Minimizing control energy in a class of bounded-control linear-quadratic regulator problems

V. Costanza[1,*,†], P. S. Rivadeneira[1] and A. H. González[2]

[1]*Nonlinear Systems Group, INTEC (UNL-CONICET), Güemes 3450, 3000 Santa Fe, Argentina*
[2]*Process Control Laboratory, INTEC (UNL-CONICET), Güemes 3450, 3000 Santa Fe, Argentina*

SUMMARY

Minimal-control-energy strategies are substantiated and illustrated for linear-quadratic problems with penalized endpoints and no state-trajectory cost, when bounds in control values are imposed. The optimal solution for a given process with restricted controls, starting at a known initial state, is shown to coincide with the saturated solution to the unrestricted problem that has the same coefficients but starts at a generally different initial state. This result reduces the searching span for the solution: from the infinite-dimensional set of admissible control trajectories to the finite-dimensional Euclidean space of initial conditions. An efficient real-time scheme is proposed here to approximate (eventually to find) the optimal control strategy, based on the detection of the appropriate initial state while avoiding as much as possible the generation and evaluation of state and control trajectories. Numerical (including model predictive control) simulations are provided, compared, and checked against the analytical solution to 'the cheapest stop of a train' problem in its pure-upper-bounded brake, flexible-endpoint setting. Copyright © 2013 John Wiley & Sons, Ltd.

1. INTRODUCTION

The linear-quadratic regulator (LQR) is probably the most studied and quoted problem in the state-space optimal control literature. The main line of work in this direction has evolved around the algebraic (for infinite-horizon problems) and differential (DRE, for finite-horizon ones) Riccati equations, together with their insertion into different model predictive control (MPC) formulations.

Since the early sixties, Hamiltonian formalism has also been at the core of the development of modern optimal control theory (see [1]). When the problem concerning an n-dimensional system and an additive cost objective is regular [2], that is, when the Hamiltonian of the problem can be uniquely optimized by a control value u^0 depending continuously on the remaining variables (t, x, λ), then a set of $2n$ ordinary differential equations (ODEs) with two-point boundary-value conditions, known as Hamilton (or Hamiltonian) canonical equations, has to be solved to obtain the optimal solution. For the LQR with a finite horizon, there exist well known methods (see for instance [3–7]) to transform the boundary-value problem into an initial-value one. In the infinite-horizon, bilinear-quadratic regulator and change of set-point servo, there also exists an attempt to find the missing initial condition for the costate variable from the data of each particular problem, which allows to integrate the equations in real time with the underlying control process [8]. For nonlinear systems, this line is in its beginnings [9–11].

Optimal control problems with hard restrictions on endpoint values, or with other constraints on states or control values, usually lack regularity and then require some version of the Pontryagin

---

*Correspondence to: V. Costanza, Nonlinear Systems Group, INTEC (UNL-CONICET), Güemes 3450, 3000 Santa Fe, Argentina.

†E-mail: tsinoli@santafe-conicet.gov.ar

maximum principle (PMP) for their solution. The PMP is a powerful result that has been systematized for very few cases. Other than some types of time-optimal problems for linear systems with bounded controls [1, 12–14], questions concerning general Lagrangians and control restrictions seem to need individual treatment. The restricted problem for linear systems with quadratic Lagrangian but with a linear final penalization has been recently discussed along controllability lines [15]. A promising iterative method to approximate the solution to restricted LQR problems has recently been published [16]. The method relies on a multi-grid method for placing a finite number of time intervals and a piecewise linear parameterization of the input within the intervals. The input values at the grid points and slopes within the time intervals are computed via nonlinear programming. The grids are gradually refined to improve the accuracy of the solution, but the possibility of controlling the rate of convergence for this process is still undefined.

The theoretical approach to minimizing control energy (MCE) under a quadratic final penalization and a bounded control adopted in this paper is, to our knowledge, original.

The 'cheapest stop of a train' problem [13] is discussed in detail to show that restrictions appear naturally in applications and also to illustrate the effect of such restrictions over the structure of the problem and its solutions. It is found that the optimal strategy for the problem with a hard restriction on the final state value lacks a realizable solution when the admissible control values are only the nonnegative numbers. To overcome this situation, a flexible but penalized endpoint condition is posed to replace the strict limitation for reaching equilibrium. The existence of feasible solutions is recuperated this way, but regularity of the problem is lost. However, it is found that, in parts of the time interval under consideration, the optimal solution behaves as the solution to a new regular problem.

The decisive theoretical finding was the following: the optimal solution to a given restricted MCE problem can be generated by saturating the solution to another unrestricted problem, with the same dynamics and cost objective as the original one, but starting at a different initial condition. Off-line and real-time schemes were developed to detect this new initial condition. In this paper, the expression 'real-time scheme' stands for a numerical procedure performed while the physical control process is evolving, in opposition to off-line schemes, which are completely solved before the process starts. The real-time algorithm is the main contribution of the manuscript from the practical point of view.

Different discrete-time MPC strategies were explored for comparison against the real-time algorithm proposed here and also to look for eventual combined schemes. In its classical 'receding horizon control' (RHC) formulation, MPC basically consists on performing a dynamical optimization on an horizon that retreats at every time step (see [17] for details). The final quadratic penalization might cause MPC not to be as effective as in fixed endpoint schemes. To avoid this inconvenience, sometimes a 'shrinking horizon control' policy is tried to fix the end time-step of predictions [18]. The continuous-time algorithm discussed in this paper can be considered as a potential tool to be used in combination with the receding or shrinking horizon policies, in an enlarged MPC context.

After the introduction, some notation and general characteristics of the problem are exposed in Section 2. The bounded-control situation, including the treatment of a case study, numerical validations, and illustrations, is discussed in detail in Section 3. The theoretical result concerning the structure of optimal solutions is given at the end of Section 3, and an off-line procedure is advanced to find the control strategy. At the beginning of Section 4, the novel real-time algorithm is described and applied to the case study chosen before. Its performance is illustrated and evaluated together with those of three different MPC versions. The whole approach and results are summarized in Section 5.

2. FORMALISM OF THE MINIMAL-CONTROL-ENERGY PROBLEM FOR LINEAR SYSTEMS

The finite-horizon, time-constant formulation of MCE problems with free final states and unconstrained controls attempts to minimize the (quadratic) cost

$$\mathcal{J}(u) = \int_0^T u'(\tau) R u(\tau) \mathrm{d}\tau \;+\; x'(T) S x(T), \tag{1}$$

with respect to all admissible (piecewise-continuous) control trajectories $u : [0, T] \to \mathbb{R}^m$ of duration T, applied to some fixed, finite-dimensional, deterministic, (here linear) plant. Then, control strategies affect the $\mathbb{R}^n$-valued states x through some initialized, autonomous, dynamical constraint

$$\dot{x} = Ax + Bu \,,\; x(0) = x_0 \neq 0. \tag{2}$$

This will be called a $(A, B, R, S, T, \mathbb{R}^m, x_0)$-problem. The general LQR problem includes another partial objective in addition to the 'minimal control energy' shown in (1), namely a quadratic state-trajectory term $x'(\tau) Q x(\tau)$ under the integral sign, penalizing state deviations from equilibrium along the process. Most results from LQR theory are valid when Q is an arbitrary positive-semidefinite matrix, so the minimal control energy inherits those properties as an LQR problem with $Q = 0$.

The (real, time-constant) matrices in (1) and (2) normally have the following properties: S is $n \times n$, symmetric and positive semi-definite, R is $m \times m$ and positive definite, A is $n \times n$, B is $n \times m$, and the pair (A, B) is controllable. The expression under the integral is usually known as the 'Lagrangian' L of the cost, which in this case is x-independent, that is,

$$L(x, u) \stackrel{\text{def}}{=} u' R u \,. \tag{3}$$

Under these conditions, the Hamiltonian of the problem, namely the $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ function defined by

$$H(x, \lambda, u) \stackrel{\text{def}}{=} L(x, u) + \lambda' f(x, u) \,, \tag{4}$$

is known to be regular, that is, that H is uniquely minimized with respect to u, and this occurs when u takes the explicit control value

$$u^0(x, \lambda) = -\frac{1}{2} R^{-1} B' \lambda \,, \tag{5}$$

(in this case, independently of x), which is usually called 'the H-minimal control'. The 'Hamiltonian' form of the problem (see for instance [4]) requires then to solve the two-point boundary-value problem

$$\dot{x} = H_\lambda^0(x, \lambda); \; x(0) = x_0 \tag{6}$$

$$\dot{\lambda} = -H_x^0(x, \lambda); \; \lambda(T) = 2 S x(T) \,, \tag{7}$$

where $H^0(x, \lambda)$ stands for $H(x, \lambda, u^0(x, \lambda))$, and H_λ^0, H_x^0 for the column vectors with i-components $\frac{\partial H^0}{\partial \lambda_i}, \frac{\partial H^0}{\partial x_i}$ respectively, that is, (6) and (7) here take the form

$$\begin{cases} \dot{x} = Ax - \frac{1}{2} W \lambda \,, \\ \dot{\lambda} = -A' \lambda \,, \end{cases} \tag{8}$$

with $W \stackrel{\text{def}}{=} B R^{-1} B'$.

