



Optimal sampled-data controls with running inequality state constraints: Pontryagin maximum principle and bouncing trajectory phenomenon

Loïc Bourdin, Gaurav Dhar

► To cite this version:

Loïc Bourdin, Gaurav Dhar. Optimal sampled-data controls with running inequality state constraints: Pontryagin maximum principle and bouncing trajectory phenomenon. Mathematical Programming, Series A, 2020. ⟨hal-02160231v2⟩

HAL Id: hal-02160231

<https://hal.science/hal-02160231v2>

Submitted on 26 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.



HAL Authorization

Optimal sampled-data controls with running inequality state constraints – Pontryagin maximum principle and bouncing trajectory phenomenon

Loïc Bourdin*, Gaurav Dhar†

June 26, 2019

Abstract

Sampled-data control systems have steadily been gaining interest for their applications in Engineering and Automation. In the present paper we derive a Pontryagin maximum principle for general nonlinear optimal sampled-data control problems in the presence of running inequality state constraints. In particular we obtain a nonpositive averaged Hamiltonian gradient condition associated to an adjoint vector being a function of bounded variations.

As a well-known challenge, theoretical and numerical difficulties may arise in optimal permanent control problems with state constraints due to the possible pathological behavior of the adjoint vector (jumps and singular part lying on parts of the trajectory in contact with the boundary of the restricted state space). However, in our case of sampled-data controls, we find that, under certain general hypotheses, the optimal trajectory only contacts the running inequality state constraints at most at the sampling times. In that context the adjoint vector only experiences jumps at most at the sampling times (and thus in a finite number and at precise instants) and its singular part vanishes. Hence, taking advantage of this bouncing trajectory phenomenon, we are able to implement a numerical indirect method which we use to solve three simple examples.

Keywords: Sampled-data control; digital control; Pontryagin maximum principle; optimal control; state constraints; Ekeland variational principle; indirect method; shooting method; function of bounded variations; Stieltjes integral.

AMS Classification: 34K35; 34H05; 49J15; 49J50; 49M15; 93C10; 93C15; 93C57.

1 Introduction

In mathematics a *dynamical system* describes the evolution of a point (usually called the *state* of the system) in an appropriate space (called the *state space*) following an evolution rule (called the *dynamics* of the system). Dynamical systems are of many different natures (continuous versus discrete systems, deterministic 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 intervenes in the dynamics and thus influences the evolution of the state. Finally an *optimal control problem* consists in determining a control which allows to steer the state of a control system from an initial condition to some desired target while minimizing a given cost and satisfying some constraints.

*Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France. loic.bourdin@unilim.fr

†Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France. gaurav.dhar@unilim.fr

Context in optimal control theory. Established in [54] by Pontryagin, Boltyanskii, Gamkrelidze and Mischenko at the end of the 1950's, the Pontryagin Maximum Principle (in short, PMP) is a fundamental result in optimal control theory with numerous theoretical and numerical applications. We refer to [18, 39, 40, 46, 59, 60, 63] and references therein. The classical PMP gives first-order necessary optimality conditions for (continuous) optimal control problems in which the dynamical system is described by a general ordinary differential equation. Roughly speaking, the PMP ensures the existence of an *adjoint vector* (also called *costate*) satisfying some terminal conditions (called *transversality conditions*) such that the optimal control maximizes the *Hamiltonian* associated to the optimal control problem. As a well-known application, if the *Hamiltonian maximization condition* allows to express the optimal control as a function of the augmented state-costate vector, then the PMP induces the so-called numerical *indirect method*¹ which consists in determining the optimal control by numerically solving the boundary value problem satisfied by the augmented state-costate vector via a shooting algorithm. 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 dynamical system is described by a general difference equation (see, e.g., [8, 37, 41]). In these discrete versions of the PMP, the Hamiltonian maximization condition does not hold in general (see a counterexample in [8, Examples 10.1-10.4 p.59-62]) and has to be replaced by a weaker condition known as the *nonpositive Hamiltonian gradient condition* (see e.g., [8, Theorem 42.1 p.330]).

An important generalization of the PMP concerns the taking into account of state constraints where the state is restricted to a certain region of the state space. Indeed it is often undesirable and even inadmissible in scientific and engineering applications that the state crosses certain limits imposed in the state space for safety or practical reasons. Many examples can be found in mechanics and aerospace engineering (e.g., an engine may overheat or overload). We refer to [11, 25, 43, 61, 62] and references therein for other examples. State constrained optimal control problems are also important in management and economics (e.g., an inventory level may be limited in a production model). We refer to [22, 50, 55, 59] and references therein for other examples. A version of the PMP for (continuous) optimal control problems with running inequality state constraints for piecewise smooth controls was given by Gamkrelidze et al. (see, e.g., [33] and [54, Theorem 25 p.311]). Other versions can be found in [49, Section 3 p.347] by Maurer and in [34, Theorem 14.1 p.106] by Girsanov. A comprehensive survey [38] of this field of research has been given in 1995 by Hartl, Sethi and Vickson. Note that the PMP for (continuous) optimal control problems is more intricate in the presence of running inequality state constraints because the adjoint vector is not absolutely continuous in general (as in the unconstrained state case), but (only) of bounded variations. Therefore theoretical and numerical difficulties may arise due to the possible pathological behavior of the adjoint vector which consists in jumps and singular part lying on parts of the trajectory in contact with the boundary of the restricted state space. As a consequence a wide portion of the literature is devoted to the analysis of the behavior of the costate and some constraint qualification conditions have been established in order to ensure that the adjoint vector has no singular part (see e.g., [6, 11, 27, 38, 42, 49]). We briefly conclude this paragraph by mentioning that the related theme of discrete optimal control problems in the presence of running state constraints has already been investigated in the literature (see, e.g., [26, Proposition 2 p.13]).

The control of a system is very often assumed to be *permanent* in the literature, in the sense that its value can be modified at any time. In the present paper we are interested in *sampled-data control systems* in which the state evolves continuously with respect to time while the control evolves discretely with respect to time. Precisely the value of the control is authorized to be modified at only a finite number of instances which are 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 in Engineering implemented by *digital controllers* which have a

¹The terminology *indirect method* is opposed to the one of *direct method* which consists in a full discretization of the optimal control problem resulting into a constrained finite-dimensional optimization problem that can be numerically solved from various standard algorithms and techniques.

finite precision (see, e.g., [58, 64]). 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., [36]). Numerous texts and articles have developed control theory for sampled-data control systems (see, e.g., [1, 2, 3, 31, 35, 44] and references therein). Recently Bourdin and Trélat have obtained in [16] a version of the PMP for general nonlinear optimal sampled-data control problems (which is valid in the even more general framework of *time scale calculus*). In the sampled-data control framework, as in the case of discrete control systems previously mentioned, 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 [16, Theorem 2.6 p.62]) which can be used to numerically compute the optimal sampled-data control by an indirect method (see, e.g., [16, Section 3.1]).

Contributions of the present paper. To the best of our knowledge optimal sampled-data control problems have never been investigated in the presence of state constraints. Our first objective in the present paper is to bridge this gap in the literature by establishing a PMP for general nonlinear optimal sampled-data control problems in the presence of running inequality state constraints (see Theorem 3.1 in Section 3). Our proof is based on the Ekeland variational principle [29]. Similarly to the PMP derived in [16, Theorem 2.6 p.62] for unconstrained optimal sampled-data control problems, we obtain a first-order necessary optimality condition described by a nonpositive averaged Hamiltonian gradient condition. Moreover, as in the case of optimal (permanent) control problems with running inequality state constraints, we find that the adjoint vector is in general (only) of bounded variations. Therefore one would expect to encounter the same difficulties as in the permanent control case when implementing an indirect method due to the jumps and the singular part of the adjoint vector. However, in the context of sampled-data controls, we find that the optimal trajectories have a common behavior that allows us to overcome these difficulties. Precisely, when we undertook studying optimal sampled-data control problems in the presence of running inequality state constraints, we first solved numerically some simple problems using a direct method. Notably we observed that in each problem the optimal trajectory “bounces” against the boundary of the restricted state space, touching the state constraint at most at exactly the sampling times. This behavior constitutes the second major focus of the present work and is referred to as the *bouncing trajectory phenomenon*. In a first place we prove that, under certain general hypotheses, any admissible trajectory necessarily bounces on the running inequality state constraints and, moreover, the rebounds occur at most at exactly the sampling times (and thus are in a finite number and at precise instants). We refer to Proposition 4.1 in Section 4. Inherent to this behavior, the singular part of the adjoint vector vanishes and its discontinuities are reduced to a finite number of jumps which are localized at most at the sampling times. Taking advantage of these informations we are able in Section 5 to implement an indirect method solving three different simple examples of optimal sampled-data control problems with running inequality state constraints.

Organization of the paper. The present paper is organized as follows. Section 2 is dedicated to notation and functional framework and gives a short recap on Cauchy-Stieltjes problems which is required in order to enunciate our main result. In Section 3 we first present the optimal sampled-data control problem with running inequality state constraints considered in this paper (see Problem (OSCP)) accompanied by some terminology and basic assumptions. The corresponding Pontryagin maximum principle is stated thereafter (see Theorem 3.1) and a list of general comments is in order. In Section 4 we give heuristic descriptions and a sufficient condition for observing the bouncing trajectory phenomenon. In Section 5 we propose an indirect method for solving numerically optimal sampled-data control problems with running inequality state constraints with the aid of our main result and of the bouncing trajectory phenomenon. Then we illustrate this method and highlight the bouncing trajectory phenomenon by solving numerically three simple examples. Finally Appendix A is devoted to technical tools needed for the proof of Theorem 3.1 which is detailed right after in Appendix B.

2 Basics and recap about linear Cauchy-Stieltjes problems

This section is dedicated to basic notions and background that are required in order to enunciate our main result (see Theorem 3.1 in Section 3.2). In the whole section we consider four fixed positive integers $m, n, q, N \in \mathbb{N}^*$ and $T > 0$ as being a fixed positive real number. Section 2.1 below is devoted to the notations and the functional framework we will encounter all along the paper. Section 2.2 is dedicated to some recalls on functions of bounded variations and on linear Cauchy-Stieltjes problems.

2.1 Notation and functional framework

In this paper we denote by:

- $L_n^1 := L^1([0, T], \mathbb{R}^n)$ the standard Lebesgue space of integrable functions defined on $[0, T]$ with values in \mathbb{R}^n , endowed with its usual norm $\|\cdot\|_{L_n^1}$;
- $L_n^\infty := L^\infty([0, T], \mathbb{R}^n)$ the standard Lebesgue space of essentially bounded functions defined on $[0, T]$ with values in \mathbb{R}^n , endowed with its usual norm $\|\cdot\|_{L_n^\infty}$;
- $C_n := C([0, T], \mathbb{R}^n)$ the standard space of continuous functions defined on $[0, T]$ with values in \mathbb{R}^n , endowed with the usual uniform norm $\|\cdot\|_\infty$;
- $AC_n := AC([0, T], \mathbb{R}^n)$ the subspace of C_n of absolutely continuous functions.

In the sequel we consider an N -partition of the interval $[0, T]$, that is, let $\mathbb{T} := \{t_i\}_{i=0, \dots, N}$ be a $(N+1)$ -tuple such that

$$0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T.$$

We denote the set of all piecewise constant functions defined on $[0, T]$ with values in \mathbb{R}^m respecting the N -partition \mathbb{T} by

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

In this paper, as usual in the Lebesgue space L_m^∞ , two functions in $PC_m^\mathbb{T}$ which are equal almost everywhere on $[0, T]$ are identified. Precisely, if $u \in PC_m^\mathbb{T}$, then u is identified to the function

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

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

2.2 Recalls on functions of bounded variations and on linear Cauchy-Stieltjes problems

In this section we refer to standard references and books such as [4, 19, 20, 32, 65] and to Appendices A.3 and A.4 for some more recalls. A function $\eta : [0, T] \rightarrow \mathbb{R}^q$ is said to be of *bounded variations* on $[0, T]$ if

$$\sup_{\{s_i\}_i} \left\{ \sum_i \|\eta(s_{i+1}) - \eta(s_i)\|_{\mathbb{R}^q} \right\} < +\infty,$$

where the supremum is taken over all finite partitions $\{s_i\}_i$ of the interval $[0, T]$. In what follows we denote by $BV_q := BV([0, T], \mathbb{R}^q)$ the space of all functions of bounded variations on $[0, T]$ with values in \mathbb{R}^q . Finally recall that the *Riemann-Stieltjes integral* defined by

$$\int_0^T \psi(\tau) d\eta(\tau) := \lim \sum_i (\eta(s_{i+1}) - \eta(s_i)) \psi(s_i),$$

exists in \mathbb{R}^n for all $\psi \in C_n$ and all $\eta \in BV_1$, where the limit is taken over all finite partitions $\{s_i\}_i$ of the interval $[0, T]$ whose length tends to zero.

We are now in a position to give a short recap on linear Cauchy-Stieltjes problems. Let $A \in L^\infty([0, T], \mathbb{R}^{n \times n})$, $B \in L_n^\infty$ and let $C_j \in C_n$ and $\eta_j \in BV_1$ for every $j = 1, \dots, q$. We say that $x \in L_n^\infty$ is a *solution to the forward linear Cauchy-Stieltjes problem* (**FCSP**) given by

$$\begin{cases} dx = (A \times x + B) dt + \sum_{j=1}^q C_j d\eta_j & \text{over } [0, T], \\ x(0) = x_0, \end{cases} \quad (\text{FCSP})$$

where $x_0 \in \mathbb{R}^n$ is fixed, if x satisfies the integral representation

$$x(t) = x_0 + \int_0^t \left(A(\tau) \times x(\tau) + B(\tau) \right) d\tau + \sum_{j=1}^q \int_0^t C_j(\tau) d\eta_j(\tau),$$

for almost every $t \in [0, T]$. Similarly we say that $p \in L_n^\infty$ is a *solution to the backward linear Cauchy-Stieltjes problem* (**BCSP**) given by

$$\begin{cases} -dp = (A \times p + B) dt + \sum_{j=1}^q C_j d\eta_j & \text{over } [0, T], \\ p(T) = p_T, \end{cases} \quad (\text{BCSP})$$

where $p_T \in \mathbb{R}^n$ is fixed, if p satisfies the integral representation

$$p(t) = p_T + \int_t^T \left(A(\tau) \times p(\tau) + B(\tau) \right) d\tau + \sum_{j=1}^q \int_t^T C_j(\tau) d\eta_j(\tau),$$

for almost every $t \in [0, T]$. From usual contraction mapping techniques, one can easily prove that Problems (**FCSP**) and (**BCSP**) both admit a unique solution. Moreover, from standard identifications in L_n^∞ , these solutions both belong to BV_n and the above integral representations are both satisfied for all $t \in [0, T]$. We refer to [12, Appendices C and D] and references therein for details.

We conclude this section with a last definition. A function $\eta \in BV_q$ is said to be *normalized* if $\eta(0) = 0_{\mathbb{R}^q}$ and η is right-continuous on $(0, T)$. We denote the subspace of BV_q of all normalized functions of bounded variations by $NBV_q := NBV([0, T], \mathbb{R}^q)$.

3 Main result

This section is dedicated to the statement of our main result (see Theorem 3.1 in Section 3.2). In Section 3.1 below we introduce the general optimal sampled-data control problem with running inequality state constraints considered in this work, and we fix the terminology and assumptions used all along the paper. In Section 3.2 we state the corresponding Pontryagin maximum principle and a list of comments is in order.

