb{N}$ . Furthermore it is clear that  $\mathcal{J}_k^R(u, x(0), T) = \varepsilon_k$  for all  $k \in \mathbb{N}$ . Therefore, from the Ekeland variational principle (see Proposition A.8), we conclude that there exists a sequence  $(u_k, x_{0,k}, T_k)_{k \in \mathbb{N}} \subset \mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  such that

$$d_{L^1([0, \tau], \mathbb{R}^m) \times \mathbb{R}^n \times \mathbb{R}}((u_k, x_{0,k}, T_k), (u, x(0), T)) \leq \sqrt{\varepsilon_k}, \quad (6)$$

and

$$-\sqrt{\varepsilon_k} d_{L^1([0, \tau], \mathbb{R}^m) \times \mathbb{R}^n \times \mathbb{R}}((u', x'_0, T'), (u_k, x_{0,k}, T_k)) \leq \mathcal{J}_k^R(u', x'_0, T') - \mathcal{J}_k^R(u_k, x_{0,k}, T_k), \quad (7)$$

for all  $(u', x'_0, T') \in \mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  and all  $k \in \mathbb{N}$ .

By contradiction let us assume that there exists  $(u', x'_0, T') \in \mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  such that  $\mathcal{J}_k^R(u', x'_0, T') = 0$ . In particular we have  $0 < T' \leq \tau$ . Let us denote by  $x' = x(\cdot, u', x'_0) \in \text{AC}([0, T'], \mathbb{R}^n)$ . Since  $u' \in \text{PC}_{N, (u, \mathbb{T})}([0, \tau], \mathbb{R}^m)$ , there exists  $\mathbb{T}' = (t'_i)_{i=0, \dots, N} \in \mathcal{P}_{N, (u, \mathbb{T})}^\tau$  such that  $u' \in \text{PC}_N^{\mathbb{T}'}([0, \tau], \mathbb{R}^m)$ . Since  $\mathbb{T}' = (t'_i)_{i=0, \dots, N} \in \mathcal{P}_{N, (u, \mathbb{T})}^\tau$ , we know that  $t'_{N-1} \leq t_{N-1} + \frac{\|\mathbb{T}\|}{4} < \tau_0 \leq T' \leq \tau$ . Then, replacing  $t'_N = \tau$  by  $t'_N = T'$ , we get that  $\mathbb{T}' \in \mathcal{P}_N^{\mathbb{T}'}([0, T'], \mathbb{R}^m)$ . Moreover it holds that  $\dot{x}'(t) = f(x'(t), u'(t), t)$  and  $u'(t) \in \Omega$  for almost every  $t \in [0, T']$ . Since  $\mathcal{J}_k^R(u', x'_0, T') = 0$ , we deduce moreover that  $g(x'(0), x'(T'), T') \in S$ . Thus the quadruple  $(T', \mathbb{T}', x', u')$  satisfies all

constraints of Problem (OSCP) and thus  $\varphi(x'(0), x'(T'), T') \geq \varphi(x(0), x(T), T)$  from optimality of the quadruple  $(T, \mathbb{T}, x, u)$ . This raises a contradiction with the equality  $\mathcal{J}_k^R(u', x'_0, T') = 0$ . We conclude that  $\mathcal{J}_k^R(u', x'_0, T') > 0$  for all  $(u', x'_0, T') \in \mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$ .

From the above paragraph we can correctly define the couple  $(\psi_k^{0R}, \psi_k^R) \in \mathbb{R} \times \mathbb{R}^j$  as

$$\psi_k^{0R} := \frac{-1}{\mathcal{J}_k^R(u_k, x_{0,k}, T_k)} \left( \varphi(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k) - \varphi(x(0), x(T), T) + \varepsilon_k \right)^+$$

and

$$\psi_k^R := \frac{-1}{\mathcal{J}_k^R(u_k, x_{0,k}, T_k)} \left( g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k) - \mathbb{P}_S(g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k)) \right),$$

for all  $k \in \mathbb{N}$ . Note that  $\psi_k^{0R} \in \mathbb{R}_-$  and  $-\psi_k^R \in \mathbb{N}_S[\mathbb{P}_S(g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k))]$  from Lemma A.1 for all  $k \in \mathbb{N}$ .