It is well known that the problem as posed earlier relies on the solution $P(\cdot)$ to the Riccati differential equation (DRE)

$$\dot{\pi} = \pi W \pi - \pi A - A' \pi \,;\; \pi(T) = S \,, \tag{9}$$

leading to a useful relationship between optimal state $x^*(\cdot)$ and costate $\lambda^*(\cdot)$ trajectories, namely

$$\lambda^*(t) = 2 P(t) x^*(t) \,, \tag{10}$$

and, based on (5), to the optimal control trajectory

$$u^*(t) = u^0(x^*(t), \lambda^*(t)) = -R^{-1}B'P(t)x^*(t) , \tag{11}$$

or equivalently to the optimal feedback law

$$u_f(t, x) = -R^{-1}B'P(t)x . \tag{12}$$

When the control values are restricted, the global regularity of the Hamiltonian cannot be assured; therefore, the search for the optimal control strategy becomes more involved, as may be observed in the following sections.

3. BOUNDED-CONTROL PROBLEMS

The manipulated variable of most control systems appearing in practical applications can only assume a bounded set of values. The term 'manipulated' indicates that a person or an instrument assigns a value to a signal generated by physical means, and therefore, this value cannot take more than a physically realizable amount. 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 some $\mathbb{R}^m$ space. The qualitative features of optimal control solutions to bounded problems are significantly different from those of unbounded ones [1, 12].

Remark 1
In what follows, the set of admissible control values $\mathbb{U}$ will be assumed to be a compact set of the form $\mathbb{U} = \{u \in \mathbb{R}^m, u_a \preceq u \preceq u_b\}$, where $X \preceq Y$ denotes the following componentwise relation among vectors of $\mathbb{R}^m : X_i \leqslant Y_i$, $i = 1, 2, \ldots, m$. In particular, for $m = 1$, this reduces to $\mathbb{U} = [a, b]$ for some closed interval of real numbers.

The adopted form of a parallelepiped for $\mathbb{U}$ is sufficiently general in the realm of Engineering applications and has the following convenient property: the treatment of LQR problems for multi-input cases ($m > 1$) is a straightforward extension of the case $m = 1$, so only scalar input versions of all results and examples will be provided for simplicity. Details for generalizing to multi-input problems can be found, for instance, in [1, Chapter 3] and in [19].

3.1. The optimal solution via the Pontryagin maximum principle

In this section, it is shown that, if the minimal-control-energy strategy $u^*_{x_0}$ for a bounded-control $(A, B, R, S, T, \mathbb{U}, x_0)$-problem (with a set of admissible control values $\mathbb{U} = [a, b] \subset \mathbb{R}$) exists, then such strategy can be constructed from the optimal solution $u^*_{y_0}$ to a related unbounded-control $(\mathbb{U} = \mathbb{R})$ $(A, B, R, S, T, \mathbb{R}, y_0)$-problem. The PMP standard formulation indicates [1, 13, 20] (see also (8)) for the minimal control-energy problem that, if there exists such an optimal control solution $u^*_{x_0}(\cdot)$, then there should also exist an optimal costate trajectory $\lambda^*_{x_0}(\cdot)$ solution to the following (linear, initial-value, ODE) problem:

$$\dot{\lambda} = -A'\lambda \; ; \; \lambda(T) = 2Sx^*_{x_0}(T) \tag{13}$$

where $x^*_{x_0}(T)$ denotes the optimal final state value, that is, the final value of the solution $x^*_{x_0}(\cdot)$ to

$$\dot{x} = Ax + Bu^*_{x_0} \; ; \; x(0) = x_0 . \tag{14}$$

From (13), the optimal costate trajectory has the form

$$\lambda^*_{x_0}(t) = 2e^{A'(T-t)}Sx^*_{x_0}(T) \; \forall t \in [0, T] . \tag{15}$$

If the Hamiltonian for this problem is denoted

$$h(x, \lambda, u) \stackrel{\text{def}}{=} L(x, u) + \lambda' f(x, u) = Ru^2 + \lambda'(Ax + Bu) ,$$

then PMP also asserts that at each time t, the function

$$h_t(u) \stackrel{\text{def}}{=} h\left(x_{x_0}^*(t), \lambda_{x_0}^*(t), u\right) \tag{16}$$

should take its minimal value for $u = u_{x_0}^*(t)$. But it is immediately deduced from standard results [2, 4] that, for each t, the control trajectory denoted by

$$\hat{u}(t) \stackrel{\text{def}}{=} -\frac{1}{2}R^{-1}B'\lambda_{x_0}^*(t) \tag{17}$$

allows to construct the optimal strategy in the following way:

$$u_{x_0}^*(t) = \begin{cases} a \text{ if } \hat{u}(t) < a \\ \hat{u}(t) \text{ if } a \leqslant \hat{u}(t) \leqslant b \\ b \text{ if } \hat{u}(t) > b \end{cases}, \tag{18}$$

independently of the value of $x_{x_0}^*(t)$. And it is also clear that $\hat{u}(\cdot)$ is the optimal control solution to some LQR problem with unrestricted control values set, that is, when allowing $u \in \mathbb{U} = \mathbb{R}$, provided that its optimal costate trajectory $\hat{\lambda}(\cdot)$ coincides with $\lambda_{x_0}^*(\cdot)$. This would happen, in turn, for the unbounded problem with the same cost objective and dynamics as the restricted one but having an optimal state trajectory $\hat{x}(\cdot)$ solution to

$$\dot{x} = Ax + B\hat{u} \,;\; x(T) = x_{x_0}^*(T). \tag{19}$$

To find such a process is always possible, as can be seen from the following identities:

$$2P(0)\hat{x}(0) = \hat{\lambda}(0) = \lambda_{x_0}^*(0) = 2e^{A'T}Sx_{x_0}^*(T), \tag{20}$$

$$y_0 \stackrel{\text{def}}{=} \hat{x}(0) = [P(0)]^{-1}e^{A'T}Sx_{x_0}^*(T), \tag{21}$$

which together with (18) completes the proof of the initial assertion. The state y_0 for the unbounded problem will be called the 'hidden' initial condition in what follows. It is clear that the optimal control for the unrestricted $(A, B, R, S, T, \mathbb{R}, y_0)$-problem is $\hat{u}$.

From now on, a control strategy that has the form of $u_{x_0}^*$ in (18) will be denoted with the superscript sat standing for 'saturation'. The result of this section is then, in short,

$$u_{x_0}^* = \hat{u}^{\text{sat}} = \left(u_{y_0}^*\right)^{\text{sat}}. \tag{22}$$

This is equivalent to transform the infinite-dimensional original optimal control problem (the solution must be found in the space of admissible control trajectories) into a finite-dimensional one (the initial condition y_0 is searched in $\mathbb{R}^n$).

The next question is how to find the unrestricted initial state y_0, or equivalently the restricted final state $x_{x_0}^*(T)$, each one allowing to solve the whole problem. There exists a final PMP prescription applicable to this situation: the Hamiltonian must remain constant along optimal trajectories. In the following case study, it is observed that the adding of this last requirement to the set of all previously stated PMP necessary conditions actually leads to the exact form of $u_{x_0}^*$ and of its related objects: $x_{x_0}^*(T), \lambda_{x_0}^*(0), y_0$, and the saturation or 'switching' instants τ_i, $i = 1, 2, \ldots$ (for which $\hat{u}(\tau_i)$ reaches or leaves any of the bounds a, b). These objects are found after translating all conditions into a set of algebraic nonlinear equations, whose unique meaningful solution is calculated off-line. But in general, this procedure is cumbersome and cannot be easily systematized. It looks reasonable to propose an alternative solving procedure, motivated just by (22), namely 'Find the state y_0 that minimizes the cost (1) amongst all trajectories u_{z_0} generated from initial conditions z_0 as follows:

$$u_{z_0}(t) \stackrel{\text{def}}{=} u_f^{\text{sat}}(t, x_{z_0}(t)), \tag{23}$$

where x_{z_0} is the state trajectory solution to the initial value ODE problem

$$\dot{x} = Ax + Bu_f^{\text{sat}} \,;\; x(0) = z_0 \tag{24}$$

and u_f^{sat} denotes the saturated feedback law

$$u_f^{\text{sat}}(t, x) \overset{\text{def}}{=} \left(-R^{-1}B'P(t)x\right)^{\text{sat}}.$$ (25)

This option will be retaken in Section 3.2.4.

3.2. An example: The cheapest stop of a train

The results of the previous subsection will be applied to a classical case study, known in the literature as 'the cheapest stop of a train' (see for instance [13], and for similar problems concerning the determination of optimal switching times, see [21] and the references therein).

The dynamics of such a problem has the following simple linear form:

$$\dot{x}_1 = x_2; \ \dot{x}_2 = u,$$ (26)

or, in matrix notation,

$$\dot{x} = f(x, u) = Ax + Bu,$$ (27)

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$ (28)

where the real-valued control u may be interpreted as a braking action over an imaginary train with position x_1 and velocity x_2.

