



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

Loïc Bourdin¹ · Gaurav Dhar¹

Received: 2 July 2019 / Accepted: 25 September 2020 / Published online: 1 October 2020
© Springer-Verlag GmbH Germany, part of Springer Nature and Mathematical Optimization Society 2020

Abstract

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. We obtain, in particular, a nonpositive averaged Hamiltonian gradient condition associated with an adjoint vector being a function of bounded variation. As a well known challenge, theoretical and numerical difficulties may arise due to the possible pathological behavior of the adjoint vector (jumps and singular part lying on parts of the optimal trajectory in contact with the boundary of the restricted state space). However, in our case with sampled-data controls, we prove that, under certain general hypotheses, the optimal trajectory activates the running inequality state constraints at most at the sampling times. Due to this so-called bouncing trajectory phenomenon, the adjoint vector experiences jumps at most at the sampling times (and thus in a finite number and at precise instants) and its singular part vanishes. Taking advantage of these informations, we are able to implement an indirect numerical method which we use to solve three simple examples.

Keywords Optimal control · Sampled-data control · Pontryagin maximum principle · State constraints · Ekeland variational principle · Indirect numerical method · Shooting method

Mathematics Subject Classification 34H05 · 49M05 · 93C10 · 93C57

✉ Gaurav Dhar
gaurav.dhar@unilim.fr

Loïc Bourdin
loic.bourdin@unilim.fr

¹ Institut de recherche XLIM, UMR CNRS 7252, Université de Limoges, Limoges, France

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.

Context in optimal control theory. Established in [63] by Pontryagin, Boltyanskii, Gamkrelidze and Mischenko at the end of the 1950s, the Pontryagin Maximum Principle (in short, PMP) is a fundamental result in optimal control theory with numerous theoretical and numerical applications. We refer to [21,48,49,55,69,70,73] 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 function* associated with 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 *indirect numerical method* 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.¹ Soon afterwards and even nowadays, the PMP has been adapted to many situations, for control systems of different natures and with various constraints. It is not our aim to give a state of the art here. 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., [10,46,50]). In these discrete versions of the PMP, the Hamiltonian maximization condition does not hold in general (see a counterexample in [10, Examples 10.1–10.4, pp. 59–62]) and has to be replaced by a weaker condition known as the *nonpositive Hamiltonian gradient condition* (see e.g., [10, Theorem 42.1, p. 330]). This is due in particular to the fact that the method of *needle-like perturbations*, used in classical continuous optimal control theory, cannot be adapted at the discrete level.

¹ The terminology *indirect numerical method* is opposed to the one of *direct numerical 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 optimization algorithms and techniques.

An important generalization of the PMP concerns *state constrained optimal control problems* in which 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 [14,29,53,71,72] and references therein for other examples. State constrained optimal control problems are also encountered in management and economics (e.g., an inventory level may be limited in a production model). We refer to [25,60,64,69] and references therein for other examples. A first version of the PMP for optimal control problems with running inequality state constraints was obtained by Gamkrelidze [42] (see also [63, Theorem 25, p. 311]) under some special assumptions on the structure of the optimal process. Later, these assumptions were somewhat excluded by Dubovitskii and Milyutin in the seminal work [37, Section 7, p. 37]. The contributions of Dubovitskii and Milyutin in the development of the PMP, in the 1960s and later years, have been the subject of a survey written by Dmitruk [31] in 2009. These contributions include, notably, general Lagrange multiplier rules for abstract optimization problems and the so-called method of *v-change of time variable* in view of generating needle-like variations by passing to a smooth control system (see more details in [31, Section 4]). These approaches have been revisited in a book of lectures on extremum problems written by Girsanov [43] in 1972, and have been extended recently in a series of papers by Dmitruk and Osmolovskii [33–36] with applications to various optimal control problems such as with integral equations, with mixed state-control constraints, etc. Other methods have been developed in the literature in order to establish versions of the PMP for state constrained optimal control problems, such as the smoothly-convex structure of the controlled system in [51], the application of the Ekeland variational principle in [73], etc. A comprehensive survey [47] of this field of research has been given in 1995 by Hartl, Sethi and Vickson. Note that the PMP for optimal control problems is more intricate in the presence of running inequality state constraints because the adjoint vector is (only) of bounded variation in general (while it is absolutely continuous in the state unconstrained case). 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 optimal 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 costate's behavior and some constraint qualification conditions have been established in order to ensure that the adjoint vector has no singular part (see e.g., [8,14,31,47,52,59]). We briefly conclude this paragraph by mentioning that the related theme of state constrained discrete optimal control problems has also been investigated in the literature (see e.g., [30, 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 control is authorized to be modified at only a finite number of fixed instants (called *sampling times*) and remains frozen elsewhere. These systems have the peculiarity of presenting a mixed continuous/discrete structure: the state evolves continuously with respect to time, while the control evolves discretely with respect to time. Sampled-data controls are modelled

by piecewise constant controls and have been considered in Engineering implemented by digital controllers which have a finite precision (see e.g., [68,74]). They 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., [45]). Numerous texts and articles have developed control theory for sampled-data control systems (see e.g., [1,3,4,40,44,54] and references therein). Recently Bourdin and Trélat have obtained in [19] a version of the PMP for (state unconstrained) optimal sampled-data control problems. In that framework, as in the case of discrete control systems (in which needle-like variations are also absent), the usual Hamiltonian maximization condition does not hold in general and has to be replaced by a weaker condition known as the *nonpositive averaged Hamiltonian gradient condition* (see [19, Theorem 2.6, p. 62]). We remark that (state unconstrained) optimal sampled-data control problems can be formulated as finite-dimensional optimization problems since the control values (to be optimized) are in finite number. We also emphasize that the PMP established in [19, Theorem 2.6, p. 62] is valid in the more general setting of *time scale calculus* which allows to encompass, in a single framework, continuous and discrete optimal sampled-data and permanent control problems and to extend to more general time structures. In particular, optimal sampled-data control problems on time scales are infinite-dimensional optimization problems. As commented in details in [17, Section 3.1], most of the standard methods used in classical optimal control theory in order to derive the PMP fail within the general time scale setting (due to a lack of convexity on the control perturbation parameters). As far as we know, the method based on the Ekeland variational principle is the only one allowing to derive a PMP in the general time scale setting.

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 Sect. 3). Note that such problems can be seen as *semi-infinite*-dimensional optimization problems since the presence of running inequality state constraints imposes an infinite number of constraints (one at each instant of time). In this paper, in the same spirit as Bourdin and Trélat in [19], our proof is based on the Ekeland variational principle.² Similarly to the PMP derived in [19, Theorem 2.6, p. 62] for state 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 state constrained optimal permanent control problems, we find that the adjoint vector is in general (only) of bounded variation. Therefore one would expect to encounter the same difficulties as in the permanent control case when implementing an indirect numerical method due to the possible jumps and singular part of the adjoint vector. However, in our context of sampled-data controls, we find that the optimal trajectories have a common behavior which allows us to overcome these difficulties. Precisely,

² In particular we have opted for the use of the Ekeland variational principle in view of generalizations to the general time scale setting in further research works.

when we undertook studying optimal sampled-data control problems in the presence of running inequality state constraints, we first numerically solved some simple problems using direct methods. 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 the sampling times. This behavior constitutes the second major focus of the present work and is referred to as the *bouncing trajectory phenomenon*. 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 the sampling times (and thus in a finite number and at precise instants). We refer to Proposition 4.1 in Sect. 4 for details. Inherent to this behavior, the singular part of the adjoint vector, associated with an optimal trajectory, vanishes and its discontinuities are reduced to a finite number of jumps which occur exactly at the sampling times. Taking advantage of these informations, we are able in Sect. 5 to implement an indirect numerical method which we use to numerically solve three simple examples of optimal sampled-data control problems with running inequality state constraints. We take this occasion to mention that a similar trajectory phenomenon has already been observed in the literature on state constrained optimal permanent control problems. Precisely, Milyutin provides an example in his doctoral dissertation [61] in 1966 (see also [31, p. 940]) in which the optimal trajectory touches the state constraint a countably infinite number of times before landing on it. As a consequence, in that example, the corresponding adjoint vector admits a countably infinite number of jumps. This example was also given independently by Robbins [66] in 1980.