Since  $(u_k, x_{0,k}, T_k) \in \mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$ , we know that  $u_k \in \text{PC}_{N, (u, \mathbb{T})}([0, \tau], \mathbb{R}^m)$  for all  $k \in \mathbb{N}$ . Let us denote by  $\mathbb{T}_k = (t_{i,k})_{i=0, \dots, N} \in \mathcal{P}_{N, (u, \mathbb{T})}^\tau$  a partition associated to  $u_k$  for all  $k \in \mathbb{N}$ . Moreover, from Inequality (6), the sequence  $(u_k)_{k \in \mathbb{N}}$  converges to  $u$  in  $L^1([0, \tau], \mathbb{R}^m)$ . Thus we can extract from Proposition A.2 a subsequence (which we do not relabel) such that  $u_k(t)$  converges to  $u(t)$  for almost every  $t \in [0, \tau]$ ,  $t_{i,k}$  converges to  $t_i$  for all  $i = 0, \dots, N$  and  $u_{i,k}$  converges to  $u_i$  for all  $i = 0, \dots, N-1$ . From Inequality (6), we know that  $x_{0,k}$  and  $T_k$  converge respectively to  $x(0)$  and  $T$ . From Proposition A.4, we deduce that  $x(T_k, u_k, x_{0,k})$  converges to  $x(T)$  and thus  $g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k)$  converges to  $g(x(0), x(T), T) \in S$ . Finally, from the definition of  $\mathcal{J}_k^R$ , it is clear that  $|\psi_k^{0R}|^2 + \|\psi_k^R\|_{\mathbb{R}^j}^2 = 1$  for all  $k \in \mathbb{N}$ . By a compactness argument, we can extract subsequences (which we do not relabel) such that  $\psi_k^{0R}$  converges to some  $\psi^{0R} \in \mathbb{R}_-$  and  $\psi_k^R$  converges to some  $\psi^R \in \mathbb{R}^j$  which satisfies  $-\psi^R \in \mathbb{N}_S[g(x(0), x(T), T)]$  from Lemma A.2. Note that  $|\psi^{0R}|^2 + \|\psi^R\|_{\mathbb{R}^j}^2 = 1$ .

### A.3.2 Crucial inequalities depending on $R$ fixed in the previous section

In this section we will use Inequality (7) along with the perturbations defined in Section A.2 to obtain four crucial inequalities (depending on  $R$  fixed in the previous section). The perturbations will be considered on  $u_k$ ,  $x_{0,k}$ ,  $t_{i,k}$ , but also on  $T_k$ .

**Lemma A.10.** *Let  $v \in \text{PC}_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$  taking values in  $\Omega \cap \overline{B}_{\mathbb{R}^m}(0_{\mathbb{R}^m}, R)$ . Then the inequality*

$$\left\langle \psi^{0R} \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi^R, w_v(T) \right\rangle_{\mathbb{R}^n} \leq 0, \quad (8)$$

where  $w_v$  is defined in Proposition A.5, holds true.

*Proof.* The proof is divided in three steps.

*First step:* For all  $k \in \mathbb{N}$ , let us define

$$v_k(t) := v_i \quad \text{if } t \in [t_{i,k}, t_{i+1,k}) \text{ for some } i \in \{0, \dots, N-1\},$$

for all  $t \in [0, \tau]$ . Then  $v_k \in \text{PC}_N^\mathbb{T}([0, \tau], \mathbb{R}^m)$  for all  $k \in \mathbb{N}$  and, since  $t_{i,k}$  converges to  $t_i$  for all  $i = 0, \dots, N$ , it is clear that the sequence  $(v_k)_{k \in \mathbb{N}}$  converges to  $v$  in  $L^1([0, \tau], \mathbb{R}^m)$ . It is also true that  $v_k$  takes its values in  $\Omega \cap \overline{B}_{\mathbb{R}^m}(0_{\mathbb{R}^m}, R)$  for all  $k \in \mathbb{N}$ .

*Second step:* Let us fix  $k \in \mathbb{N}$ . We define as in Proposition A.5 the convex perturbation

$$u_{k, v_k}(\cdot, \alpha) := \begin{cases} u_k + \alpha(v_k - u_k) & \text{over } [0, \tau), \\ u_k = u & \text{over } [\tau, +\infty), \end{cases}$$

for all  $0 \leq \alpha \leq 1$ . First of all, note that  $u_{k,v_k}(\cdot, \alpha) \in \text{PC}_N^{\mathbb{T}^k}([0, \tau], \mathbb{R}^m) \subset \text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$  and, since  $\Omega$  is convex, that  $u_{k,v_k}(\cdot, \alpha)$  takes its values in  $\Omega$  for all  $0 \leq \alpha \leq 1$ . Moreover, it holds that  $\|u_{k,v_k}(\cdot, \alpha)\|_{L^\infty} \leq R$  and

$$\begin{aligned} \|u_{k,v_k}(\cdot, \alpha) - u\|_{L^1([0, \tau], \mathbb{R}^m)} &\leq \|u_{k,v_k}(\cdot, \alpha) - u_k\|_{L^1([0, \tau], \mathbb{R}^m)} + \|u_k - u\|_{L^1([0, \tau], \mathbb{R}^m)} \\ &\leq \alpha \|v_k - u_k\|_{L^1([0, \tau], \mathbb{R}^m)} + \sqrt{\varepsilon_k}. \end{aligned}$$