3.2.1. Unbounded controls. Fixed endpoints.
The objective (in the original formulation of the case study) is to optimize the 'braking energy' needed to 'stop the train', that is, to arrive at the desired final states

$$x_1(T) = x_2(T) = 0,$$ (29)

then the optimal control problem will be defined by (26), (29), and (1), and

$$L(x, u) \overset{\text{def}}{=} \frac{u^2}{2}, \ Q = 0, \ R = \frac{1}{2}, \ T = 1.$$ (30)

Unless indicated, the nominal initial conditions $x(0) = x_0$ chosen for illustration will be kept fixed at

$$x_1(0) = 1; \ x_2(0) = -1.$$ (31)

When the admissible control values are all the real numbers, the treatment of this problem along the lines of the PMP is given in the following. The Hamiltonian H of the problem reads

$$H(x, \lambda, u) = L + \lambda' f = \frac{u^2}{2} + \lambda_1 x_2 + \lambda_2 u,$$ (32)

which admits a global u-minimization, with explicit forms for the H-minimal control u^0 and the minimized Hamiltonian H^0, that is,

$$\frac{\partial H}{\partial u} = u + \lambda_2,$$ (33)

$$u^0(x, \lambda) = \arg \min_u H(x, \lambda, u) = -\lambda_2,$$ (34)

$$H^0(x, \lambda) = H(x, \lambda, u^0(x, \lambda)) = \lambda_1 x_2 - \frac{\lambda_2^2}{2},$$ (35)

leading to the canonical adjoint equations

$$\begin{cases} \dot\lambda_1 = -\frac{\partial H^0}{\partial x_1} = 0 \,, \\ \dot\lambda_2 = -\frac{\partial H^0}{\partial x_2} = -\lambda_1 \,. \end{cases} \tag{36}$$

Then, there should exist real constants a, b such that the optimal costate λ^* and control trajectory u^* have the form

$$\lambda_1^*(t) = -b \,, \tag{37}$$

$$\lambda_2^*(t) = a + bt \,, \tag{38}$$

$$u^*(t) = u^0(x^*(t), \lambda^*(t)) = -a - bt \,, \tag{39}$$

and consequently the dynamics' equations (26) can be symbolically integrated to obtain the form of the optimal state trajectories, namely

$$\begin{cases} x_2^*(t) = -at - b\frac{t^2}{2} - 1 \,, \\ x_1^*(t) = -a\frac{t^2}{2} - b\frac{t^3}{6} - t + 1 \,. \end{cases} \tag{40}$$

Final conditions are compatible (29), and the values of the constants and the optimal control strategy can be uniquely determined

$$a = 2, \ b = -6, \tag{41}$$

$$u^*(t) = -2 + 6t. \tag{42}$$

In Figure 1, the optimal control and state trajectories are shown. The unbounded-control solution found above uses (intuitively unexpected) negative control values, that is, the optimal solution prescribes that the train should be accelerated before beginning the purely braking process. So, an obvious question comes to mind: what would the optimal solution be if only (positive) braking is admitted. That will shortly be discussed in detail, but first it will be shown that the solution to the fixed endpoint problem is the limit of flexible-endpoint solutions corresponding to increasing final penalties.
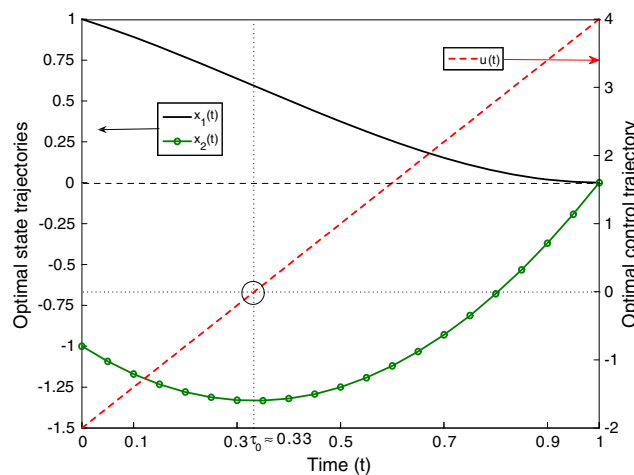

Figure 1. Optimal states and control trajectories for the 'cheapest stop of a train' problem with unbounded control values and hard endpoint restriction $x(T) = (0,0)'$.

3.2.2. Unbounded controls. Flexible endpoints. The 'flexible endpoint' problem for the same system and the same Lagrangian replaces the hard constraints in (29) by an alternative 'quadratic final penalty' $K(x(T)) = x'(T)Sx(T)$ in the cost objective function, as announced in (1). In this example, only scalar matrices $S = sI$ will be considered for simplicity. The components of the Riccati matrix $P(\cdot)$ solution to (9) can be expressed as functions of t,

$$
\begin{cases}
P_{11}(t) = \frac{3s + 6s^2(1-t)}{3 - 6s(t-1) - 2s^2(t-1)^3 + s^2(t-1)^4}, \\
P_{12}(t) = P_{21}(t) = \frac{-3s(t-1) + 3s^2(t-1)^2}{3 - 6s(t-1) - 2s^2(t-1)^3 + s^2(t-1)^4}, \\
P_{22}(t) = \frac{3s + 3s(t-1)^2 - 2s^2(t-1)^3}{3 - 6s(t-1) - 2s^2(t-1)^3 + s^2(t-1)^4},
\end{cases}
\tag{43}
$$

and therefore the dynamics corresponding to the optimal feedback in (11) and (12) can also be analytically integrated to ascertain the limiting behavior

$$
\lim_{s \to \infty} x_s^*(t) = x^*(t), \, t \in [0, T],
\tag{44}
$$

where $x^*(t)$ denotes the solution to the fixed-end problem resulting from (40) and (41), that is,

$$
x^*(t) = \begin{pmatrix} t^3 - t^2 - t + 1 \\ 3t^2 - 2t - 1 \end{pmatrix}.
\tag{45}
$$

3.2.3. Only pure braking admitted. In what follows, only nonnegative controls will be allowed (pure braking action), and the initial conditions will remain as in (31). Because the Hamiltonian has to be minimized at each point, and given that the optimal control begins with a negative value in the unbounded case (or equivalently, $\lambda_2(0) > 0$), then from (32), there should exist an initial time interval where the control u assumes the value of the lower bound, namely

$$
u(t) \equiv 0 \, \forall \, t \in [0, \tau) \subset [0, T],
\tag{46}
$$

where τ is still to be determined. Now, while the control variable is kept at its lower bound, the system evolves along a state trajectory denoted $\{x(t), t \in [0, \tau]\}$. As time increases, it is possible to pose successive unbounded control problems starting at $x(t)$ and with optimization horizon $[t, T]$. By continuity, near $t = 0$, the optimal solutions to those unbounded control problems will remain negative. If the optimal control for the bounded problem were nontrivial, there should exist a switching time τ where the optimal control $\tilde{u}_\tau(t)$ for the unbounded problem corresponding to (i) the remaining horizon $\{t \in [\tau, T]\}$ and (ii) the initial condition $x(\tau)$ turns nonnegative, that is, $\tilde{u}_\tau(\tau) \geqslant 0$ (for alternative arguments leading to the same statement, see for instance [22] and [23]). Let us assume that such a $\tau \in (0, T)$ exists; thus, during the interval $[0, \tau)$, the dynamics can be integrated, to obtain

$$
x_1(t) = 1 - t, \quad x_2(t) = -1.
$$

After time τ, the optimal control (denoted u_τ) must be linear, because it behaves as the u^0 of a regular problem ((33)–(36) must be met), that is, real numbers c, d, m, n exist such that $\forall t \in [\tau, T]$,

$$
\begin{aligned}
\lambda_1(t) &\equiv d, \\
\lambda_2(t) &= -c - dt,
\end{aligned}
\tag{47}
$$

$$
u_\tau(t) = c + dt,
\tag{48}
$$

and then, after integration, the corresponding state trajectories take the forms

$$
\begin{aligned}
x_2(t) &= ct + d\frac{t^2}{2} + m, \\
x_1(t) &= c\frac{t^2}{2} + d\frac{t^3}{6} + mt + n.
\end{aligned}
\tag{49}
$$

The concatenation of the lower-bound control followed by $\tilde{u}_\tau$ is an admissible control strategy, so the states must match at $t = \tau$. In view of (31) and (49), this amounts to say

$$x_2(\tau) = c\tau + d\frac{\tau^2}{2} + m = -1 ,$$

$$x_1(\tau) = c\frac{\tau^2}{2} + d\frac{\tau^3}{6} + m\tau + n = 1 - \tau . \tag{50}$$

Continuity of the control values $u_\tau(\tau)$ with respect to τ follows from the regularity of the unbounded optimal control problem and the smoothness of their governing equations, which implies in this case

$$0 = u_\tau(\tau) = c + d\tau \Longrightarrow \tau = -\frac{c}{d} . \tag{51}$$

Then, (50) and (51) require that the unknowns must be solutions to the following pair of equations:

$$m = \frac{c^2}{2d} - 1 , \quad n = \frac{c^3}{6d^2} + 1 . \tag{52}$$

Let us denote $\tilde{x}_0 = \phi(-\tau, x(\tau), \tilde{u}_\tau)$, where ϕ is the transition map of the system and $\tilde{u}_\tau(t) \stackrel{\text{def}}{=} c + dt$, $t \in [0, \tau)$, that is, the continuation of u_τ to the initial 'saturation' interval (where the optimal control is $u(t) \equiv 0$). This $\tilde{x}_0$ results then the proper initial condition of an unbounded-control problem with (full) optimization horizon $[0, T]$, whose optimal control solution is

$$\tilde{u}^*(t) \stackrel{\text{def}}{=} \begin{cases} \tilde{u}_\tau(t), \ t \in [0, \tau) \\ u_\tau(t), \ t \in [\tau, T] \end{cases} . \tag{53}$$

