A Cost Reduction Procedure for Control-Restricted Nonlinear Systems

John A. Gómez Múnera*, Pablo S. Rivadeneira*,**,[1] and Vicente Costanza *

*Instituto de Desarrollo Tecnológico para la Industria Química
(INTEC, UNL-CONICET) Güemes 3450, 3000 Santa Fe, Argentina

**Universidad Nacional de Colombia - Sede Medellín, Facultad de Minas,
Grupo GITA, Cra. 80 №65-223, Medellín, Colombia

Abstract This paper describes a numerical scheme to approximate the solution of the optimal control problem for nonlinear systems with restrictions on the manipulated variable. The method proposed here systematically reduces the cost associated with successively updated control strategies, after proposing an initial seed trajectory based on approximation arguments. The numerical scheme follows two main lines of reasoning: the first one relying on linearizations around a seed state/control trajectory, exploiting a theoretical expression for the increment of the cost, and mainly valid in regular situations, although after some adaptations can be used when saturations occur. One of its advantages is that the decreasing of the cost can be assessed without integrating numerically the nonlinear dynamics. However, and because of the constraints, eventually this method fails, and an alternative approach must be activated to keep decreasing the cost. The second approach, based on specific control variations of the current control strategy, and it is activated depending on two theoretical criteria (failure alert) developed here. The initial control variation is derived from the differential Riccati equation for the linearized system and appropriate quadratic cost functions. Other variations, similar to those used in Pontryagin classical theorem for generating the final cone of states, are proposed by modifying the locations of 'switching times', and so producing oscillations in the interior of regular periods. The performance of the numerical combined method and the related mathematical objects are illustrated by optimizing some two-dimensional nonlinear systems with scalar bounded controls.

Keywords: Optimal control, restricted controls, nonlinear systems, control variations, LQR problems.

1 Introduction

The Hamiltonian formalism has been at the core of the development of modern optimal control [1, 2, 8, 19, 24]. For regular problems the standard mathematical results towards optimal solutions are: (i) the Hamilton-Jacobi-Bellman equation [3, 4], which is a first-order nonlinear partial differential equation (PDE), or equivalently (ii) Hamilton Canonical Equations (HCE), which are a set of $2n$ ordinary differential equations (ODE) subject to two-point mixed boundary conditions [5, 6, 11, 21]. The bounded-control context may lead to non regular optimal control problems, for whose solutions there are not standard

[1]Corresponding author. E-mail: psrivade@santafe-conicet.gov.ar

recipes [2, 20, 22, 23]. Since the early sixties, the Pontryagin Maximum Principle (PMP) has been the standard theoretical setup to treat such non regular situations [19], which despite being a theoretically powerful tool is difficult to use in practical situations. PMP has been systematized for very few cases, especially in time-optimal problems for linear systems with bounded controls [1, 2, 17, 19].

When the optimal control is a smooth trajectory only in a nontrivial subinterval $I \subseteq [0, t_f]$, then the optimal control problem is called 'partially regular'. Clearly, if $I = [0, t_f]$, then the problem is also 'totally regular'. The treatment of partially regular problems will be the main concern here when considering nonlinear systems with input constraints. The standard procedure for nonlinear systems requires the integration of the Hamiltonian Canonical Equations (HCE) by using indirect shooting methods, which because of HCE's inherent instability leads to unpredictable results, and therefore cannot be applied in real-time.

In contrast to linear systems, there are very few control strategies dealing with the constrained nonlinear optimal control reported in the literature. Recently [16] a suboptimal strategy was devised based on the optimal solution of a linear time-varying (LTV) linearization, updated successively until reaching some predefined tolerance. However, only bang-bang type solutions were obtained this way. In [25] a penalty function method and an improved Newton algorithm were designed for solving the problem. Although the proposed method was globally convergent with quadratic rate, the strategy needs the calculation of the solution of the nonlinear HCEs by numerical integration.

The PMP brings forth open-loop optimal control trajectories when regularity is lacking. Therefore, in control-restricted, non-regular situations, these open-loop situations often present saturations of the control variable. On the contrary, regulars problems conduct to smooth optimal trajectories. Consequently, methods based on improving solutions to regular approximate problems will eventually need to resort to other, "non-regular" control variations, in order to force the cost to decrease below the value obtained through regular solutions.

In this paper, a new numerical method is proposed to approximate, during a first stage, the optimal control solution for nonlinear systems with input constraints by using successive LTV 'regular' approximations along with a theoretical expression for the cost (valid in regular situations). After proposing an initial seed trajectory u_{seed}, the cost is reduced by gradient-based iterations until some point, when this regular method fails due to the influence of the control constraints. When this happens, another numerical strategy is activated, achieving further reductions of the cost. The 'alternative strategy' generates control variations of the last updated control trajectory obtained during the regular procedure. Several types of variations are proposed, the first one constructed from the differential Riccati equation (DRE) for the linearized system under an appropriated quadratic cost functional, obtained after disturbing the final state value (as it will be explained in the text). The second one used here comes from Pontryagin 'spatial variations', and the third one are oscillations created in the interior of regular periods, by modifying the locations of switching-times τ_i (time instants when the control trajectory meets the bounds). One main advantage of the proposed method is that the cost decrement can be computed without need for the numerical integration of the nonlinear HCEs. The approximation of the optimal control is obtained by choosing the variation which produces the minimal cost in each iteration.

The existence of the derivatives of the cost functional involved in this strategy imply unconstrained controls and regularity of the problem, which are not guaranteed in the proposed method. However, the approach has the following main positive features: (i) the resulting formula for the derivatives produce in general acceptable results during a certain number of iterations, and (ii) the expression of the cost increment due to control variations is valid as far as the control variation is small, regardless of being constrained or not, which allows to activate an alert condition when this increment fails to be negative (a clear indication that the validity of the general setup has collapsed). Two criteria to give the alert signal are developed and discussed. If such an alert occurs, then the alternative control variations are generated and tested to force further decreasing of the cost.

In general the resulting control will at least be suboptimal, although in the test-cases illustrated here the optimal trajectories were reached through the proposed procedure.

The article has the following structure: In Section 2, the optimal control problem and its formalism for finding solutions under regular situations are presented. Also, the bounded-control scenario is settled and the problem with input constraints posed. In Section 3, the gradient-based algorithm to reduce the cost is substantiated, including the formula for the cost increment in terms of a time-variant linearization of the augmented dynamics. The criteria to detect and activate the alternative strategy are also proposed and discussed. In Section 4, a nonlinear case-studies are treated and its optimal solution computed applying the proposed method. Finally, conclusions and perspectives are exposed.

2 Optimal control for nonlinear systems

2.1 The Hamiltonian formalism. Unconstrained controls

The finite-horizon, time-constant formulation of a nonlinear problem with free final states and unconstrained controls attempts to minimize the cost

$$\mathcal{J}(u) = \int_0^{t_f} L(x(\tau), u(\tau))\, d\tau \; + \; \mathcal{K}(x(t_f)), \tag{1}$$

subject to the following dynamic constraint (a nonlinear control system)

$$\dot{x} = f(x, u), \quad x(0) = x_0, \tag{2}$$

with the state x evolving in some open set $\mathcal{O}$ of $\mathbb{R}^n$ containing x_0, and assuming that the vector field f and $\mathcal{K}$ is (at least) a C^1 function in its domain. The admissible control trajectories are the piecewise continuous functions $u : [0, t_f] \to \mathbb{R}^m$, $t_f < \infty$. $\mathcal{J}(u)$ is a short notation for $\mathcal{J}(0, t_f, x_0, u(\cdot))$. To be precise, the notation $x(\cdot)$ inside the Lagrangian L of Eq. (1) is the state-trajectory corresponding to the control strategy $u(\cdot)$ via $\{x(t) = \varphi(t, x_0, u(\cdot)), t \in [0, t_f]\}$, with φ denoting the transition function or transition map of the system (2), (see [22]), assumed to exist in the appropriate range of its variables, and to coincide with the unique solution to the dynamics (2).