Since  $\sqrt{\varepsilon_k} < \varepsilon$ , it follows that there exists  $0 < \alpha_0 \leq 1$  small enough such that  $(u_{k,v_k}(\cdot, \alpha), x_{0,k}) \in \mathcal{N}_\varepsilon^R$  for all  $\alpha \in [0, \alpha_0]$ . From Inequality (7) we obtain

$$-\sqrt{\varepsilon_k} \|u_{k,v_k}(\cdot, \alpha) - u_k\|_{L^1([0, \tau], \mathbb{R}^m)} \leq \mathcal{J}_k^R(u_{k,v_k}(\cdot, \alpha), x_{0,k}, T_k) - \mathcal{J}_k^R(u_k, x_{0,k}, T_k),$$

and thus

$$\begin{aligned} &-\sqrt{\varepsilon_k} \|v_k - u_k\|_{L^1([0, \tau], \mathbb{R}^m)} \\ &\leq \frac{1}{\mathcal{J}_k^R(u_{k,v_k}(\cdot, \alpha), x_{0,k}, T_k) + \mathcal{J}_k^R(u_k, x_{0,k}, T_k)} \times \frac{\mathcal{J}_k^R(u_{k,v_k}(\cdot, \alpha), x_{0,k}, T_k)^2 - \mathcal{J}_k^R(u_k, x_{0,k}, T_k)^2}{\alpha}, \end{aligned}$$

for all  $\alpha \in (0, \alpha_0]$ . Taking the limit as  $\alpha$  tends to 0 and using the definitions of  $\psi_k^{0R}$  and  $\psi_k^R$ , we obtain from Proposition A.5 that

$$\begin{aligned} &\left\langle \psi_k^{0R} \partial_2 \varphi(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k) + \partial_2 g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k)^\top \times \psi_k^R, w_{v_k}^k(T_k) \right\rangle_{\mathbb{R}^n} \\ &\leq \sqrt{\varepsilon_k} \|v_k - u_k\|_{L^1([0, \tau], \mathbb{R}^m)}. \end{aligned}$$

where  $w_{v_k}^k$  is defined in Lemma A.7.

*Third step:* We take the limit of the above inequality as  $k$  tends to  $+\infty$ . Since  $\varphi$  and  $g$  are both of class  $C^1$  and from the uniform convergence of  $(w_{v_k}^k)_{k \in \mathbb{N}}$  to  $w_v$  over  $[0, \tau]$  (see Lemma A.7), it holds that

$$\left\langle \psi^{0R} \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi^R, w_v(T) \right\rangle_{\mathbb{R}^n} \leq 0.$$

The proof is complete.  $\square$

**Lemma A.11.** *Let  $y \in \mathbb{R}^n$  be fixed. Then the inequality*

$$\begin{aligned} &\left\langle \psi^{0R} \partial_1 \varphi(x(0), x(T), T) + \partial_1 g(x(0), x(T), T)^\top \times \psi^R, y \right\rangle_{\mathbb{R}^n} \\ &+ \left\langle \psi^{0R} \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi^R, w_y(T) \right\rangle_{\mathbb{R}^n} \leq 0, \end{aligned} \quad (9)$$

where  $w_y$  is defined in Proposition A.6, holds true.

*Proof.* The proof is standard and left to the reader. For similar statements with detailed proofs, we refer to [11, Lemma 17 p.3807-3808] or [13, Lemma 4.20 p.87].  $\square$

**Lemma A.12.** *Let  $i \in \{1, \dots, N-1\}$  such that  $u_{i-1} \neq u_i$  and let  $\mu \in \{-1, 1\}$ . Then the inequality*

$$\left\langle \psi^{0R} \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi^R, w_{t_i}^\mu(T) \right\rangle_{\mathbb{R}^n} \leq 0, \quad (10)$$

where  $w_{t_i}^\mu$  is defined in Proposition A.7, holds true.

*Proof.* The proof is divided in two steps.

*First step:* Since  $t_{i,k}$  converges to  $t_i$  and since  $t_i - \frac{\|\mathbb{T}\|}{4} \leq t_{i,k} \leq t_i + \frac{\|\mathbb{T}\|}{4}$ , we fix  $k \in \mathbb{N}$  sufficiently large in order to guarantee that  $t_i - \frac{\|\mathbb{T}\|}{8} \leq t_{i,k} \leq t_i + \frac{\|\mathbb{T}\|}{8}$ . Since  $u_k \in \text{PC}_N^{\mathbb{T}_k}([0, \tau], \mathbb{R}^m)$ , the point  $t_{i,k}$  is a switching time of  $u_k$  with  $\eta_{t_{i,k}} = \min(t_{i,k} - t_{i-1,k}, t_{i+1,k} - t_{i,k}) > 0$ . We define the perturbation  $u_{k,t_{i,k}}^\mu(\cdot, \alpha)$  as