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 state our main result. In Sect. 3 we first present the optimal sampled-data control problem with running inequality state constraints considered in this work (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 follows. In Sect. 4 we give heuristic descriptions and a sufficient condition for observing the bouncing trajectory phenomenon. In Sect. 5 we propose an indirect method for numerically solving optimal sampled-data control problems with running inequality state constraints based on our main result and with the aid of the bouncing trajectory phenomenon. Then we illustrate this method and highlight the bouncing trajectory phenomenon by numerically solving 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 state our main result (see Theorem 3.1 in Sect. 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 variation 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

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

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 variation and on linear Cauchy–Stieltjes problems

In this section we refer to standard references and books such as [5,22,23,41,75] 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 variation* 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

variation 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 now 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

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

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

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

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

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

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

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

for a.e. $t \in [0, T]$. 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 [15, 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 variation 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 Sect. 3.2). In Sect. 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 Sect. 3.2 we state the corresponding Pontryagin maximum principle and a list of comments follows.

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

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

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 its 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 optimal control theory (see e.g., [21,24,49,63,69,70,73] and references therein), the control u usually can be any function in L_m^∞ (with values in U). In that case we say that the control is *permanent* in the sense that its value can be modified at any 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 control u in Problem (OSCP) is a *sampled-data control* (and the elements t_i of \mathbb{T} are called the *sampling times*).

Control theory for sampled-data control systems has already been considered in several texts (see e.g., [1,4,40,44,54]), often in the context of *digital control*. Optimal sampled-data control problems have also been investigated in the literature (see e.g., [3,9,16,18–20]) 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 references mentioned above.

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 [27, 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 [24, Theorem 9.2.i, p. 311]) and Rockafellar (see [67, 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 [19, 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 case of sampled-data controls

with relaxed assumptions with respect to the case of permanent controls, and thus it 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 unconstrained linear-quadratic problems, the authors of [20] 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 Sect. 5, an interesting research perspective would be to extend these convergence results to the context of the present work.
- (ii) In the paper [16] 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.
- (iii) Several papers in the literature consider optimal permanent control problems with constraints of different natures, for instance with state constraints where the constraint is given as an inclusion into a general subset of the state space (see e.g., [24,67]) or with mixed state-control constraints of the form $h(x(t), u(t), t) \leq 0$ (see e.g., [33,47]). A possible challenge would be to extend the present work to the previous mentioned contexts.
- (iv) 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 [6] 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 recall 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. We refer to Introduction for a brief bibliographic recap and we refer to [2,11–13,13,33,34,36,73] for recent contributions with various generalizations.

The novelty of the present work is to deal with nonpermanent controls, precisely, with sampled-data controls. As usual in the literature we introduce the Hamiltonian function $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ associated with 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*

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

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

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

Moreover the complementary slackness condition:

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

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 presented hereafter.

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*. We observe that, if the running inequality state constraints are never activated (that is, if $h_j(x(t), t) < 0$ for all $t \in [0, T]$ and all $j = 1, \dots, q$), then the case is normal. Outside of this trivial situation, sufficient conditions ensuring normality is a difficult topic which has been widely developed in the literature on constrained optimal control problems, mostly in case of permanent controls (see e.g., [\[7,26,58,65\]](#) and references therein). To the best of our knowledge, the extension of such results to the present sampled-data control setting is an open challenge in the literature, and thus perspectives for further research works are possible in that direction. Since it is not our aim in this paper to discuss this point in more depth, we will opt in practice (as in Examples 1, 2 and 3 in Sect. 5) for proofs by contradiction in order to show that the case is normal.

Remark 3.5 Our strategy in Appendices A and B in order to prove Theorem 3.1 is based on the Ekeland variational principle [38] 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 Sect. 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 Sect. A.2). Nevertheless, the convexity of U can be removed by considering the concept of stable U -dense directions (see e.g., [19]).

Remark 3.6 If there is no running inequality state constraint in Problem (OSCP), that is, considering $h_j = -1$ for all $j = 1, \dots, q$ for example, then Theorem 3.1 recovers the standard Pontryagin maximum principle for optimal sampled-data control problems obtained for example in [18,19]. Let us recall that the Hamiltonian maximization condition, which holds in case of permanent controls, is not true in general when considering sampled-data controls, and has to be replaced by the nonpositive averaged Hamiltonian gradient condition. This is due in particular to the fact that the method of needle-like perturbations used in classical optimal permanent control theory cannot be adapted to the sampled-data control context. 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 a.e. $t \in [0, T]$, 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., [73, Theorem 9.3.1]) to find the adjoint vector $p \in BV_n$ written as the sum $p = p_1 + p_2$ where $p_1 \in AC_n$ 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 $p_2 \in BV_n$ 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 Sect. 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 with η_j (see Sect. 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 $\partial_1 h_j(x(t), t) \neq 0_{\mathbb{R}^n}$ whenever $h_j(x(t), t) = 0$ for all $j = 1, \dots, q$. We refer to [73, Remark (b), p. 330] for a similar remark in the classical case of permanent controls.

Remark 3.13 In the classical case of state unconstrained optimal permanent control problems, the Pontryagin maximum principle induces an indirect numerical method based on the resolution by a shooting method of the boundary value problem satisfied by the augmented state-costate vector (see e.g., [70, pp. 170–171] for details). Recall that:

- (i) In the presence of state constraints, the indirect numerical method can be adapted. However, some theoretical and numerical difficulties may appear due to the possible pathological behavior of the adjoint vector (see Sect. 5 for more details).
- (ii) The indirect numerical method has also been adapted to the case of (state unconstrained) optimal sampled-data control problems in [18, 19], and also in case of free sampling times in [16].

To the best of our knowledge, the indirect numerical method has never been adapted to optimal sampled-data control problems in the presence of running inequality state constraints. Our aim in Sect. 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 $p \in BV_n$ is pathological and/or admits an infinite number of discontinuities, but it will be shown in Sects. 4 and 5 that, under certain (quite unrestrictive) hypotheses, the implementation of the indirect numerical method is simplified due to the particular behavior of the optimal trajectory.

Remark 3.14 In this paper, as explained in Introduction, the proof of Theorem 3.1 is based on the Ekeland variational principle [38] in view of generalizations to the general time scale setting in further research works. Let us note that an alternative proof of Theorem 3.1 can be obtained by adapting a remarkable technique exposed in the paper [32] 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 inequality state constraints (see e.g., [15, Theorem 1] and [73, Theorem 9.5.1, pp. 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., [17, 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 numerically solve some simple problems using a direct method (see Sect. 5 for some details on direct methods in optimal control theory). On this occasion we observed that the optimal trajectories returned by the algorithm had a common behavior with respect to the running inequality state constraints. Precisely the optimal trajectories were “bouncing” on them. We refer to Fig. 3 and Sect. 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 Sect. 5, this feature presents some benefits from a numerical point of view.