Clearly the desired final (target) state is 0, so the problem aims to send x_0 as near as possible to 0 with optimal cost (regulation problem). The Hamiltonian of the problem is defined as usual [8]

$$H(x, \lambda, u) \triangleq L + \lambda' f = L(x, u) + \lambda' f(x, u) \ , \tag{3}$$

and seen as a function that map $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, where λ is the adjoint or costate variable, with values in $\mathbb{R}^n$, in particular

$$\lambda(t_f) = \left(\frac{\partial \mathcal{K}}{\partial x} \right)' (x(t_f)). \tag{4}$$

The problem (or the Hamiltonian H) is said to be regular [18] when for each pair (x, λ) there exists a unique H-minimizing control value u^0, with u^0 a continuous function of its variables, called the 'H-minimal' control. In this case it is possible to define the 'optimal' Hamiltonian $\mathcal{H}^0$ as

$$\mathcal{H}^0(x, \lambda) \triangleq H(x, \lambda, u^0(x, \lambda)) \ , \tag{5}$$

Since in that case H results smooth with respect to all its variables, the definition of u^0 clearly implies

$$\frac{\partial H}{\partial u}(x, \lambda, u^0(x, \lambda)) \equiv 0 \ , \tag{6}$$

and the 'value function' (or the 'Bellman function') V of the problem, i.e.

$$V(x, t) \triangleq \inf_{u(\cdot)} \mathcal{J}(u) \, , \tag{7}$$

satisfies the partial differential equation known as the 'Hamilton-Jacobi-Bellman' (HJB) equation, namely:

$$\frac{\partial V}{\partial t} + \mathcal{H}^0 \left[x, \left(\frac{\partial V}{\partial x} \right)' \right] = 0 \ , \ V(x, t_f) = \mathcal{K}(x(t_f)), \tag{8}$$

for regular problems, and from (5 - 8) it is clear that the optimal costate trajectory λ^* verifies

$$\lambda^*(t) = \left(\frac{\partial V}{\partial x}(x^*(t), t) \right)' , \tag{9}$$

'the Hamiltonian Equations' HE with their respective boundary conditions (associated to the HJB problem) [22]

$$\begin{cases} \dot{x} = \left(\frac{\partial \mathcal{H}^0}{\partial \lambda} \right)' \ ; \ x(0) = x_0 \ , \\ \dot{\lambda} = -\left(\frac{\partial \mathcal{H}^0}{\partial x} \right)' ; \ \lambda(t_f) = \left(\frac{\partial \mathcal{K}}{\partial x} \right)' (x(t_f)), \end{cases} \tag{10}$$

give rise to a $2n$-dimensional two-point boundary-value problem in (x, λ) [19, 22]. If the HE are solved (their solution denoted as $(x^*(\cdot), \lambda^*(\cdot))$), then the optimal control strategy $u^*(\cdot)$ is calculated at each time instant t through

$$u^*(t) = u^0(x^*(t), \lambda^*(t)) \ . \tag{11}$$

The immediate (and well-known) drawback to solve this optimal problem is the lack of information on the initial costate value $\lambda(0)$, which precludes the ODEs (10) to be integrated from initial conditions and applied to in real-time. The final value $x(t_f)$ also is unknown. For the nonlinear and unconstrained optimal control problem, a couple of partial differential equations (which are solved off-line) allows to find these missing boundary conditions [11]. Recently [9], efficient numerical methods are provided to solve the linear and constrained problem.

2.2 The bounded-control case

Commonly, the manipulated variable can move inside and on the boundary of some bounded subset of a metric space, then it is natural to assume that the admissible set of control values is a compact subset of $\mathbb{R}^m$. In the scalar case,

$$u(t) \in \mathbb{U} \triangleq [u_{\min}, u_{\max}] \ . \tag{12}$$

The qualitative features of optimal control solutions to bounded problems are significantly different from those of unbounded ones [19]. But questions about how much they actually differ, which classes of problems lead to bang-bang controls, and whether their solutions are just saturations of the optimal trajectories of unbounded problems, are still open.

The search for solutions to restricted problems most frequently falls in the domains of the Pontryagin Principle (PMP) [19]. However, even when solved, PMP is not flexible enough to treat state perturbations: no optimal feedback laws arise from the application of PMP equations, just open-loop control strategies.

Let us assume from the beginning that there exists a time instant τ in $(0, t_f)$ for which the optimal control of the restricted problem posed above takes the value: $u^*_{x_0}(\tau) \in (u_{\min}, u_{\max}) = Int(\mathbb{U})$.

The PMP standard formulation for the original problem indicates [1, 12, 19, 24] that, if there exists an optimal control solution $u^*_{x_0}(\cdot)$ (for the bounded-control problem and all initial conditions x_0), then there should also exist an optimal costate trajectory $\lambda^*_{x_0}(\cdot)$, solution to the following (final-value, ODE) problem:

$$\dot{\lambda} = -\left(\frac{\partial H}{\partial x}\right)'(x^*_{x_0}, \lambda, u^*_{x_0}) \ ; \ \lambda'(t_f) = \nabla\mathcal{K}\left(x^*_{x_0}(t_f)\right) \ , \tag{13}$$

where $x^*_{x_0}(t_f)$ denotes the optimal final state value, *i.e.* the final value of the trajectory $x^*_{x_0}(\cdot)$ with dynamics

$$\dot{x} = f(x, u^*_{x_0}) \ ; \ x(0) = x_0 \ . \tag{14}$$

PMP also guarantees that the related functions of u

$$h_t(u) \triangleq H(x^*_{x_0}(t), \lambda^*_{x_0}(t), u), \tag{15}$$

defined for each $t \in [0, t_f]$ after the Hamiltonian H in Eq. (3), must assume their minimal values at $u^*_{x_0}(t)$. In addition, for the class of autonomous problems at hand,

$$h_t(u^*_{x_0}(t)) \equiv \bar{h}_{x_0} \ , \tag{16}$$

a constant in the whole optimization interval $[0, t_f]$. But from standard results [18, 22] it follows that for each t in some neighborhood of τ, the control trajectory $\tilde{u}$ defined by

$$\tilde{u}(t) \triangleq u^0(x^*_{x_0}(t), \lambda^*(t)) \;, \tag{17}$$

allows to construct the corresponding optimal strategy $u^*_{x_0}$ in the following way:

$$u^*_{x_0}(t) = \tilde{u}^{sat}(t) = \begin{cases} u_{\min} & \text{if} \quad \tilde{u}(t) \le u_{\min} \\ \tilde{u}(t) & \text{if} \quad u_{\min} < \tilde{u}(t) < u_{\max} \\ u_{\max} & \text{if} \quad \tilde{u}(t) \ge u_{\max} \end{cases} . \tag{18}$$

It is clear from the assumptions that $u^*_{x_0}(\tau) \in (u_{\min}, u_{\max})$. Then by continuity of regular controls this situation should extend to a maximal nontrivial interval $\mathcal{I} := [\tau_1, \tau_2] \subset [0, t_f]$ containing τ (these time instants are called 'switching times'), where the optimal state and costate variables $\left(x^*_{x_0}, \lambda^*_{x_0}\right)$ verify the following ODEs:

(i) from Eqs. (6, 13, 17, 18):

$$\begin{aligned} \dot{\lambda} &= -\left(\frac{\partial H}{\partial x}\right)'(x^*_{x_0}, \lambda, u^*_{x_0}) = -\left(\frac{\partial H}{\partial x}\right)'(x^*_{x_0}, \lambda, \tilde{u}) \\ &= -\left(\frac{\partial \mathcal{H}^0}{\partial x}\right)'(x^*_{x_0}, \lambda) \;;\;\; \lambda(\tau) = \lambda^*_{x_0}(\tau) \;; \end{aligned} \tag{19}$$

(ii) and by replacing $\tilde{u}$ from Eq. (17) into the dynamics and resorting to Eq. (6) again:

$$\dot{x} = f(x, u^*_{x_0}) = f(x, \tilde{u}) \tag{20}$$

$$= f(x, u^*_{x_0}) = f(x, u^0(x, \lambda^*_{x_0})) \tag{21}$$

$$= \left(\frac{\partial \mathcal{H}^0}{\partial \lambda}\right)'(x, \lambda^*_{x_0}) \;;\;\; x(\tau) = x^*_{x_0}(\tau) \;. \tag{22}$$

3 Gradient-based cost reduction methods

3.1 Cost reduction in regular problems

Let us assume that a solution $x_{seed}(\cdot)$ of the Eq. (2), corresponding to some suboptimal piecewise-continuous control $u_{seed}(\cdot)$, is known. Then the dynamics of the system will be approximated in the neighborhood of $x_{seed}(t)$ and $u_{seed}(t)$. It is convenient to define the new variables

$$\begin{aligned} X(t) &\triangleq x(t) - x_{seed}(t) \\ U(t) &\triangleq u(t) - u_{seed}(t), \end{aligned} \tag{23}$$

where the state vector $x(t)$ is assumed to deviate from the seed state vector $x_{seed}(t)$ due to small control corrections $U(t)$, allowing that the error $X(t)$ can be kept small.