$$u_{k,t_{i,k}}^\mu(\cdot, \alpha) := \begin{cases} u_k(t_{i,k}^-) = u_{i-1,k} & \text{over } [t_{i,k} - \eta_{t_{i,k}}, t_{i,k} + \mu\alpha), \\ u_k(t_{i,k}^+) = u_{i,k} & \text{over } [t_{i,k} + \mu\alpha, t_{i,k} + \eta_{t_{i,k}}), \\ u_k & \text{otherwise,} \end{cases}$$

for all  $0 \leq \alpha \leq \frac{\eta_{t_{i,k}}}{2}$ . Considering  $\mathbb{T}_k^{i,\alpha}$  the  $N$ -partition given by

$$0 = t_{0,k} < t_{1,k} < \dots < t_{i-1,k} < t_{i,k} + \mu\alpha < t_{i+1,k} < \dots < t_{N-1,k} < t_{N,k} = \tau,$$

it is clear that  $u_{k,t_{i,k}}^\mu(\cdot, \alpha) \in \text{PC}_N^{\mathbb{T}_k^{i,\alpha}}([0, \tau], \mathbb{R}^m)$  for all  $0 \leq \alpha \leq \frac{\eta_{t_{i,k}}}{2}$ . Since  $t_i - \frac{\|\mathbb{T}\|}{8} \leq t_{i,k} \leq t_i + \frac{\|\mathbb{T}\|}{8}$ , then  $t_i - \frac{\|\mathbb{T}\|}{4} \leq t_{i,k} + \mu\alpha \leq t_i + \frac{\|\mathbb{T}\|}{4}$  and thus  $\mathbb{T}_k^{i,\alpha} \in \mathcal{P}_{N,(u,\mathbb{T})}^\tau$  and  $u_{k,t_{i,k}}^\mu(\cdot, \alpha) \in \text{PC}_{N,(u,\mathbb{T})}([0, \tau], \mathbb{R}^m)$  for small enough  $0 \leq \alpha \leq \frac{\eta_{t_{i,k}}}{2}$ . Note that  $u_{k,t_{i,k}}^\mu(\cdot, \alpha)$  takes its values in  $\Omega$  for all  $0 \leq \alpha \leq \frac{\eta_{t_{i,k}}}{2}$ . It holds that  $\|u_{k,t_i}^\mu(\cdot, \alpha)\|_{L^\infty} \leq R$  and

$$\|u_{k,t_i}^\mu(\cdot, \alpha) - u\|_{L^1([0,\tau], \mathbb{R}^m)} \leq \|u_{k,t_i}^\mu(\cdot, \alpha) - u_k\|_{L^1([0,\tau], \mathbb{R}^m)} + \|u_k - u\|_{L^1([0,\tau], \mathbb{R}^m)} \leq 2R\alpha + \sqrt{\varepsilon_k},$$

for all  $0 \leq \alpha \leq \frac{\eta_{t_{i,k}}}{2}$ . Since  $\sqrt{\varepsilon_k} < \varepsilon$ , we conclude that there exists  $0 < \alpha_0 \leq \frac{\eta_{t_{i,k}}}{2}$  small enough such that  $(u_{k,t_i}^\mu(\cdot, \alpha), x_{0,k}) \in \mathcal{N}_\varepsilon^R$  for all  $0 \leq \alpha \leq \alpha_0$ . From Inequality (7) we obtain

$$-\sqrt{\varepsilon_k} \|u_{k,t_i}^\mu(\cdot, \alpha) - u_k\|_{L^1([0,\tau], \mathbb{R}^m)} \leq \mathcal{J}_k^R(u_{k,t_i}^\mu(\cdot, \alpha), x_{0,k}, T_k) - J_k^R(u_k, x_{0,k}, T_k),$$

and thus

$$\begin{aligned} & -2R\sqrt{\varepsilon_k} \\ & \leq \frac{1}{\mathcal{J}_k^R(u_{k,t_i}^\mu(\cdot, \alpha), x_{0,k}, T_k) + J_k^R(u_k, x_{0,k}, T_k)} \times \frac{\mathcal{J}_k^R(u_{k,t_i}^\mu(\cdot, \alpha), x_{0,k}, T_k)^2 - J_k^R(u_k, x_{0,k}, T_k)^2}{\alpha}, \end{aligned}$$

for all  $\alpha \in (0, \alpha_0]$ . Taking the limit as  $\alpha$  tends to 0 and using the definitions of  $\psi_k^{0R}$  and  $\psi_k^R$ , we obtain from Proposition A.7 that

$$\left\langle \psi_k^{0R} \partial_2 \varphi(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k) + \partial_2 g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k)^\top \times \psi_k^R, w_{t_{i,k}}^{\mu,k}(T_k) \right\rangle_{\mathbb{R}^n} \leq 2R\sqrt{\varepsilon_k},$$

where  $w_{t_{i,k}}^{\mu,k}$  is defined in Lemma A.9.