Consequently, the optimal costates $\tilde{\lambda}(\cdot)$ for such unbounded-control problem will obey the Hamilton canonical equations in $[0, \tau]$, resulting in

$$\tilde{\lambda}_1(t) = \tilde{\lambda}_1(\tau) = \lambda_1(\tau) = d , \tag{54}$$

$$\tilde{\lambda}_2(t) = -\tilde{u}_\tau(t) = -c - dt . \tag{55}$$

Then, the costate trajectories $\lambda(\cdot)$ in (47) can be seen as the continuation of trajectories $\tilde{\lambda}(\cdot)$, that is, the trajectories

$$\left\{ \begin{pmatrix} d \\ -c - dt \end{pmatrix}, t \in [0, T] \right\} \tag{56}$$

are the optimal costate trajectories corresponding to the initial condition $\tilde{x}_0$ and control $\tilde{u}^*$. This implies, in particular, that at the final time T, the boundary condition for the unbounded problem should be met, that is,

$$\lambda(T) = 2sx(T) . \tag{57}$$

The analysis can then be continued as follows:

$$d = \lambda_1(1) = 2sx_1(1) \Longrightarrow x_1(1) = \frac{d}{2s} ,$$

$$\lambda_2(1) = -c - d = 2sx_2(1) \Longrightarrow x_2(1) = -\frac{c + d}{2s} , \tag{58}$$

which combined with (49) and (52), render the desired pair of new conditions:

$$c + \frac{d}{2} + \frac{c^2}{2d} - 1 = -\frac{c + d}{2s} ,$$

$$\frac{c}{2} + \frac{d}{6} + \frac{c^2}{2d} + \frac{c^3}{6d^2} = \frac{d}{2s} . \tag{59}$$
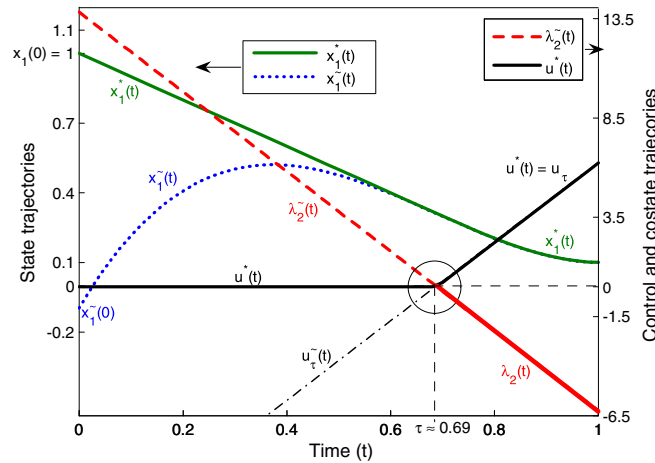
Figure 2. Optimal first state, second costate, and control trajectories associated with the pure braking situation, that is, with admissible control values $0 \leqslant u < \infty$. Idem for the related unbounded problem starting at the initial condition $\tilde{x}(0)$.

At this point, it should be noticed that, whatever the final values of the costates, (58) already reflects the expected limiting behavior

$$\lim_{s \to \infty} x^{T,s}(T) = 0.$$

To approximate the fixed-endpoint situation, a large enough value of s should eventually be chosen. Taking for instance $s = 100$, there exist five solutions to (59) for (c, d), but only one of them lies in $\mathbb{R}^2$, namely $c = -13.8331$, $d = 20.069$. The crossing time results

$$\tau = -\frac{c}{d} = 0.6893 ; \tag{60}$$

and then, by using (58), the final states for this case are obtained: $x_1(1) = 0.1004$, $x_2(1) = -0.0312$. The train does not arrive to rest, but the solution is more like what can be found in practice: at time T, the velocity $x_2(T)$ is small enough as to let the train inertially continue from $x_1(T)$ to its stoppage without any braking action ($u = 0$).

In Figure 2, the relations between (some components of) the state–costate and control trajectories are depicted.

In the literature (see [22, 23] and the references therein), little theoretical advances have been made in the treatment of these situations, other than including extra Lagrange multipliers to take into account the control bounds (a major inconvenience of such an approach is the appearing of inequalities, difficult to handle analytically).

3.2.4. Extension to the treatment of both lower and upper bounds in control values. Let us consider now the existence of two bounds on control values, for instance, $0 \leqslant u(t) \leqslant D$. If the control becomes saturated at both constrains, that will happen at least for two switching times τ_1, τ_2. Assume $0 < \tau_1 < \tau_2 < T$, and applying PMP's methodology, which implies the constancy of the Hamiltonian along the optimal trajectory, the following relations are found:

$$\lambda_1(t) \equiv 2s\bar{x}_1 ; \quad \bar{x}_1 \overset{\text{def}}{=} x_1(1) ; \tag{61}$$

$$\lambda_2(t) = 2s[\bar{x}_2 + \bar{x}_1(1 - t)] ; \quad \bar{x}_2 \overset{\text{def}}{=} x_2(1) ; \tag{62}$$

$$u(\tau_1) = 0 = -\lambda_2(\tau_1) \Longrightarrow \bar{x}_2 + \bar{x}_1(1 - \tau_1) = 0 ; \tag{63}$$

$$u(\tau_2) = D \implies -2s \left[\bar{x}_2 + \bar{x}_1 (1 - \tau_2) \right] = D \; ; \tag{64}$$

$$\dot{x}_2 = u \implies \bar{x}_1 = \frac{D}{2} \left[\frac{(\tau_2 - \tau_1)^2}{3} + (1 - \tau_1)(1 - \tau_2) \right] \tag{65}$$

$$\dot{x}_1 = x_2 \implies \bar{x}_2 = D \left[1 - \frac{\tau_1 + \tau_2}{2} \right] - 1 \; ; \tag{66}$$

Because the 'pure braking' optimal control reaches the value $\tilde{u}^*(1) = c + d = -13.8331 + 20.069 = 6.235\,9$, then any value for D below $\tilde{u}^*(1)$ will force a change in the optimal solution.

Consistently, in this section, the upper bound $D = 5$ was adopted, although changes will be made in Section 4.2 to better illustrate the behavior of the crossing times τ_i. Under these conditions, the only meaningful solution to (63)–(66) turns to be the following:

$$\bar{x}_1 = 0.105282 \; ; \; \bar{x}_2 = -0.0328644 \; ; \; \tau_1 = 0.687844 \; ; \; \tau_2 = 0.925301 \; ; \tag{67}$$

The fact that $\tau_1 < \tau$ shows that previous results become invalid when a new restriction is imposed on the control values. But still, the form of the costate equations and their relation with the optimal control remain the same. In short, the new optimal control adopts the 'saturated' form announced in (18) for $a = 0$, $b = D$, where the portion of the strategy with control values $0 < u^*(t) < D$ is also a portion of the optimal strategy $u_{y_0}^*(t)$ corresponding to the unrestricted problem with the same dynamics and cost objective but starting at the *hidden* initial condition

$$y_0 = [P(0)]^{-1} e^{A'T} \begin{pmatrix} s\bar{x}_1 \\ s\bar{x}_2 \end{pmatrix} = \begin{pmatrix} -0.0848165 \\ 3.92247 \end{pmatrix} \neq x_0 \; . \tag{68}$$

These results are illustrated in Figure 3. In practice, the optimal final values of the state components are unknown. A straightforward off-line procedure to treat these situations would be the following:

(i) Solve the DRE for the problem and denote its solution $P(\cdot)$.
(ii) For enough different initial conditions z_0 around x_0, integrate

$$\dot{x}_{z_0} = A x_{z_0} + B u_{z_0}^{\text{sat}} \; , \; x_{z_0}(0) = z_0 \; , \tag{69}$$

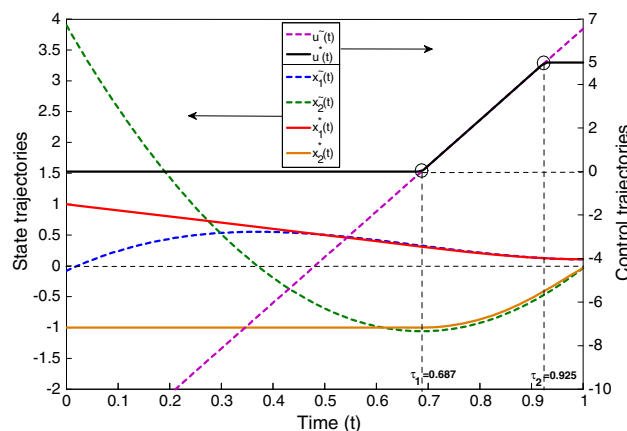$$\dot{J}_{z_0} = L(x, \tilde{u}) \; , \; J_{z_0}(0) = 0 \; , \tag{70}$$

Figure 3. Optimal control and states trajectories for the original problem and for the unrestricted one starting at the *hidden* initial condition, when both a lower bound and an upper bound are present.
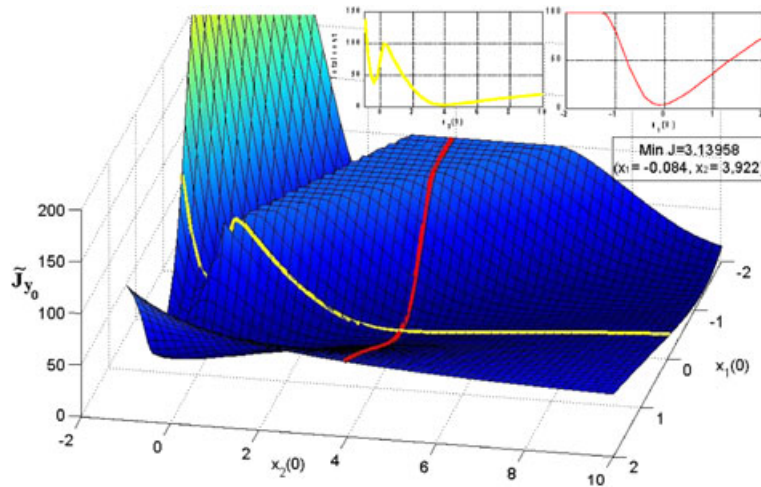
Figure 4. Costs associated with different initial conditions $y_0 = (x_1(0), x_2(0))'$ and saturated control strategies. The upper bound in control values is $D = 5$.