The time-dependent matrices

$$
\begin{aligned}
A(t) &\triangleq \frac{df}{dx}(x_{seed}(t), u_{seed}(t)) \\
B(t) &\triangleq \frac{df}{du}(x_{seed}(t), u_{seed}(t)),
\end{aligned} \tag{24}
$$

characterize the time-variant linear system

$$
\begin{aligned}
\dot{X}(t) &= f(x(t), u(t)) - f(x_{seed}(t), u_{seed}(t)) \\
&\approx A(t)X(t) + B(t)U(t)\ .
\end{aligned} \tag{25}
$$

In this Subsection the original unconstrained optimal control problem posed before will be slightly transformed into an equivalent one. In the new formulation two state variables x_e and z (associated with the original cumulative Lagrangian and with the final penalization, respectively) are added to the initial system. The new state

$$
x^{\#} \triangleq \begin{pmatrix} x \\ z \\ x_e \end{pmatrix}, \tag{26}
$$

will evolve in $\mathbb{R}^{n+2}$ with dynamics

$$
\dot{x}^{\#} = f^{\#}(x^{\#}, u) = \begin{pmatrix} f(x,u) \\ 0 \\ L(x,u) + \frac{z}{t_f} \end{pmatrix},\ x^{\#}(0) = \begin{pmatrix} x_0 \\ \mathcal{K}(x(t_f)) \\ 0 \end{pmatrix}. \tag{27}
$$

Thus, in these variables, the cost function (1) can be expressed in a pure final penalization form, namely:

$$
\mathcal{J}(u) = x_e(t_f). \tag{28}
$$

The expressions of the relevant objects for the augmented system result in [13, 24]

$$
A^{\#}(t) = \begin{pmatrix} \frac{\partial f(x,\tilde{u})}{\partial x} & 0 & 0 \\ 0 & 0 & 0 \\ \frac{\partial L(x,\tilde{u})}{\partial x} & \frac{1}{t_f} & 0 \end{pmatrix}, \tag{29}
$$

$$
B^{\#}(t) = \begin{pmatrix} \frac{\partial f(x,\tilde{u})}{\partial u} \\ 0 \\ \frac{\partial L(x,\tilde{u})}{\partial u} \end{pmatrix}, \tag{30}
$$

$$
[\lambda^{\#}(t_f)]' = (\nabla\mathcal{K}(x(t_f)), -1, 1)\ , \tag{31}
$$

where $\lambda^{\#}(t_f)$ is the final value of the costate variable for the augmented system. In what follows the # sign will be avoided for the sake of simplicity, assuming that the meaning of the new augmented state and cost expressions are clear. The numerical approach for the minimization of $\mathcal{J}(u)$ for a nonlinear system in the unconstrained context [22] attempts to compute a control variation (or 'perturbation') $\mu : [0, t_f] \subset \mathbb{R}^m$ of u_{seed} such that

$$
\mathcal{J}(u_{seed} + h\mu) < \mathcal{J}(u_{seed}) \tag{32}
$$

for a sufficiently small 'step size' $h > 0$. A suitable variation can be chosen [22] as:

$$\mu(t) \triangleq -B(t)'\Phi(\tau,t)'\lambda(t_f), \tag{33}$$

where $\Phi(\tau,t)$ is the fundamental matrix associated with $A(\cdot)$, *i.e.* the solution to the linear equation:

$$\frac{\partial\Phi(\tau,t)}{\partial\tau} = A(\tau)\Phi(\tau,t),\ \Phi(\tau,\tau) = I\ . \tag{34}$$

The cost increment then can be estimated by the equation [22]

$$\Delta\mathcal{J}_S = \lambda(t_f)'\int_0^{t_f}\Phi(t_f,s)B(s)\Delta u(s)ds\ , \tag{35}$$

where the control increment is, for regular problems,

$$\Delta u \triangleq h\mu\ . \tag{36}$$

3.2 Updating the numerical method to the constrained control scenario

First approximations of the switching times $\tau_{i,0}$ (defined in Section 2.2, and where the second index represents the iteration number) of the optimal saturation points τ_i can be estimated after obtaining the solution to the unbounded-control problem via the HJB (8) and then saturating the resulting control, namely $\tilde{u}$ as in Eq. (17). Thus, the initial switching times $\tau_{i,0}$ can be adopted as the time instants where $\tilde{u}$ meets the boundary of the admissible control-values set $\mathbb{U}$. Also, the feasible version $u^*_{x_0} = \tilde{u}^{sat}$ can be adopted as the initial candidate u_0 (called the 'seed' trajectory and denoted by u_{seed}) for the solution to the constrained problem, *i.e.*

$$u_{seed}(t) = u_0(t) \triangleq u^*_{x_0}(t) = \tilde{u}^{sat}(t) = [u^0(x(t),\lambda(t))]^{sat}\ . \tag{37}$$

Other controls can be chosen as seeds trajectories, according to intuition or knowledge about the particular problem at hand, especially when the HJB equation turns too difficult to be solved. In any case, the seed trajectory must be updated through successive iterations $u_j\ ;\ j = 1,2,\dots$ under the condition that

$$\mathcal{J}(u_{j+1}) \le \mathcal{J}(u_j) \le \cdots \mathcal{J}(u_{seed})\ ;\ j = 1,2,\dots, \tag{38}$$

which will be ascertained by calculating the cost decrements through Eq. (35), where the control increment needed verifies

$$\Delta u_{j+1} \triangleq u_{j+1} - u_j\ . \tag{39}$$

Taking into account constrained controls, the first goal is to adapt the algorithm described for regular problems in Subsection 3.1 to the constrained scenario ensuring the reduction of the cost. This procedure will consist of:

(i) Compute μ_j using Eq. (33) for $j = 0,1,2,\dots$, where $B(t), \Phi(\tau,t), \lambda(t_f)$ are evaluated from the seed trajectories x_{seed}, u_{seed}, then

(ii) saturate the regular control obtained, *i.e.* the updated control must verify

$$u_{j+1}(t) \triangleq [u_j(t) + h\mu_j]^{sat}, \tag{40}$$

for a small $h > 0$. In this way, all the updated controls u_j will always take admissible values. The h value must be chosen to obtain a negative value for $\Delta\mathcal{J}_S$, calculated with Eq. (35). If after reducing h, $\Delta\mathcal{J}_S$ remains positive, then the method fails and should be stopped.

In the case that the method starts to reduce the cost it is necessary to have an indication saying when the method must be stopped. This alert should verify

$$\Delta\mathcal{J}_S \approx \Delta\mathcal{J}_A \ , \tag{41}$$

where $\Delta\mathcal{J}_A$ is a measure of the cost deviation produced by the saturated control trajectory generated, and where the symbol $\approx$ means that both terms are of the same order of magnitude (these two assertions will be discussed in subsection 3.3). At the moment, if the condition (41) is satisfied, then the method finishes. If not,

(iii) Rename $u_{j+1} \to u_{seed}$, $j+1 \to j$, and update the seed state trajectory, *i.e.* evaluate x_{j+1} through numerical integration of

$$\dot{x} = f(x, u_{j+1}) \ , x(0) = x_0 \ , \tag{42}$$

and, thus update $x_{j+1} \to x_{seed}$.

(iv) Repeat the steps (i) to (iii) until $\Delta\mathcal{J}_S$ is considered to be small enough and reaches a defined tolerance.

3.3 Failure alert criteria for the constrained procedure

When restrictions on the control values are met, then the derivatives of the Hamiltonian or Lagrangian with respect to u are not defined on the boundaries of the admissible control-values set, and thus the formula for the cost increment (35) is no longer valid. It is expected that, when the saturated control trajectory is used, the procedure described in the previous subsection can eventually fail to reduce the cost. Here, two criteria will be generated to assess an eventual failure of the modified method for constrained controls:

(i) The Pontryagin main theorem related to PMP allows to analyze the effect of control perturbations on the final reached state. The truncation (saturation) of regular control strategies due to bounds can be regarded as such control perturbations, and their effects assessed via the sensitivity matrices as it is described in [19, 24].