*Second step:* We take the limit of the above inequality as  $k$  tends to  $+\infty$ . Since  $\varphi$  and  $g$  are of class  $C^1$  and, since  $t_i < \tau_0$ , from the uniform convergence of  $(w_{t_{i,k}}^{\mu,k})_{k \in \mathbb{N}}$  to  $w_{t_i}^\mu$  over  $[\tau_0, \tau]$  (see Lemma A.9), it holds that

$$\left\langle \psi^{0R} \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi^R, w_{t_i}^\mu(T) \right\rangle_{\mathbb{R}^n} \leq 0.$$

The proof is complete.  $\square$

**Lemma A.13.** *The equality*

$$\begin{aligned} & \left\langle \psi^{0R} \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi^R, f(x(T), u_{N-1}, T) \right\rangle_{\mathbb{R}^n} \\ & + \psi^{0R} \partial_3 \varphi(x(0), x(T), T) + \partial_3 g(x(0), x(T), T)^\top \times \psi^R = 0, \end{aligned} \quad (11)$$

holds.

*Proof.* The proof is divided in two steps.

*First step:* Let  $\mu \in \{-1, 1\}$ . Since  $(T_k)_{k \in \mathbb{N}}$  converges to  $T \in (\tau_0, \tau)$ , then  $T_k \in (\tau_0, \tau)$  for  $k \in \mathbb{N}$  sufficiently large. Let us fix such an integer  $k \in \mathbb{N}$ . Thus there exists  $\alpha_0 > 0$  small enough such that  $(x_{0,k}, u_k, T_k + \mu\alpha) \in \mathcal{N}_\varepsilon^R \times [\tau_0, \tau]$  for all  $0 \leq \alpha \leq \alpha_0$ . From Inequality (7) we obtain

$$-\sqrt{\varepsilon_k} |T_k + \mu\alpha - T_k| \leq \mathcal{J}_k^R(u_k, x_{0,k}, T_k + \mu\alpha) - \mathcal{J}_k^R(u_k, x_{0,k}, T_k),$$

and thus

$$-\sqrt{\varepsilon_k} \leq \frac{1}{\mathcal{J}_k^R(u_k, x_{0,k}, T_k + \mu\alpha) + \mathcal{J}_k^R(u_k, x_{0,k}, T_k)} \times \frac{\mathcal{J}_k^R(u_k, x_{0,k}, T_k + \mu\alpha)^2 - \mathcal{J}_k^R(u_k, x_{0,k}, T_k)^2}{\alpha},$$

for all  $\alpha \in (0, \alpha_0]$ . Taking the limit as  $\alpha$  tends to 0 and using the definitions of  $\psi_k^{0R}$  and  $\psi_k^R$ , we obtain from the differentiability of  $x(\cdot, u_k, x_{0,k})$  at  $T_k$  (since  $u_k$  is constant over the interval  $[\tau_0, \tau] \subset [t_{N-1,k}, t_{N,k}]$  and since  $T_k \in (\tau_0, \tau)$ ) that

$$\begin{aligned} & \mu \left\langle \psi_k^{0R} \partial_2 \varphi(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k) + \partial_2 g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k)^\top \times \psi_k^R \right. \\ & \quad \left. , f(x(T_k, u_k, x_{0,k}), u_k(T_k), T_k) \right\rangle_{\mathbb{R}^n} \\ & + \mu \psi_k^{0R} \partial_3 \varphi(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k) + \mu \partial_3 g(x_{0,k}, x(T_k, u_k, x_{0,k}), T_k)^\top \times \psi_k^R \leq \sqrt{\varepsilon_k}, \end{aligned}$$

where  $u_k(T_k) = u_{N-1,k}$ .

*Second step:* We take the limit of the above inequality as  $k$  tends to  $+\infty$ . Let us recall that  $u_{N-1,k}$  converges to  $u_{N-1}$ . Furthermore, since  $f$  is continuous, since  $\varphi$  and  $g$  are of class  $C^1$ , and since  $u_k(T_k)$  converges to  $u(T)$  from Proposition A.2, it holds that

$$\begin{aligned} & \mu \left\langle \psi^{0R} \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi^R, f(x(T), u_{N-1}, T) \right\rangle_{\mathbb{R}^n} \\ & + \mu \psi^{0R} \partial_3 \varphi(x(0), x(T), T) + \mu \partial_3 g(x(0), x(T), T)^\top \times \psi^R \leq 0. \end{aligned}$$

Since  $\mu$  can be chosen arbitrarily in  $\{-1, 1\}$ , the proof is complete.  $\square$