3.1 A general optimal sampled-data control problem with running inequality state constraints

Let $n, m, q, N \in \mathbb{N}^*$ be four fixed positive integers. Let us fix a positive real number $T > 0$, as well as an N -partition $\mathbb{T} = \{t_i\}_{i=0, \dots, N}$ of the interval $[0, T]$. In the present work we focus on the

general optimal sampled-data control problem with running inequality state constraints given by

$$\left. \begin{array}{ll} \text{minimize} & g(x(T)) + \int_0^T L(x(\tau), u(\tau), \tau) d\tau, \\ \text{subject to} & x \in \text{AC}_n, u \in \text{PC}_m^{\mathbb{T}}, \\ & \dot{x}(t) = f(x(t), u(t), t) \text{ for a.e. } t \in [0, T], \\ & x(0) = x_0, \\ & h_j(x(t), t) \leq 0 \text{ for all } t \in [0, T] \text{ and all } j = 1, \dots, q, \\ & u(t) \in U \text{ for all } t \in [0, T]. \end{array} \right\} \quad (\text{OSCP})$$

In Problem (OSCP), x is called the *state function* (also called the *trajectory*) and u is called the *control function*. A couple (x, u) is said to be *admissible* for Problem (OSCP) if it satisfies all the above constraints. A *solution* to Problem (OSCP) is an admissible couple (x, u) which minimizes the *Bolza cost* given by $g(x(T)) + \int_0^T L(x(\tau), u(\tau), \tau) d\tau$ among all admissible couples.

Throughout the paper we will make use of the following regularity and topology assumptions:

- the function $g : \mathbb{R}^n \rightarrow \mathbb{R}$, that describes the *Mayer cost* $g(x(T))$, is of class C^1 ;
- the function $L : \mathbb{R}^n \times \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}$, that describes the *Lagrange cost* $\int_0^T L(x(t), u(t), t) dt$, is continuous and of class C^1 with respect to its first two variables;
- the *dynamics* $f : \mathbb{R}^n \times \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}^n$, that drives the *state equation* $\dot{x}(t) = f(x(t), u(t), t)$, is continuous and of class C^1 with respect to its first two variables;
- the function $h = (h_j)_{j=1, \dots, q} : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^q$, that describes the *running inequality state constraints* $h_j(x(t), t) \leq 0$, is continuous and of class C^1 in its first variable;
- the set $U \subset \mathbb{R}^m$, that describes the *control constraint* $u(t) \in U$, is a nonempty closed convex subset of \mathbb{R}^m ;
- the initial condition $x_0 \in \mathbb{R}^n$ is fixed.

In the classical literature about Pontryagin maximum principle (see, e.g., [18, 21, 40, 54, 59, 60, 63] and references therein), the control u usually can be any function in L_m^∞ (with values in U). In that situation we say that the control is *permanent* in the sense that its value can be modified at any real time $t \in [0, T]$. In the present paper the control u is constrained to be a piecewise constant function respecting the fixed N -partition $\mathbb{T} = \{t_i\}_{i=0, \dots, N}$. In other words, the value of the control u is authorized to be modified at most $N - 1$ times and at precise fixed instants. In that situation we say that the control is *nonpermanent*. The standard terminology adopted in the literature is to say that the elements t_i of \mathbb{T} are the *sampling times* and that the control u in Problem (OSCP) is a *sampled-data control*.

Control theory for sampled-data control systems has already been considered in several texts (see, e.g., [1, 3, 31, 35, 44]), often in the context of *digital control*. Optimal sampled-data control problems have also been investigated in the literature (see, e.g., [2, 7, 13, 15, 16, 17]) with general terminal constraints on $x(0)$ and $x(T)$, free final time, free sampling times, etc. To the best of our knowledge, running inequality state constraints have never been investigated with sampled-data controls. Our aim in this paper is to fill this gap in the literature, and thus we will essentially focus in the present work on the running inequality state constraints $h_j(x(t), t) \leq 0$ in Problem (OSCP). As a consequence, for the sake of simplicity, we took the decision not to consider general terminal constraints in Problem (OSCP). Indeed we only consider the basic case in which the initial condition $x(0) = x_0$ is fixed and the final condition $x(T)$ is free. Similarly we also chose to consider

that the final time $T > 0$ and the partition $\mathbb{T} = \{t_i\}_{i=0,\dots,N}$ are fixed. If the reader is interested in techniques allowing to handle general terminal constraints, free final time, free sampling times, etc., we refer to the last above references.

Remark 3.1. Let (x, u) be an admissible couple 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 sampling interval $[t_i, t_{i+1}]$.

Remark 3.2. Existence theorems for optimal permanent control problems with running inequality state constraints can be found in a text by Clarke (see [23, Theorem 5.4.4 p.222]). Furthermore, existence theorems for related problems such as optimal permanent control problems with state constraints where the constraint is given as an inclusion into a general subset of the state space can be found in works of Cesari (see [21, Theorem 9.2.i p.311]) and Rockafellar (see [57, Theorem 2 p.696]). A Filippov-type theorem for the existence of a solution to Problem (OSCP) without running inequality state constraints was derived in [16, Theorem 2.1 p.61]. The present work only focuses on necessary optimality conditions and thus it is not our aim to discuss the extension of the previously mentioned result to the case with running inequality state constraints. Nevertheless we note that in the standard Filippov's theorem, as usually established for optimal permanent control problems, the controls belong to the infinite dimensional space L_m^∞ , while the sampled-data control framework considered here can be seen as a finite dimensional problem. This fundamental difference could potentially lead to existence results in the case of sampled-data controls with relaxed assumptions and constitutes an interesting challenge for forthcoming works.

Remark 3.3. We take this occasion to provide a (nonexhaustive) list of perspectives:

- (i) In the context of linear-quadratic problems, the authors of [17] prove that optimal sampled-data controls 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. Since these phenomenons are suggested to hold true in the numerical simulations presented in Section 5, an interesting research perspective would be to extend these convergence results to the context of the present work.
- (ii) Chattering phenomenon has been investigated for optimal permanent control problems with running inequality state constraints where, in particular, the optimal trajectory touches the constraint at an infinite sequence of isolated points in rapid succession (see, e.g., [56]). In view of handling chattering phenomenon, it would be relevant to extend the present framework to the case of infinite partitions. As a first step one should consider the case of countably infinite partitions with exactly one accumulation point.
- (iii) In the paper [13] the authors consider optimal sampled-data control problems with free sampling times and obtain a corresponding necessary optimality condition which happens to coincide with the continuity of the Hamiltonian function. It would be relevant to extend the scope of Problem (OSCP) to study optimal sampling times in the presence of running inequality state constraints.
- (iv) Several papers in the literature consider optimal permanent control problems with constraints of different natures, for instance with mixed constraints of the form $h(x(t), u(t), t) \leq 0$ (see, e.g., [38]) or with state constraints where the constraint is given as an inclusion into a general subset of the state space (see, e.g., [21, 57]). A possible challenge would be to extend the present work to the previous mentioned contexts.
- (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 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 [5] 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.

3.2 Pontryagin maximum principle

The main objective of the present paper is to derive a Pontryagin maximum principle for Problem (OSCP). Let us mention here that establishing a consensual version of the Pontryagin maximum principle for optimal permanent control problems in the presence of running inequality state constraints still constitutes a wonderful mathematical challenge. Necessary optimality conditions have been obtained by Jacobson, Lele and Speyer [42] including the maximization condition of the Hamiltonian. Similar conditions can be found for example in the articles of Lastman [45], Maurer [49, 51], Norris [53], etc. This field of research has been the subject of a comprehensive survey by Hartl, Sethi and Vickson [38] in 1995. More recent versions of necessary optimality conditions for various generalizations have been claimed in the literature (see, e.g., [9, 10, 63]).

The novelty of the present work is to deal with running inequality state constraints with nonpermanent controls, precisely, with sampled-data controls. As usual in the literature we introduce the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ associated to Problem (OSCP) 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 $(x, u, p, p^0, t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times [0, T]$. The main result of this article is given by the following theorem.

Theorem 3.1 (Pontryagin maximum principle). *Let (x, u) be a solution to Problem (OSCP). Then there exists a nontrivial couple (p^0, η) , where $p^0 \leq 0$ and $\eta = (\eta_j)_{j=1, \dots, q} \in \text{NBV}_q$, such that the nonpositive averaged Hamiltonian gradient condition*

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

holds for all $v \in U$ and all $i = 0, \dots, N-1$, where the adjoint vector $p \in \text{BV}_n$ (also called costate) is the unique solution to the backward linear Cauchy-Stieltjes problem given by

$$\begin{cases} -dp = \left(\partial_1 f(x, u, \cdot)^\top \times p + p^0 \partial_1 L(x, u, \cdot) \right) dt - \sum_{j=1}^q \partial_1 h_j(x, \cdot) d\eta_j & \text{over } [0, T], \\ p(T) = p^0 \nabla g(x(T)). \end{cases} \quad (2)$$

Moreover the complementary slackness condition:

$$\eta_j \text{ is monotonically increasing on } [0, T] \text{ and } \int_0^T h_j(x(\tau), \tau) d\eta_j(\tau) = 0, \quad (3)$$

is satisfied for each $j = 1, \dots, q$.

Appendices A and B are dedicated to the detailed proof of Theorem 3.1. A list of comments is in order.

Remark 3.4. The nontrivial couple (p^0, η) provided in Theorem 3.1, which corresponds to 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 $p^0 = -1$. The case $p^0 = 0$ is usually called the abnormal case.

Remark 3.5. Our strategy in Appendices A and B in order to prove Theorem 3.1 is based on the Ekeland variational principle [29] applied to an appropriate penalized functional which requires the closedness of U in order to be defined on a complete metric set (see details in Section B.1.1). The closedness of U is therefore a crucial assumption in the current paper. On the other hand,

the convexity of U is useful in order to consider convex L^∞ -perturbations of the control for the sensitivity analysis of the state equation (see Proposition A.4 in Section A.2). Nevertheless, the convexity of U can be removed by considering the concept of U -dense directions (see, e.g., [16]).

Remark 3.6. If there is no running inequality state constraint in Problem (OSCP), that is, considering $h_j = 0$ for all $j = 1, \dots, q$, then Theorem 3.1 recovers the standard Pontryagin maximum principle for optimal sampled-data control problems derived for instance in [15, 16]. Let us recall that the classical Hamiltonian maximization condition (which holds in the case of optimal permanent controls) is not true in general in the case of optimal sampled-data controls, and has to be replaced by the nonpositive averaged Hamiltonian gradient condition. We refer to the above references for a detailed discussion about this feature.

Remark 3.7. The nonpositive averaged Hamiltonian gradient condition in Theorem 3.1 can be rewritten as

$$\int_{t_i}^{t_{i+1}} \partial_2 H(x(\tau), u_i, p(\tau), p^0, \tau) d\tau \in N_U[u_i]$$

for all $i = 0, \dots, N-1$, where $N_U[u_i]$ stands for the classical normal cone to U at the point $u_i \in U$. We deduce that

$$u_i = \text{proj}_U \left(u_i + \int_{t_i}^{t_{i+1}} \partial_2 H(x(\tau), u_i, p(\tau), p^0, \tau) d\tau \right),$$

for all $i = 0, \dots, N-1$, where proj_U stands for the classical projection operator onto U . In particular, if $U = \mathbb{R}^m$ (that is, if there is no control constraint in Problem (OSCP)), then the nonpositive averaged Hamiltonian gradient condition can be rewritten as

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

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

Remark 3.8. Following the proof in Appendices A and B, one can easily see that Theorem 3.1 is still valid for a couple (x, u) which is a solution to Problem (OSCP) in (only) a local sense to be precised. For the ease of statement, we took the decision to establish Theorem 3.1 for a couple (x, u) which is a solution to Problem (OSCP) in a global sense.

Remark 3.9. In the context of Theorem 3.1 and using the definition of the Hamiltonian, note that the state equation can be written as

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

for almost every $t \in [0, T]$, and that the adjoint equation can be written as

$$-dp = \partial_1 H(x, u, p, p^0, \cdot) dt - \sum_{j=1}^q \partial_1 h_j(x, \cdot) d\eta_j,$$

over $[0, T]$.

Remark 3.10. It is frequent in the literature (see, e.g., [63, Theorem 9.3.1]) to find the adjoint vector $p \in \text{BV}_n$ written as the sum $p = p_1 + p_2$ where $p_1 \in \text{AC}_n$ is the unique solution to the backward linear Cauchy problem

$$\begin{cases} \dot{p}_1(t) = -\partial_1 H(x(t), u(t), p(t), p^0, t) \text{ for a.e. } t \in [0, T], \\ p_1(T) = p^0 \nabla g(x(T)), \end{cases}$$

and where $p_2 \in \text{BV}_n$ is defined by

$$p_2(t) := - \sum_{j=1}^q \int_t^T \partial_1 h_j(x(\tau), \tau) d\eta_j(\tau),$$

for all $t \in [0, T]$. This decomposition easily follows from the integral representation of the solutions to backward linear Cauchy-Stieltjes problem recalled in Section 2.2.

Remark 3.11. In the context of Theorem 3.1, note that the complementary slackness condition implies that, for all $j = 1, \dots, q$, the function η_j remains constant on any open subinterval $(\tau_1, \tau_2) \subset \{t \in [0, T] \mid h_j(x(t), t) < 0\}$ with $\tau_1 < \tau_2$. Denoting by $d\eta_j$ the finite nonnegative Borel measure associated to η_j (see Section A.3 for more details), we deduce that

$$\text{supp}(d\eta_j) \subset \{t \in [0, T] \mid h_j(x(t), t) = 0\},$$

for all $j = 1, \dots, q$, where $\text{supp}(d\eta_j)$ stands for the classical notion of support of the measure $d\eta_j$.

Remark 3.12. Note that the necessary optimality conditions of Theorem 3.1 are of interest only when the running inequality state constraints are nondegenerate, in the sense that $\partial_1 h_j(x(t), t) \neq 0$ whenever $h_j(x(t), t) = 0$ for all $j = 1, \dots, q$. We refer to [63, Remark (b) p.330] for a similar remark in the classical case of permanent controls.

Remark 3.13. In the classical case of optimal permanent control problems without running inequality state constraint, the Pontryagin maximum principle induces a numerical indirect method based on the resolution by a shooting method of the boundary value problem satisfied by the augmented state-costate vector (see, e.g., [60, p.170-171] for details). Recall that:

- (i) In the presence of running inequality state constraints, the adaptation of a numerical indirect method can be considered using an appropriate version of the Pontryagin maximum principle, such as those mentioned at the beginning of this section. Nevertheless, some theoretical and/or numerical difficulties may appear due to the behavior of the adjoint vector $p \in \text{BV}_n$ which can be pathological and can admit an infinite number of discontinuities (see Section 5 for more details).
- (ii) A numerical indirect method has also been adapted to the case of optimal sampled-data controls (without running inequality state constraint) in [15, 16], and also in the case of free sampling times in [13].

To the best of our knowledge, the numerical indirect method has never been adapted to optimal sampled-data control problems in the presence of running inequality state constraints. Our aim in Section 5 is to fill this gap in the literature by using the Pontryagin maximum principle derived in Theorem 3.1. Of course, in the context of Theorem 3.1, it might be possible that the adjoint vector $p \in \text{BV}_n$ is pathological and/or admits an infinite number of discontinuities, but it will be shown in Sections 4 and 5 that, under certain (quite unrestrictive) hypotheses, the implementation of a numerical indirect method is simplified due to the particular behavior of the admissible trajectories.

Remark 3.14. In this paper the proof of Theorem 3.1 is based on the Ekeland variational principle [29]. An alternative proof of Theorem 3.1 can be obtained by adapting a remarkable technique exposed in the paper [28] by Dmitruk and Kaganovich that consists of mapping each sampling interval $[t_i, t_{i+1}]$ to the interval $[0, 1]$ and by taking the values u_i of the sampled-data control to be additional parameters. Then, through the application of the Pontryagin maximum principle for optimal permanent control problems with running state constraints (see, e.g., [12, Theorem 1] and [63, Theorem 9.5.1 p. 339-340]), one obtains the adjoint equation (2) and complementary slackness condition (3) given in Theorem 3.1. Moreover the application of a “Pontryagin maximum principle with parameters” (see, e.g., [14, Remark 5 p.3790]) leads to a necessary optimality condition written in integral form which coincides with the nonpositive averaged Hamiltonian gradient condition (1).

4 Bouncing trajectory phenomenon

When we undertook to study optimal sampled-data control problems in the presence of running inequality state constraints, one of our first actions was to solve numerically some simple problems using a direct method (see Section 5 for some details on direct methods in optimal control

theory). At this occasion we observed that the optimal trajectories returned by the algorithm had a common behavior with respect to the running inequality state constraint. Precisely the optimal trajectories were “bouncing” on it. We refer to Figure 3 and Section 5 for some examples illustrating this observation which we refer to as the *bouncing trajectory phenomenon*. Actually, when dealing with sampled-data controls and running inequality state constraints, the bouncing trajectory phenomenon concerns, not only the optimal trajectories, but all admissible trajectories.