In Sect. 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 Sect. 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 [47, p. 183] or [69, p. 105]. Let x be an admissible trajectory of Problem (OSCP). An element $t \in [0, T]$ is called an *activating time* if it satisfies $h(x(t), t) = 0$. 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 $t \in [\tau_1, \tau_2]$. Note that any point of a boundary interval is an activating time, while the reverse is not true in general. In what follows, we say that the trajectory x exhibits 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 common 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 only (see (ii) below). The mathematical framework and rigorous justifications will be provided in Sect. 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 Fig. 1 for an illustration. Therefore, since we deal with piecewise constant controls in Problem (OSCP), one should expect

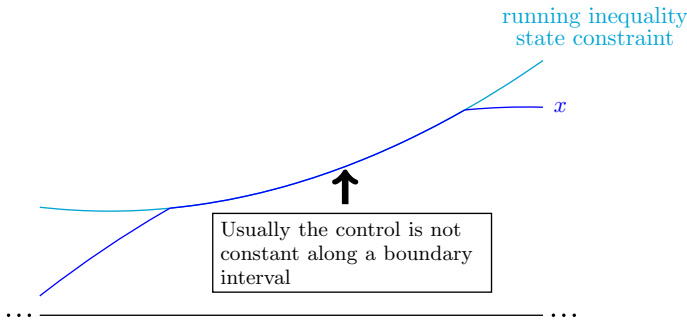


Fig. 1 In the classical theory (that is, with permanent controls), a boundary interval is usually associated with a nonconstant control

that an admissible trajectory of Problem (OSCP) does not contain any boundary interval and thus exhibits 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 Sect. 4.2 (see Hypothesis (H1)).

- (ii) Let $t \in [0, T]$ be a left isolated (resp. right isolated) activating time of an admissible trajectory x of Problem (OSCP). In what follows 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” (resp. “exits”) the running inequality state constraint transversally 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 (resp. immediately before t), which contradicts the admissibility of x . We refer to Fig. 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 no left or right isolated activating time outside of the sampling times t_i . 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” or “exits” the running inequality state constraint tangentially. This will be done in Sect. 4.2 (see Hypothesis (H2)). Actually our Hypothesis (H2) will even guarantee that an admissible trajectory of Problem (OSCP) has no activating time outside of the sampling times t_i .

We conclude from (i) and (ii) that one should expect the admissible trajectories of Problem (OSCP) to exhibit the bouncing trajectory phenomenon and, moreover, so 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 Fig. 3 for an illustration of this feature. Note that, even if activating times are sampling times, the reverse is not true in general.

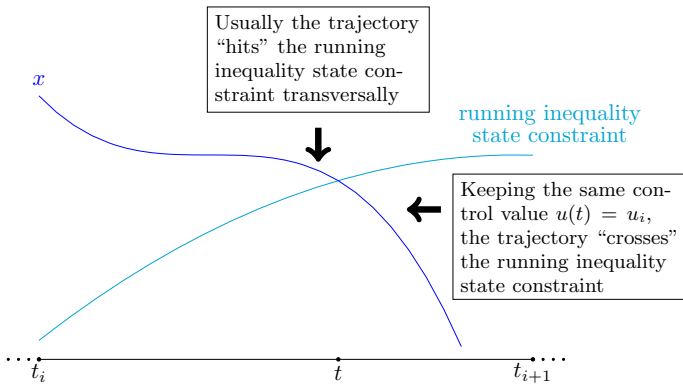


Fig. 2 Illustration of a trajectory x hitting transversally the running inequality state constraint at some left isolated activating time t which belongs to the interior (t_i, t_{i+1}) of a sampling interval

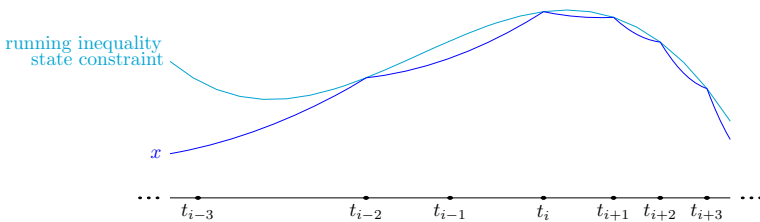


Fig. 3 Illustration of the expected behavior of an admissible trajectory x of Problem (OSCP)

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 Fig. 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 Sect. 4.2 below and as illustrated by the examples numerically solved in Sect. 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 Sect. 4.1. In particular we will formulate a sufficient condition (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 [47, 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 $v \in U$. Finally we introduce the set

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

for all $(y, t) \in \mathcal{M}$. We now state 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*

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

and

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

are both satisfied, then the activating times of an admissible trajectory x of Problem (OSCP) are sampling times. In particular x exhibits the bouncing trajectory phenomenon and the rebounds 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 $u_i \in U$, 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 guarantee the validity of the arguments presented heuristically in the items (i) and (ii) of Sect. 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 Sect. 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 numerical 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 optimization algorithms and techniques. The second strategy is called *indirect numerical 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 numerical methods consist in the numerical resolution by a shooting method of the boundary value problem satisfied by the augmented state-costate vector (x, p) . We emphasize that neither direct nor indirect methods are fundamentally better than the other. We refer for instance to [70, pp. 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 numerical methods can be adapted easily. On the contrary, solving optimal permanent control problems with running inequality state constraints might be more intricate when using indirect numerical methods. Indeed, in that situation, the adjoint vector p is not absolutely continuous in general, but (only) of bounded variation. From the Lebesgue decomposition (see e.g., [23, 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 on intervals with no activating time of the optimal trajectory 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 costate's behavior and some constraint qualification conditions have been established. We refer for instance to [8,14,47,52,59].

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 numerically 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 variation and we will *a priori* encounter the same difficulties outlined above. Nevertheless, as detailed in Sect. 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 Sect. 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 these informations, we will propose in Sect. 5.1 a simple indirect method in order to numerically solve optimal sampled-data control problems with running inequality state constraints of the form of Problem (OSCP) under Hypotheses (H1) and (H2).

In Sects. 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 [20] in the case of unconstrained linear-quadratic problems.

We conclude this paragraph by noting that the indirect numerical method proposed in Sect. 5.1 (and its implementation in Sects. 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 for more details).

5.1 A shooting function for an indirect method

In this section 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. This numerical method can be implemented in the normal case as well as in the abnormal case (in the sense of Remark 3.4).

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. As explained at the beginning of Sect. 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 Sect. 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:

$$\eta^{[0]} := \eta(t_0^+) - \eta(t_0) = \eta(t_0^+), \quad \eta^{[1]} := \eta(t_1) - \eta(t_0^+)$$

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

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

$$p^{[0]} := p(t_0^+) - p(t_0) = \eta^{[0]} \partial_1 h(x(t_0), t_0)$$

$$\text{and } \forall i = 1, \dots, N, \quad p^{[i]} := p(t_i) - p(t_i^-) = \eta^{[i]} \partial_1 h(x(t_i), t_i).$$

The general indirect numerical 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 $x(T) = x_T$ and of the nonnegative jumps $\eta^{[i]}$ for all $i = 0, \dots, N$;
- (ii) we compute $p(T) = p^0 \nabla g(x(T))$;
- (iii) we numerically solve the state and adjoint equations 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 indirect numerical method, we solve three simple examples in Sects. 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. We also mention that we used the basic forward Euler method in order to numerically solve the state and adjoint equations (but numerous other approaches can be considered, such as Runge-Kutta methods for example). Finally we emphasize that the numerical results obtained and presented hereafter have all been confirmed by direct numerical approaches (using a basic forward Euler discretization of the whole problem resulting into a constrained finite-dimensional optimization problem solved numerically 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

$$\left. \begin{aligned}
 &\text{minimize } \int_0^4 x(\tau) + \frac{1}{4}u(\tau)^2 d\tau \\
 &\text{subject to } x \in AC_1, u \in PC_1^{\mathbb{T}}, \\
 &\quad \dot{x}(t) = u(t) \text{ for a.e. } t \in [0, 4], \\
 &\quad x(0) = 6, \\
 &\quad \frac{1}{2}(t - 2)^2 + 2 - x(t) \leq 0 \text{ for all } t \in [0, 4], \\
 &\quad u(t) \in [-3, +\infty) \text{ for all } t \in [0, 4],
 \end{aligned} \right\} \tag{E1}$$

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).