### A.3.3 Crucial inequalities letting $R$ go to $+\infty$

In the previous section we have obtained Inequalities (8), (9) and (10) and Equality (11) which are valid for  $R \in \mathbb{N}$  being fixed such that  $R \geq \|u\|_{L^\infty}$ . In particular Inequality (8) is satisfied only for  $v \in \text{PC}_N^{\mathbb{T}}([0, \tau], \mathbb{R}^m)$  taking values in  $\Omega \cap \bar{B}_{\mathbb{R}^m}(0_{\mathbb{R}^m}, R)$ . Our goal in this section is to get rid of the dependence in  $R$ . From the equality  $|\psi^{0R}|^2 + \|\psi^R\|_{\mathbb{R}^j}^2 = 1$  (see the end of Section A.3.1), we can extract subsequences (that we do not relabel) such that  $(\psi^{0R})_{R \in \mathbb{N}}$  converges to some  $\psi^0$  in  $\mathbb{R}$  and  $(\psi^R)_{R \in \mathbb{N}}$  converges to some  $\psi$  in  $\mathbb{R}^j$  when  $R \rightarrow \infty$ . It clearly holds that  $|\psi^0|^2 + \|\psi\|_{\mathbb{R}^j}^2 = 1$  and, since  $\mathbb{R}_-$  and  $\text{N}_S[g(x(0), x(T), T)]$  are closed, that  $\psi^0 \in \mathbb{R}_-$  and  $-\psi \in \text{N}_S[g(x(0), x(T), T)]$ .

Now let us fix  $v \in \text{PC}_N^{\mathbb{T}}([0, \tau], \mathbb{R}^m)$  taking values in  $\Omega$ . Considering  $R \in \mathbb{N}$  large enough in order to get that  $R \geq \|u\|_{L^\infty}$  and  $R \geq \|v\|_{L^\infty}$ , we know from Lemma A.10 that Inequality (8) is satisfied. Taking the limit as  $R$  tends to  $+\infty$  we conclude that

$$\langle \psi^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi, w_v(T) \rangle_{\mathbb{R}^n} \leq 0. \quad (12)$$

Similarly, letting  $R$  go to  $+\infty$  in Inequalities (9) and (10) and in Equality (11), we get that

$$\begin{aligned} & \langle \psi^0 \partial_1 \varphi(x(0), x(T), T) + \partial_1 g(x(0), x(T), T)^\top \times \psi, y \rangle_{\mathbb{R}^n} \\ & + \langle \psi^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi, w_y(T) \rangle_{\mathbb{R}^n} \leq 0, \end{aligned} \quad (13)$$

for any  $y \in \mathbb{R}^n$ , that

$$\langle \psi^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi, w_{t_i}^\mu(T) \rangle_{\mathbb{R}^n} \leq 0, \quad (14)$$

for any  $i \in \{1, \dots, N-1\}$  such that  $u_{i-1} \neq u_i$  and any  $\mu \in \{-1, 1\}$ , and that

$$\begin{aligned} & \langle \psi^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \psi, f(x(T), u_{N-1}, T) \rangle_{\mathbb{R}^n} \\ & + \psi^0 \partial_3 \varphi(x(0), x(T), T) + \partial_3 g(x(0), x(T), T)^\top \times \psi = 0. \end{aligned} \quad (15)$$

#### A.3.4 End of the proof

Now we can end the proof of Theorem 2.1 with the introduction of the adjoint vector  $p$ . Before coming to this point, let us first define  $p^0 := \psi^0$  and  $\Psi := \psi$ . In particular note that  $p^0 \in \mathbb{R}_-$ , that  $\Psi \in \mathbb{R}^j$  is such that  $-\Psi \in \text{NS}[g(x(0), x(T), T)]$  and that  $|p^0|^2 + \|\Psi\|_{\mathbb{R}^j}^2 = 1$ .

We define the adjoint vector  $p \in \text{AC}([0, T], \mathbb{R}^n)$  as the unique solution (that is global) to the backward linear Cauchy problem given by

$$\begin{cases} \dot{p}(t) = -\partial_1 f(x(t), u(t), t)^\top \times p(t), & \text{a.e. } t \in [0, T], \\ p(T) = p^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \Psi. \end{cases}$$

From the Duhamel formula, recall that

$$p(t) = Z(T, t)^\top \times \left( p^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \Psi \right),$$

for all  $t \in [0, T]$ , where  $Z(\cdot, \cdot) : [0, T]^2 \rightarrow \mathbb{R}^{n \times n}$  stands for the state transition matrix associated to the matrix function  $t \mapsto \partial_1 f(x(t), u(t), t)$ .