where $u_{z_0}^{\text{sat}}(t)$ denotes the 'saturated' form of the control strategy $u_{z_0}(t) \stackrel{\text{def}}{=} -R^{-1}B' P(t)x_{z_0}(t)$, that is,

$$u_{z_0}^{\text{sat}}(t) \stackrel{\text{def}}{=} \begin{cases} a, \forall t \mid u_{z_0}(t) \leqslant a \\ u_{z_0}(t), \forall t \mid a < u_{z_0}(t) < b \\ b, \forall t \mid u_{z_0}(t) > b \end{cases}, \tag{71}$$

and compute the cost corresponding to each process, that is,

$$\tilde{J}_{z_0} = J_{z_0}(T) + x'_{z_0}(T)Sx_{z_0}(T). \tag{72}$$

(iii) Find the optimal initial condition y_0 through

$$y_0 = \arg \min_{z_0} \tilde{J}_{z_0}, \tag{73}$$

and then the optimal control for the original problem will be

$$u^*(t) = u_{y_0}^{\text{sat}}(t) \; \forall t \in [0, T]. \tag{74}$$

The interpolated results of an off-line search conducted along these lines are depicted in Figure 4.

4. NUMERICAL TREATMENT

4.1. A new algorithm to approximate the optimal control

The purely off-line procedure described in the previous subsection, oriented to find the abstract initial condition from which the optimal saturated control is constructed, can be regarded as an alternative to existing numerical methods to find PMP solutions (most usually *via* 'shooting' or 'steepest descent' schemes for approximately solving boundary-value problems in ODE systems; see for instance [12, 15], [20]).

In this section, an alternative scheme, more in the line of standard numerical software for Engineering applications, is presented. Before starting the search for the hidden initial condition, the calculations needed concern the following objects, in their analytic or interpolated approximate forms:

(i) $P(t)$, the DRE solution for the (A, B, R, S)-problem (for the case study already found in analytical terms, see (43)).

(ii) $\Phi(t, 0)$, the fundamental matrix associated with the closed-loop form of the unrestricted problem, that is, the solution in $[0, T]$ to the (actually ODE) equation

$$\frac{\partial \Phi(t, 0)}{\partial t} = [A - WP(t)]\Phi(t, 0) \; ; \; \Phi(0, 0) = I \; . \tag{75}$$

Because this equation is linear, the effort to solve it numerically should in principle be smaller than the one required to obtain $P(\cdot)$ from the nonlinear DRE of the problem. For the 'train' problem, $\Phi(t, 0)$ can even be solved analytically:

$$\Phi(t, 0) = \begin{pmatrix} 1 - 2.805t^2 + 1.860t^3 & t - 1.906t^2 + 0.935t^3 \\ -5.610t + 5.582t^2 & t - 3.814t^2 + 2.805t^3 \end{pmatrix} \; . \tag{76}$$

(iii) $\Psi(t, \tau) \stackrel{\text{def}}{=} \int_\tau^t e^{A(t-s)} \mathrm{d}s = e^{At} \int_\tau^t e^{-As} \mathrm{d}s$, needed for calculating the state-transition map in the saturation periods, that is, for piecewise-constant control trajectories. For the case study,

$$\Psi(t, \tau) = \begin{pmatrix} t - \tau & \frac{(t-\tau)^2}{2} \\ 0 & t - \tau \end{pmatrix} \; . \tag{77}$$

Then, the approximation scheme to the optimal control involves the following steps:

- For an adopted time grid $t_0 = 0 < t_1 < t_2 \cdots < t_N = T$, calculate, by numerical integration, an initial discrete estimation (called 'the seed') of the optimal state trajectory, namely the values of the states $x^i \stackrel{\text{def}}{=} x(t_i)$, $i = 1, 2, \ldots, N$ where $x(t)$ is the solution to

$$\dot{x} = Ax + Bu_f^{\text{sat}} \; , \; x(0) = x_0 \; , \tag{78}$$

and where, analogous to the notation used in (71),

$$u_f^{\text{sat}}(t, x) \stackrel{\text{def}}{=} \begin{cases} a, & \text{if } u_f(t, x) \leqslant a \\ u_f(t, x), & \text{if } a < u_f(t, x) < b \\ b, & \text{if } u_f(t, x) \geqslant b \end{cases} \; , \tag{79}$$

$$u_f(t, x) \stackrel{\text{def}}{=} -R^{-1}B'P(t)x \; , \tag{80}$$

$u_f(t, x)$ being the feedback control law of the unrestricted LQR problem. Then, initiate a search for the optimal unrestricted process, that is, for the *hidden* initial condition y_0 necessary to construct the optimal $\tilde{u}_{y_0}(t)$ as in (71). This search proceeds through the following:

- For $i = 1, 2, \ldots, N$, find the initial conditions y^i corresponding to unrestricted processes arriving to x^i in time t_i, that is,

$$y^i \stackrel{\text{def}}{=} \Phi^{-1}(t_i, 0)x^i \; . \tag{81}$$

Because the state trajectory $x(t)$ is an approximation to the optimal one, the calculated y^i values can be regarded as approximations to the desired initial condition y_0.

- Calculate the corresponding costs J^i associated to the trajectory pairs $\left\{ \left(x_{y^i}(t), u_{y^i}^{\text{sat}}(t) \right), \; t \in [0, T] \right\}$, where each x_{y^i} is the solution to

$$\dot{x} = Ax + Bu_f^{\text{sat}} \; , \; x(0) = y^i \; , \tag{82}$$

The corresponding control strategies $u_{yi}(t)$ could directly be evaluated from the analytic form of the costate trajectory $\lambda_{yi}(t)$. The necessary expressions follow from (13) and (8), namely

$$u_{yi}(t) = -\frac{1}{2}R^{-1}B'\lambda_{yi}(t), \tag{83}$$

$$\lambda_{yi}(t) = e^{-A't}\lambda_{yi}(0), \tag{84}$$

$$\lambda_{yi}(0) = 2P(0)y^i. \tag{85}$$

Then, the costs J^i can be efficiently obtained by first detecting the switching times $t_{ij}, j = 1, 2, \ldots, N_i$, involved in the construction of u_{yi}^{sat} through

$$\{t_{ij}, j = 0, 1, 2, \ldots, N_i\} = \{0\} \cup \{t \mid u_{yi}(t) = a\} \cup \{t \mid u_{yi}(t) = b\}, \tag{86}$$

where it is understood that the set is ordered chronologically, that is,

$$0 = t_{i0} \leqslant t_{i1} \leqslant t_{i2} \leqslant \cdots \leqslant t_{iN_i} = T, \tag{87}$$

then continuing with the corresponding intercepting states, namely for $j = 0, 1, \ldots, N_i$,

$$x_{yi}(t_{ij}) = \Phi(t_{ij}, 0)y^i \tag{88}$$

and for the solution to the saturated control, the corresponding intercepting states may be calculated, in chronologically ascending order, by

$$x_{yi}^{\text{sat}}(t_{i,0}) = x_0, \tag{89}$$

and for $j = 0, 1, \ldots, N_i - 1$,

$$x_{yi}^{\text{sat}}(t_{i(j+1)}) = \begin{cases} e^{A\Delta_j^a}x_{yi}^{\text{sat}}(t_{ij}) + \Psi\left(t_{i(j+1)}, t_{ij}\right)Ba, & \text{if } u_{yi}(t) \leqslant a \\ e^{A\Delta_j}\left[x_{yi}^{\text{sat}}(t_{ij}) - x_{yi}(t_{ij})\right] + x_{yi}(t_{ij}), & \text{if } a < u_{yi}(t) < b \\ e^{A\Delta_j^b}x_{yi}^{\text{sat}}(t_{ij}) + \Psi\left(t_{i(j+1)}, t_{ij}\right)Bb, & \text{if } u_{yi}(t) \geqslant b \end{cases} \tag{90}$$

where

$$\Delta_k^a \overset{\text{def}}{=} t_{i(k+1)} - t_{ik} \text{ whenever } u_{yi}(t) \leqslant a \ \forall t \in [t_{ik}, t_{i(k+1)}], \tag{91}$$

$$\Delta_k \overset{\text{def}}{=} t_{i(k+1)} - t_{ik} \text{ whenever } a \leqslant u_{yi}(t) \leqslant b \ \forall t \in [t_{ik}, t_{i(k+1)}], \tag{92}$$

$$\Delta_k^b \overset{\text{def}}{=} t_{i(k+1)} - t_{ik} \text{ whenever } u_{yi}(t) \geqslant b \ \forall t \in [t_{ik}, t_{i(k+1)}], \tag{93}$$

and analogously for Δ_k^b; and finally computing

$$J^i = R\left(a^2\sum_k\Delta_k^a + b^2\sum_k\Delta_k^b\right) - \sum_k\Delta_k^V + \left(x_{yi}^{\text{sat}}\right)'(T)Sx_{yi}^{\text{sat}}(T), \tag{94}$$

where

$$\Delta_k^V \overset{\text{def}}{=} x_{yi}'\left(t_{i(k+1)}\right)P\left(t_{i(k+1)}\right)x_{yi}\left(t_{i(k+1)}\right) - x_{yi}'(t_{ik})P(t_{ik})x_{yi}(t_{ik}), \tag{95}$$

because for the unsaturated periods, the increments in cost can be calculated from the value function V, known to be in this case

$$V(t, x) = x'P(t)x. \tag{96}$$

- Find

$$\tilde{y} = \arg\min_{y^i} J^i \,. \tag{97}$$

- Osculate around $\tilde{y}$ (for instance, calculating the costs $\tilde{J}^l$ for saturated trajectories starting at $\tilde{y}_l, l = 1, 2, \ldots$ lying in a small circle centered at $\tilde{y}$, then selecting the new

$$\tilde{y} = \arg\min_{\tilde{y}_l} \tilde{J}^l \,, \tag{98}$$

and so on), until no significant improvements are found in the cost of saturated strategies.
- If desired, a refinement of the circle around $\tilde{y}$ may be tried, and the previous steps repeated. Otherwise, the last $\tilde{y}$ found is the best numerical approximation to the desired y_0 that can be obtained through this scheme.