In this section our aim is to give a detailed description of this new observation (which does not appear in general in the classical theory, that is, with permanent controls). Precisely we will show that, under certain hypotheses, an admissible trajectory of Problem (OSCP) necessarily bounces on the running inequality state constraints and, moreover, the activating times occur at most at the sampling times t_i (and thus in a finite number and at precise instants). As detailed later in Section 5, this feature presents some benefits from a numerical point of view.

In Section 4.1 below we initiate an heuristic discussion allowing to understand why, usually, the admissible trajectories of Problem (OSCP) bounce on the running inequality state constraints and, moreover, at most at the sampling times t_i . Then we provide in Section 4.2 a mathematical framework and rigorous justifications which allow us to specify a sufficient condition ensuring this behavior (see Proposition 4.1).

Throughout this section, for simplicity, we will assume that $q = 1$, that is, there is only one running inequality state constraint in Problem (OSCP) denoted by $h(x(t), t) \leq 0$. Nevertheless the results and comments of this section can be extended to multiple running inequality state constraints, that is, for $q \geq 2$. Furthermore we will assume that the dynamics f and the running inequality state constraint function h are of class C^∞ in all variables. In particular note that any admissible trajectory of Problem (OSCP) is thus piecewise smooth of class C^∞ , in the sense that it is of class C^∞ over each sampling interval $[t_i, t_{i+1}]$.

4.1 Expected behavior of an admissible trajectory

We start this section by recalling some standard terminology from [38, p.183] or [59, p.105]. Let x be an admissible trajectory of Problem (OSCP). An interval $[\tau_1, \tau_2] \subset [0, T]$ with $\tau_1 < \tau_2$ is called a *boundary interval* if $h(x(t), t) = 0$ for all $t \in [\tau_1, \tau_2]$. An element $t \in [0, T]$ is called an *activating time* if it satisfies $h(x(t), t) = 0$. Note that any point of a boundary interval is an activating time, while the reverse is not true in general. In particular an activating time $t \in [0, T]$ which is isolated from all other activating times is called a *contact time*. Finally, in the sequel, we say that the trajectory x displays the *bouncing trajectory phenomenon* if the set of activating times contains no boundary interval.

Our aim in this section is to give some heuristic descriptions (and illustrative figures) of the main reason why a bouncing trajectory phenomenon is usually displayed when dealing with sampled-data controls in the presence of running inequality state constraints (see (i) below) and why, moreover, the activating times occur at most at the sampling times t_i (see (ii) below). The mathematical framework and rigorous justifications will be provided in Section 4.2.

- (i) In the classical theory (that is, with permanent controls), a boundary interval may correspond to a *feedback control*, that is, to an expression of the control as a function of the state. Such an expression usually leads to a nonconstant control. More generally, a running inequality state constraint usually cannot be activated by a trajectory on an interval $[\tau_1, \tau_2]$ with $\tau_1 < \tau_2$ on which the associated (permanent) control is constant. We refer to Figure 1 for an illustration. Therefore, since we deal with piecewise constant controls in Problem (OSCP), one should expect that an admissible trajectory of Problem (OSCP) does not possess any boundary interval and thus displays a bouncing trajectory phenomenon. In order to guarantee the validity of this remark, it is sufficient to make an assumption on f and h which prevents the existence of an admissible trajectory x of Problem (OSCP) and an interval $[\tau_1, \tau_2] \subset [0, T]$ with $\tau_1 < \tau_2$ for which $\varphi^{(\ell)}(t) = 0$ for all $\ell \in \mathbb{N}$ and all $t \in [\tau_1, \tau_2]$, where φ is defined by $\varphi(t) := h(x(t), t)$ for all $t \in [\tau_1, \tau_2]$. This will be done in Section 4.2 (see Hypothesis (H1)).

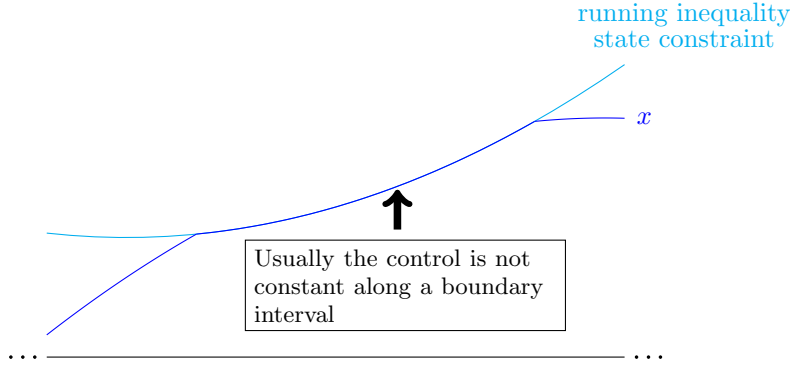


Figure 1: In the classical theory (that is, with permanent controls), a boundary interval is usually associated to a nonconstant control

- (ii) Let $t \in [0, T]$ be a contact time of an admissible trajectory x of Problem (OSCP). We denote by u the corresponding control. Let us assume that t is not a sampling time, that is, $t \in (t_i, t_{i+1})$ for some $i \in \{0, \dots, N-1\}$. Usually the trajectory x “hits” the running inequality state constraint transversely at t . Since the control value $u(t) = u_i$ is fixed all along the sampling interval $[t_i, t_{i+1}]$, the trajectory x then “crosses” the running inequality state constraint immediately after t , which contradicts the admissibility of x . We refer to Figure 2 for an illustration. Hence, in order to preserve the admissibility of x , we understand that the control value must change at t , that is, since u is a sampled-data control, that t must be one of the sampling times t_i . From this simple heuristic discussion, one should expect that an admissible trajectory of Problem (OSCP) has contact times occurring at most at the sampling times t_i (and thus in a finite number). In order to guarantee the validity of this remark, it is sufficient to make an assumption on f and h which prevents the existence of an admissible trajectory of Problem (OSCP) which “hits” the running inequality state constraint tangentially. This will be done in Section 4.2 (see Hypothesis (H2)).

Finally, in case of absence of boundary interval, since the set of activating times coincides then with the closure of the set of contact times which is a finite set in the above context, one should expect that an admissible trajectory of Problem (OSCP) has activating times occurring at most at the sampling times t_i (and thus in a finite number and at precise instants).

We conclude from (i) and (ii) that one should expect from admissible trajectories of Problem (OSCP) to display the bouncing trajectory phenomenon and, moreover, such that the activating times occur at most at the sampling times t_i (and thus in a finite number and at precise instants). We refer to Figure 3 for an illustration of this feature. Note that, even if activating times are all sampling times, the reverse is not true in general.

We conclude this section by mentioning that the above descriptions are only heuristic and, of course, one can easily find counterexamples in which the behavior of Figure 3 is not observed. Nonetheless we emphasize that the bouncing trajectory phenomenon is quite ordinary when dealing with sampled-data controls and running inequality state constraints, as guaranteed by the mathematical justifications provided in Section 4.2 below and as illustrated by the examples numerically solved in Section 5.

4.2 A sufficient condition for the bouncing trajectory phenomenon

Our aim in this section is to provide a rigorous mathematical framework describing the heuristic discussion provided in the previous subsection. In particular we will formulate a sufficient condition

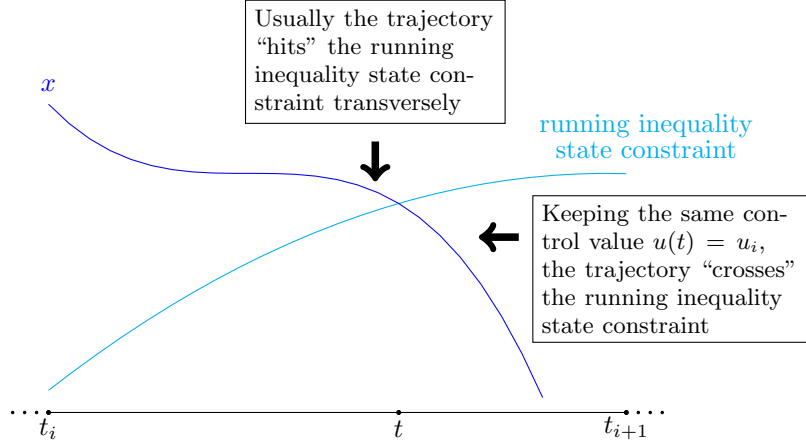


Figure 2: Illustration of an admissible trajectory x hitting transversely the running inequality state constraint at some contact time t which belongs to the interior (t_i, t_{i+1}) of a sampling interval.

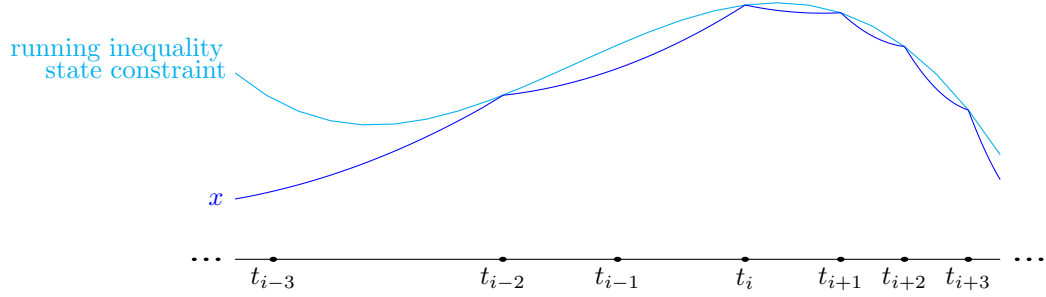


Figure 3: Illustration of the expected behavior of an admissible trajectory x of Problem (OSCP).

(see Proposition 4.1 below) ensuring the bouncing trajectory phenomenon and that the rebounds occur at most at the sampling times t_i .

To this aim, and similarly to [38, p.183], we introduce the functions $h^{[\ell]} : \mathbb{R}^n \times \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}$ defined by the induction

$$\begin{cases} h^{[0]}(y, v, t) := h(y, t), \\ \forall \ell \in \mathbb{N}, \quad h^{[\ell+1]}(y, v, t) := \langle \partial_1 h^{[\ell]}(y, v, t), f(y, v, t) \rangle_{\mathbb{R}^n} + \partial_3 h^{[\ell]}(y, v, t), \end{cases}$$

for all $(y, v, t) \in \mathbb{R}^n \times \mathbb{R}^m \times [0, T]$. We introduce the subset

$$\mathcal{M} := \{(y, t) \in \mathbb{R}^n \times ([0, T] \setminus \mathbb{T}) \mid h(y, t) = 0\},$$

and we denote by

$$\ell'(y, v, t) := \min\{\ell \in \mathbb{N} \mid h^{[\ell]}(y, v, t) \neq 0\} \in \mathbb{N}^* \cup \{+\infty\},$$

for all $(y, t) \in \mathcal{M}$ and all $v \in \mathbb{U}$. Finally we introduce the set

$$\mathcal{N}(y, t) := \{v \in \mathbb{U} \mid \ell'(y, v, t) \text{ is finite and even}\},$$

for all $(y, t) \in \mathcal{M}$. We are now in a position to establish the main result of this section.

Proposition 4.1. Assume that $q = 1$ and that f and h are of class C^∞ in all their variables. If the hypotheses

$$\forall (y, t) \in \mathcal{M}, \quad \forall v \in \mathcal{U}, \quad \ell'(y, v, t) < +\infty, \quad (\text{H1})$$

and

$$\forall (y, t) \in \mathcal{M}, \quad \forall v \in \mathcal{N}(y, t), \quad h^{[\ell'(y, v, t)]}(y, v, t) > 0, \quad (\text{H2})$$

are both satisfied, then the activating times of an admissible trajectory x of Problem (OSCP) are sampling times. In particular x displays the bouncing trajectory phenomenon and the activating times occur at most at the sampling times t_i (and thus in a finite number and at precise instants).

Proof. Let (x, u) be an admissible couple of Problem (OSCP). Let $t \in [0, T]$ be an activating time and assume by contradiction that $t \in (t_i, t_{i+1})$ for some $i = 0, \dots, N-1$. In particular we have $(x(t), t) \in \mathcal{M}$. Since $u_i \in \mathcal{U}$, from Hypothesis (H1), we know that $\ell' := \ell'(x(t), u_i, t) < +\infty$ and it holds that $h^{[\ell']}(x(t), u_i, t) \neq 0$. From Taylor's theorem it holds that

$$h(x(t+\varepsilon), t+\varepsilon) = \varepsilon^{\ell'} \left(\frac{h^{[\ell']}(x(t), u_i, t)}{\ell'!} + R(\varepsilon) \right),$$

for all $\varepsilon \in \mathbb{R}$ such that $t+\varepsilon \in (t_i, t_{i+1})$, where the remainder term R satisfies $\lim_{\varepsilon \rightarrow 0} R(\varepsilon) = 0$. Thus there exists $\bar{\varepsilon} > 0$ such that $(t-\bar{\varepsilon}, t+\bar{\varepsilon}) \subset (t_i, t_{i+1})$ and $h(x(t'), t')$ has the same sign than $(t'-t)^{\ell'} h^{[\ell']}(x(t), u_i, t)$ for all $t' \in (t-\bar{\varepsilon}, t+\bar{\varepsilon})$ with $t' \neq t$. We now distinguish two cases: ℓ' odd and ℓ' even. If ℓ' is odd, then there clearly exists $t' \in (t-\bar{\varepsilon}, t+\bar{\varepsilon})$ with $t' \neq t$ such that $h(x(t'), t') > 0$ which raises a contradiction with the admissibility of (x, u) . If ℓ' is even, then $u_i \in \mathcal{N}(x(t), t)$ and, from Hypothesis (H2), it holds that $h^{[\ell']}(x(t), u_i, t) > 0$. We easily deduce that there exists $t' \in (t-\bar{\varepsilon}, t+\bar{\varepsilon})$ with $t' \neq t$ such that $h(x(t'), t') > 0$ which raises the same contradiction. The proof is complete. \square

Remark 4.1. We emphasize that Hypotheses (H1) and (H2) are assumptions which exactly guarantee the validity of the arguments presented heuristically in the items (i) and (ii) of Section 4.1.

In the context of Proposition 4.1, it is ensured that an admissible trajectory of Problem (OSCP) activates the running inequality state constraint at most at the sampling times t_i (and thus in a finite number and at precise instants). We will see in Section 5 below that this bouncing trajectory phenomenon (with localized rebounds) presents some benefits from a numerical point of view. Taking this advantage we will numerically solve some simple examples in which Hypotheses (H1) and (H2) are both satisfied and we will observe optimal trajectories bouncing on the running inequality state constraint considered.

5 Numerical experiments

Two predominant kinds of numerical methods are known in classical optimal control theory (that is, with permanent controls) without running inequality state constraints. The first kind is usually called *direct methods* and they consist in making a full discretization of the optimal control problem which results into a constrained finite-dimensional optimization problem that can be numerically solved from various standard algorithms and techniques. The second strategy is called *indirect methods* because they are based on the Pontryagin maximum principle. Precisely, if the Hamiltonian maximization condition allows to express the optimal control u as a function of the state x and of the (absolutely continuous) adjoint vector p , then the indirect methods consist in the numerical resolution by a shooting method of the boundary value problem on the augmented vector (x, p) derived in the Pontryagin maximum principle. We emphasize that neither direct nor indirect methods are fundamentally better than the other. We refer for instance to [60, p.170-171] for details and discussions on the advantages and drawbacks of each kind of methods.

In the presence of running inequality state constraints, direct methods can easily be adapted. On the other hand, solving optimal permanent control problems with running inequality state

constraints using indirect methods might be more difficult than the unconstrained state case. Indeed, in that situation, the adjoint vector p is not absolutely continuous in general, but (only) of bounded variations. From the Lebesgue decomposition (see, e.g., [20, Corollary 20.20 p.373]), we can write

$$p = p_{ac} + p_{sc} + p_s,$$

where p_{ac} is the *absolutely continuous part*, p_{sc} is the *singularly continuous part* and p_s is the *saltus* or *pure jump part* of p . From the complementary slackness condition, it is well-known that the adjoint vector p is absolutely continuous outside the activating times of x . On the other hand, on boundary intervals, the adjoint vector p may have an infinite number of unlocalized jumps or a pathological behavior due to its singular part. As a consequence, an important part of the literature is devoted to the analysis of the behavior of the adjoint vector p and some constraint qualification conditions have been established. We refer for instance to [6, 11, 38, 42, 49].