For a saturated control strategy $\tilde{u}^{sat} \in \mathbb{U}$ in $[0, t_f]$ and considering that truncations occur at the switching-time instants $\tau_i \in (0, t_f)$, a variation control $u_s(\cdot)$ with respect to the feasible control $\tilde{u}^{sat}(\cdot)$ can be computed as follows:

$$u_s(t) \triangleq \begin{cases} w & \forall t \in [\tau', \tau) \\ \tilde{u}^{sat}(t) & \forall t \in [0, \tau') \cup [\tau, t_f) \end{cases} \ , \tag{43}$$

where w is the control value generated by the regular procedure at the time instant τ' near τ. Taking into account that $x_{\tilde{u}^{sat}}(\cdot)$ is the state trajectory corresponding to $\tilde{u}^{sat}(\cdot)$, and $x_{u_s}(\cdot)$ is associated to the variation control $u_s(\cdot)$, then its final state value verifies (see [24]):

$$x_{u_s}(t_f) = x_{\tilde{u}^{sat}}(t_f) + \varepsilon c Z_\tau(t_f)\Delta f(\tau, w) + \varepsilon l(\varepsilon), \tag{44}$$

where ε is a small number, $\varepsilon c = \tau - \tau'$, and $l(\varepsilon)$ is an (infinitesimal) vector whose norm tends to 0 when $\varepsilon \to 0$.

The change in the final value is then, approximately,

$$\Delta x(t_f) = x_{u_s}(t_f) - x_{\tilde{u}^{sat}}(t_f) = \varepsilon c Z_\tau(t_f)\Delta f(\tau, w). \tag{45}$$

From Eqs. (27, 28), the cost variations are reflected in the last $(n+2)$ component, whose expression is

$$\Delta \mathcal{J}_Z = [\ \Delta t^j Z_\tau(t_f)\Delta f(\tau, w)]_{n+2}\ , \tag{46}$$

where $\Delta t^j \triangleq \tau^{j+1} - \tau^j$; $j = 1, 2, ...$ replaces the (previously perturbation time accounted for) εc.

The sensitivity matrix Z_τ comes from spatial variations of the control, *i.e.* from the difference between the control proposed in Eq. (33) and the control that is truly applied after saturation. $Z_\tau(\cdot)$ is a $(n+2) \times (n+2)-$matrix, solution to the following initial-value problem defined in $[\tau, t_f]$:

$$\dot{Z} = A(t) \cdot Z\ ;\ Z(\tau) = I\ ,$$

and $\Delta f(\tau, w)$ is the difference between the function f evaluated at (τ, w) and at $(\tau, \tilde{u}^{sat})$, namely

$$\Delta f(\tau, w) \triangleq f(x_{\tilde{u}^{sat}}(\tau), w) - f(x_{\tilde{u}^{sat}}(\tau), \tilde{u}^{sat}(\cdot)). \tag{47}$$

Finally, considering now the sum of the effects of the variations for a finite number of switching times (k), the cost increment results

$$\Delta \mathcal{J}_Z = \sum_{i=1}^{k} [\Delta t_i^j Z_{\tau_i}(t_f)\Delta f(\tau_i, w_i)]_{n+2}. \tag{48}$$

(ii) The second criterion for failure detection arise after calculating the cost variation by using the regular control in the interval $[\tau^{'}, \tau)$. This new computation of the cost variation (called $\Delta \mathcal{J}_T$) is obtained by approximating Eq. (35) during the interval considered, more precisely

$$\Delta \mathcal{J}_T = \lambda(t_f)'\tilde{\Phi}\tilde{B}\Delta u \Delta t\ , \tag{49}$$

where $\Delta u = w$, and $\tilde{\Phi}$ and $\tilde{B}$ are average values for the fundamental matrix and the control matrix in $[\tau^{'}, \tau)$, respectively. $\Delta t = \Delta t^j$ maintains the same meaning as in the previous criterion.

Then, by considering the sum effect of the variations for a finite number of switching-times (k), finally $\Delta \mathcal{J}_T$ verifies

$$\Delta \mathcal{J}_T = \lambda(t_f)' \sum_{i=1}^{k} \tilde{\Phi}(\tau_i', \tau_i)\tilde{B}(\tau_i', \tau_i)\Delta u_i \Delta t_i^j. \tag{50}$$

Both criteria can be compared against the formula (35), and thus to decide if it is necessary to stop the regular procedure. When the order of magnitude of any of the two criteria is similar to the total cost reduction (35), then the failure alert suggest to stop the regular cost reduction (via the cost-differential method (33)), because it will probably diverge in one of the next iterations. In the example treated below both criteria yielded similar results, so at least numerically it can be stated that

$$\Delta\mathcal{J}_A \triangleq \Delta\mathcal{J}_Z \approx \Delta\mathcal{J}_T \quad , \tag{51}$$

where $\Delta\mathcal{J}_A$ accounts for the alert condition (41), as it was announced.

3.4 How to proceed when a failure alert is detected

Several types of control variations will be used to force additional cost reductions after the regular procedure is eventually stopped.

3.4.1 Variations based on the Differential Riccati Equation.

The first control variation proposed comes from the solution of the time-varying LQR problem defined by Eq. (25), and the following quadratic cost function:

$$\mathcal{J}_{lin}(u) \triangleq \int_0^{t_f} [X'(\tau)QX(\tau) + U'(\tau)RU(\tau)]d\tau \; + \; X'(t_f)SX(t_f), \tag{52}$$

where Q is a $n \times n$ positive semi-definite matrix, R is a $m \times m$ positive-definite matrix, and the final penalization S is also a semidefinite $n \times n$ matrix as usual in finite-time LQR problems. Here, and for the sake of simplicity, the dimension of the control space will be maintained in $m = 1$. .

The matrix $P_\triangle(\cdot)$ is the solution of the Differential Riccati Equation (DRE) for the optimal control problem posed by Eqs. (25, 52), namely

$$\dot{P}_\triangle(t) = P_\triangle(t)W(t)P_\triangle(t) - P_\triangle(t)A(t) - A'(t)P_\triangle(t) - Q, \;\; P_\triangle(t_f) = S, \tag{53}$$

where

$$W(t) \triangleq B(t)R^{-1}B'(t), \tag{54}$$

and the optimal corresponding of control is

$$U(t) = -R^{-1}B'(t)P_\triangle(t)X(t). \tag{55}$$

where $X(t)$ is the solution of Eq. (25), with the final conditions $X_i(t_f) = \pm\upsilon_i$, for some small $\upsilon_i > 0$, $i = 1, ...n$.

This type of variations result in small deviations from the seed trajectories (due to the penalizations involved in the cost functional). In turn, this approach ensures that the linearizations used in the definitions of X, U are also good approximations. Based on these properties, the Riccati type control variations are preferred to the variations described below, under theoretical grounds. Numerically they are slow, since their application involve solving the matrix DRE each time, so the resorting to the other variations may turn mandatory when the cost decrement becomes too small. The control variation is updated in this

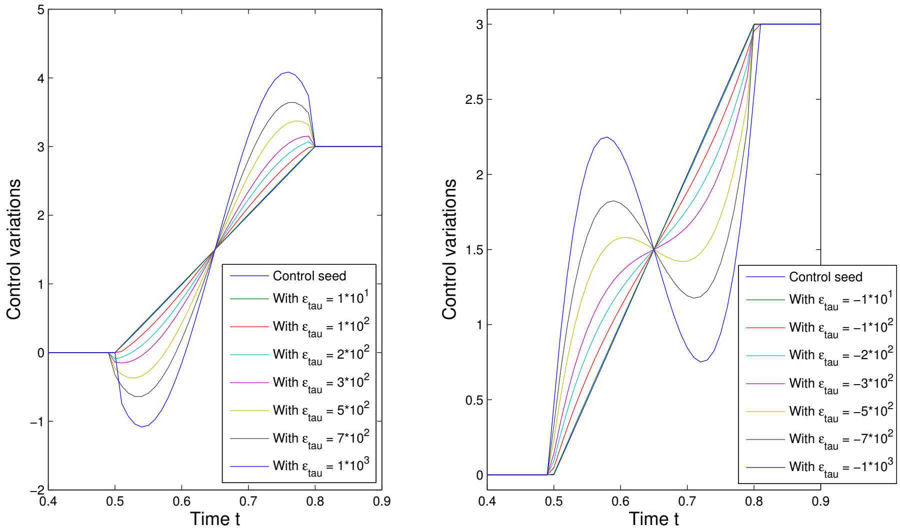

Figure 1: Effect of variations ε_τ with the δ_i constants for other types of control variations. $\delta_1 = \delta_2 = 0.01$.