It should be noted that ODEs integrations are systematically avoided (with the exception of the first approximation, at the beginning of the procedure).

4.2. Optimal and suboptimal results in continuous time

4.2.1. Applying the algorithm to the case study. In this subsection, the upper bound of the admissible control values, for the 'cheapest stop of a train' problem, will be changed from $D = 5$ to $D = 3$, to illustrate the behavior of an optimal solution with switching times more around the center of the time interval in consideration. Precisely, the values of the relevant objects for the new case $\mathbb{U} = [0, 3]$ were found by solving (61)–(66) to give

$$\begin{aligned} \tau_1 &= 0.6349, \ \tau_2 = 0.7348, \ J^* = 3.89762 \,, \\ x(T) &= \bar{x} = (0.150146, -0.054805)', \ y_0 = (0.6806, 3.9986)' \,. \end{aligned} \tag{99}$$

These results were reproduced, up to negligible deviations, by following the search described at the end of Section 3.2.4, which required a significant computing effort: 297.952 CPUTime.

Suboptimal values were also generated by following the algorithm of Section 4.1. After calculating the y^i, J^i, $i = 1, 2, \ldots, N$ (with $N = 30$), as prescribed by (81) and (94), the best initial condition of (97) was $\tilde{y} = (0.8173, 3.7747)'$ found for $\tilde{t} = t_{\min} = 0.933$; and the corresponding cost was $\tilde{J} = 3.91$. The final states reached from $x(0) = \tilde{y}$ were $x(T) = (0.151494, -0.056590)'$, already a good approximation to the optimal solution calculated off-line.

Up to this moment, the required computation effort was of 1.1516 CPUTime. The cost $\tilde{J} = 3.91$ was already much better than the cost of the initial strategy in (78) and (80), which amounted to $J = 10.95$.

4.2.2. Comparison against shooting methods. Shooting methods have been frequently attempted for (off-line) numerically solving mixed-type boundary-conditions problems such as the one proposed by the Hamiltonian equations (6) and (7). Because these equations are known to be highly sensitive to changes in initial (or final) conditions, because of the mixed signs of the eigenvalues associated with Hamiltonian systems, the performance of shooting methods in optimal control problems is unpredictable. For instance, if we choose the 'natural' approach to the problem treated in Section 4.2.1, the method diverges (Table I). By 'natural', it is meant to start with the known initial state $x(0) = (1, -1)'$ and an initial costate $\lambda(0) = \beta(T, S)\alpha^{-1}(T, S)x_0 = 2P(0)x_0$ for $S = 100$ and $T = 1$, and saturating the control when needed (its results correspond to those under Iteration #1 in I). Notice that the associated cost $J = 14.35$ is different to the $J = 10.95$ associated with the 'seed' trajectory because the shooting method cannot work in a strict feedback fashion but instead by using a control strategy calculated as $u(t) = \left(u^0(x(t), \lambda(t))\right)^{\text{sat}} = \left(-\frac{1}{2}R^{-1}B'\lambda(t)\right)^{\text{sat}}$, in fact unaware of the real value of the state $x(t)$. The returning path (Iteration #2) is calculated from 'corrected' final conditions, that is, averaging the final state obtained in Iteration #1 and the desired final state $(0, 0)'$, to obtain $x(T) = (0.14, 0.11)'$. The final costate $\lambda(T) = 2Sx(T) = (28, 22)'$ was the obvious choice. For Iteration #3, in the forward direction of time, the initial state was restored to its

Table I. Diverging iteration of the shooting method.

Iteration #	$x(0)$	$\lambda(0)$	$x(T)$	$\lambda(T)$	J
$1 \rightarrow$	1	5.55	0.28	5.55	14.35
	−1	1.8	0.22	−3.76	
$2 \leftarrow$	0.032	28	0.14	28	
	0.11	50	0.11	22	
$3 \rightarrow$	1	17	0	17	100
	−1	23	−1	6.33	
$4 \leftarrow$	2	0	0	0	
	−3.5	−100	−0.5	−100	

known value $x(0) = (1, -1)'$ and for $\lambda(0)$ was recalculated as the average of its values for the two preceding iterations. The cost associated with this iteration was already $J = 100$, and the final state obtained was $x(T) = (0, -1)'$. As it can be observed, the method leads far from the expected values for states and costates, and the cost of the trajectories increases, showing the approach useless in this version.

A second try was performed under 'helped' conditions. Results are shown in Table II. For backward iterations, the final costate was corrected in this case by using the choice $\lambda(T) = \frac{(\lambda_p(T) + \lambda_d(T))}{2}$. Here, $\lambda_p(T)$ denotes the final costate obtained from the previous iteration, and $\lambda_d(T) = (30.03, -10.96)'$ was kept constant and equal to the 'desired' value extracted from the optimal solution reported in (99), that is, $\lambda_d(T) = 2S\bar{x}$, with $\bar{x} = (0.150146, -0.054805)'$. This value of $\lambda_d(T)$ helps the evolution of the method, making it convergent. Even with this help, the

Table II. 'Helped' iteration of the shooting method.

Iteration #	$x(0)$	$\lambda(0)$	$x(t)$	$\lambda(T)$	J
$1 \rightarrow$	1	5.55	0.28	5.55	14.35
	−1	1.8	0.22	−3.76	
$2 \leftarrow$	0.85	17.79	0.14	17.79	
	−0.87	10.43	0.11	10.43	
$3 \rightarrow$	1	11.67	0.19	11.67	5.16
	−1	6.11	0.043	−5.56	
$4 \leftarrow$	0.8123	20.85	0.095	20.85	
	−0.88	12.6	0.022	−8.26	
$5 \rightarrow$	1	16.26	0.17	16.29	4.26
	−1	9.35	−0.002	−6.9	
$6 \leftarrow$	0.81	23.15	0.085	23.15	
	−0.88	14.21	−0.001	−8.93	
$7 \rightarrow$	1	19.70	0.1622	19.70	4.03
	−1	11.8	−0.02	−7.9	
$8 \leftarrow$	0.803	24.87	0.081	24.87	
	−0.877	15.43	−0.01	−9.44	
$9 \rightarrow$	1	22.29	0.1580	22.29	3.96
	−1	13.60	−0.03	−8.68	
$10 \leftarrow$	0.80	26.16	0.08	26.16	
	−0.877	16.34	−0.015	−9.82	
$11 \rightarrow$	1	24.22	0.156	24.22	3.93
	−1	14.97	−0.04	−9.25	
$12 \leftarrow$	0.79	27.13	0.077	27.13	
	−0.87	17.02	−0.02	−10.11	
$13 \rightarrow$	1	25.67	0.1539	25.67	3.91
	−1	15.99	−0.044	−9.68	
$14 \leftarrow$	0.79	27.85	0.077	27.85	
	−0.87	17.53	−0.022	−10.31	
$15 \rightarrow$	1	26.76	0.1528	26.76	3.90
	−1	16.76	−0.047	−9.99	

shooting method needs 15 iterations to reach a cost similar to the one obtained from the application of the new algorithm described earlier, in Section 4.2.1. The computing effort needed by the helped shooting version was of 1.966 CPUTime, significantly greater than the 1.1516 CPUTime required by the new procedure.

4.2.3. Refining the search. In practice, given the significant cost reduction, the algorithm could have been stopped at this stage. For illustration, a refinement was pursued through an osculating circle of radius 0.01 around $\tilde{y}$, resulting in a new best angle of 2.71 rad, a new best $\tilde{y}$ (Equation (98)) of $(0.5858, 4.2014)'$; and the corresponding final states $x(T) = (0.150541, -0.0539935)'$ and cost $J = 3.89782$. Again, because the new reduction in cost was negligible, calculations should have ended in a real application. A final refined osculating circle of radius 0.001 was essayed, giving the following updated best values: angle 0.81 rad, $\tilde{y} = (0.6840, 3.9915)'$, $x(T) = (0.150132, -0.0548345)'$, and $J = 3.89762$. These values, practically coinciding with those of the exact optimal solution, were obtained after a 2.12 CPUTime effort.

4.2.4. Perspectives for real-time application. Because the treatment of the problem has been maintained in the continuous-time domain, a procedure to update control actions during the period between two successive sampling times could be envisaged. The algorithm might then be regarded as a potential approach to transfer previous results to the real-time context.

It can be assumed that the continuous-time model can be run (eventually using parallel computing) in synchronization with the physical system. Whenever at a sampling time t the state $\bar{x}(t)$ of the physical system differs from the state $x(t)$ predicted by the model, then a new osculating circle can be explored after resetting its center at $\bar{y} = \Phi^{-1}(t, 0)\bar{x}(t)$, and the new $\tilde{y}$ can be adopted as the initial condition that minimizes the cost around $\bar{y}$ (Equation (97)).