In this paper we have established a Pontryagin maximum principle (Theorem 3.1) and our aim in this section is to propose an indirect method for solving optimal sampled-data control problems with running inequality state constraints. As in the classical theory (with permanent controls), it appears that the adjoint vector obtained in Theorem 3.1 is (only) a function of bounded variations and we will *a priori* encounter the same difficulties outlined above. Nevertheless, as detailed in Section 4, we have proved in Proposition 4.1 that, under (quite unrestrictive) Hypotheses (H1) and (H2), the optimal trajectory x of Problem (OSCP) activates the running inequality state constraint at most at the sampling times t_i . As detailed in Section 5.1 below, it follows that the corresponding adjoint vector p has no singular part and admits a finite number of jumps which are localized at most at the sampling times t_i . Taking advantage of this knowledge, we will propose in Section 5.1 a simple indirect method in order to solve numerically optimal sampled-data control problems with running inequality state constraints of the form of Problem (OSCP) under Hypotheses (H1) and (H2).

In Sections 5.2, 5.3 and 5.4, this indirect method is implemented in order to numerically solve three simple examples. We precise that the parameters of these examples have been chosen in order to obtain figures which illustrate and highlight the bouncing trajectory phenomenon and the jumps of the adjoint vector. Furthermore note that the numerical results returned by the indirect method suggests the convergence of the optimal sampled-data controls to the optimal permanent control as N tends to $+\infty$. This provides a very interesting perspective to investigate in future works. We mention that such a result has already been established in [17] in the case of linear-quadratic optimal sampled-data control problems and without running inequality state constraint.

We conclude this paragraph by remarking that the indirect method proposed in Section 5.1 (and its implementation in Sections 5.2, 5.3 and 5.4) is based on the assumption that there exists a solution to Problem (OSCP). This question of existence has not been addressed in the present paper and constitutes an open question for future works, see Remark 3.2.

5.1 A shooting function for an indirect method

In this subsection our aim is to provide an indirect method, based on the Pontryagin maximum principle given in Theorem 3.1, which allows to numerically solve some optimal sampled-data control problems with running inequality state constraints.

Let (x, u) be a solution to Problem (OSCP). We denote by p^0, η, p the elements provided by the Pontryagin maximum principle given in Theorem 3.1. In what follows we assume that the case is normal and we normalize $p^0 = -1$ (see Remark 3.4). As explained at the beginning of Section 5, the adjoint vector p may have a pathological behavior which would imply some theoretical and/or numerical difficulties. Our aim in the sequel is to take advantage of Proposition 4.1 established in Section 4. To this aim, we assume in the sequel that $q = 1$, that f and h are of class C^∞ in all variables and that Hypotheses (H1) and (H2) are satisfied. As a consequence, it follows from Proposition 4.1 that x activates the running inequality state constraint at most at the sampling times t_i . From the complementary slackness condition in Theorem 3.1, we deduce that η admits exactly $(N + 1)$ nonnegative jumps localized exactly at the sampling times t_i , and that η remains

constant over (t_0, t_1) and over all $[t_i, t_{i+1})$ with $i = 1, \dots, N-1$. In what follows we denote the nonnegative jumps of η as follows:

$$\eta^{[0]} := \eta(t_0^+) - \eta(t_0), \quad \eta^{[1]} := \eta(t_1) - \eta(t_0^+) \quad \text{and} \quad \forall i = 2, \dots, N, \quad \eta^{[i]} := \eta(t_i) - \eta(t_{i-1}).$$

From the adjoint equation in Theorem 3.1, it follows that the adjoint vector p has no singular part, that it admits $(N+1)$ jumps localized exactly at the sampling times t_i , and that p remains absolutely continuous over (t_0, t_1) and over all $[t_i, t_{i+1})$ with $i = 1, \dots, N-1$. Moreover, from the integral representation of p , the jumps of the adjoint vector are given by

$$p^{[i]} := -\eta^{[i]} \partial_1 h(x(t_i), t_i),$$

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

The general numerical indirect method proposed in this paper is based on the shooting map

$$\left(x(T), (\eta^{[i]})_{i=0, \dots, N} \right) \mapsto \left(x(0) - x_0, \left(\eta^{[i]} h(x(t_i), t_i) \right)_{i=0, \dots, N} \right),$$

where:

- (i) we provide a guess of the final value $x(T)$ and of the nonnegative jumps $\eta^{[i]}$ for all $i = 0, \dots, N$;
- (ii) we compute $p(T) = -\nabla g(x(T))$;
- (iii) we solve numerically the state equation and the adjoint equation in a backward way (from $t = T$ to $t = 0$), by using the nonpositive averaged Hamiltonian gradient condition in order to compute the control values u_i for all $i = 0, \dots, N-1$;
- (iv) we finally compute $x(0) - x_0$ and $\eta^{[i]} h(x(t_i), t_i)$ for all $i = 0, \dots, N$.

As illustrations of the above numerical indirect method, we solve three simple examples in Sections 5.2, 5.3 and 5.4 below. We precise that we used the MATLAB function *fsolve* in order to find the zeros of the above shooting function. Let us mention that the numerical results obtained and presented hereafter have all been confirmed by direct numerical approaches (using a discretization of the whole problem resulting into a constrained finite-dimensional optimization problem solved by the MATLAB function *fmincon*).

5.2 Example 1: a problem with a parabolic running inequality state constraint

We first consider the following optimal sampled-data control problem with running inequality state constraint given by

$$\left. \begin{array}{ll} \text{minimize} & \int_0^4 x(\tau) + \frac{1}{4} u(\tau)^2 d\tau \\ \text{subject to} & x \in \text{AC}_1, u \in \text{PC}_1^{\mathbb{T}}, \\ & \dot{x}(t) = u(t) \text{ for a.e. } t \in [0, 4], \\ & x(0) = 6, \\ & \frac{1}{2}(t-2)^2 + 2 - x(t) \leq 0 \text{ for all } t \in [0, 4], \\ & u(t) \in [-3, +\infty) \text{ for all } t \in [0, 4], \end{array} \right\} \quad (\text{E1})$$

for fixed uniform N -partitions \mathbb{T} of the interval $[0, 4]$ with different values of $N \in \mathbb{N}^*$. This simple problem coincides with a calculus of variations problem (with running inequality state constraints on the trajectory and its derivative, and also constraining the derivative to be piecewise constant).

We first verify that Problem (E1) satisfies Hypotheses (H1) and (H2). To this aim we follow the notations introduced in Section 4.2. We compute that for all $(y, t) \in \mathcal{M}$ and all $v \in [-3, +\infty)$ it holds that $h^{[2]}(y, v, t) = 1$ and so Hypothesis (H1) is satisfied. Furthermore, for all $(y, t) \in \mathcal{M}$ and all $v \in \mathcal{N}(y, t)$, we also deduce that $\ell'(y, v, t) = 2$ and that $h^{[\ell'(y, v, t)]}(y, v, t) = 1 > 0$, so Hypothesis (H2) is satisfied as well. We conclude from Proposition 4.1 that all admissible trajectories activate the running inequality state constraint at most at the sampling times t_i .

In the sequel we assume that there exists an optimal couple (x, u) for Problem (E1) and we denote by p^0, η, p the elements provided by the Pontryagin maximum principle given in Theorem 3.1. Let us verify that the case is normal (in the sense of Remark 3.4). Assume by contradiction that $p^0 = 0$. We have the adjoint equation $-dp = d\eta$ over $[0, 4]$ with $p(4) = 0$. Therefore $p(t) = \int_t^4 d\eta(\tau) = \eta(4) - \eta(t)$ for all $t \in [0, 4]$. Then, from the nontriviality of the couple (p^0, η) , it follows that $\eta \neq 0_{\text{NBV}_1}$ and thus, from the complementary slackness condition, we deduce that x necessarily activates the running inequality state constraint. Let $\bar{t} \in [0, 4]$ denote the first activating time. From Proposition 4.1, we know that $\bar{t} = t_{\bar{i}}$ for some $\bar{i} \in \{0, \dots, N\}$. Since $x(0) = 6$, we have $\bar{i} \geq 1$. It follows that $p(t) > 0$ for all $t \in [0, t_1)$. Finally, from the nonpositive averaged Hamiltonian gradient condition at $i = 0$, it follows that $u_0 \geq v$ for all $v \in [-3, +\infty)$ which is absurd.

From the previous paragraph, we normalize $p^0 = -1$ (see Remark 3.4). Since we are in the context of Proposition 4.1, we can now apply the shooting method detailed in Section 5.1. As expected, we observe in Figure 4 (with $N = 5$) that the optimal trajectory returned by the algorithm activates the running inequality state constraint at most at the sampling times t_i (represented by dashed lines). As also expected, the jumps of the adjoint vector occur at the same activating times. Figures 5 and 6 continue to illustrate this bouncing trajectory phenomenon for larger values of N (respectively with $N = 10$ and $N = 40$). Furthermore, in Figures 4, 5 and 6, we observe that the adjoint vector has no jump at sampling times which are not activating times.

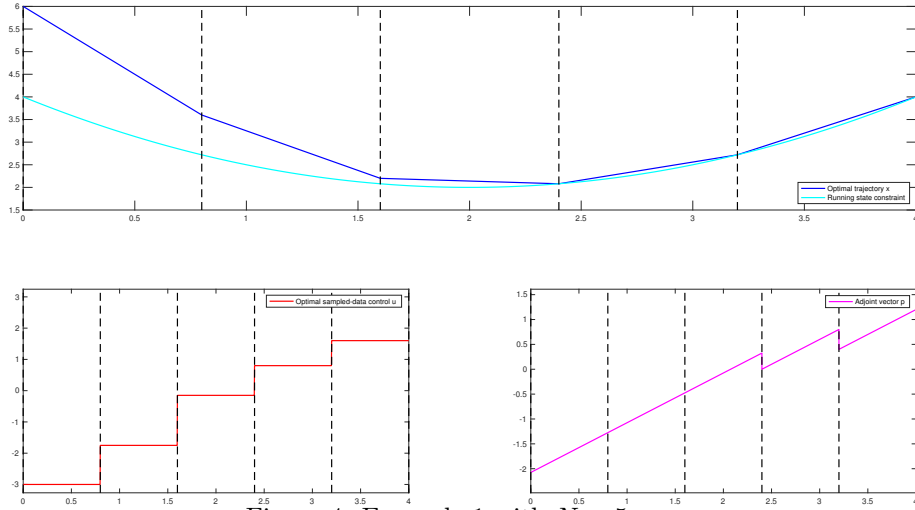


Figure 4: Example 1 with $N = 5$.

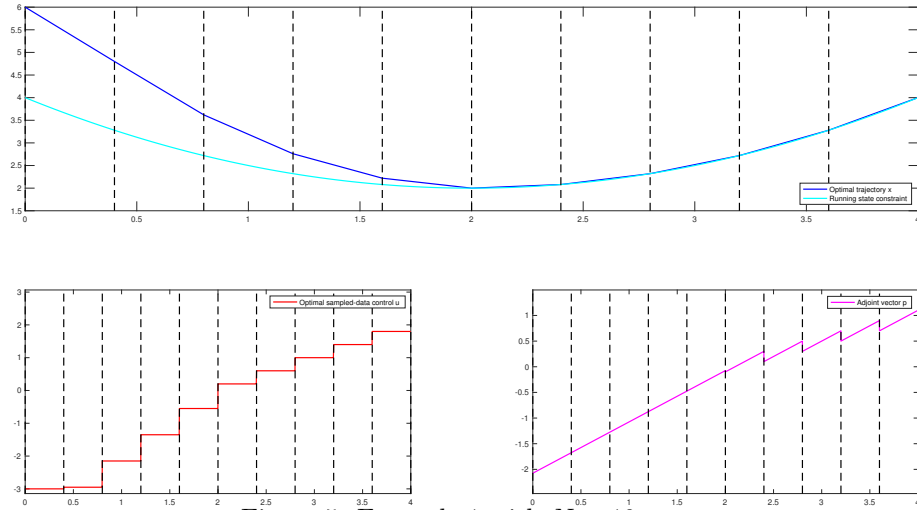


Figure 5: Example 1 with $N = 10$.

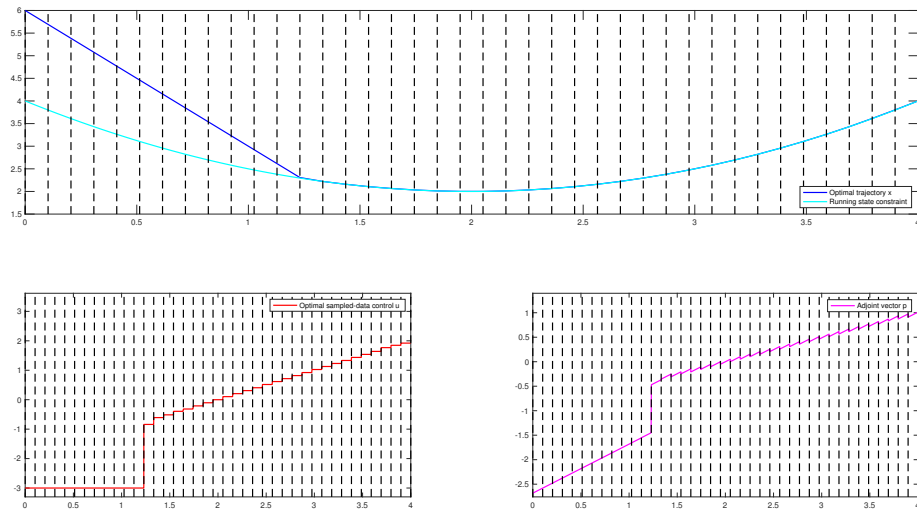


Figure 6: Example 1 with $N = 40$.