Let us check that Problem (E1) satisfies Hypotheses (H1) and (H2). To this aim we follow the notations introduced in Sect. 4.2. 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. We deduce that, for all $(y, t) \in \mathcal{M}$ and all $v \in \mathcal{N}(y, t)$, we have $\ell'(y, v, t) = 2$ and $h^{[\ell'(y,v,t)]}(y, v, t) = 1 > 0$, 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 what follows 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 check 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 $\eta \neq 0_{NBV_1}$ 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 $\bar{t} = \hat{t}_i$ for some $\hat{i} \in \{0, \dots, N\}$. Since $x(0) = 6$, we have $\hat{i} \geq 1$. 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 Sect. 5.1. As expected, we observe in Fig. 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 $N = 10$

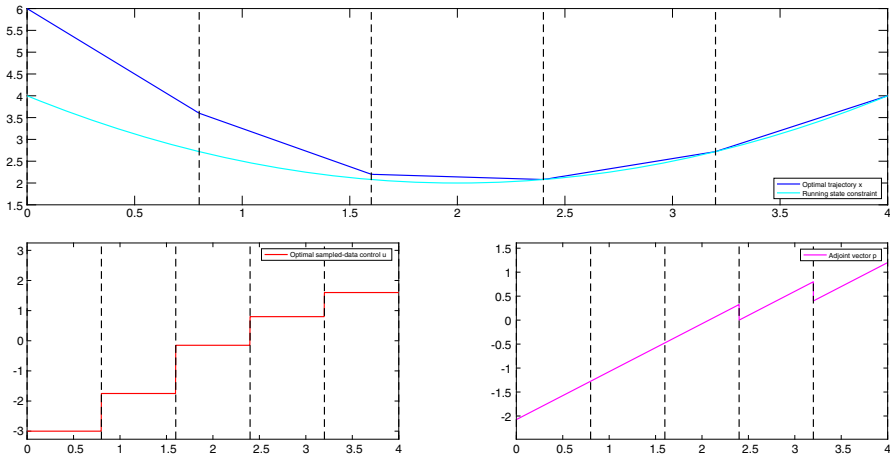


Fig. 4 Example 1 with $N = 5$

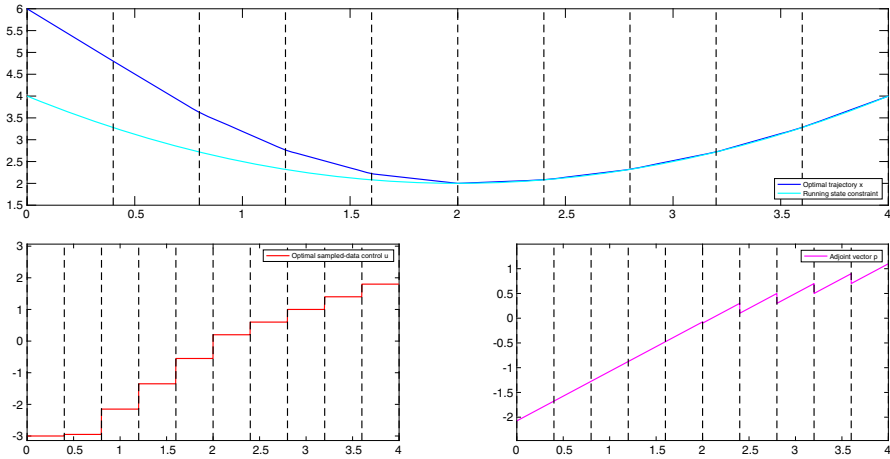


Fig. 5 Example 1 with $N = 10$

and $N = 40$). Furthermore, in Figs. 4, 5 and 6, we observe that the adjoint vector has no jump at sampling times which are not activating times.

Remark 5.1 Actually, in that simple Example 1, the state and adjoint equations are very simple and can be solved explicitly. As a consequence, the shooting map can even be expressed in the closed form given by

$$\left(x_T, (\eta^{[i]})_{i=0, \dots, N} \right) \mapsto \left(x_T - x_0 - \sum_{i=0}^{N-1} u_i(t_{i+1} - t_i), (\eta^{[i]} \left[\frac{1}{2}(t_i - 2)^2 + 2 - x_T + \sum_{k=i}^{N-1} u_k(t_{k+1} - t_k) \right])_{i=0, \dots, N-1} \right),$$

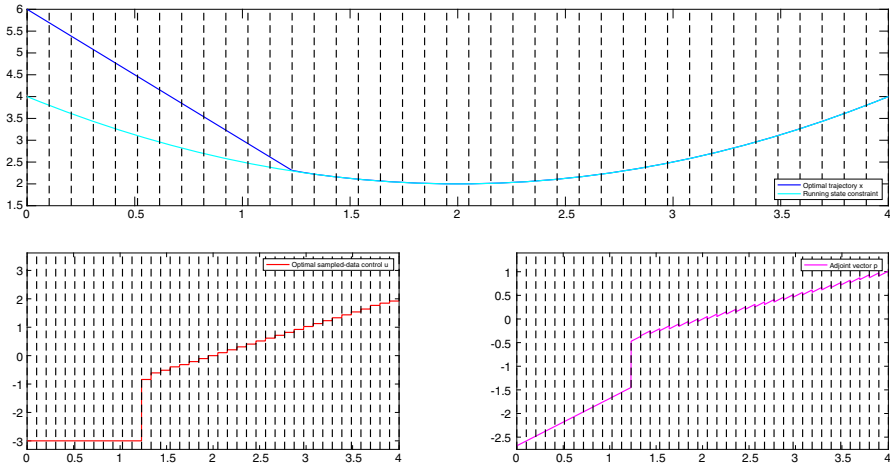


Fig. 6 Example 1 with $N = 40$

where

$$u_i = \max \left\{ 2 \left(\frac{t_i + t_{i+1}}{2} - T + \sum_{k=i+1}^N \eta^{[k]} \right), -3 \right\},$$

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

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

$$\left. \begin{aligned} & \text{minimize } \int_0^{12} (u(\tau) - 1)x(\tau) d\tau \\ & \text{subject to } x \in AC_1, u \in PC_1^{\mathbb{T}}, \\ & \quad \dot{x}(t) = u(t)x(t) \text{ for a.e. } t \in [0, 12], \\ & \quad x(0) = 1, \\ & \quad x(t) - 10t - 2 \leq 0 \text{ for all } t \in [0, 12], \\ & \quad u(t) \in [0, 1] \text{ for all } t \in [0, 12], \end{aligned} \right\} \tag{E2}$$

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., [39, p. 5]) revisited with sampled-data controls and running inequality state constraint.