All equations (78)–(98) can in principle be processed during the current sampling period, that is, a (parallel) computer is doing all calculations while the process is running. The control to be applied during the next period is calculated by using the last best 'hidden' initial condition $\tilde{y}$ found up to that moment. In case early sampling times arrive before the first $\tilde{y}$ is found, then the last-available suboptimal solution is applied in the meantime. Afterwards, the control can be recalculated from (83)–(85), with y^i replaced by $\tilde{y}$ and saturating the result. As times goes on, the $\tilde{y}$ can be refined, still along the lines described above, until the improvement in total cost is considered negligible.

All cumulative costs, switching points, and related objects are calculated through algebraic formulae derived from optimal control theory, which gives the method the necessary speed to predict corrections to the control strategy while the physical process is evolving.

Figures 5–7 illustrate the objects appearing in a trial application of this prospective approach, where the physical system has been replaced by model simulations.
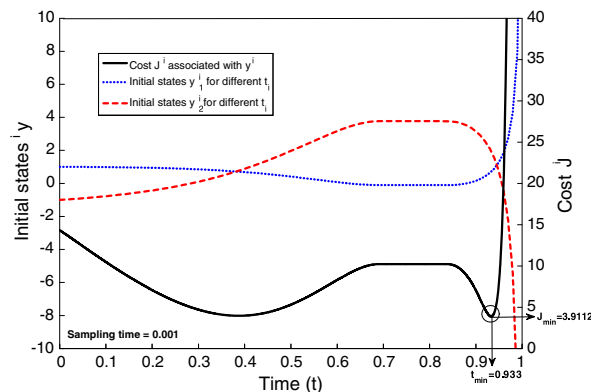

Figure 5. Initial states $y^i = \Phi^{-1}(t_i, 0)x(t_i)$, with $t_i = 0.001i$, $i = 1, \ldots, 1000$, of unrestricted suboptimal problems, computed from the first (seed) estimation of the optimal control, and the total costs J^i corresponding to the saturated strategies starting at each initial condition y^i.
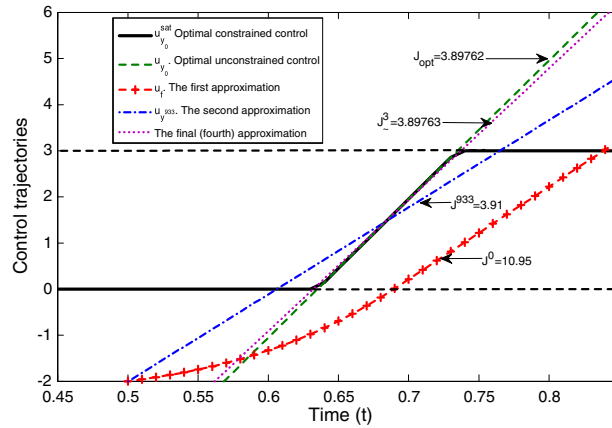
Figure 6. Optimal solution and suboptimal control strategies computed during the application of the real-time numerical scheme. The values J of the costs corresponding to the saturated-control versions are attached to each trajectory.
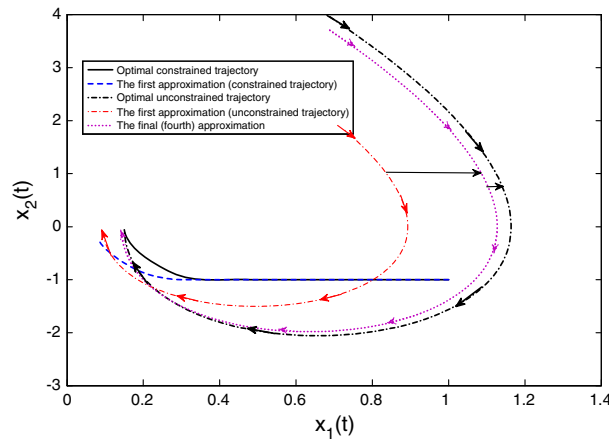

Figure 7. Phase-space states trajectories corresponding to saturated and to their corresponding unsaturated control strategies.

4.3. Discrete-time MPC approximations

This section is devoted to treat an approximation of the case study by using a MPC strategy. The MPC is an RHC scheme, in the sense that it optimizes an objective function along a horizon that recedes as the current time goes on. In fact, when an RHC strategy is implemented, the endpoint of the prediction horizon is corrected at each time step, and so, any action performed on it will recede as well. The discrete version of model (26), for a selected sampling time of $\Delta t = 0.05$, is

$$x_{k+1} = Ax_k + Bu_k. \tag{100}$$

where

$$A = \begin{pmatrix} 1 & 0.05 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0.0013 \\ 0.05 \end{pmatrix}. \tag{101}$$

The input constraints are maintained: $\mathbb{U} = \{u \in \mathbb{R} : 0 \leqslant u \leqslant 3\}$.

On the basis of this discretization of the original continuous model, three different stable MPC formulations were found useful to make a comparison with the proposed optimal control strategy of previous sections.

4.3.1. Classical MPC. The general MPC formulation includes both a terminal state constraint of the type

$$x_{k+N} \in \mathcal{X}_f \, , \tag{102}$$

and a terminal state penalty, which should be chosen in an appropriate form as to ensure closed-loop stability ([17, 24]). This classical MPC solves at each time step $k = 1, 2, \ldots, k_T$, where $k_T \overset{\text{def}}{=} \frac{T}{\Delta t}$, the following optimization problem:

$$\min_{\mathbf{u}} V_N(x_k, \mathbf{u}) = \frac{1}{2} \sum_{i=k}^{k+N-1} \left(x_i' Q x_i + u_i' R u_i \right) + \frac{1}{2} x_{k+N}' S x_{k+N} \tag{103}$$

subject to the dynamics

$$\begin{aligned} x_{i+1} &= A x_i + B u_i, & \text{for } i = k, \ldots, k + N - 1 \, , \\ u_i &\in \mathbb{U}, & \text{for } i = k, \ldots, k + N - 1 \, , \end{aligned} \tag{104}$$

where $\mathbf{u} = \left[u_k', \ldots, u_{k+N-1}' \right]'$. In the objective function, Q and R represent the penalizing matrices corresponding to the states and the input, respectively, and S is a symmetric, positive semi-definite terminal state matrix penalization, where now *terminal* means the final point of a horizon that recedes at each time step. N is the 'control horizon', and the expression (102) represents a general terminal state constraint. The conditions under which the present formulation derives in a closed-loop stable controller can be summarized as follows:

- The model function $f(x, u) = Ax + Bu$, the stage cost function $L(x, u) = x'Qx + u'Ru$, and the terminal function $V_f(x) = x'Sx$ are continuous, and $f(0, 0) = 0$, $L(0, 0) = 0$, and $V_f(0) = 0$. These conditions are clearly satisfied for the selected quadratic cost (103).
- The sets $\mathbb{X}$ and $\mathcal{X}_f$ are closed, $\mathcal{X}_f \subseteq \mathbb{X}$, and $\mathbb{U}$ is compact; each set containing the origin.
- The set $\mathcal{X}_f$ is *control invariant* for the system $x_{k+1} = Ax_k + Bu_k$.
- $V_f(x)$ is a Lyapunov function for the system $x_{k+1} = (A - BF)x_k$,

where the usual choice is to adopt $F = -R^{-1}B'\pi$, π being the matrix solution to the algebraic Riccati equation, and the set $\mathcal{X}_f$ is selected as the *maximal invariant set* for the closed-loop system $x_{k+1} = (A - BF)x_k$ [17]. With this formulation, the closed loop is stable with a domain of attraction given by the set of states that can be admissibly steered to $\mathcal{X}_f$ in N steps. The parameters in the MPC cost are selected as $N = 9$, $Q = 0$, $R = \frac{1}{2}$, and $S = \text{diag}(100, 100)$. To compare the behavior of this controller with the proposed strategy and with other formulations of MPC, the following indexes were used:

$$J_u = \frac{T}{2} \sum_{k=1}^{k_T} u_k^2 \, , \tag{105}$$

$$J_{\text{ter}} = x_{k_T}' S x_{k_T} \, , \tag{106}$$

$$J_{\text{tot}} = J_u + J_{\text{ter}} \, , \tag{107}$$

$$\text{CPUTime} = \text{computer time.} \tag{108}$$

Their values are reported in Table III for all MPC versions. The meaning of J_u and J_{ter} for the results of the algorithm used in Section 4.2 refer to the first and second terms of the right-hand side of (1), respectively. Even when this last interpretation convey a continuous-time model evaluation whereas those in (105) and (106) respond to a discrete-time approximation, the symbols J_u and J_{ter} reflect the cumulative control energy and the final penalization, respectively, in both cases. Thus, they may serve as a rough comparison for the suboptimal partial costs delivered by the different methods.

Table III. Cost values for several indexes and the different control strategies.