5.3 Example 2: an optimal consumption problem with an affine running inequality state constraint

The second example is the optimal sampled-data control problem with running inequality state constraint given by

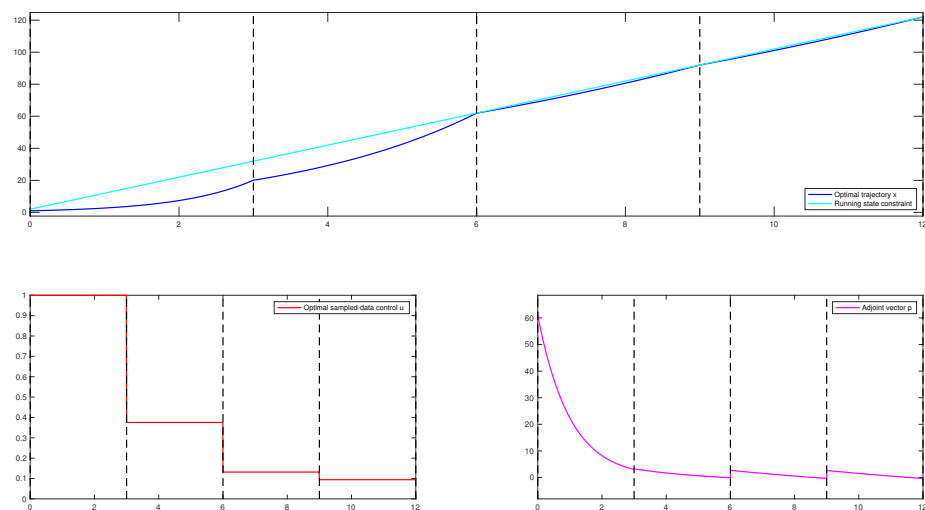
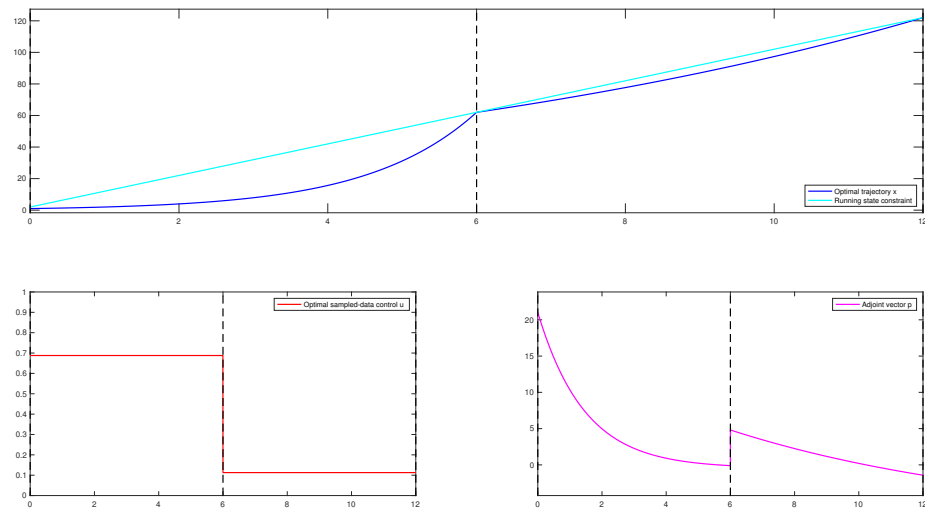
$$\left. \begin{array}{ll} \text{minimize} & \int_0^{12} (u(\tau) - 1)x(\tau) d\tau \\ \text{subject to} & x \in \text{AC}_1, u \in \text{PC}_1^{\mathbb{T}}, \\ & \dot{x}(t) = u(t)x(t) \text{ for a.e. } t \in [0, 12], \\ & x(0) = 1, \\ & x(t) - 10t - 2 \leq 0 \text{ for all } t \in [0, 12], \\ & u(t) \in [0, 1] \text{ for all } t \in [0, 12], \end{array} \right\} \quad (\text{E2})$$

for fixed uniform N -partitions \mathbb{T} of the interval $[0, 12]$ with different values of $N \in \mathbb{N}^*$. This problem corresponds to a classical optimal consumption problem (see, e.g., [30, p.5]) revisited with sampled-data controls and running inequality state constraint.

Let us verify that Problem (E2) satisfies Hypotheses (H1) and (H2). To this aim we follow the notations introduced in Section 4.2. Let us assume by contradiction that there exist $(y, t) \in \mathcal{M}$ and $v \in [0, 1]$ such that $\ell'(y, v, t) = +\infty$. Then it follows that $h^{[0]}(y, v, t) = h^{[1]}(y, v, t) = h^{[2]}(y, v, t) = 0$. From $h^{[0]}(y, v, t) = 0$, it holds that $y > 0$ and, from $h^{[1]}(y, v, t) = 0$, it holds that $v > 0$. Therefore $h^{[2]}(y, v, t) = v^2 y > 0$ which raises a contradiction. Thus Hypothesis (H1) is satisfied. From a similar reasoning, we prove that Hypothesis (H2) is also satisfied. We conclude from Proposition 4.1 that all admissible trajectories activate the running inequality state constraint at most at the sampling times t_i .

In what follows we assume that there exists an optimal couple (x, u) for Problem (E2) and we denote by p^0, η, p the elements provided by the Pontryagin maximum principle given in Theorem 3.1. Note that necessarily $x(t) > 0$ for all $t \in [0, 12]$. Let us verify that the case is normal (in the sense of Remark 3.4). Assume by contradiction that $p^0 = 0$. We have the adjoint equation $dp = d\eta$ over $[0, 12]$ with $p(12) = 0$. Therefore $p(t) = -\int_t^{12} d\eta(\tau) = \eta(t) - \eta(12)$ for all $t \in [0, 12]$. Then, from the nontriviality of the couple (p^0, η) , it follows that $\eta \neq 0_{\text{NBV}_1}$ and thus, from the complementary slackness condition, we deduce that x necessarily activates the running inequality state constraint. Let $\bar{t} \in [0, 12]$ denote the first activating time. From Proposition 4.1, we know that $\bar{t} = t_{\bar{i}}$ for some $\bar{i} \in \{0, \dots, N\}$. Since $x(0) = 1$, we know that $\bar{i} \geq 1$. It follows that $p(t) < 0$ for all $t \in [0, \bar{t}]$. Finally, since $x(t) > 0$ for all $t \in [0, 12]$ and from the nonpositive averaged Hamiltonian gradient condition at $i = 0, \dots, \bar{i} - 1$, we get that $u_0 = \dots = u_{\bar{i}-1} = 0$, which gives $x(t) = 1$ for all $t \in [0, \bar{t}]$. This raises a contradiction since $x(\bar{t}) = 10\bar{t} + 2 > 1$.

From the previous paragraph, we normalize $p^0 = -1$ (see Remark 3.4). Since we are in the context of Proposition 4.1, we can now apply the shooting method detailed in Section 5.1. In Figure 7 (with $N = 2$) we observe that the optimal trajectory returned by the algorithm activates the running inequality state constraint at most at the sampling times t_i (represented by dashed lines). Figures 8 and 9 continue to illustrate this bouncing trajectory phenomenon for larger values of N (respectively with $N = 4$ and $N = 6$). Furthermore, in Figures 8 and 9, we observe that the adjoint vector has no jump at sampling times which are not activating times.



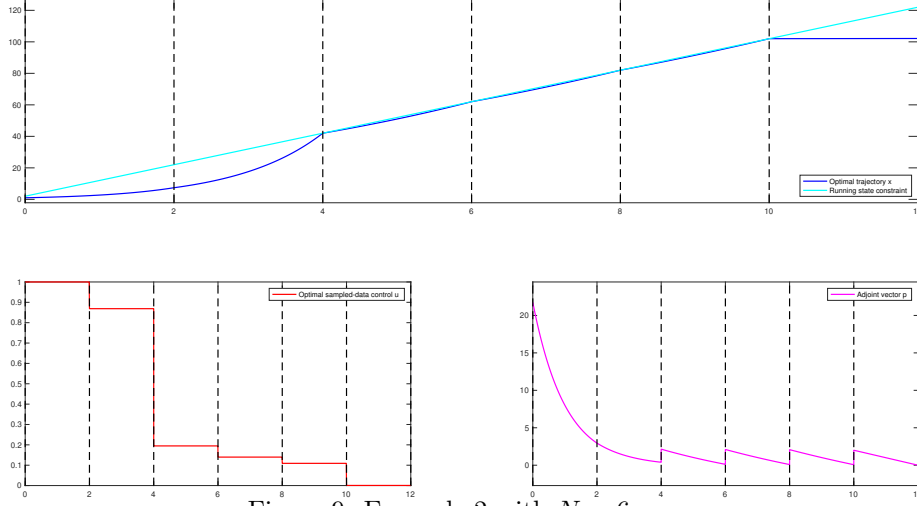


Figure 9: Example 2 with $N = 6$.

5.4 Example 3: a two-dimensional problem with a linear running inequality state constraint

As a third and last example we consider the optimal sampled-data control problem with running inequality state constraint given by

$$\left. \begin{aligned}
 &\text{minimize} && \int_0^2 x_2(\tau) + \frac{1}{4}u(\tau)^2 d\tau \\
 &\text{subject to} && x \in \text{AC}_2, u \in \text{PC}_1^{\mathbb{T}}, \\
 &&& \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix}(t) = \begin{pmatrix} x_2(t) - u(t) \\ x_1(t) + x_2(t) + u(t) \end{pmatrix} \text{ for a.e. } t \in [0, 2], \\
 &&& \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(0) = \begin{pmatrix} 0.05 \\ -0.1 \end{pmatrix}, \\
 &&& x_1(t) - 16x_2(t) - 2 \leq 0 \text{ for all } t \in [0, 2], \\
 &&& u(t) \in [-0.1, +\infty) \text{ for all } t \in [0, 2],
 \end{aligned} \right\} \quad (\text{E3})$$

for fixed uniform N -partitions \mathbb{T} of the interval $[0, 2]$ with different values of $N \in \mathbb{N}^*$. This problem constitutes a two-dimensional problem with a linear running inequality state constraint.

Let us verify that Problem (E3) satisfies Hypotheses (H1) and (H2). We denote by $y := (y_1, y_2) \in \mathbb{R}^2$ and we follow the notations introduced in Section 4.2. Let us assume by contradiction that there exist $(y, t) \in \mathcal{M}$ and $v \in [-0.1, +\infty)$ such that $\ell'(y, v, t) = +\infty$. Then it follows the system of linear equalities given by

$$\begin{aligned}
 h^{[0]}(y, v, t) &= y_1 - 16y_2 - 2 = 0, \\
 h^{[1]}(y, v, t) &= -16y_1 - 15y_2 - 17v = 0, \\
 h^{[2]}(y, v, t) &= -15y_1 - 31y_2 + v = 0,
 \end{aligned}$$

which has a unique solution given by $(y_1, y_2, v) = \frac{1}{9}(2, -1, -1)$ which raises a contradiction since $v \in [-0.1, +\infty)$. Thus Hypothesis (H1) is satisfied. If $h^{[0]}(y, v, t) = h^{[1]}(y, v, t) = 0$ for some $(y, t) \in \mathcal{M}$ and some $v \in [-0.1, +\infty)$, it follows that $-15y_1 - 31y_2 = 2 + 17v$. Therefore $h^{[2]}(y, v, t) = -15y_1 - 31y_2 + v = 2 + 18v \geq 0.2 > 0$. Thus Hypothesis (H2) is also satisfied. We conclude from Proposition 4.1 that all admissible trajectories activate the running inequality state constraint at most at the sampling times t_i .

In the sequel we assume that there exists an optimal couple (x, u) for Problem (E3) and we denote by p^0, η, p the elements provided by the Pontryagin maximum principle given in Theorem 3.1. We also assume that the case is normal and we normalize $p^0 = -1$ (see Remark 3.4). Since we are in the context of Proposition 4.1, we can now apply the shooting method detailed in Section 5.1. In Figure 10 (with $N = 4$) we observe as expected a bouncing trajectory phenomenon. Figures 11 and 12 give illustrations for larger values of N (respectively with $N = 5$ and $N = 8$).

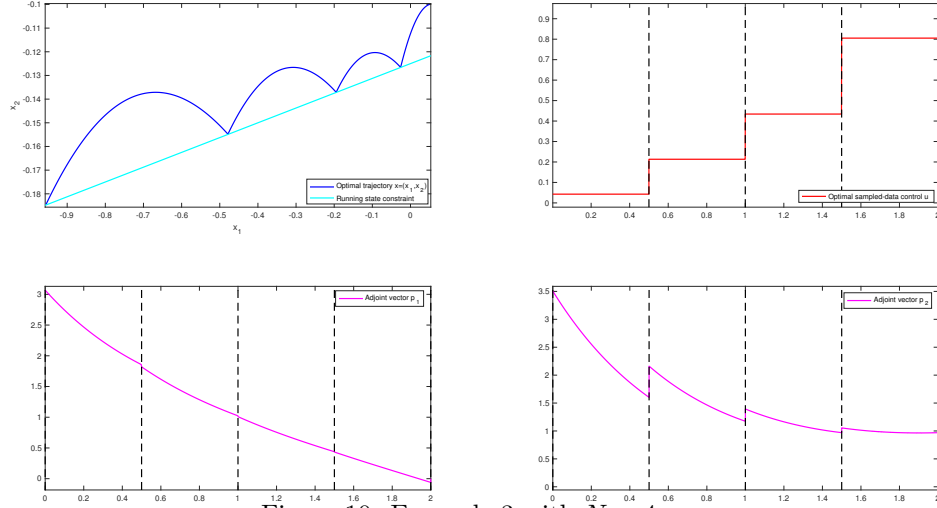


Figure 10: Example 3 with $N = 4$.

A Preliminaries for the proof of Theorem 3.1

This appendix is devoted to some required preliminaries for the proof of Theorem 3.1 found in Appendix B. In Section A.1 we give some recalls on renorming Banach spaces and on the regularity of distance functions. Section A.2 is concerned with the sensitivity analysis of the state equation in Problem (OSCP). Then, in Section A.3, we give some recalls on Stieltjes integrations and on Fubini formulas. Finally Section A.4 is devoted to Duhamel formulas for Cauchy-Stieltjes Problems (FCSP) and (BCSP).

A.1 About renorming Banach spaces and regularity of distance functions

Let $(Y, \|\cdot\|)$ be a normed space. We recall that the *dual space* of $(Y, \|\cdot\|)$, which we denote by $Y^* := \mathcal{L}((Y, \|\cdot\|), \mathbb{R})$, is the space of linear continuous forms on $(Y, \|\cdot\|)$. We recall that Y^* can

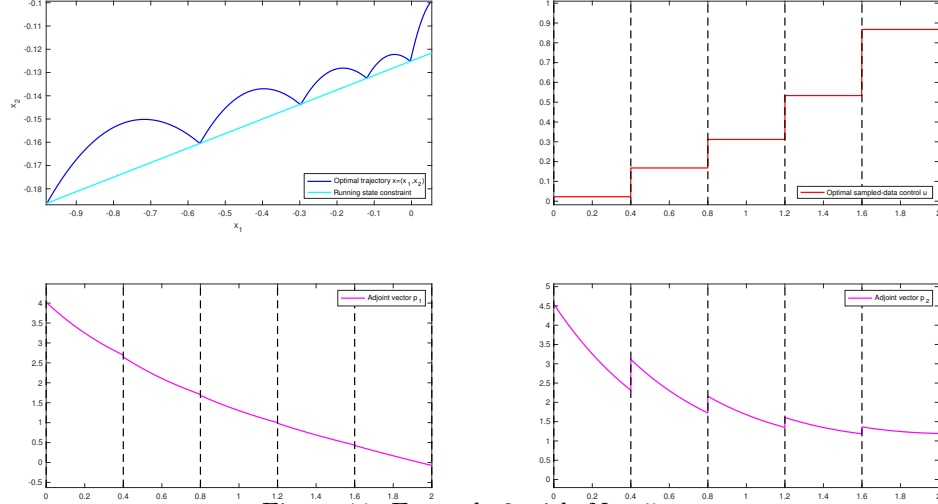


Figure 11: Example 3 with $N = 5$.

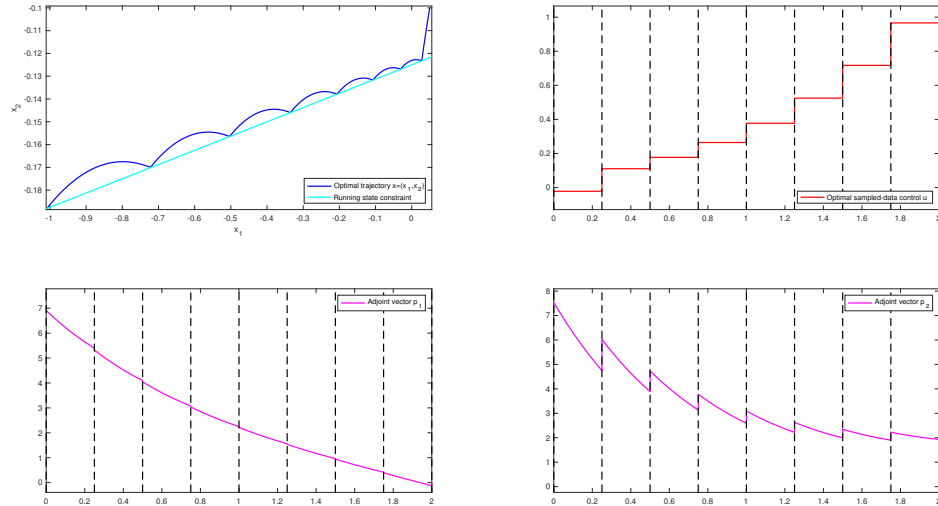


Figure 12: Example 3 with $N = 8$.

be endowed with the *dual norm* $\|\cdot\|_*$ defined by

$$\begin{aligned} \|\cdot\|_* : Y^* &\longrightarrow \mathbb{R}_+ \\ y^* &\longmapsto \|y^*\|_* := \sup_{\substack{y \in Y \\ \|y\| \leq 1}} |\langle y^*, y \rangle_{Y^* \times Y}|. \end{aligned}$$

In this situation we denote by $(Y^*, \|\cdot\|_*) := \text{dual}(Y, \|\cdot\|)$. We recall the following proposition on renorming separable Banach spaces.