case as

$$u_{j+1}(t) = [u_j(t) + U_j(t)]^{sat},\ j = 1, 2, \ldots \tag{56}$$

3.4.2 Other types of control variations

Let us assume that the saturated seed control has M_j switching times in an iteration j. The interval between each pair k $(k = 1, 2, ..., M_j/2)$ of subsequent switching-time instants will be denoted as (τ_{1k}, τ_{2k}). The following third-order curves can then be proposed as alternative control variations:

$$\mu_{jk}(t) \triangleq \begin{cases} 0 \text{ if } t \leq \tau_{1k} - \delta_1 \\ 0 \text{ if } t \geq \tau_{2k} + \delta_2 \\ (t - \tau_{1k} + \delta_1)(\tau^* - t)(t - \tau_{2k} - \delta_2)\varepsilon_\tau \quad \text{otherwise} \end{cases} , k = 1, 2, ..., M_j \tag{57}$$

where the inflection instant is

$$\tau^* = \frac{\tau_{1k} + \tau_{2k} + \delta_2 - \delta_1}{2} , \tag{58}$$

with δ_1, δ_2 perturbations of the switching times τ_1 and τ_2 respectively, and ε_τ is the amplitude of the deviation (see Figs. 1, 2). The control trajectories are updated then through

$$u_{j+1}(t) = [u_j(t) + \mu_{jk}]^{sat},\ j = 1, 2, \ldots \tag{59}$$

These are geometrically motivated variations. Other types may be generated along similar lines, as needed.

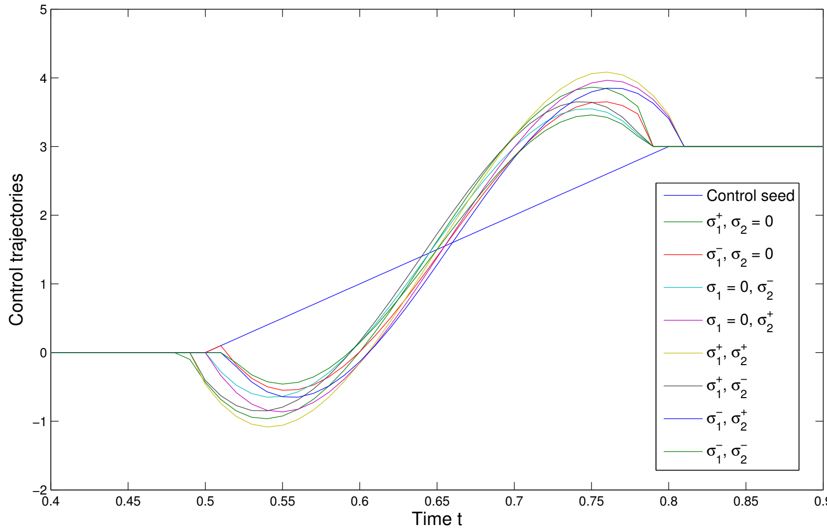

Figure 2: Effect of variations δ_i with the ε_τ constant for other types of control variations. $\varepsilon_\tau = 1 \times 10^3$.

3.4.3 Numerical algorithm for the gradient-based nonlinear method

Here, the proposed numerical algorithm is summarized:

(i) Generate the first approximation of the state and control trajectories, by solving the HJB equation or by any other criterion. After choosing the 'seed control trajectory' u_{seed} (for instance, that corresponding to the saturation of the optimal control of the unrestricted regular problem), x_{seed} is computed from the differential equation

$$\dot{x} = f(x, u_{seed}) \; ; \; x(0) = x_0 \; . \tag{60}$$

The trajectory u_{seed} provides the initial values for the switching times τ_i (the time instants where u_{seed} meets any of the bounds $u_{\min},\ u_{\max}$).

(ii) Update the matrix parameters $A(t), B(t), \lambda(t_f)$ by using the linearization proposed in 3.1, and after that update the control trajectory as in Eq. (40). This updated control will be taken as the new u_{seed} in the next iteration.

(iii) Applied the regular method while the cost keeps decreasing. The cost reduction is evaluated via Eq. (35).

(iv) While the regular method is working, the failure criteria must be checked in parallel (see subsection 3.3). If none of these criteria reach the same magnitude order than the cost deviation, *i.e.* $\Delta\mathcal{J}_S \approx \Delta\mathcal{J}_A$, then steps (ii) and (iii) are repeated. If there is a failure alert, then

(v) apply the control variations described by subsections 3.4.1 and 3.4.2, the control trajectory being modified in each case as in Eqs. (56, and 59), respectively, and their corresponding cost variations being calculated through Eq. (35). The variation producing the smallest cost is finally chosen to update the present seed control.

(vi) The last step is repeated until the cost reduction is too small (or meets some predefined tolerance) or becomes positive.

4 Applications and numerical results

This case-study is a modification of the 'cheapest stop of a train' problem already considered in [1, 7, 10]. Here, a friction term is included in the model as proposed elsewhere [14, 15] to obtain the following nonlinear system:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= \alpha u(t) + \beta u(t) x_2(t), \end{aligned} \tag{61}$$

where the states x_1 and x_2 are the position and the velocity of the train, respectively. The parameter α is a constant which normalizes the units of the control action, and β is a constant associated to the friction of the train over the rails. The manipulated variable $u(t)$ is a scalar-valued function, interpreted as the braking effort of a train. The feasible values of the control are assumed as $u(t) \in [0, 3] \subset \mathbb{R}$.

The initial conditions are chosen as $x_1(0) = 1$ and $x_2(0) = -1$. The parameters are taken as $\alpha = 1$ and $\beta = 0.15$. The objective is to minimize a quadratic cost functional during a finite time-horizon formulation. The matrices of the cost function are $Q = 10I$, $R = 0.5$, and the final penalization (also quadratic) is $\mathcal{K}(x(t_f)) = x(t_f)' S x(t_f)$ with $S = 100I$. The time horizon is settled in $t_f = 1$.

The linearization of the nonlinear system (61) around the seed trajectories (the subindex 'seed' is omitted) is described by the matrices

$$A(t) \triangleq \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.15 u^{sat}(t) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 20 x_1(t) & 20 x_2(t) & 1 & 0 \end{pmatrix}, \tag{62}$$

$$B(t) \triangleq \begin{pmatrix} 0 \\ 1 + 0.15 x_2(t) \\ 0 \\ u^{sat}(t) \end{pmatrix}, \tag{63}$$

The final costate of the linearized system is given by

$$\lambda(t_f)' = (200 x_1(t_f), 200 x_2(t_f), -1, 1). \tag{64}$$

The first h value adopted in the simulations was $h = 0.01$. A failure alert was detected when (35) attained the order of (50) at around the 30-th iteration. The cost reached was $\mathcal{J}_s = 15.4947$ and the regular method was stopped. Fig. 3a shows the reduction of the total cost using the 'regular' method.

Fig. 4a shows the corresponding evolution of $\Delta\mathcal{J}$ obtained from the procedure formulated in subsection 3.3. The result of Eq. (35) was compared against Eqs. (48,50) to generate the alert condition (41). Eq. (43) was used to generate

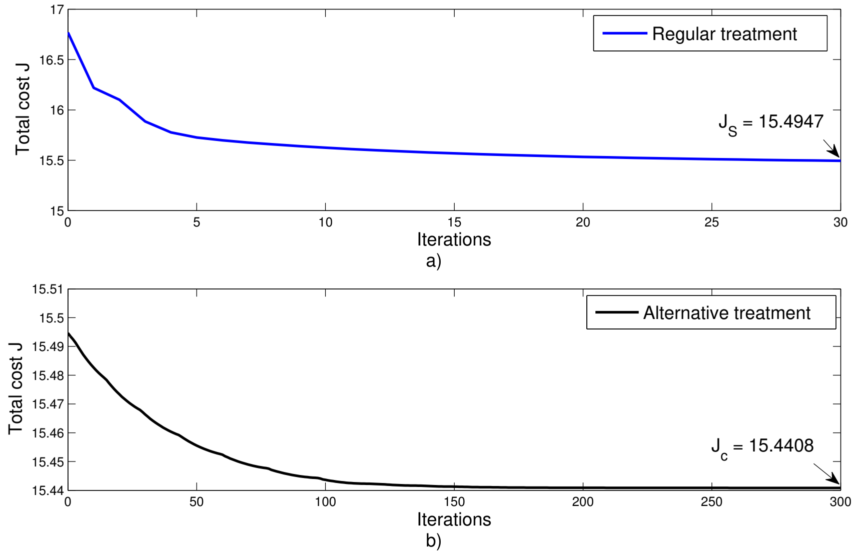

Figure 3: a) Evolution of the total cost resulting from regular method, b) evolution of the total cost resulting from alternative method.

an appropiatte control variation $u_s(\cdot)$

$$u_s(t) = \begin{cases} w_1 & \forall t \in [\tau_1', \tau_1) \\ w_2 & \forall t \in [\tau_2', \tau_2) \\ u^{sat}(t) & \forall t \in [0, \tau_1') \cup [\tau_1, \tau_2') \cup [\tau_2, t_f) \end{cases}, \tag{65}$$

where τ_1', τ_2' are small perturbations of the switching times.