Let us check that Problem (E2) satisfies Hypotheses (H1) and (H2). To this aim we follow the notations introduced in Sect. 4.2. Let us assume by contradiction that

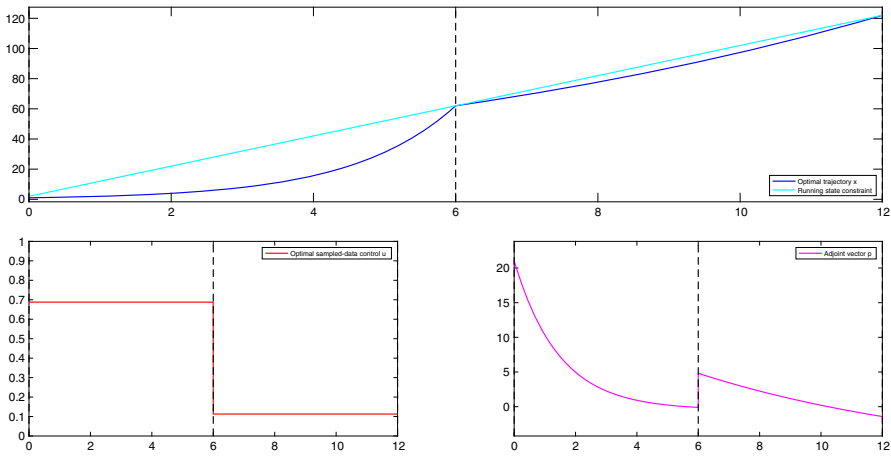


Fig. 7 Example 2 with $N = 2$

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 check 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 $\bar{t} = t_{\hat{i}}$ for some $\hat{i} \in \{0, \dots, N\}$. Since $x(0) = 1$, we know that $\hat{i} \geq 1$. It follows that $p(t) < 0$ for all $t \in [0, t_{\hat{i}}]$. Finally, since $x(t) > 0$ for all $t \in [0, 12]$ and from the nonpositive averaged Hamiltonian gradient condition at $i = 0, \dots, \hat{i} - 1$, we get that $u_0 = \dots = u_{\hat{i}-1} = 0$, 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 Sect. 5.1. In Fig. 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 $N = 4$

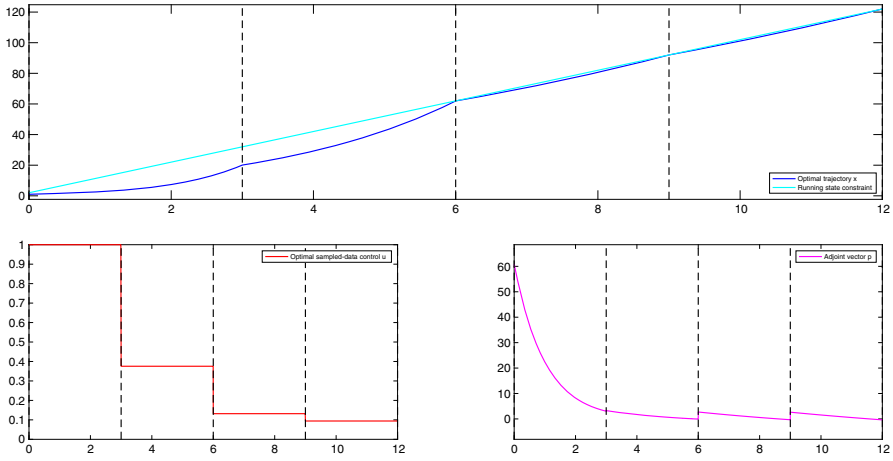


Fig. 8 Example 2 with $N = 4$

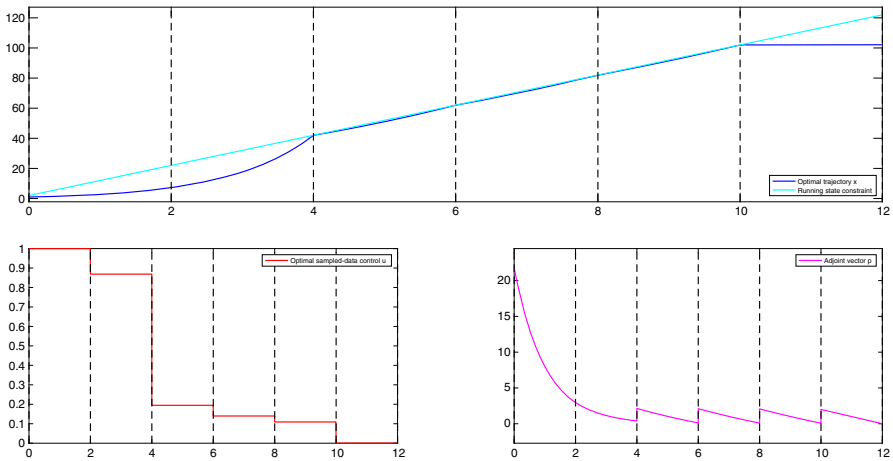


Fig. 9 Example 2 with $N = 6$

and $N = 6$). Furthermore, in Figs. 8 and 9, we observe that the adjoint vector has no jump at sampling times which are not activating times.

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

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

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 check 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 Sect. 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 what follows 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. Let us check that the case is normal (in the sense of Remark 3.4). Assume by contradiction that $p^0 = 0$. From the previous paragraph, we know 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 $i = 1, \dots, N - 1$. Similarly to Sect. 5.1, we denote by $\eta^{[i]}$, for all $i = 0, \dots, N$, the $N + 1$ jumps of η . From the nontriviality of the couple (p^0, η) , we know that the jumps $\eta^{[i]}$ are not all zero. Since the initial condition $x(0)$ does not activate the running inequality state constraint, we know that $\eta^{[0]} = 0$. Now take $\hat{i} \in \{1, \dots, N\}$ such that $\eta^{[\hat{i}]} > 0$ is the last nonzero jump of η . On the other hand, since $p(T) = 0_{\mathbb{R}^2}$ and from the adjoint equation considered over the time interval $[t_{\hat{i}}, T]$, we obtain that $p(t_{\hat{i}}) = 0_{\mathbb{R}^2}$. From

the adjoint equation considered over the time interval $[t_{\hat{\tau}_-1}, t_{\hat{\tau}}]$, we obtain that

$$p(t_{\hat{\tau}}^-) = p(t_{\hat{\tau}}) - \eta^{\widehat{I}1} \begin{pmatrix} 1 \\ -16 \end{pmatrix} = \eta^{\widehat{I}1} \begin{pmatrix} -1 \\ 16 \end{pmatrix},$$

and $p(t) = e^{(t-t_{\hat{\tau}})A} \times p(t_{\hat{\tau}}^-)$ for all $t \in [t_{\hat{\tau}_-1}, t_{\hat{\tau}}]$ where

$$A := \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}.$$

We get that

$$\begin{aligned} & \int_{t_{\hat{\tau}_-1}}^{t_{\hat{\tau}}} \partial_2 H(x(\tau), u_{\hat{\tau}_-1}, p(\tau), p^0, \tau) \, d\tau \\ &= \int_{t_{\hat{\tau}_-1}}^{t_{\hat{\tau}}} p_2(\tau) - p_1(\tau) \, d\tau = \int_{t_{\hat{\tau}_-1}}^{t_{\hat{\tau}}} \left\langle p(\tau), \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle_{\mathbb{R}^2} \, d\tau \\ &= \eta^{\widehat{I}1} \left\langle \int_{t_{\hat{\tau}_-1}}^{t_{\hat{\tau}}} e^{(\tau-t_{\hat{\tau}})A} \, d\tau \times \begin{pmatrix} -1 \\ 16 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle_{\mathbb{R}^2} \\ &= \eta^{\widehat{I}1} \left\langle A^{-1} \times (\text{Id}_2 - e^{-(t_{\hat{\tau}}-t_{\hat{\tau}_-1)A})} \right\rangle_{\mathbb{R}^2} \times \begin{pmatrix} -1 \\ 16 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle_{\mathbb{R}^2}. \end{aligned}$$

The sign of this term is independent of the value of $\eta^{\widehat{I}1} > 0$. We compute numerically the above term with different positive values of $t_{\hat{\tau}} - t_{\hat{\tau}_-1}$ belonging to $(0, 2]$ (in particular $\frac{2}{4}, \frac{2}{5}$ and $\frac{2}{8}$ which are the values used in the next paragraph) and we always obtain a positive value which raises a contradiction with the nonpositive averaged Hamiltonian gradient condition provided in Theorem 3.1.

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 Sect. 5.1. In Fig. 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 $N = 8$).