Proposition A.1. *Let $(Y, \|\cdot\|)$ be a separable Banach space and let $(Y^*, \|\cdot\|_*) = \text{dual}(Y, \|\cdot\|)$. Then there exists a norm N on Y equivalent to $\|\cdot\|$ such that:*

- (i) N^* is equivalent to $\|\cdot\|_*$;

(ii) N^* is strictly convex;

where $(Y^*, N^*) = \text{dual}(Y, N)$.

Proof. We refer to [47, Theorem 2.18 p.42] or to [12, Proposition 4 p.16] for a complete proof. \square

Let $F : Y \rightarrow \mathbb{R}$ be a convex function. Recall that the *subdifferential* of F at a point $y \in Y$ is defined to be the set

$$\partial F(y) := \{y^* \in Y^* \mid \langle y^*, y' - y \rangle_{Y^* \times Y} \leq F(y') - F(y) \text{ for all } y' \in Y\}.$$

We recall that a function $F : Y \rightarrow \mathbb{R}$ is said to be *strictly Hadamard-differentiable* at a point $y \in Y$ with the strict Hadamard derivative $DF(y) \in Y^*$ if

$$\lim_{\substack{y' \rightarrow y \\ t \searrow 0}} \left[\sup_{y'' \in K} \left| \frac{F(y' + ty'') - F(y')}{t} - \langle DF(y), y'' \rangle_{Y^* \times Y} \right| \right] = 0,$$

for every compact set $K \subset Y$. We refer to [52, p.312-313] for more details. Finally we denote by $d_S : Y \rightarrow \mathbb{R}$ the *distance function* to a nonempty subset $S \subset Y$ defined by $d_S(y) := \inf_{y' \in S} \|y - y'\|$ for all $y \in Y$, and by $d_S^2 : Y \rightarrow \mathbb{R}$ the *squared distance function* defined by $d_S^2(y) := d_S(y)^2$ for all $y \in Y$. We recall the following proposition on the regularity of distance functions.

Proposition A.2. *Let $(Y, \|\cdot\|)$ be a normed space. Let $S \subset Y$ be a nonempty closed convex subset and let us assume that $\|\cdot\|^*$ is strictly convex, where $(Y^*, \|\cdot\|^*) := \text{dual}(Y, \|\cdot\|)$. Then it holds that:*

- (i) d_S is convex and 1-Lipschitz continuous;
- (ii) d_S is strictly Hadamard-differentiable on $Y \setminus S$ with $\|Dd_S(y)\|^* = 1$ and $\partial d_S(y) = \{Dd_S(y)\}$ for all $y \in Y \setminus S$;
- (iii) d_S^2 is strictly Hadamard-differentiable on $Y \setminus S$ with $Dd_S^2(y) = 2d_S(y)Dd_S(y)$ for all $y \in Y \setminus S$;
- (iv) d_S^2 is Fréchet-differentiable on S with $Dd_S^2(y) = 0_{Y^*}$ for all $y \in S$.

Proof. The proof of (i) is a standard result. We refer to [52, Theorem 3.54 p.313] and [12, Appendix B.2] for the proof of (ii). The proofs of (iii) and (iv) are straightforward. \square

A.2 About sensitivity analysis of the state equation in Problem (OSCP)

For all $u \in L_m^\infty$ we consider the Cauchy problem (CP_u) given by

$$\begin{cases} \dot{x}(t) = f(x(t), u(t), t) \text{ for a.e. } t \in [0, T], \\ x(0) = x_0. \end{cases} \quad (\text{CP}_u)$$

Before proceeding to the sensitivity analysis of the Cauchy problem (CP_u) with respect to the control u , we first recall some definitions and results from the classical Cauchy-Lipschitz (or Picard-Lindelöf) theory (see, e.g., [24]).

Definition A.1. *Let $u \in L_m^\infty$. A (local) solution to the Cauchy problem (CP_u) is a couple (x, I) such that:*

- (i) I is an interval such that $\{0\} \subsetneq I \subset [0, T]$;
- (ii) $x \in \text{AC}([0, T'], \mathbb{R}^n)$, with $\dot{x}(t) = f(x(t), u(t), t)$ for almost every $t \in [0, T']$, for all $T' \in I$;
- (iii) $x(0) = x_0$.

Let (x_1, I_1) and (x_2, I_2) be two (local) solutions to the Cauchy problem (CP_u) . 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_u) is a (local) solution that does not admit any strict extension. Finally a global solution to the Cauchy problem (CP_u) is a solution (x, I) such that $I = [0, T]$.

Proposition A.3. For all $u \in L_m^\infty$, the Cauchy problem (CP_u) admits a unique maximal solution, denoted by $(x(\cdot, u), I(u))$, which is an extension to any other local solution.

We now introduce the notion of controls admissible for globality.

Definition A.2. A control $u \in L_m^\infty$ is said to be admissible for globality if the corresponding maximal solution $(x(\cdot, u), I(u))$ is global, that is, if $I(u) = [0, T]$. In what follows we denote by $\mathcal{AG} \subset L_m^\infty$ the set of all controls admissible for globality.

In the following lemma we state a continuous dependence result for the trajectory $x(\cdot, u)$ with respect to the control u . In particular we prove that \mathcal{AG} is open.

Lemma A.1. For all $u \in \mathcal{AG}$, there exists $\varepsilon_u > 0$ such that $\bar{B}_{L_m^\infty}(u, \varepsilon_u) \subset \mathcal{AG}$, where $\bar{B}_{L_m^\infty}(u, \varepsilon_u)$ stands for the standard closed ball in L_m^∞ centered at u and of radius ε_u . Moreover the map

$$u' \in (\bar{B}_{L_m^\infty}(u, \varepsilon_u), \|\cdot\|_{L_m^\infty}) \mapsto x(\cdot, u') \in (C_n, \|\cdot\|_\infty),$$

is Lipschitz continuous.

Proof. This proof is standard and essentially based on the classical Gronwall lemma. We refer to [14, Lemmas 1 and 3 p.3795-3797], [16, Lemmas 4.3 and 4.5 p.73-74] (in the general framework of time scale calculus) or to [12, Propositions 1 and 2 p.4-5] (in a more classical framework, closer to the present considerations) for similar statements with detailed proofs. \square

Remark A.1. Let $u \in \mathcal{AG}$ and $\varepsilon_u > 0$ as given in Lemma A.1. Let $u' \in \bar{B}_{L_m^\infty}(u, \varepsilon_u)$ and $(u_k)_{k \in \mathbb{N}}$ be a sequence in $\bar{B}_{L_m^\infty}(u, \varepsilon_u)$ converging to u' in L_m^∞ . From Lemma A.1, we deduce that the sequence $(x(\cdot, u_k))_{k \in \mathbb{N}}$ uniformly converges to $x(\cdot, u')$ over $[0, T]$.

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

Proposition A.4. Let $u \in \mathcal{AG}$ and let $z \in L_m^\infty$. We consider the convex L^∞ -perturbation of u given by

$$u_z(\cdot, \rho) := u + \rho(z - u),$$

for all $\rho \in [0, 1]$. Then:

- (i) there exists $0 < \rho_0 \leq 1$ such that $u_z(\cdot, \rho) \in \mathcal{AG}$ for all $\rho \in [0, \rho_0]$;
- (ii) the map

$$\rho \in ([0, \rho_0], |\cdot|) \mapsto x(\cdot, u_z(\cdot, \rho)) \in (C_n, \|\cdot\|_\infty),$$

is differentiable at $\rho = 0$ and its derivative is equal to the variation vector $w_z(\cdot, u) \in \text{AC}_n$ being the unique solution (that is global) to the linearized Cauchy problem given by

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

Proof. This proof is standard and essentially based on the classical Gronwall lemma. We refer to [14, Lemma 4 and Proposition 1 p.3797-3798] for a similar statement with detailed proof. \square

We conclude this section by a technical lemma on the convergence of variation vectors which is required in the proof of our main result.

Lemma A.2. *Let $u \in \mathcal{AG}$ and $\varepsilon_u > 0$ as in Lemma A.1. Let $z \in L_m^\infty$. Let $u' \in \overline{B}_{L_m^\infty}(u, \varepsilon_u)$ and $(u_k)_{k \in \mathbb{N}}$ be a sequence in $\overline{B}_{L_m^\infty}(u, \varepsilon_u)$ converging to u' in L_m^∞ . Then the sequence $(w_z(\cdot, u_k))_{k \in \mathbb{N}}$ uniformly converges to $w_z(\cdot, u')$ over $[0, T]$.*

Proof. This proof is standard and essentially based on the classical Gronwall lemma. We refer to [14, Lemmas 4.8 and 4.9 p.77-78] for a similar statement with detailed proof. \square

A.3 About Stieltjes integrations and Fubini formulas

In this section our aim is to recall some notions on Stieltjes integrations and to recall some Fubini formulas. We refer to standard references and books such as [4, 19, 20, 32, 65] for more details. We also refer to [12, Appendix C] and references therein. In the sequel we denote by $C_n^+ := C([0, T], (\mathbb{R}_+)^n)$ where $\mathbb{R}_+ := [0, +\infty)$. We denote by C_n^* as the dual space of C_n (see Section A.1 for some details on dual spaces). We first recall the following Riesz representation theorem (see [48, Theorem 14.5 p.245-246] or [12, Proposition 7 p.19]).

Proposition A.5. (Riesz representation theorem). *Let $\psi^* \in C_1^*$. Then there exists a unique $\eta \in \text{NBV}_1$ such that:*

$$\langle \psi^*, \psi \rangle_{C_1^* \times C_1} = \int_0^T \psi(\tau) d\eta(\tau),$$

for all $\psi \in C_1$. Moreover it holds that:

- (i) $\langle \psi^*, \psi \rangle_{C_1^* \times C_1} \geq 0$ for all $\psi \in C_1^+$ if and only if η is monotonically increasing on $[0, T]$;
- (ii) $\psi^* = 0_{C_1^*}$ if and only if $\eta = 0_{\text{NBV}_1}$.

Recall that if $\eta \in \text{NBV}_1$ is monotonically increasing on $[0, T]$, then η induces a finite nonnegative Borel measure $d\eta$ on $[0, T]$ by defining $d\eta(\{0\}) := \eta(0^+)$ and $d\eta((a, b]) := \eta(b) - \eta(a)$ for all semiopen intervals $(a, b] \subset [0, T]$ and by using the Carathéodory extension theorem. Furthermore, for all $\psi \in C_1$, the Riemann-Stieltjes integral $\int_a^b \psi(\tau) d\eta(\tau)$ coincides with the Lebesgue-Stieltjes integral $\int_{(a, b]} \psi(\tau) d\eta(\tau)$ for all $0 \leq a \leq b \leq T$. We refer to [32, p.83] and [65, p.288] for more details. Consequently the Fubini formula

$$\int_0^T \int_0^\tau \Psi(\tau, s) ds d\eta(\tau) = \int_0^T \int_s^T \Psi(\tau, s) d\eta(\tau) ds, \quad (4)$$

holds for all $\Psi \in L^\infty([0, T]^2, \mathbb{R})$ such that Ψ is continuous in its first variable.

We now introduce some notations for Riemann-Stieltjes integrals with respect to vectorial functions of bounded variations. We denote by

$$\int_0^T \langle \psi(\tau), d\eta(\tau) \rangle := \sum_{j=1}^q \int_0^T \psi_j(\tau) d\eta_j(\tau) \in \mathbb{R},$$

for all $\psi = (\psi_j)_{j=1, \dots, q} \in C_q$ and all $\eta = (\eta_j)_{j=1, \dots, q} \in \text{BV}_q$. Moreover we denote by

$$\int_0^T M(\tau) \times d\eta(\tau) := \left(\sum_{j=1}^q \int_0^T m_{rj}(\tau) d\eta_j(\tau) \right)_{r=1, \dots, n} \in \mathbb{R}^n,$$

and

$$\int_0^T \langle \psi(\tau), M(\tau) \times d\eta(\tau) \rangle := \int_0^T \langle M(\tau)^\top \times \psi(\tau), d\eta(\tau) \rangle \in \mathbb{R},$$

for all $\psi = (\psi_r)_{r=1,\dots,n} \in \mathbb{C}_n$, all $\eta = (\eta_j)_{j=1,\dots,q} \in \text{BV}_q$ and all continuous matrices $M = (m_{rj})_{rj} : [0, T] \rightarrow \mathbb{R}^{n \times q}$. In particular one can easily prove that, if $\psi \in \mathbb{R}^n$ (i.e. $\psi \in \mathbb{C}_n$ constant), then

$$\int_0^T \langle \psi, M(\tau) \times d\eta(\tau) \rangle = \left\langle \psi, \int_0^T M(\tau) \times d\eta(\tau) \right\rangle_{\mathbb{R}^n},$$

for all $\eta = (\eta_j)_{j=1,\dots,q} \in \text{BV}_q$ and all continuous matrices $M = (m_{rj})_{rj} : [0, T] \rightarrow \mathbb{R}^{n \times q}$.

Finally, from the Fubini formula (4) and the above notations, one can easily deduce that the Fubini formula

$$\int_0^T \left\langle \int_0^\tau \Psi(\tau, s) ds, d\eta(\tau) \right\rangle = \int_0^T \int_s^T \langle \Psi(\tau, s), d\eta(\tau) \rangle ds,$$

holds for all $\Psi \in L^\infty([0, T]^2, \mathbb{R}^q)$ being continuous in its first variable and for all $\eta = (\eta_j)_{j=1,\dots,q} \in \text{NBV}_q$ such that η_j is monotonically increasing on $[0, T]$ for each $j = 1, \dots, q$.

A.4 About Problems (FCSP) and (BCSP) and Duhamel formulas

Let us consider the framework and the notations introduced in Section 2.2. Our aim in this section is to provide Duhamel formulas for the solutions to Problems (FCSP) and (BCSP). To this aim, we recall that the *state-transition matrix* $\Phi(\cdot, \cdot) : [0, T]^2 \rightarrow \mathbb{R}^{n \times n}$ associated to $A \in L^\infty([0, T], \mathbb{R}^{n \times n})$ is defined as follows. For all $s \in [0, T]$, $\Phi(\cdot, s)$ is the unique solution (that is global) to the linear forward/backward Cauchy problem given by

$$\begin{cases} \dot{\Phi}(t) = A(t) \times \Phi(t) \text{ for a.e. } t \in [0, T], \\ \Phi(s) = \text{Id}_n. \end{cases}$$

The equalities

$$\Phi(t, s) = \text{Id}_n + \int_s^t A(\tau) \times \Phi(\tau, s) d\tau = \text{Id}_n + \int_s^t \Phi(t, \tau) \times A(\tau) d\tau,$$

both hold for all $(t, s) \in [0, T]^2$. From these two equalities and the Fubini formulas from Section A.3, one can easily derive the following proposition. We also refer to [12, Appendix D] for some details.

Proposition A.6 (Duhamel formulas). *The solutions to (FCSP) and (BCSP) are given by*

$$x(t) = \Phi(t, 0) \times x_0 + \int_0^t \Phi(t, \tau) \times B(\tau) d\tau + \sum_{j=1}^q \int_0^t \Phi(t, \tau) \times C_j(\tau) d\eta_j(\tau),$$

and

$$p(t) = \Phi(T, t)^\top \times p_T + \int_t^T \Phi(\tau, t)^\top \times B(\tau) d\tau + \sum_{j=1}^q \int_t^T \Phi(\tau, t)^\top \times C_j(\tau) d\eta_j(\tau),$$

for all $t \in [0, T]$, where $\Phi(\cdot, \cdot)$ stands for the state-transition matrix associated to A .

B Proof of Theorem 3.1

This appendix is dedicated to the detailed proof of Theorem 3.1. Section B.1 deals with the case $L = 0$ (the case $L \neq 0$ is treated in Section B.2 with a simple change of variable). In Section B.1.1 the Ekeland variational principle is applied on an appropriate penalized functional in order to derive a crucial inequality (see Inequality (9)). In Section B.1.2 we conclude the proof of Theorem 3.1 by introducing the adjoint vector p .