Fig. 5 shows the alternative variations generated from Eq. (33). When the alert was active, the alternative control variations were essayed. For the variations the type (55), the disturbance imposed at the final state value was $X_i(t_f) = \pm 0.01$ in all directions i, and for the variations of the type described in subsection 3.4.2, $\varepsilon_\tau = 1$. Fig. 3b shows the reduction of the total cost until obtaining a minimum value through the alternative variations, and Fig. 3b illustrates the evolution of its corresponding cost variation along with the alert criteria. The relative reduction of the total cost after applying the regular and the alternative variations resulted in $\frac{J_{seed} - J_c}{J_c} * 100\% = 8.67\%$.

Figs. 6 and 7 show some control variations used during the numerical treatment, generated from Eq. (55). Fig. 6 depicts variations obtained by disturbing the final states, and the Fig. 7 the variations related to the Riccati equation. Fig. 8 illustrates different controls obtained after using different methods (saturation of the unbounded control strategy, Sontag's method, and the method proposed here). For the last iteration, the Hamiltonian $\bar{h}_{x_0} \equiv -23.56$ was constant in the whole time horizon interval $[0, t_f]$, confirming the optimatility of the finally computed control trajectory. Although in this example the optimal control was found, normally the numerical method developed here can just lead to suboptimal solutions.

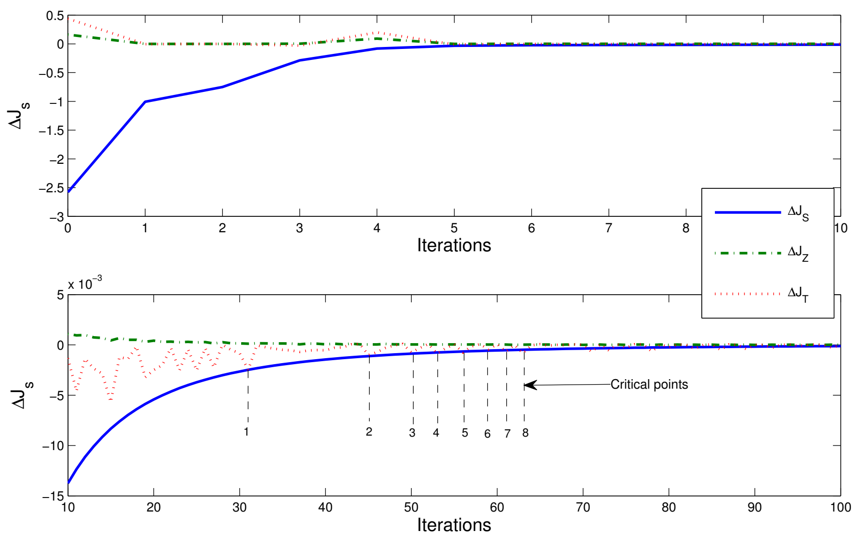

Figure 4: $\Delta\mathcal{J}$ obtained from the formulation 3.3.

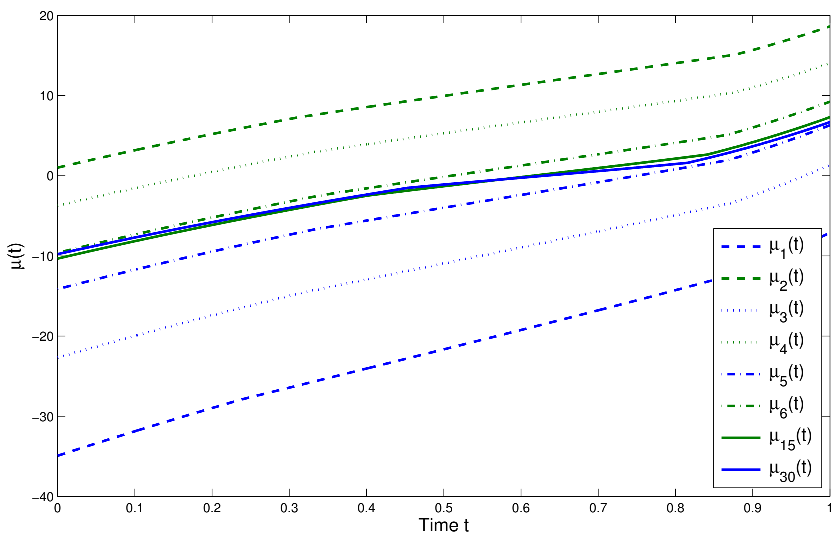

Figure 5: Control variations for regular strategy from Equation (33).

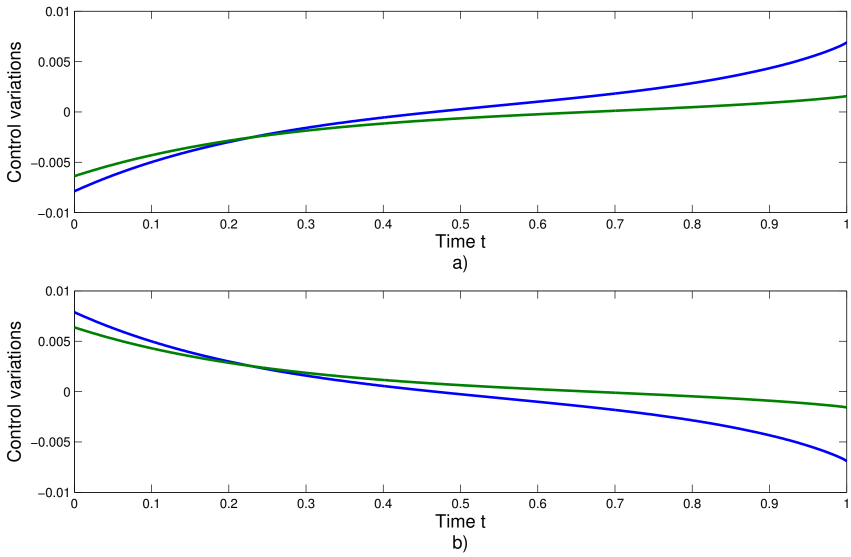

Figure 6: Control variations for all directions of the states, a) positive variations, b) negative variations.

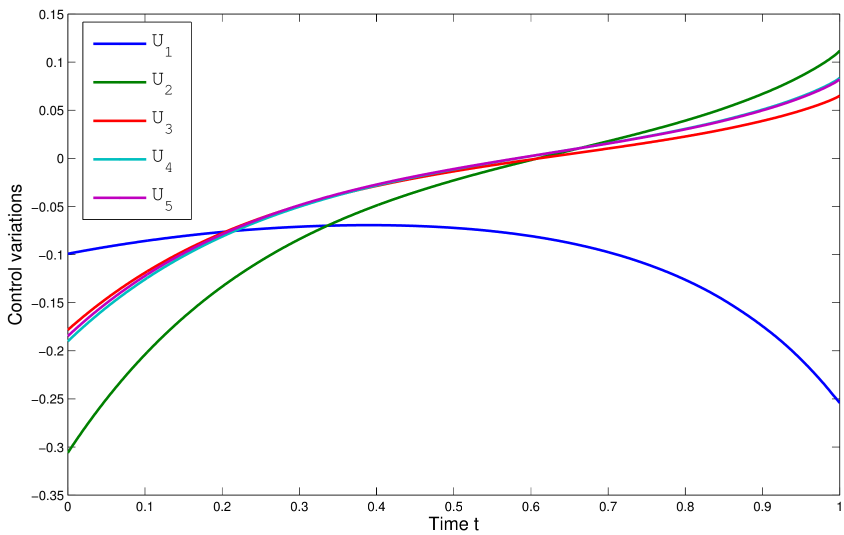

Figure 7: Control variations generated by solving the DRE.