Performance indexes	Continuous-time algorithm	Classical MPC	IHMPC	SHMPC
J_u	1.3428	0.4565	0.7094	0.5823
J_{ter}	2.5548	21.7694	21.0992	9.0672
J_{tot}	3.8976	22.2259	21.8086	9.6494
CPUTime	2.12	4.1964	5.1108	3.6504

MPC, model predictive control; IHMPC, infinite horizon MPC; SHMPC, shrinking horizon MPC.
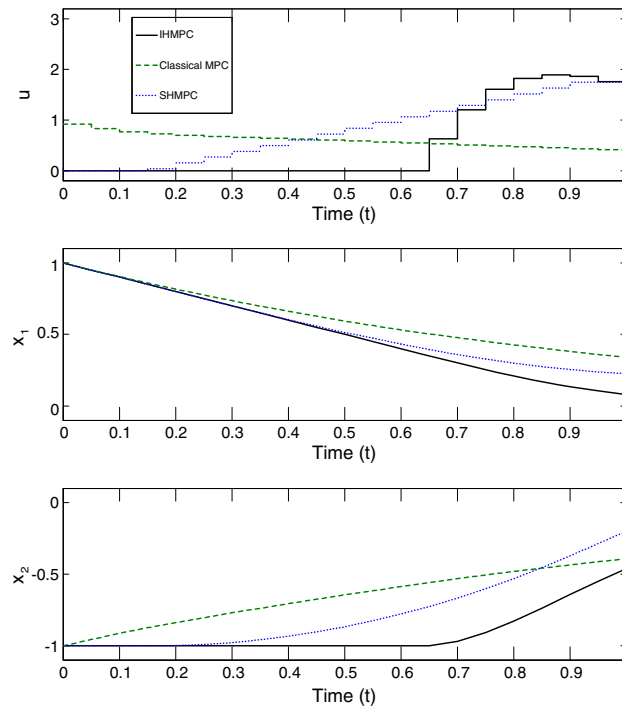

Figure 8. Approximations to the optimal control of the case study obtained through different MPC schemes (first subplot) and their corresponding reached final states (in the remaining subplots). Admissible control values in [0, 3].

Figure 8 shows the discrete time input and states response corresponding to the classical MPC. During the first 12 time steps, that is, up to 0.6, the input remains on the lower limit. Then, at time step 12, MPC begins to act, to steer the velocity of the train, x_2, to zero.

The resulting final state is $x_{k_T} = (0.0819, -0.4606)'$. As can be seen, the reached final state is far from the target $(0, 0)'$ because the input strategy is quite conservative, also reflected by the value of the index J_u.

4.3.2. Infinite horizon MPC with modes separation. A more sophisticated alternative is the so-called infinite horizon MPC *with modes separation* (IHMPC) [25, 26]. The transition matrix A of system (100) has an eigenvalue 1 of algebraic multiplicity 2 and geometrical multiplicity 1, which means that only one eigenvector is associated to it. In a phase portrait of the discrete time system, it can be observed that any state in the horizontal coordinate axis remains there indefinitely, whereas any state in the vertical coordinate axis moves to the right if positive and to the left if negative.

The purpose of IHMPC is to separate the modes of the system in order to control them individually. This is attempted by imitating an infinite state horizon: that is, (i) by replacing the finite

sum $\sum\limits_{i=k}^{k+N-1}$ in the cost function of (103) by the infinite one $\sum\limits_{i=k}^{\infty}$, while (ii) making the input sequence to remain null from N to infinity, and (iii) substituting the terminal constraint in (102) by $x_{k+N} = (\alpha, 0)'$, for a feasible real value of α. The states with zero second component become a stationary set for the system of (100). Notice that α operates in this formulation as a slack variable, which is penalized in (103) by the term $(x_{k+N})'S x_{k+N}$.

The results of applying the constraint $x_{k+N} = (\alpha, 0)'$ should be interpreted as the only approximation IHMPC can provide to the solution of the original problem and therefore the only one available for comparison.

The selected parameters for the IHMPC with modes separation are $N = 9$, $Q = 0$, $R = \frac{1}{2}$, and $S = 100I$. Figure 8 shows the input and state evolution for the chosen parameters.

4.3.3. Shrinking horizon MPC. An alternative to the RHC strategy is the so-called *shrinking horizon MPC* (SHMPC). In this formulation, the cost horizon shrinks when the current time is increased in such a way that the terminal penalization is applied always to the actual terminal state. With this fixed-endpoint formulation, the nominal stability conditions are easily established without the need for a terminal constraint [18].

Figure 8 shows that the control action starts earlier in the SHMPC case, and consequently, the second state comes nearer to the null value at the end of the simulation. The SHMPC performance is nearer to the one obtained with the continuous-time methodology than those of other MPC schemes. However, the SHMPC control does not reach the saturation value $u_{\max} = 3$ (assumed at the end of the time interval by the optimal and suboptimal continuous-time solutions).

5. CONCLUSIONS

A novel theoretical result and derived numerical procedures to solve minimal-control-energy LQR problems, when restrictions on the control values are present, have been substantiated and illustrated through the flexible-endpoint version of a classical example. It is shown how the optimal control solution can be generated off-line after detecting the initial condition of a related problem, which actually contains all the relevant information. Also, an efficient real-time algorithm is devised and applied to the case study. The resulting strategy in both cases is quite different from the saturated form of the optimal control corresponding to the unrestricted problem with same parameters and initial condition, sometimes naively adopted in Engineering practice (and used here as a seed to initiate the real-time numerical procedure).

Although the discrete-time approximation leads to a theoretically different problem from the original one in continuous time, the application of several MPC schemes to the case study has also been attempted, just for illustration and insight, and the results discussed. The discrete-time versions reveal lower performances, probably as a result of the crude approximations of the dynamics essayed. A combination of MPC concepts with the main lines of the continuous-time algorithm proposed here might be useful for dynamically optimizing industrial processes, as some batch processes of long duration, for which the receding-horizon approach may become mandatory.

REFERENCES

1. Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mischenko EF. *The Mathematical Theory of Optimal Processes*. Wiley: New York, 1962.
2. Kalman RE, Falb PL, Arbib MA. *Topics in Mathematical System Theory*. McGraw-Hill: New York, 1969.
3. Bernhard P. *Introducción a la Teoría de Control Óptimo*, Cuaderno Nro. 4. Instituto de Matemática "Beppo Levi", Rosario: Argentina, 1972.
4. Sontag ED. *Mathematical Control Theory*. Springer: New York, 1998.
5. Costanza V. Finding initial costates in finite-horizon nonlinear-quadratic optimal control problems. *Optimal Control Applications & Methods* 2008; **29**:225–242.
6. Costanza V, Rivadeneira PS. Feedback óptimo del problema lineal-cuadrático con condiciones flexibles. Paper A001VP. *Proceedings XXIº Congreso Argentino de AADECA*, AADECA, Buenos Aires, Sept 10, 2008; AVP 1–8.

7. Costanza V, Neuman CE. Partial differential equations for missing boundary conditions in the linear-quadratic optimal control problem. *Latin American Applied Research* 2009; **39**(3):207–212.
8. Costanza V, Neuman CE. Optimal control of nonlinear chemical reactors via an initial-value Hamiltonian problem. *Optimal Control Applications & Methods* 2006; **27**:41–60.
9. Costanza V, Rivadeneira PS. Finite-horizon dynamic optimization of nonlinear systems in real time. *Automatica* 2008; **44**:2427–2434.
10. Costanza V, Rivadeneira PS, Spies RD. Equations for the missing boundary values in the Hamiltonian formulation of optimal control problems. *Journal of Optimization Theory and Applications* 2011; **149**:26–46.
11. Costanza V, Rivadeneira PS. Initial values for Riccati ODEs from variational PDEs. *Computational and Applied Mathematics* 2011; **30**(2):331–347.
12. Athans M, Falb P. *Optimal Control. An Introduction to the Theory and its Applications*. Mc-Graw-Hill: New York, 1966.
13. Agrachev A, Sachkov Y. *Control Theory from the Geometric Viewpoint*. Springer-Verlag: Berlin-Heidelberg, 2004.
14. Jurdjevic V. *Geometric Control Theory*. Cambridge University Press: Cambridge, UK, 2006. (digital printing of 1997 edition).
15. Speyer JL, Jacobson DH. *Primer on Optimal Control Theory*. SIAM Books: Philadelphia, 2010.
16. Pannocchia G, Rawlings JB, Mayne DQ, Marquardt W. On computing solutions to the continuous time constrained linear quadratic regulator. *IEEE Transactions on Automatic Control* 2010; **55**(9):2192–2198.
17. Rawlings JB, Mayne DQ. *Model Predictive Control: Theory and Design*. Nob Hill Publishing: Madison, WI, USA, 2009.
18. Russell SA, Kesavan P, Lee JH, Ogunnaike BA. Recursive data-based prediction and control of batch product quality. *AIChE Journal* 1998; **44**:2442–2458.
19. Herzog R, Kunisch K. Algorithms for PDE-constrained optimization. *GAMM-Mitteilungen* 2010; **33**(2):163–176.
20. Troutman JL. *Variational Calculus and Optimal Control*. Springer: New York, 1996.
21. Howlett PG, Pudney PJ, Vu X. Local energy minimization in optimal train control. *Automatica* 2009; **45**(11):2692–2698.
22. Alt W. Approximation of optimal control problems with bound constraints by control parameterization. *Control and Cybernetics* 2003; **32**(3):451–472.
23. Bryson AE, Ho YC. *Applied Optimal Control*. Blaisdell: New York, 1969.
24. Mayne DQ, Rawlings JB, Rao CV, Scokaert POM. Constrained model predictive control: stability and optimality. *Automatica* 2000; **36**:789–814.
25. González AH, Odloak D. Enlarging the domain of attraction of stable MPC controllers, maintaining the output performance. *Automatica* 2009; **45**:1080–1085.
26. González AH, Adam E, Marcovecchio M, Odloak D. Stable MPC for tracking with maximal domain of attraction. *Journal of Process Control* 2011; **21**:573–584.