We first remark that the running inequality state constraints in Problem (OSCP) can be written as $\mathfrak{h}(x) \in \text{S}$ where:

- $\mathfrak{h} : C_n \rightarrow C_q$ is defined as $\mathfrak{h}(x) := h(x, \cdot)$ for all $x \in C_n$. Note that \mathfrak{h} is of class C^1 with $D\mathfrak{h}(x)(x') = \partial_1 h(x, \cdot) \times x'$ for all $x, x' \in C_n$;
- $S := C([0, T], (\mathbb{R}_-)^q)$ where $\mathbb{R}_- := (-\infty, 0]$. We emphasize that $S \subset C_q$ is a nonempty closed convex cone of C_q with a nonempty interior.

Recall that $(C_q, \|\cdot\|_\infty)$ is a separable Banach space. Applying Proposition A.1, we endow C_q with an equivalent norm $\|\cdot\|_{C_q}$ such that the associated dual norm $\|\cdot\|_{C_q^*}$ is strictly convex. We denote by $d_S : C_q \rightarrow \mathbb{R}$ the 1-Lipschitz continuous distance function to S (see Section A.1). Then, from Proposition A.2, we know that d_S and d_S^2 are strictly Hadamard-differentiable on $C_q \setminus S$ with $Dd_S^2(x) = 2d_S(x)Dd_S(x)$ and $\|Dd_S(x)\|_{C_q^*} = 1$ for all $x \in C_q \setminus S$, and that d_S^2 is Fréchet-differentiable on S with $Dd_S^2(x) = 0_{C_q^*}$ for all $x \in S$.

B.1 The case $L = 0$

In the whole section we will assume that $L = 0$ in Problem (OSCP) (see Section B.2 for the case $L \neq 0$). Let $(x, u) \in AC_n \times PC_m^\mathbb{T}$ be a solution to Problem (OSCP). Following the notation introduced in Section A.2, it holds that $u \in \mathcal{AG}$ and that $x = x(\cdot, u)$. In what follows we will also consider the positive real number $\varepsilon_u > 0$ given in Lemma A.1.

B.1.1 Application of the Ekeland variational principle

Let us recall a simplified version (but sufficient for our purposes) of the Ekeland variational principle (see [29]).

Proposition B.1 (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 $e \in E$ such that $\mathcal{J}(e) = \varepsilon$. Then there exists $e_\varepsilon \in E$ such that $d_E(e_\varepsilon, e) \leq \sqrt{\varepsilon}$, and $-\sqrt{\varepsilon} d_E(e', e_\varepsilon) \leq \mathcal{J}(e') - \mathcal{J}(e_\varepsilon)$ for every $e' \in E$.*

We introduce the set

$$E_u := \{u' \in \overline{B}_{L_m^\infty}(u, \varepsilon_u) \mid u' \in PC_m^\mathbb{T} \text{ and } u'(t) \in U \text{ for all } t \in [0, T]\}.$$

From the closedness assumption on U , one can easily prove that $(E_u, \|\cdot\|_{L_m^\infty})$ is a complete metric set. Let us choose a sequence $(\varepsilon_k)_{k \in \mathbb{N}}$ such that $0 < \sqrt{\varepsilon_k} < \varepsilon_u$ for all $k \in \mathbb{N}$ and satisfying $\lim_{k \rightarrow \infty} \varepsilon_k = 0$. We introduce the penalized functional

$$\begin{aligned} \mathcal{J}_k : E_u &\longrightarrow \mathbb{R}_+ \\ u' &\longmapsto \mathcal{J}_k(u') := \sqrt{\left(\left(g(x(T, u')) - g(x(T)) + \varepsilon_k \right)^+ \right)^2 + d_S^2(\mathfrak{h}(x(\cdot, u')))}, \end{aligned}$$

for all $k \in \mathbb{N}$. From Lemma A.1, note that \mathcal{J}_k is correctly defined for all $k \in \mathbb{N}$. Also, from Lemma A.1 and from the continuities of g , \mathfrak{h} and d_S^2 (see Proposition A.2), it follows that \mathcal{J}_k is continuous as well for all $k \in \mathbb{N}$. Note that \mathcal{J}_k is nonnegative and, since the constraint $\mathfrak{h}(x) \in S$ is satisfied, it holds that $\mathcal{J}_k(u) = \varepsilon_k$ for all $k \in \mathbb{N}$. Therefore, from the Ekeland variational principle (see Proposition B.1), we conclude that there exists a sequence $(u_k)_{k \in \mathbb{N}} \subset E_u$ such that

$$\|u_k - u\|_{L_m^\infty} \leq \sqrt{\varepsilon_k}, \quad (5)$$

and

$$-\sqrt{\varepsilon_k} \|u' - u_k\|_{L_m^\infty} \leq \mathcal{J}_k(u') - \mathcal{J}_k(u_k), \quad (6)$$

for all $u' \in E_u$ and all $k \in \mathbb{N}$. In particular, from Inequality (5), note that the sequence $(u_k)_{k \in \mathbb{N}}$ converges to u in L_m^∞ . From optimality of the couple (x, u) , note that $\mathcal{J}_k(u') > 0$ for all $u' \in E_u$ and all $k \in \mathbb{N}$. We thus define correctly the couple $(\lambda_k, \psi_k^*) \in \mathbb{R} \times C_q^*$ as

$$\lambda_k := \frac{1}{\mathcal{J}_k(u_k)} \left(g(x(T, u_k)) - g(x(T)) + \varepsilon_k \right)^+ \geq 0,$$

and

$$\psi_k^* := \begin{cases} \frac{1}{\mathcal{J}_k(u_k)} \mathrm{d}_S(\mathfrak{h}(x(\cdot, u_k))) D\mathrm{d}_S(\mathfrak{h}(x(\cdot, u_k))) & \text{if } \mathfrak{h}(x(\cdot, u_k)) \notin S, \\ 0_{C_q^*} & \text{if } \mathfrak{h}(x(\cdot, u_k)) \in S, \end{cases}$$

for all $k \in \mathbb{N}$. From Proposition A.2 it holds that $|\lambda_k|^2 + \|\psi_k^*\|_{C_q^*}^2 = 1$ for all $k \in \mathbb{N}$. As a consequence, we can extract subsequences (which we do not relabel) such that $(\lambda_k)_{k \in \mathbb{N}}$ converges to some $\lambda \geq 0$ and $(\psi_k^*)_{k \in \mathbb{N}}$ weakly* converges to some $\psi^* \in C_q^*$. In particular it holds that $|\lambda|^2 + \|\psi^*\|_{C_q^*}^2 \leq 1$. At this step note that we cannot ensure that the couple (λ, ψ^*) is not trivial. The nontriviality is guaranteed by the next proposition.

Proposition B.2. *The couple $(\lambda, \psi^*) \in \mathbb{R} \times C_q^*$ is nontrivial and it holds that*

$$\langle \psi^*, \psi - \mathfrak{h}(x) \rangle_{C_q^* \times C_q} \leq 0, \quad (7)$$

for all $\psi \in S$.

Proof. Let $k \in \mathbb{N}$ be fixed. From Proposition A.2, if $\mathfrak{h}(x(\cdot, u_k)) \notin S$, then $D\mathrm{d}_S(\mathfrak{h}(x(\cdot, u_k))) \in \partial \mathrm{d}_S(\mathfrak{h}(x(\cdot, u_k)))$. Hence, if $\mathfrak{h}(x(\cdot, u_k)) \notin S$, it holds that

$$\left\langle D\mathrm{d}_S(\mathfrak{h}(x(\cdot, u_k))), \psi - \mathfrak{h}(x(\cdot, u_k)) \right\rangle_{C_q^* \times C_q} \leq \mathrm{d}_S(\psi) - \mathrm{d}_S(\mathfrak{h}(x(\cdot, u_k))) \leq 0,$$

for all $\psi \in S$. As a consequence, in both cases $\mathfrak{h}(x(\cdot, u_k)) \in S$ and $\mathfrak{h}(x(\cdot, u_k)) \notin S$, it holds that

$$\langle \psi_k^*, \psi - \mathfrak{h}(x(\cdot, u_k)) \rangle_{C_q^* \times C_q} \leq 0, \quad (8)$$

for all $\psi \in S$. Using Lemma A.1 and taking the limit as k tends to $+\infty$, we get Inequality (7). Now let us prove that the couple $(\lambda, \psi^*) \in \mathbb{R} \times C_q^*$ is nontrivial. Since S has a nonempty interior, there exists $\xi \in S$ and $\delta > 0$ such that $\xi + \delta\psi \in S$ for all $\psi \in \overline{B}_{C_q}(0_{C_q}, 1)$. Hence we obtain from Inequality (8) that

$$\delta \langle \psi_k^*, \psi \rangle_{C_q^* \times C_q} \leq \langle \psi_k^*, \mathfrak{h}(x(\cdot, u_k)) - \xi \rangle_{C_q^* \times C_q},$$

for all $\psi \in \overline{B}_{C_q}(0_{C_q}, 1)$ and all $k \in \mathbb{N}$. We deduce that

$$\delta \|\psi_k^*\|_{C_q^*} = \delta \sqrt{1 - |\lambda_k|^2} \leq \langle \psi_k^*, \mathfrak{h}(x(\cdot, u_k)) - \xi \rangle_{C_q^* \times C_q},$$

for all $k \in \mathbb{N}$. Using Lemma A.1 and taking the limit as k tends to $+\infty$, we obtain that

$$\delta \sqrt{1 - |\lambda|^2} \leq \langle \psi^*, \mathfrak{h}(x) - \xi \rangle_{C_q^* \times C_q}.$$

Since $\delta > 0$, the last inequality implies that the couple (λ, ψ^*) is nontrivial which completes the proof. \square

Finally, in the next result, we use Inequality (6) with convex L^∞ -perturbations of the control u_k in order to establish a crucial inequality.

Proposition B.3. *The inequality*

$$\lambda \left\langle \nabla g(x(T)), w_z(T, u) \right\rangle_{\mathbb{R}^n} + \left\langle \psi^*, \partial_1 h(x, \cdot) \times w_z(\cdot, u) \right\rangle_{C_q^* \times C_q} \geq 0, \quad (9)$$

holds for all $z \in \mathrm{PC}_m^{\mathbb{T}}$ with values in U , where $w_z(\cdot, u)$ is the variation vector defined in Proposition A.4.

Proof. Let $z \in \text{PC}_m^\mathbb{T}$ with values in U . We fix $k \in \mathbb{N}$. Since U is convex, it is clear that the convex L^∞ -perturbation of the control u_k associated to z , defined by $u_{k,z}(t, \rho) := u_k(t) + \rho(z(t) - u_k(t))$ for all $t \in [0, T]$ and all $0 \leq \rho \leq 1$, belongs to $\text{PC}_m^\mathbb{T}$ and takes values in U . Furthermore it holds that $\|u_{k,z}(\cdot, \rho) - u\|_{L_m^\infty} \leq \rho\|z - u_k\|_{L_m^\infty} + \|u_k - u\|_{L_m^\infty} \leq \rho\|z - u_k\|_{L_m^\infty} + \sqrt{\varepsilon_k}$. Since $\sqrt{\varepsilon_k} < \varepsilon_u$, we deduce that $u_{k,z}(\cdot, \rho) \in E_u$ for small enough $\rho > 0$. From Inequality (6) we get that

$$-\sqrt{\varepsilon_k} \|z - u_k\|_{L_m^\infty} \leq \frac{1}{\mathcal{J}_k(u_{k,z}(\cdot, \rho)) + \mathcal{J}_k(u_k)} \times \frac{\mathcal{J}_k(u_{k,z}(\cdot, \rho))^2 - \mathcal{J}_k(u_k)^2}{\rho},$$

for small enough $\rho > 0$. From Proposition A.4 and from strict Hadamard-differentiability of d_S^2 over $C_q \setminus S$ and Fréchet-differentiability of d_S^2 over S (see Proposition A.2), taking the limit as ρ tends to 0, we get that

$$\begin{aligned} -\sqrt{\varepsilon_k} \|z - u_k\|_{L_m^\infty} &\leq \frac{1}{2\mathcal{J}_k(u_k)} \left[2 \left(g(x(T, u_k)) - g(x(T)) + \varepsilon_k \right)^+ \left\langle \nabla g(x(T, u_k)), w_z(T, u_k) \right\rangle_{\mathbb{R}^n} \right. \\ &\quad \left. + \left\langle 2d_S(h(x(\cdot, u_k))) Dd_S(h(x(\cdot, u_k))), \partial_1 h(x(\cdot, u_k), \cdot) \times w_z(\cdot, u_k) \right\rangle_{C_q^* \times C_q} \right], \end{aligned}$$

with the convention that the second term on the right-hand side is zero if $h(x(\cdot, u_k)) \in S$. Using the definition of λ_k and ψ_k^* , we deduce that

$$-\sqrt{\varepsilon_k} \|z - u_k\|_{L_m^\infty} \leq \lambda_k \left\langle \nabla g(x(T, u_k)), w_z(T, u_k) \right\rangle_{\mathbb{R}^n} + \left\langle \psi_k^*, \partial_1 h(x(\cdot, u_k), \cdot) \times w_z(\cdot, u_k) \right\rangle_{C_q^* \times C_q}.$$

We take the limit of this inequality as k tends to $+\infty$. From the smoothness of g and h and from Lemmas A.1 and A.2, Inequality (9) is proved. \square

B.1.2 Introduction of the adjoint vector

We are now in a position to conclude the proof of Theorem 3.1 (in the case $L = 0$) by introducing the adjoint vector p . We refer to Sections 2.2, A.3 and A.4 for notations and background concerning Stieltjes integrations and linear Cauchy-Stieltjes problems.

Introduction of the nontrivial couple (p^0, η) and complementary slackness condition.

We introduce $p^0 := -\lambda \leq 0$ and we write $\psi^* = (\psi_j^*)_{j=1, \dots, q}$ where $\psi_j^* \in C_1^*$ for every $j = 1, \dots, q$. From the Riesz representation theorem (see Proposition A.5), there exists a unique $\eta_j \in \text{NBV}_1$ such that

$$\langle \psi_j^*, \psi \rangle_{C_1^* \times C_1} = \int_0^T \psi(\tau) d\eta_j(\tau),$$

for all $\psi \in C_1$ and all $j = 1, \dots, q$. Furthermore $\psi_j^* = 0_{C_1^*}$ if and only if $\eta_j = 0_{\text{NBV}_1}$. Thus it follows from Proposition B.2 that the couple (p^0, η) is not trivial, where $\eta := (\eta_j)_{j=1, \dots, q} \in \text{NBV}_q$. Moreover, from Inequality (7) (and the fact that S is a cone containing $h(x)$), one can easily deduce that $\langle \psi_j^*, h_j(x) \rangle_{C_1^* \times C_1} = 0$, that is,

$$\int_0^T h_j(x(\tau), \tau) d\eta_j(\tau) = 0,$$

for all $j = 1, \dots, q$. Finally one can similarly deduce from Inequality (7) that $\langle \psi_j^*, \psi \rangle_{C_1^* \times C_1} \geq 0$ for all $\psi \in C_1^+$ and all $j = 1, \dots, q$. From Proposition A.5, it follows that η_j is monotonically increasing on $[0, T]$ for all $j = 1, \dots, q$.