**Adjoint equation and transversality conditions on the adjoint vector.** From the above definition of the adjoint vector  $p$ , it is clear that the adjoint equation in Theorem 2.1 and the transversality condition  $p(T) = p^0 \partial_2 \varphi(x(0), x(T), T) + \partial_2 g(x(0), x(T), T)^\top \times \Psi$  are satisfied. Moreover, from the Duhamel formula, it holds that  $w_y(T) = Z(T, 0) \times y$  and thus Inequality (13) can be rewritten as

$$\langle p^0 \partial_1 \varphi(x(0), x(T), T) + \partial_1 g(x(0), x(T), T)^\top \times \Psi + p(0), y \rangle_{\mathbb{R}^n} \leq 0,$$

for all  $y \in \mathbb{R}^n$ . Thus we conclude that the transversality condition  $-p(0) = p^0 \partial_1 \varphi(x(0), x(T), T) + \partial_1 g(x(0), x(T), T)^\top \times \Psi$  holds.

**Nonpositive averaged Hamiltonian gradient condition.** Let us fix  $\omega \in \Omega$  and  $i \in \{0, \dots, N-1\}$ . Let us consider  $v \in \text{PC}_N^{\mathbb{T}}([0, T], \mathbb{R}^m)$  defined by

$$v(t) := \begin{cases} \omega & \text{if } t \in [t_i, t_{i+1}), \\ u(t) & \text{otherwise,} \end{cases}$$

for all  $t \in [0, T]$ . From the Duhamel formula given by

$$w_v(T) = \int_0^T Z(T, t) \times \partial_2 f(x(t), u(t), t) \times (v(t) - u(t)) dt,$$

Inequality (12) can be rewritten as

$$\int_0^T \langle \partial_2 f(x(t), u(t), t)^\top \times p(t), v(t) - u(t) \rangle_{\mathbb{R}^m} dt \leq 0,$$

that is

$$\left\langle \int_{t_i}^{t_{i+1}} \partial_2 H(x(t), u_i, p(t), p^0, t) dt, \omega - u_i \right\rangle_{\mathbb{R}^m} \leq 0.$$

**Transversality conditions on the optimal sampling times.** Let us fix some  $i \in \{1, \dots, N-1\}$  and  $\mu \in \{-1, 1\}$ . If  $u_{i-1} = u_i$ , then the transversality condition (3) in Theorem 2.1 is obviously satisfied. Now let us assume that  $u_{i-1} \neq u_i$ . From the Duhamel formula given by

$$w_{t_i}^\mu(T) = \mu Z(T, t_i) \times \left( f(x(t_i), u_{i-1}, t_i) - f(x(t_i), u_i, t_i) \right),$$

Inequality (14) can be rewritten as

$$\mu \langle p(t_i), f(x(t_i), u_{i-1}, t_i) - f(x(t_i), u_i, t_i) \rangle_{\mathbb{R}^n} \leq 0.$$

Since  $\mu$  can be arbitrarily chosen in  $\{-1, 1\}$  and from the definition of the Hamiltonian  $H$ , we get that

$$H(x(t_i), u_{i-1}, p(t_i), p^0, t_i) = H(x(t_i), u_i, p(t_i), p^0, t_i).$$

**Transversality condition on the optimal final time.** Equality (15) can directly be rewritten as

$$-H(x(T), u_{N-1}, p(T), p^0, T) = p^0 \partial_3 \varphi(x(0), x(T), T) + \partial_3 g(x(0), x(T), T)^\top \times \Psi.$$

**Nontriviality of the couple  $(p, p^0)$ .** Let us assume by contradiction that the couple  $(p, p^0)$  is trivial. Then  $p(0) = p(T) = 0_{\mathbb{R}^n}$  and  $p^0 = 0$ . We get from the transversality conditions on the adjoint vector and on the optimal final time that  $Dg(x(0), x(T), T)^\top \times \Psi = 0_{\mathbb{R}^{2n+1}}$ . From the submersion property, we deduce that  $\Psi = 0_{\mathbb{R}^j}$  which raises a contradiction with the equality  $|p^0|^2 + \|\Psi\|_{\mathbb{R}^j}^2 = 1$ .

## References

- [1] J. E. Ackermann. *Sampled-data control systems: analysis and synthesis, robust system design*. Springer-Verlag, 1985.
- [2] K. J. Aström. On the choice of sampling rates in optimal linear systems. *IBM Research: Engineering Studies*, 1963.