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 Sect. A.1 we give some recalls on renorming Banach spaces and on the regularity of distance functions. Sect. A.2 is concerned with the sensitivity analysis of the state equation in Problem (OSCP). Then, in Sect. A.3, we give some recalls on Stieltjes integrations and on Fubini formulas. Finally Sect. A.4 is devoted to Duhamel formulas for Cauchy-Stieltjes Problems (FCSP) and (BCSP).

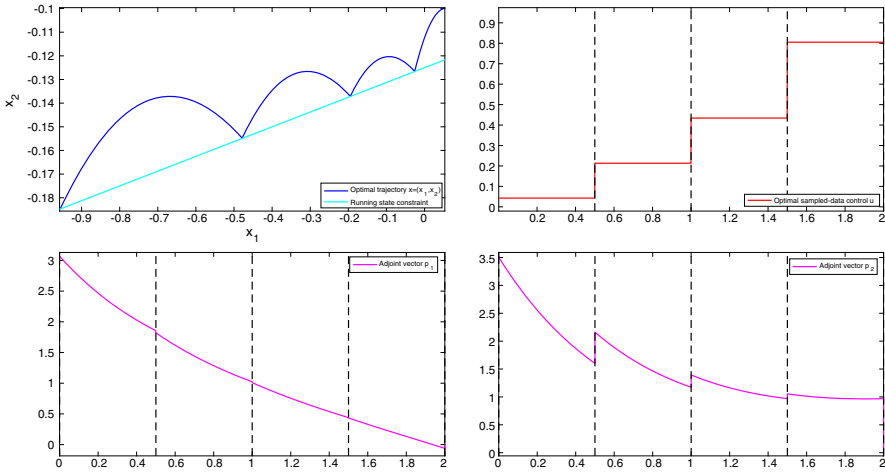


Fig. 10 Example 3 with $N = 4$

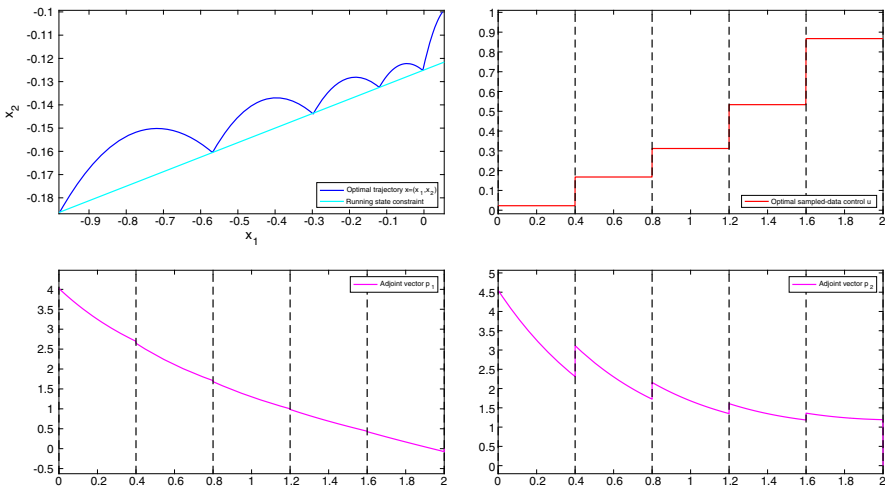


Fig. 11 Example 3 with $N = 5$

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

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

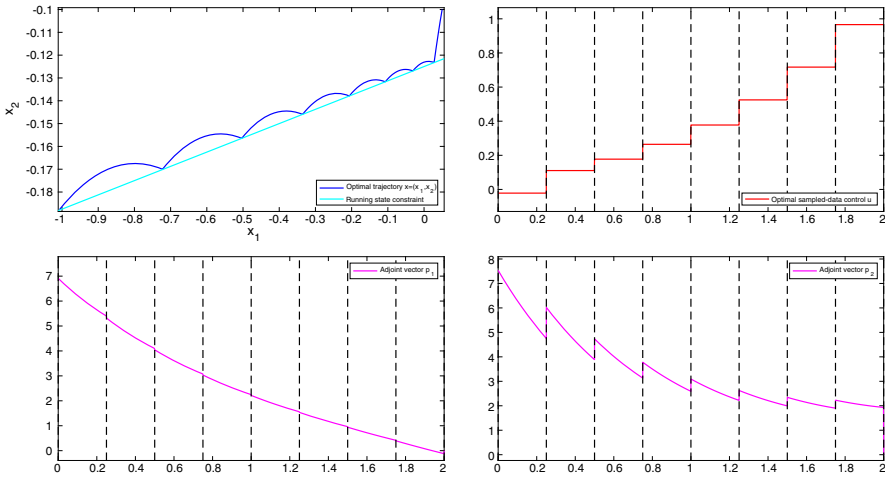


Fig. 12 Example 3 with $N = 8$

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 [56, Theorem 2.18, p. 42] or to [15, Proposition 4, p. 16] for a complete proof. □

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 [62, pp. 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 [62, Theorem 3.54, p. 313] and [15, Appendix B.2] for the proof of (ii). The proofs of (iii) and (iv) are straightforward. □

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

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

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., [28]).

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) $x \in AC([0, T'], \mathbb{R}^n)$, with $\dot{x}(t) = f(x(t), u(t), t)$ for a.e. $t \in [0, T']$, 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.1 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.

Remark A.1 Using the standard combination of the Gronwall lemma with the blow-up theorem for nonglobal solutions in ordinary differential equations theory, we can establish the following sufficient condition. Given a control $u \in L_m^\infty$, if there exist a nonnegative coercive mapping $\Theta : \mathbb{R}^n \rightarrow \mathbb{R}_+$ of class C^1 with two nonnegative constants $c_1, c_2 \geq 0$ such that $\langle f(x, u(t), t), \nabla \Theta(x) \rangle_{\mathbb{R}^n} \leq c_1 \Theta(x) + c_2$ for all $x \in \mathbb{R}^n$ and for a.e. $t \in [0, T]$, then $u \in \mathcal{AG}$. This sufficient condition covers, not only some typical situations for which $\mathcal{AG} = L_m^\infty$ (such as global Lipschitz dynamics, or more generally dynamics with a sublinear growth, taking $\Theta(x) := \|x\|_{\mathbb{R}^n}^2$ for all $x \in \mathbb{R}^n$), but also some dynamics with polynomial growth for which $\mathcal{AG} \subsetneq L_m^\infty$. As an illustration, take the scalar case $n = m = 1$ and the dynamics $f(x, u, t) := x - ux^3$ for all $(x, u, t) \in \mathbb{R} \times \mathbb{R} \times [0, T]$. In that example, if a scalar control $u \in L_m^\infty$ takes only nonnegative values on $[0, T]$, by considering $\Theta(x) := x^2$ for all $x \in \mathbb{R}$, we prove that $u \in \mathcal{AG}$.

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 $\overline{B}_{L_m^\infty}(u, \varepsilon_u) \subset \mathcal{AG}$, where $\overline{B}_{L_m^\infty}(u, \varepsilon_u)$ stands for the standard closed ball in L_m^∞ centered at u and of radius ε_u . Moreover the map*

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

is Lipschitz continuous.

Proof This proof is standard and essentially based on the classical Gronwall lemma. We refer to [17, Lemmas 1 and 3, pp. 3795–3797], [19, Lemmas 4.3 and 4.5, pp. 73–74] (in the general framework of time scale calculus) or to [15, Propositions 1 and 2, pp. 4–5] (in a more classical framework, closer to the present considerations) for similar statements with detailed proofs. \square

Remark A.2 Let $u \in \mathcal{AG}$ and $\varepsilon_u > 0$ as given in Lemma A.1. 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^∞ . 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 $w_z(\cdot, u) \in AC_n$ 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 [17, Lemma 4 and Proposition 1, pp. 3797–3798] for a similar statement with detailed proof. □

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 [17, Lemmas 4.8 and 4.9, pp. 77–78] for a similar statement with detailed proof. □

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 [5,22,23, 41,75] for more details. We also refer to [15, 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 Sect. A.1 for some details on dual spaces). We first recall the following Riesz representation theorem (see [57, Theorem 14.5, pp. 245–246] or [15, Proposition 7, p. 19]).