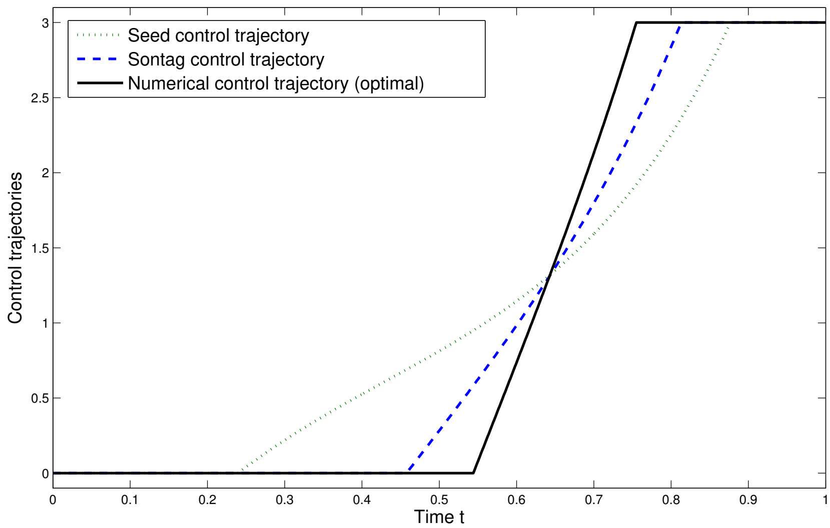

Figure 8: Control strategy for the bidimensional example.

Table 1: Final numerical values

	Seed trajectory	The best 'regular' trajectory	Optimal solution (proposed method)
$x_1(t_f)$	0.2187	0.17370	0.1657952
$x_2(t_f)$	-0.0137	-0.0485	-0.0450726
τ_1	0.23	0.4570	0.5440
τ_2	0.88	0.8130	0.7549
$\mathcal{J}$	16.779	15.4947	15.4408

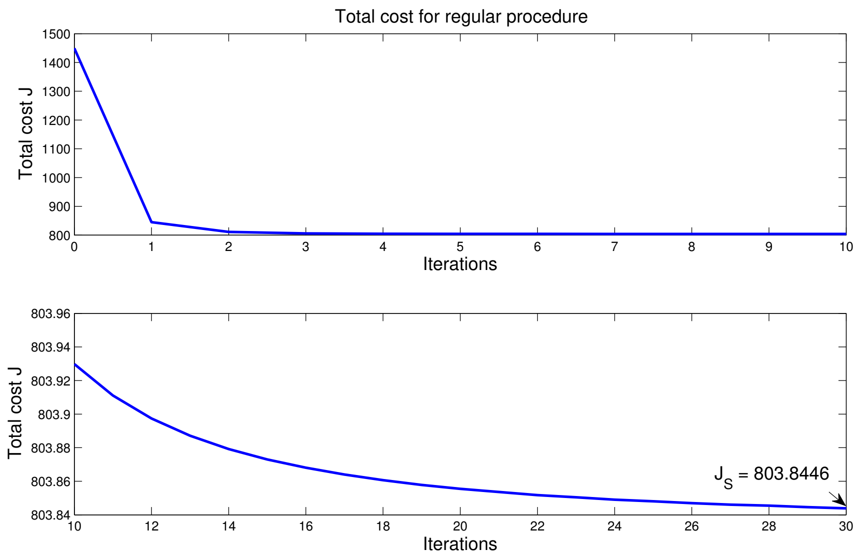

Figure 9: Evolution of the total cost resulting from regular method.

The second example discussed below is a modification of the previous problem. The nonlinear two-dimensional model was modified to produce an unstable linearization, and also different bounds in the control values were imposed. The proposed dynamics read

$$\begin{array}{rcl} \dot{x}_1(t) & = & 1.5x_1(t) + x_2(t) \\ \dot{x}_2(t) & = & \alpha u(t) + \beta u(t) x_2(t) \\ u(t) & \in & [-1, 1] \subset \mathbb{R}, \end{array} \tag{66}$$

with parameters $\alpha = 1$, $\beta = 0.15$, $Q = 10I$, $R = 0.5$, and $S = 100I$. The initial conditions are chosen now as $x_1(0) = 1$ and $x_2(0) = -0.5$. When the failure alert was activated, the cost was $\mathcal{J}_s = 803.8446$, at the 10-th iteration. Fig. 9 shows the reduction of total cost for the regular procedure.

After that, the alternative control variations were essayed. For variations of the type (55), the disturbance imposed in the final state were $X_i(t_f) = \pm 0.01$ in all directions i, and for the other variations described in subsection 3.4.2, $\varepsilon_\tau = 1 \times 10^6$ and $\delta_1 = \delta_2 = \pm 0.01$. Fig. 10 shows the reduction of the total cost until obtaining the minimum value after resorting to the alternative variations. The relative reduction of the total cost after applying the regular and the alternative variations resulted in $\frac{J_{seed} - J_c}{J_c} * 100\% = 80.7\%$.

The behavior of this example preserved some features of the previous problem: the same number of switching-times and a single regular period. The control variations associated with Equation (65) generate cost reductions $\Delta\mathcal{J}_S$ (35) significantly bigger than errors $\Delta\mathcal{J}_Z$ (48), i.e. acceptable during the first iterations. Around iteration 40 it was considered that the absolute value of $\Delta\mathcal{J}_S$ and $\Delta\mathcal{J}_Z$ were too close (see Figure 11) as to activate the failure alert. The Fig. 12 shows different controls obtained along the different methods used in

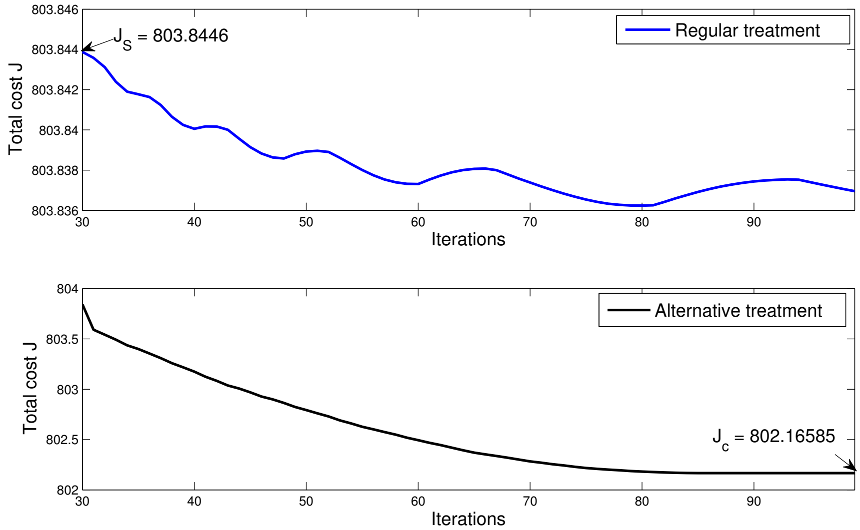

Figure 10: Method failure, and evolution of the total cost resulting from alternative method.

Table 2: Final numerical values

	Seed trajectory	The best ‘regular’ trajectory	Optimal solution (proposed method)
$x_1(t_f)$	3.7361	2.6442	2.617883
$x_2(t_f)$	0.2661	-0.8175	-0.887606
τ_1	0.035	0.676	0.715
τ_2	0.204	0.6850	0.718
$\mathcal{J}$	1449.228	803.8445	802.1658

the sequel.

The Hamiltonian arrived to a constant in the whole optimization interval $[0, t_f]$, $\bar{h}_{x_0} = -1438$, confirming the optimatility of the limit control trajectory. In table 2 the relevant numerical results are reported.