- [3] K. J. Aström and B. Wittenmark. *Computer-Controlled Systems*. Prentice Hall, 1997.
- [4] T. Bakir, B. Bonnard, L. Bourdin, and J. Rouot. Pontryagin-type conditions for optimal muscular force response to functional electrical stimulations. *In revision, reference HAL-01854551*, 2019.
- [5] M. Bergounioux and L. Bourdin. Pontryagin maximum principle for general Caputo fractional optimal control problems with Bolza cost and terminal constraints. *Accepted for publication in ESAIM Control Optim. Calc. Var.*, 2019.
- [6] E. Bini and G. M. Buttazzo. The optimal sampling pattern for linear control systems. *IEEE Trans. Automat. Control*, 59(1):78–90, 2014.
- [7] R. S. Biryukov. Generalized  $H_\infty$ -optimal control of linear continuous-discrete plant. *Avtomat. i Telemekh.*, 77(3):33–51, 2016.
- [8] V. G. Boltyanski and A. S. Poznyak. *The robust maximum principle*. Systems & Control: Foundations & Applications. Birkhäuser/Springer, New York, 2012. Theory and applications.
- [9] V. G. Boltyanskii. *Optimal control of discrete systems*. John Wiley & Sons, New York-Toronto, Ont., 1978.
- [10] L. Bourdin. Note on Pontryagin maximum principle with running state constraints and smooth dynamics - Proof based on the Ekeland variational principle. *Research notes*, hal-01302222, 2016.
- [11] L. Bourdin and E. Trélat. Pontryagin maximum principle for finite dimensional nonlinear optimal control problems on time scales. *SIAM J. Control Optim.*, 51(5):3781–3813, 2013.
- [12] L. Bourdin and E. Trélat. Pontryagin maximum principle for optimal sampled-data control problems. In *Proceedings of the IFAC workshop CAO*, 2015.
- [13] L. Bourdin and E. Trélat. Optimal sampled-data control, and generalizations on time scales. *Mathematical Control & Related Fields*, 6(1):53–94, 2016.
- [14] L. Bourdin and E. Trélat. Linear-quadratic optimal sampled-data control problems: Convergence result and Riccati theory. *Automatica*, 79:273–281, 2017.
- [15] T. Chen and B. Francis. *Optimal sampled-data control systems*. Springer-Verlag London, Ltd., London, 1996.
- [16] A. V. Dmitruk and A. M. Kaganovich. Maximum principle for optimal control problems with intermediate constraints. *Comput. Math. Model.*, 22(2):180–215, 2011. Translation of Nelineinaya Din. Upr. No. 6 (2008), 101–136.
- [17] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.
- [18] M. S. Fadali and A. Visioli. *Digital control engineering: Analysis and design*. Elsevier, 2013.
- [19] H. O. Fattorini. *Infinite-dimensional optimization and control theory*, volume 62 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999.
- [20] K. A. Grasse and H. J. Sussmann. Global controllability by nice controls. In *Nonlinear controllability and optimal control*, volume 133 of *Monogr. Textbooks Pure Appl. Math.*, pages 33–79. Dekker, New York, 1990.

- [21] L. Grüne and J. Pannek. *Nonlinear model predictive control*. Communications and Control Engineering Series. Springer, Cham, 2017. Theory and algorithms, Second edition [of MR3155076].
- [22] T. Haberkorn and E. Trélat. Convergence results for smooth regularizations of hybrid nonlinear optimal control problems. *SIAM J. Control Optim.*, 49(4):1498–1522, 2011.
- [23] H. Halkin. A maximum principle of the Pontryagin type for systems described by nonlinear difference equations. *SIAM J. Control*, 4(1):90–111, 1966.
- [24] J. M. Holtzman and H. Halkin. Discretional convexity and the maximum principle for discrete systems. *SIAM J. Control*, 4:263–275, 1966.
- [25] I. D. Landau and G. Zito. *Digital Control Systems: Design, Identification and Implementation*. Springer, 2006.
- [26] A. H. Levis and R. A. Schlueter. On the behaviour of optimal linear sampled-data regulators. *International Journal of Control*, 13(2):343–361, 1971.
- [27] M. Margaliot. Stability analysis of switched systems using variational principles: an introduction. *Automatica J. IFAC*, 42(12):2059–2077, 2006.
- [28] S. M. Melzer and B. C. Kuo. Sampling period sensitivity of the optimal sampled data linear regulator. *Automatica J. IFAC*, 7:367–370, 1971.
- [29] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. John Wiley & Sons, Inc., 1962.
- [30] M. S. Santina and A. R. Stubberud. Basics of sampling and quantization. In *Handbook of networked and embedded control systems*, Control Eng., pages 45–69. Birkhauser Boston, Boston, MA, 2005.
- [31] R. A. Schlueter. The optimal linear regulator with constrained sampling times. *IEEE Trans. Automatic Control*, AC-18(no.5):515–518, 1973.
- [32] R. A. Schlueter and A. H. Levis. The optimal linear regulator with state-dependent sampling. *IEEE Trans. Automatic Control*, AC-18(5):512–515, 1973.
- [33] S. P. Sethi and G. L. Thompson. *Optimal control theory*. Kluwer Academic Publishers, Boston, MA, second edition, 2000. Applications to management science and economics.
- [34] H. J. Sussmann. A maximum principle for hybrid optimal control problems. In *Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No. 99CH36304)*, volume 1, pages 425–430. Ieee, 1999.
- [35] R. Vinter. *Optimal control*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2010. Paperback reprint of the 2000 edition.
- [36] R. A. Volz and L. F. Kazda. Design of a digital controller for a tracking telescope. *IEEE Transactions on Automatic Control*, AC-12(4):359–367, August 1966.