Proposition A.5 (Riesz representation theorem). *Let $\psi^* \in C_1^*$. Then there exists a unique $\eta \in NBV_1$ 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 $\eta = 0_{NBV_1}$.

Recall that if $\eta \in NBV_1$ 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 [41, p. 83] and [75, p. 288] for more details. Consequently the Fubini formula

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

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 variation. 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 $\eta = (\eta_j)_{j=1, \dots, q} \in BV_q$. 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 $\psi = (\psi_r)_{r=1, \dots, n} \in C_n$, all $\eta = (\eta_j)_{j=1, \dots, q} \in BV_q$ 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. $\psi \in C_n$ 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 $\eta = (\eta_j)_{j=1, \dots, q} \in BV_q$ 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 Sect. 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 with $A \in L^\infty([0, T], \mathbb{R}^{n \times n})$ 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 Sect. A.3, one can easily derive the following proposition. We also refer to [15, 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 with 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 Sect. B.2 with a simple change of variable). In Sect. 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 Sect. 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 $\mathfrak{h}(x) \in S$ 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 Sect. 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 Sect. 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 Sect. 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 [38]).

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

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

and

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

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

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

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*

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

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 $Dd_S(\mathfrak{h}(x(\cdot, u_k))) \in \partial d_S(\mathfrak{h}(x(\cdot, u_k)))$. Hence, if $\mathfrak{h}(x(\cdot, u_k)) \notin S$, it holds that

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

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

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

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. □

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*

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

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

Proof Let $z \in PC_m^\mathbb{T}$ 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 with 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 $PC_m^\mathbb{T}$ 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

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

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. □

B.1.2 Introduction of the adjoint vector

We can now conclude the proof of Theorem 3.1 (in the case $L = 0$) by introducing the adjoint vector p . We refer to Sects. 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 $p \in BV_n$ 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 Sect. 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 with $\partial_1 f(x, u, \cdot) \in L^\infty([0, T], \mathbb{R}^{n \times n})$.

Nonpositive averaged Hamiltonian gradient condition. From Inequality (9) and using notations introduced in Sect. 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 $z \in PC_m^\mathbb{T}$ 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 Sect. A.3, it follows that

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

for all $z \in PC_m^\mathbb{T}$ 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 $v \in U$, 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 $v \in U$ 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

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

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

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

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

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

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. Ackermann, J.E.: *Sampled-Data Control Systems: Analysis and Synthesis*, Robust System Design. Springer, Berlin (1985)
2. Aronna, M.S., Bonnans, J.F., Goh, B.S.: Second order analysis of control-affine problems with scalar state constraint. *Math. Program.* **160**(1–2, Ser. A), 115–147 (2016)
3. Åström, K.J.: On the choice of sampling rates in optimal linear systems. *IBM Res. Eng. Stud.* (1963)
4. Åström, K.J., Wittenmark, B.: *Computer-Controlled Systems*. Prentice Hall, Upper Saddle River (1997)
5. Bachman, G., Narici, L.: *Functional Analysis*. Dover Publications Inc., Mineola (2000). (Reprint of the 1966 original)
6. Bakir, T., Bonnard, B., Bourdin, L., Rouot, J.: Pontryagin-type conditions for optimal muscular force response to functional electrical stimulations. *J. Optim. Theory Appl.* **184**(2), 581–602 (2020)
7. Bettiol, P., Frankowska, H.: Normality of the maximum principle for nonconvex constrained Bolza problems. *J. Differ. Equ.* **243**(2), 256–269 (2007)
8. Bettiol, P., Frankowska, H.: Hölder continuity of adjoint states and optimal controls for state constrained problems. *Appl. Math. Optim.* **57**(1), 125–147 (2008)
9. Bini, E., Buttazzo, G.M.: The optimal sampling pattern for linear control systems. *IEEE Trans. Automat. Control* **59**(1), 78–90 (2014)
10. Boltyanskii, V.G.: *Optimal Control of Discrete Systems*. Wiley, New York (1978)
11. Bonnans, J.F., de la Vega, C.: Optimal control of state constrained integral equations. *Set Valued Var. Anal.* **18**(3–4), 307–326 (2010)
12. Bonnans, J.F., de la Vega, C., Dupuis, X.: First- and second-order optimality conditions for optimal control problems of state constrained integral equations. *J. Optim. Theory Appl.* **159**(1), 1–40 (2013)
13. Bonnans, J.F., Hermant, A.: No-gap second-order optimality conditions for optimal control problems with a single state constraint and control. *Math. Program.* **117**(1–2, Ser. B), 21–50 (2009)

14. Bonnard, B., Faubourg, L., Launay, G., Trélat, E.: Optimal control with state constraints and the space shuttle re-entry problem. *J. Dyn. Control Syst.* **9**(2), 155–199 (2003)
15. Bourdin, L.: 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)
16. Bourdin, L., Dhar, G.: Continuity/constancy of the hamiltonian function in a pontryagin maximum principle for optimal sampled-data control problems with free sampling times. *Math. Control Signals Syst.* **31**(4), 503–544 (2019)
17. Bourdin, L., Trélat, E.: Pontryagin maximum principle for finite dimensional nonlinear optimal control problems on time scales. *SIAM J. Control Optim.* **20**(4), 526–547 (2013)
18. Bourdin, L., Trélat, E.: Pontryagin maximum principle for optimal sampled-data control problems. In: 16th IFAC Workshop on Control Applications of Optimization CAO'2015 (2015)
19. Bourdin, L., Trélat, E.: Optimal sampled-data control, and generalizations on time scales. *Math. Control Relat. Fields* **6**(1), 53–94 (2016)
20. Bourdin, L., Trélat, E.: Linear-quadratic optimal sampled-data control problems: convergence result and Riccati theory. *Autom. J. IFAC* **79**, 273–281 (2017)
21. Bressan, A., Piccoli, B.: Introduction to the Mathematical Theory of Control, volume 2 of AIMS Series on Applied Mathematics. American Institute of Mathematical Sciences (AIMS), Springfield (2007)
22. Burk, F.E.: A Garden of Integrals, volume 31 of the Dolciani Mathematical Expositions. Mathematical Association of America, Washington (2007)
23. Carothers, N.L.: Real Analysis. Cambridge University Press, Cambridge (2000)
24. Cesari, L.: Optimization—Theory and Applications, volume 17 of Applications of Mathematics (New York). Springer, New York (1983). (Problems with ordinary differential equations)
25. Cho, D.I., Abad, P.L., Parlar, M.: Optimal production and maintenance decisions when a system experience age-dependent deterioration. *Optim. Control Appl. Methods* **14**(3), 153–167 (1993)
26. Clarke, F.H.: The generalized problem of Bolza. *SIAM J. Control Optim.* **14**(4), 682–699 (1976)
27. Clarke, F.H.: Optimization and Nonsmooth Analysis, volume 5 of Classics in Applied Mathematics, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1990)
28. Coddington, E.A., Levinson, N.: Theory of Ordinary Differential Equations. McGraw-Hill Book Company Inc, New York (1955)
29. Cots, O.: 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)
30. Cots, O., Gergaud, J., Goubinat, D.: Direct and indirect methods in optimal control with state constraints and the climbing trajectory of an aircraft. *Optim. Control Appl. Methods* **39**(1), 281–301 (2018)
31. Dmitruk, A.V.: On the development of Pontryagin's maximum principle in the works of A. Ya. Dubovitskii and A. A. Milyutin. *Control Cybern.* **38**(4A), 923–957 (2009)
32. Dmitruk, A.V., Kaganovich, A.M.: 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
33. Dmitruk, A.V., Osmolovskii, N.P.: Necessary conditions for a weak minimum in optimal control problems with integral equations subject to state and mixed constraints. *SIAM J. Control Optim.* **52**(6), 3437–3462 (2014)
34. Dmitruk, A.V., Osmolovskii, N.P.: Necessary conditions for a weak minimum in a general optimal control problem with integral equations on a variable time interval. *Math. Control Relat. Fields* **7**(4), 507–535 (2017)
35. Dmitruk, A.V., Osmolovskii, N.P.: A general Lagrange multipliers theorem and related questions. In: *Control Systems and Mathematical Methods in Economics*, volume 687 of *Lecture Notes in Economics and Mathematical Systems*, pp. 165–194. Springer, Cham (2018)
36. Dmitruk, A.V., Osmolovskii, N.P.: Proof of the maximum principle for a problem with state constraints by the V-change of time variable. *Discrete Contin. Dyn. Syst. Ser. B* **24**(5), 2189–2204 (2019)
37. Dubovitskii, A.Y., Milyutin, A.A.: Extremum problems in the presence of restrictions. *USSR Comput. Math. Math. Phys.* **5**(3), 1–80 (1965)
38. Ekeland, I.: On the variational principle. *J. Math. Anal. Appl.* **47**, 324–353 (1974)
39. Evans, L.C.: An introduction to mathematical optimal control theory. Version 0.2, Lecture notes
40. Fadali, M.S., Visioli, A.: Digital Control Engineering: Analysis and Design. Elsevier, New York (2013)
41. Faraut, J.: Calcul intégral (L3M1). EDP Sciences, Les Ulis (2012)
42. Gamkrelidze, R.V.: Optimal control processes for bounded phase coordinates. *Izv. Akad. Nauk SSSR. Ser. Mat.* **24**, 315–356 (1960)