Adjoint equation. We define the adjoint vector $p \in \text{BV}_n$ as the unique solution to the backward linear Cauchy-Stieltjes problem given by

$$\begin{cases} -dp = \left(\partial_1 f(x, u, \cdot)^\top \times p \right) dt - \sum_{j=1}^q \partial_1 h_j(x, \cdot) d\eta_j & \text{over } [0, T], \\ p(T) = p^0 \nabla g(x(T)). \end{cases}$$

From the Duhamel formula for backward linear Cauchy-Stieltjes problems (see Proposition A.6) and using notations introduced in Section A.3, it holds that

$$p(t) = \Phi(T, t)^\top \times \left(p^0 \nabla g(x(T)) \right) - \int_t^T \Phi(\tau, t)^\top \times \partial_1 h(x(\tau), \tau)^\top \times d\eta(\tau),$$

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

Nonpositive averaged Hamiltonian gradient condition. From Inequality (9) and using notations introduced in Section A.3, it holds that

$$\lambda \left\langle \nabla g(x(T)), w_z(T, u) \right\rangle_{\mathbb{R}^n} + \int_0^T \left\langle \partial_1 h(x(\tau), \tau) \times w_z(\tau, u), d\eta(\tau) \right\rangle \geq 0,$$

for all $z \in \text{PC}_m^\mathbb{T}$ with values in U . From the definition of the variation vector $w_z(\cdot, u)$ and the classical Duhamel formula for standard forward linear Cauchy problems, it holds that

$$w_z(\tau, u) = \int_0^\tau \Phi(\tau, s) \times \partial_2 f(x(s), u(s), s) \times (z(s) - u(s)) ds,$$

for all $\tau \in [0, T]$. Substituting this expression into the previous inequality and using the last Fubini formula given in Section A.3, it follows that

$$\begin{aligned} & \int_0^T \left\langle \Phi(T, s)^\top \times \left(p^0 \nabla g(x(T)) \right), \partial_2 f(x(s), u(s), s) \times (z(s) - u(s)) \right\rangle_{\mathbb{R}^n} ds \\ & - \int_0^T \left\langle \partial_2 f(x(s), u(s), s) \times (z(s) - u(s)), \int_s^T \Phi(\tau, s)^\top \times \partial_1 h(x(\tau), \tau)^\top \times d\eta(\tau) \right\rangle_{\mathbb{R}^n} ds \leq 0, \end{aligned}$$

for all $z \in \text{PC}_m^\mathbb{T}$ with values in U . Finally, grouping like terms, we exactly obtain

$$\int_0^T \left\langle p(s), \partial_2 f(x(s), u(s), s) \times (z(s) - u(s)) \right\rangle_{\mathbb{R}^n} ds \leq 0,$$

for all $z \in \text{PC}_m^\mathbb{T}$ with values in U . For all $i = 0, \dots, N-1$ and all $v \in \text{U}$, let us consider $z_{i,v} \in \text{PC}_m^\mathbb{T}$ with values in U as

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

for all $s \in [0, T]$. Substituting z by $z_{i,v}$ in the above inequality and from the definition of the Hamiltonian H , we exactly get that

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

for all $v \in \text{U}$ and all $i = 0, \dots, N-1$. The proof of Theorem 3.1 is complete (in the case $L = 0$).

B.2 The case $L \neq 0$

In the previous section we have proved Theorem 3.1 in the case $L = 0$ (without Lagrange cost). This section is dedicated to the case $L \neq 0$. Let $(x, u) \in \text{AC}_n \times \text{PC}_m^{\mathbb{T}}$ be a solution to Problem (OSCP). Let us introduce

$$X(t) := \int_0^t L(x(\tau), u(\tau), \tau) d\tau,$$

for all $t \in [0, T]$. We see that the augmented couple $((x, X), u) \in \text{AC}_{n+1} \times \text{PC}_m^{\mathbb{T}}$ is a solution to the *augmented* optimal sampled-data control problem with running inequality state constraints of Mayer form given by

$$\left. \begin{aligned} & \text{minimize} && \tilde{g}((x, X)(T)) \\ & \text{subject to} && (x, X) \in \text{AC}_{n+1}, u \in \text{PC}_m^{\mathbb{T}}, \\ & && \begin{pmatrix} \dot{x} \\ X \end{pmatrix}(t) = \begin{pmatrix} f(x(t), u(t), t) \\ L(x(t), u(t), t) \end{pmatrix} \text{ for a.e. } t \in [0, T], \\ & && \begin{pmatrix} x \\ X \end{pmatrix}(0) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}, \\ & && \tilde{h}_j((x, X)(t), t) \leq 0 \text{ for all } t \in [0, T] \text{ and all } j = 1, \dots, q, \\ & && u(t) \in U \text{ for all } t \in [0, T], \end{aligned} \right\} \quad (\text{OSCP}_{aug})$$

where $\tilde{g} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is defined by $\tilde{g}(x_1, X_1) := g(x_1) + X_1$ for all $(x_1, X_1) \in \mathbb{R}^{n+1}$ and where $\tilde{h} : \mathbb{R}^{n+1} \times [0, T] \rightarrow \mathbb{R}^q$ is defined by $\tilde{h}((x_1, X_1), t) := h(x_1, t)$ for all $(x_1, X_1) \in \mathbb{R}^{n+1}$ and all $t \in [0, T]$. Note that Problem (OSCP_{aug}) satisfies all of the assumptions of Theorem 3.1 and is without Lagrange cost. We introduce the augmented Hamiltonian $\tilde{H} : \mathbb{R}^{n+1} \times \mathbb{R}^m \times \mathbb{R}^{n+1} \times [0, T] \rightarrow \mathbb{R}$ defined as

$$\tilde{H}((x, X), u, (p, P), t) := \left\langle \begin{pmatrix} p \\ P \end{pmatrix}, \begin{pmatrix} f(x, u, t) \\ L(x, u, t) \end{pmatrix} \right\rangle_{\mathbb{R}^{n+1}},$$

for all $((x, X), u, (p, P), t) \in \mathbb{R}^{n+1} \times \mathbb{R}^m \times \mathbb{R}^{n+1} \times [0, T]$. Applying Theorem 3.1 (without Lagrange cost, proved in the previous section), we deduce the existence of a nontrivial couple (p^0, η) , where $p^0 \leq 0$ and $\eta = (\eta_j)_{j=1, \dots, q} \in \text{NBV}_q$, such that all conclusions of Theorem 3.1 are satisfied. In particular, the adjoint vector $(p, P) \in \text{BV}_{n+1}$ satisfies the backward linear Cauchy-Stieltjes problem given by

$$\left\{ \begin{aligned} - \begin{pmatrix} dp \\ dP \end{pmatrix} &= \partial_{(x, X)} \begin{pmatrix} f \\ L \end{pmatrix} (x, u, \cdot)^\top \times \begin{pmatrix} p \\ P \end{pmatrix} dt - \sum_{j=1}^q \partial_{(x, X)} \tilde{h}_j((x, X), \cdot) d\eta_j && \text{over } [0, T], \\ \begin{pmatrix} p \\ P \end{pmatrix} (T) &= p^0 \nabla_{(x, X)} \tilde{g}((x, X)(T)). \end{aligned} \right.$$

We deduce that $P(T) = p^0$ and $dP = 0$ over $[0, T]$. Thus $P(t) = p^0$ for all $t \in [0, T]$, and we obtain that $p \in \text{BV}_n$ satisfies the backward linear Cauchy-Stieltjes problem

$$\left\{ \begin{aligned} -dp &= \left(\partial_1 f(x, u, \cdot)^\top \times p + p^0 \partial_1 L(x, u, \cdot) \right) dt - \sum_{j=1}^q \partial_1 h_j(x, \cdot) d\eta_j && \text{over } [0, T], \\ p(T) &= p^0 \nabla g(x(T)). \end{aligned} \right.$$

The rest of the proof is straightforward from all the necessary conditions obtained from the version of Theorem 3.1 without Lagrange cost.

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] G. Bachman and L. Narici. *Functional analysis*. Dover Publications, Inc., Mineola, NY, 2000. Reprint of the 1966 original.
- [5] 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.
- [6] P. Bettiol and H. Frankowska. Hölder continuity of adjoint states and optimal controls for state constrained problems. *Appl. Math. Optim.*, 57(1):125–147, 2008.
- [7] E. Bini and G. M. Buttazzo. The optimal sampling pattern for linear control systems. *IEEE Trans. Automat. Control*, 59(1):78–90, 2014.
- [8] V. G. Boltyanskii. *Optimal control of discrete systems*. John Wiley & Sons, New York-Toronto, Ont., 1978.
- [9] J. F. Bonnans and C. de la Vega. Optimal control of state constrained integral equations. *Set-Valued Var. Anal.*, 18(3-4):307–326, 2010.
- [10] J. F. Bonnans, C. de la Vega, and X. Dupuis. First- and second-order optimality conditions for optimal control problems of state constrained integral equations. *J. Optim. Theory Appl.*, 159(1):1–40, 2013.
- [11] B. Bonnard, L. Faubourg, G. Launay, and E. Trélat. Optimal control with state constraints and the space shuttle re-entry problem. *J. Dynam. Control Systems*, 9(2):155–199, 2003.
- [12] L. Bourdin. Note on Pontryagin maximum principle with running state constraints and smooth dynamics - Proof based on the Ekeland variational principle. *Research notes – available on HAL*, 2016.
- [13] L. Bourdin and G. Dhar. Continuity/constancy of the Hamiltonian function in a Pontryagin maximum principle for optimal sampled-data control problems with free sampling times. *In revision process – available on HAL*, 2019.
- [14] L. Bourdin and E. Trélat. Pontryagin maximum principle for finite dimensional nonlinear optimal control problems on time scales. *SIAM Journal on Control and Optimization*, 20(4):526–547, 2013.
- [15] L. Bourdin and E. Trélat. Pontryagin maximum principle for optimal sampled-data control problems. *16th IFAC workshop on control applications of optimization CAO’2015*, 2015.
- [16] L. Bourdin and E. Trélat. Optimal sampled-data control, and generalizations on time scales. *Mathematical Control & Related Fields*, 6(1):53–94, March 2016.
- [17] L. Bourdin and E. Trélat. Linear-quadratic optimal sampled-data control problems: convergence result and Riccati theory. *Automatica J. IFAC*, 79:273–281, 2017.
- [18] A. Bressan and B. Piccoli. *Introduction to the mathematical theory of control*, volume 2 of *AIMS Series on Applied Mathematics*. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

- [19] F. E. Burk. *A garden of integrals*, volume 31 of *The Dolciani Mathematical Expositions*. Mathematical Association of America, Washington, DC, 2007.
- [20] N. L. Carothers. *Real analysis*. Cambridge University Press, Cambridge, 2000.
- [21] L. Cesari. *Optimization—theory and applications*, volume 17 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1983. Problems with ordinary differential equations.
- [22] D. I. Cho, P. L. Abad, and M. Parlar. Optimal production and maintenance decisions when a system experience age-dependent deterioration. *Optimal Control Applications and Methods*, 14(3):153–167, 1993.
- [23] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.
- [24] E. A. Coddington and N. Levinson. *Theory of ordinary differential equations*. McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.
- [25] O. Cots. Geometric and numerical methods for a state constrained minimum time control problem of an electric vehicle. *ESAIM Control Optim. Calc. Var.*, 23(4):1715–1749, 2017.
- [26] O. Cots, J. Gergaud, and D. Goubinat. Direct and indirect methods in optimal control with state constraints and the climbing trajectory of an aircraft. *Optimal Control Appl. Methods*, 39(1):281–301, 2018.
- [27] A. V. Dmitruk. On the development of Pontryagin’s maximum principle in the works of A. Ya. Dubovitskii and A. A. Milyutin. *Control Cybernet.*, 38(4A):923–957, 2009.
- [28] 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.
- [29] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.
- [30] L. C. Evans. An introduction to mathematical optimal control theory. *Version 0.2*, Lecture notes.
- [31] M. S. Fadali and A. Visioli. *Digital control engineering: Analysis and design*. Elsevier, 2013.
- [32] J. Faraut. *Calcul intégral (L3M1)*. EDP sciences, 2012.
- [33] R. V. Gamkrelidze. Optimal control processes for bounded phase coordinates. *Izv. Akad. Nauk SSSR. Ser. Mat.*, 24:315–356, 1960.
- [34] I. V. Girsanov. *Lectures on mathematical theory of extremum problems*. Springer-Verlag, Berlin-New York, 1972. Edited by B. T. Poljak, Translated from the Russian by D. Louvish, Lecture Notes in Economics and Mathematical Systems, Vol. 67.
- [35] 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.
- [36] L. Grüne and J. Pannek. *Nonlinear model predictive control*. Communications and Control Engineering Series. Springer, Cham, 2017. Theory and algorithms, Second edition [of MR3155076].
- [37] H. Halkin. A maximum principle of the pontryagin type for systems described by nonlinear difference equations. *SIAM Journal on control*, 4(1):90–111, 1966.
- [38] R. F. Hartl, S. P. Sethi, and R. G. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev.*, 37(2):181–218, 1995.

- [39] M. R. Hestenes. *Calculus of variations and optimal control theory*. Robert E. Krieger Publishing Co., Inc., Huntington, N.Y., 1980. Corrected reprint of the 1966 original.
- [40] J. B. Hiriart-Urruty. Les mathématiques du mieux faire, vol. 2: La commande optimale pour les débutants. *Collection Opuscles*, 2008.
- [41] J. M. Holtzman and H. Halkin. Discretional convexity and the maximum principle for discrete systems. *SIAM Journal on Control*, 4(2):263–275, 1966.
- [42] D. H. Jacobson, M. M. Lele, and J. L. Speyer. New necessary conditions of optimality for control problems with state-variable inequality constraints. *J. Math. Anal. Appl.*, 35:255–284, 1971.
- [43] N. Kim, A. Rousseau, and D. Lee. A jump condition of PMP-based control for PHEVs. *Journal of Power Sources*, 196(23):10380–10386, 2011.
- [44] I. D. Landau and G. Zito. *Digital Control Systems: Design, Identification and Implementation*. Springer, 2006.
- [45] G. J. Lastman. Two formulations for optimization problems with state-variable inequality constraints. *AIAA Journal*, 7(7):1378–1380, 1969.
- [46] E. B. Lee and L. Markus. *Foundations of optimal control theory*. Robert E. Krieger Publishing Co., Inc., Melbourne, FL, second edition, 1986.
- [47] X. Li and J. Yong. *Optimal control theory for infinite-dimensional systems*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1995.
- [48] B. V. Limaye. *Functional analysis*. New Age International Publishers Limited, New Delhi, second edition, 1996.
- [49] H. Maurer. On optimal control problems with bounded state variables and control appearing linearly. *SIAM J. Control Optimization*, 15(3):345–362, 1977.
- [50] H. Maurer, J. R. Kim, and G. Vossen. On a state-constrained control problem in optimal production and maintenance. In *Optimal Control and Dynamic Games*, pages 289–308. Springer, 2005.
- [51] H. Maurer and H. D. Mittelman. The non-linear beam via optimal control with bounded state variables. *Optimal Control Applications and Methods*, 12(1):19–31, 1991.
- [52] B. S. Mordukhovich. *Variational analysis and generalized differentiation. I*, volume 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006. Basic theory.
- [53] D. O. Norris. Nonlinear programming applied to state-constrained optimization problems. *J. Math. Anal. Appl.*, 43:261–272, 1973.
- [54] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. John Wiley & Sons, Inc., 1962.
- [55] A. Puchkova, V. Rehbock, and K. L. Teo. Closed-form solutions of a fishery harvesting model with state constraint. *Optimal Control Appl. Methods*, 35(4):395–411, 2014.
- [56] H. Robbins. Junction phenomena for optimal control with state-variable inequality constraints of third order. *J. Optim. Theory Appl.*, 31(1):85–99, 1980.
- [57] R. T. Rockafellar. State constraints in convex control problems of Bolza. *SIAM J. Control*, 10:691–715, 1972.

- [58] 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.
- [59] S. P. Sethi and G. L. Thompson. *Optimal control theory*. Kluwer Academic Publishers, Boston, MA, second edition, 2000. Applications to management science and economics.
- [60] E. Trélat. *Contrôle optimal: théorie & applications*. Vuibert Paris, 2005.
- [61] T. van Keulen, J. Gillot, B. de Jager, and M. Steinbuch. Solution for state constrained optimal control problems applied to power split control for hybrid vehicles. *Automatica J. IFAC*, 50(1):187–192, 2014.
- [62] V. Van Reeve, T. Hofman, F. Willems, R. Huisman, and M. Steinbuch. Optimal control of engine warmup in hybrid vehicles. *Oil & Gas Science and Technology—Revue d'IFP Energies nouvelles*, 71(1):14, 2016.
- [63] R. Vinter. *Optimal control*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2010. Paperback reprint of the 2000 edition.
- [64] 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.
- [65] R. L. Wheeden and A. Zygmund. *Measure and integral*. Pure and Applied Mathematics (Boca Raton). CRC Press, Boca Raton, FL, second edition, 2015. An introduction to real analysis.