5 Conclusions and perspectives

Feedback control is desirable in industrial applications when perturbations are expected to appear, but an explicit expression is hardly attainable for nonlinear systems subject to general performance criteria and restricted control values. Besides, in these situations existing PMP solutions are obtained as open-loop strategies, so it should not be expected for optimal controls to be reachable after a finite sequence of closed-loop approximations. With these limitations in mind, here an efficient algorithm has been devised to approximate the open-loop optimal control, based on recent theoretical results and alternative control

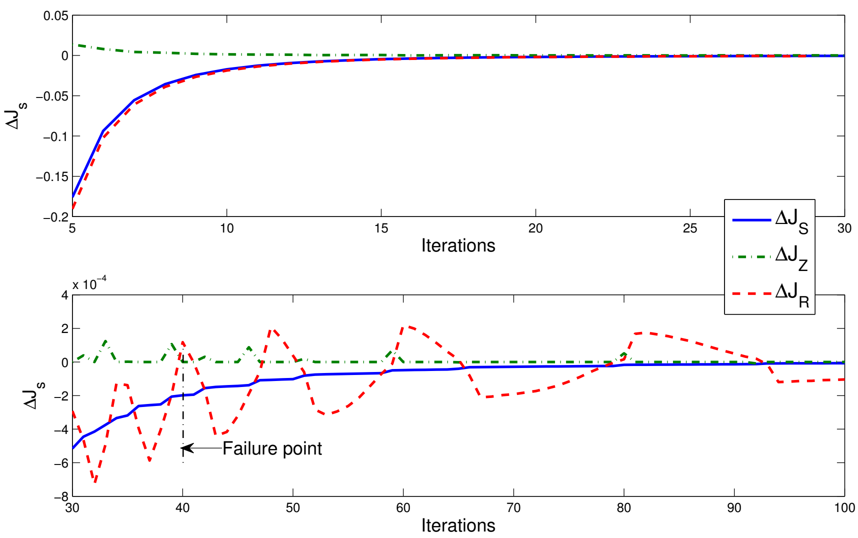

Figure 11: $\Delta\mathcal{J}$ obtained from the formulation Equations (35, 48).

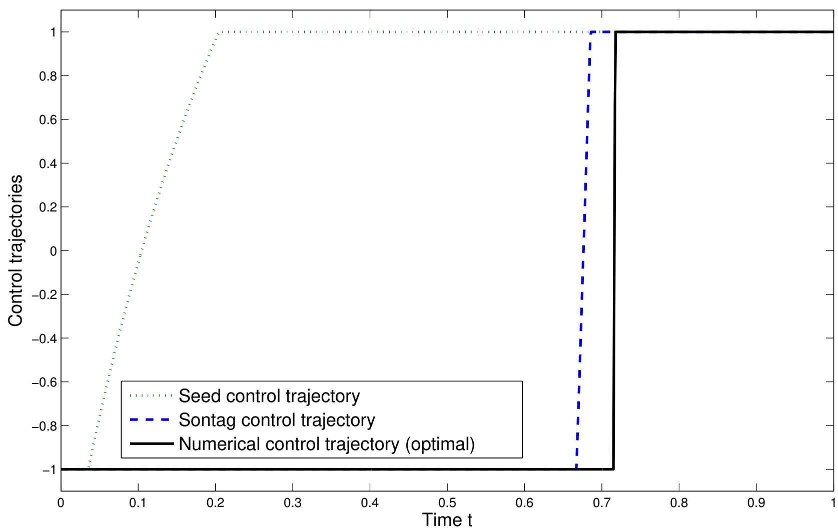

Figure 12: Control strategy for the modified bidimensional example.

variations. The resulting strategies are quite different from the saturated form of the optimal control corresponding to the unrestricted problem with the same parameters and initial condition. This last control strategy was used in the worked examples just as a first approximation, and so called a 'seed' strategy, although the method does not strictly depend on the initial approximation. The method then proceeds recursively, based on the linearization of the dynamics around the last approximate trajectories and a formula devised for evaluating the cost reduction in the regular context [22]. Since the problem at hand is generally non regular, the iterative procedure eventually fails in reaching the optimum, i.e., the cost ceases to decrease, or it enters in a dangerous zone were the error of saturation is of the same order as the cost reduction achieved. Then a second series of iterations is started, by using special classes of control variations: solutions to the DRE for the linearized dynamics subject to a new cost objective, cubic variations of regular arcs, and Pontryagin-type spatial variations generated from perturbations of the switching time-instants.

In general, the resulting control after these iterations will be suboptimal, although in the test-cases illustrated here the optimal values were reached at the end. In the first numerical example the cost reduction relative to the cost of the seed trajectory was approximately 9%, and in the second one was more than 80%, showing the unpredictable behavior of seed trajectories coming from regular solutions.

The stability of the method is guaranteed since (*i*) the regular procedure is stopped after an efficient alert condition is detected, (*ii*) the approximation to the differential cost $\Delta\mathcal{J}_S$ is not allowed to change sign, and (*iii*) the total cost is bounded from below.

References

[1] A. A. Agrachev and Yu. L. Sachkov. *Control Theory from the Geometric Viewpoint*. Springer-Verlag, Berlin-Heidelberg, 2004.

[2] M. Athans and P.L. Falb. *Optimal Control: An Introduction to the Theory and Its Applications*. Dover, New York, USA, 2006.

[3] R. E. Bellman. *Dynamic Programming*. Dover, Princeton, NJ, USA, 1957.

[4] P. Bernhard. Introducción a la teoría de control Óptimo. Technical Report Cuaderno Nro. 4, Instituto de Matemática "Beppo Levi", Rosario, Argentina, 1972.

[5] V. Costanza and C. E. Neuman. Optimal control of nonlinear chemical reactors via an initial-value hamiltonian problem. *Optimal Control Applications & Methods*, 27:41–60, 2006.

[6] V Costanza and P. S. Rivadeneira. Finite-horizon dynamic optimization of nonlinear systems in real time. *Automatica*, 44:2427–2434, 2008.

[7] V. Costanza and P. S. Rivadeneira. Optimal satured feedback laws for LQR problems with bounded controls. *Computational and Applied Mathematics*, 32:355–371, 2013.

[8] V. Costanza and P. S. Rivadeneira. *Enfoque Hamiltoniano al control óptimo de sistemas dinámicos.* OmniScriptum, Saarbrücken, Alemania, 2014.

[9] V. Costanza, P. S. Rivadeneira, and J. A. Gómez Múnera. An Efficient Cost Reduction Procedure for Bounded-Control LQR Problems. *Computational and Applied Mathematics*, 2016.

[10] V. Costanza, P. S. Rivadeneira, and A. H. González. Minimizing control-energy in a class of bounded-control lqr problems. *Optimal Control Applications & Methods*, 35(3):361–382, 2013.

[11] V. Costanza, P. S Rivadeneira, and R. D. Spies. Equations for the missing boundary values in the hamiltonian formulation of optimal control problems. *Journal of Optimization Theory and Applications*, 149:26–46, 2009.

[12] Lawrence C Evans. An introduction to mathematical optimal control theory version 0.2.

[13] W. H. Fleming and R.W. Rishel. *Deterministic and Sthocastic Optimal Control.* Dover, New York, USA, 1975.

[14] Phil Howlett. The optimal control of a train. *Annals of Operations Research*, 98(1-4):65–87, 2000.

[15] Phil G Howlett, Peter J Pudney, and Xuan Vu. Local energy minimization in optimal train control. *Automatica*, 45(11):2692–2698, 2009.

[16] M. Itik. Optimal control of nonlinear systems with input constraints using linear time varying approximations. *NONLINEAR ANALYSIS-MODELLING AND CONTROL*, 21(3):400–412, 2016.

[17] V. Jurdjievic. *Geometric Control Theory.* Cambridge University Press, Cambridge, 2006.

[18] R. E. Kalman, P. L. Falb, and M. A. Arbib. *Topics in Mathematical System Theory.* McGraw-Hill, New York, 1969.

[19] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes.* Macmillan, New York, USA, 1964.

[20] S. J. Qin and T. A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11:733–764, 2003.

[21] A. V. Rao, D. A. Benson, G. T. Huntington, C. Francolin, C. L. Darby, and M. A. Patterson. User's manual for gpops: A matlab package for dynamic optimization using the gauss pseudospectral method. Technical report, University of Florida, August 2008.

[22] E. D. Sontag. *Mathematical Control Theory.* Springer, New York, 1998.

[23] J. L. Speyer and D. H. Jacobson. *Primer on Optimal Control Theory.* SIAM Books, Philadelphia, USA, 2010.

[24] J. L. Troutman. *Variational Calculus and Optimal Control.* Springer, New York, USA, 1996.

[25] X. Wu and K. Zhang. Constrained optimal control problems of nonlinear systems based on improved newton algorithms. In *3rd International Conference on Informative and Cybernetics for Computational Social*, 2016.