43. Girsanov, I.V.: Lectures on Mathematical Theory of Extremum Problems. Springer, Berlin (1972). Edited by B. T. Poljak, Translated from the Russian by D. Louvish, Lecture Notes in Economics and Mathematical Systems, vol. 67
44. Grasse, K.A., Sussmann, H.J.: Global controllability by nice controls. In: Nonlinear Controllability and Optimal Control, volume 133 of Monographs and Textbooks in Pure and Applied Mathematics, pp. 33–79. Dekker, New York (1990)
45. Grüne, L., Pannek, J.: Nonlinear Model Predictive Control. Communications and Control Engineering Series, 2nd edn. Springer, Cham (2017). (Theory and algorithms)
46. Halkin, H.: A maximum principle of the pontryagin type for systems described by nonlinear difference equations. *SIAM J. Control* **4**(1), 90–111 (1966)
47. Hartl, R.F., Sethi, S.P., Vickson, R.G.: A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev.* **37**(2), 181–218 (1995)
48. Hestenes, M.R.: Calculus of Variations and Optimal Control Theory. Robert E. Krieger Publishing Co., Inc., Huntington (1980). Corrected reprint of the 1966 original
49. Hiriart-Urruty, J.B.: Les mathématiques du mieux faire, vol. 2: La commande optimale pour les débutants. Collection Opuscules (2008)
50. Holtzman, J.M., Halkin, H.: Discretional convexity and the maximum principle for discrete systems. *SIAM J. Control* **4**(2), 263–275 (1966)
51. Ioffe, A.D., Tihomirov, V.M.: Theory of Extremal Problems, volume 6 of Studies in Mathematics and its Applications. North-Holland Publishing Co., Amsterdam (1979). (Translated from the Russian by Karol Makowski)
52. Jacobson, D.H., Lele, M.M., Speyer, J.L.: New necessary conditions of optimality for control problems with state-variable inequality constraints. *J. Math. Anal. Appl.* **35**, 255–284 (1971)
53. Kim, N., Rousseau, A., Lee, D.: A jump condition of PMP-based control for PHEVs. *J. Power Sources* **196**(23), 10380–10386 (2011)
54. Landau, I.D., Zito, G.: Digital Control Systems: Design: Identification and Implementation. Springer, Berlin (2006)
55. Lee, E.B., Markus, L.: Foundations of Optimal Control Theory, 2nd edn. Robert E. Krieger Publishing Co., Inc, Melbourne (1986)
56. Li, X., Yong, J.: Optimal Control Theory for Infinite-Dimensional Systems. Systems and Control: Foundations and Applications. Birkhäuser Boston Inc, Boston (1995)
57. Limaye, B.V.: Functional Analysis, second edn. New Age International Publishers Limited, New Delhi (1996)
58. Malanowski, K.: On normality of Lagrange multipliers for state constrained optimal control problems. *Optimization* **52**(1), 75–91 (2003)
59. Maurer, H.: On optimal control problems with bounded state variables and control appearing linearly. *SIAM J. Control Optim.* **15**(3), 345–362 (1977)
60. Maurer, H., Kim, J.R., Vossen, G.: On a state-constrained control problem in optimal production and maintenance. In: Optimal Control and Dynamic Games, pp. 289–308. Springer (2005)
61. Milyutin, A.A.: Extremum problems in the presence of constraints. Ph.D. thesis, Doctoral Dissertation, Institute of Applied Mathematics, Moscow (1966)
62. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. I, volume 330 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Berlin (2006). (Basic theory)
63. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: The Mathematical Theory of Optimal Processes. Wiley, New York (1962)
64. Puchkova, A., Rehbock, V., Teo, K.L.: Closed-form solutions of a fishery harvesting model with state constraint. *Optim. Control Appl. Methods* **35**(4), 395–411 (2014)
65. Rampazzo, F., Vinter, R.B.: A theorem on existence of neighbouring trajectories satisfying a state constraint, with applications to optimal control. *IMA J. Math. Control Inf.* **16**(4), 335–351 (1999)
66. Robbins, H.: Junction phenomena for optimal control with state-variable inequality constraints of third order. *J. Optim. Theory Appl.* **31**(1), 85–99 (1980)
67. Rockafellar, R.T.: State constraints in convex control problems of Bolza. *SIAM J. Control* **10**, 691–715 (1972)
68. Santina, M.S., Stubberud, A.R.: Basics of sampling and quantization. In: Handbook of Networked and Embedded Control Systems, Control Engineering, pp. 45–69. Birkhauser, Boston (2005)

69. Sethi, S.P., Thompson, G.L.: *Optimal Control Theory*, 2nd edn. Kluwer Academic Publishers, Boston (2000). (Applications to management science and economics)
70. Trélat, E.: *Contrôle optimal: théorie & applications*. Vuibert, Paris (2005)
71. van Keulen, T., Gillot, J., de Jager, B., Steinbuch, M.: Solution for state constrained optimal control problems applied to power split control for hybrid vehicles. *Automat. J. IFAC* **50**(1), 187–192 (2014)
72. Van Reeve, V., Hofman, T., Willems, F., Huisman, R., Steinbuch, M.: Optimal control of engine warmup in hybrid vehicles. *Oil Gas Sci. Technol.* **71**(1), 14 (2016)
73. Vinter, R.: *Optimal Control*. Modern Birkhäuser Classics. Birkhäuser, Boston (2010). (Paperback reprint of the 2000 edition)
74. Volz, R.A., Kazda, L.F.: Design of a digital controller for a tracking telescope. *IEEE Trans. Automat. Control* **AC-12**(4), 359–367 (1966)
75. Wheeden, R.L., Zygmund, A.: *Measure and Integral*. Pure and Applied Mathematics (Boca Raton), 2nd edn. CRC Press, Boca Raton (2015). (An introduction to real analysis)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.