Comp. Appl. Math. (2013) 32:355–371
DOI 10.1007/s40314-013-0025-7

Optimal saturated feedback laws for LQR problems with bounded controls

Vicente Costanza · Pablo S. Rivadeneira

Received: 5 October 2012 / Revised: 6 December 2012 / Accepted: 12 February 2013 /
Published online: 15 March 2013

Abstract Optimal and suboptimal strategies are substantiated and illustrated for linear-quadratic problems with penalized endpoints, 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 some unrestricted problem that has the same coefficients, except for the final penalization matrix S, and 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 space of symmetric positive semi-definite symmetric matrices $\hat{S}$ and initial states $\hat{x}_0$. An efficient scheme is also proposed to approximate (and eventually to find) the optimal feedback strategy on-line, based on the updating of $\hat{S}$ at successive sampling times t_k, and on the possibility to generate the corresponding Riccati matrix $P(t, T, \hat{S})$ for $t_k < t \leq t_{k+1}$ from auxiliary matrices stored in memory. Numerical simulations are provided, compared, and checked against the analytical solutions of two classical case-studies.

Keywords Optimal control · Constrained controls · Linear-quadratic problem

Mathematics Subject Classification (2000) Primary 93C10; Secondary 49N10

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

Communicated by José Eduardo Souza de Cursi.

V. Costanza (✉) · P. S. Rivadeneira
Nonlinear System Group, INTEC, CONICET-UNL, Santa Fe, Argentina
e-mail: tsinoli@santafe-conicet.gov.ar

P. S. Rivadeneira
e-mail: psrivade@santafe-conicet.gov.ar

around the algebraic (ARE, 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 1960s, Hamiltonian formalism has been at the core of the development of modern optimal control theory (see Pontryagin et al. 1962). When the problem concerning an n-dimensional system and an additive cost objective is regular (Kalman et al. 1969), i.e. 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 (HCEs), has to be solved to obtain the optimal solution. For the LQR with a finite horizon, there exist well-known methods (see for instance Bernhard 1972; Sontag 1998; Costanza 2008; Costanza and Rivadeneira 2008a; Costanza and Neuman 2009) to transform the boundary-value problem into an initial-value one. In the infinite-horizon, bilinear-quadratic regulator, and change of set-point servo problems, 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 on-line with the underlying control process (Costanza and Neuman 2006). For nonlinear systems, this line of work is in its beginnings (Costanza and Rivadeneira 2011a, 2008b; Costanza et al. 2011).

Optimal control problems with hard restrictions on endpoint values, or with other constraints on states or control values, usually lack regularity and they require some version of the Pontryagin Maximum Principle (PMP) for their solution. PMP is a powerful result that has been systematized for very few cases. Other than in some types of time-optimal problems for linear systems with bounded controls (Pontryagin et al. 1962; Athans and Falb 1966; Agrachev and Sachkov 2004; Jurdjievic 2006), 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 (Speyer and Jacobson 2010). The theoretical approach to bounded-control flexible-endpoint LQR problems adopted in this paper is, to our knowledge, original.

The 'cheapest stop of a train' problem (Agrachev and Sachkov 2004) 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 end-point 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 some subsets of the time-horizon, the optimal solution behaves as the solution to a new regular problem.

The decisive theoretical finding is as follows: the optimal solution to a given restricted LQR problem can be generated by saturating the solution to another unrestricted LQR problem, with same dynamics and cost objective as the original one, but starting at a different initial condition and subject to a quadratic final penalization with a changed matrix coefficient. Off-line and on-line schemes were developed to detect this new initial condition and final penalization matrix. The on-line algorithm in this direction is the main contribution of the manuscript from the practical point of view. The numerical scheme takes advantage of the availability of Riccati matrices, generated from the solutions to a pair of first order partial differential equations (Costanza and Rivadeneira 2008a; Costanza et al. 2011), and provides the suboptimal control in feedback form. For its simplicity and small computational effort, the on-line algorithm can be considered as a potential tool to be used in combination with the receding or shrinking horizon policies, in an enlarged MPC context allowing for strict finite-horizon problems.

2 Formalism of the bounded-control LQR problem with a flexible endpoint

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

$$\mathcal{J}(u) = \int_0^T [x'(\tau)Qx(\tau) + u'(\tau)Ru(\tau)]\mathrm{d}\tau + x'(T)Sx(T), \tag{1}$$

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

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

This will be called a $(A, B, Q, R, S, T, \mathbb{R}^m, x_0)$-problem.

The (real, time-constant) matrices in Eqs. (1, 2) will be assumed to have the following properties: Q and S are positive-semidefinite $n \times n$ matrices, 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, namely

$$L(x, u) := x'Qx + u'Ru. \tag{3}$$

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

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

is known to be regular, i.e. 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 Sontag 1998) requires then to solve the two-point boundary-value problem for the HCEs

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

$$\dot{\lambda} = -H^0_x(x, \lambda); \quad \lambda(T) = 2Sx(T), \tag{7}$$

where $H^0(x, \lambda)$, usually called the minimized (or control) Hamiltonian, stands for

$$H^0(x, \lambda) := H(x, \lambda, u^0(x, \lambda)), \tag{8}$$

and H^0_λ, H^0_x for the column vectors with i-components $\frac{\partial H^0}{\partial \lambda_i}$, $\frac{\partial H^0}{\partial x_i}$ respectively, i.e. Eqs. (6, 7) here take the form

$$\begin{cases} \dot{x} = Ax - \frac{1}{2}W\lambda, \\ \dot{\lambda} = -2Qx - A'\lambda, \end{cases} \tag{9}$$

with $W := BR^{-1}B'$.

It is well known that the solution to the unrestricted regular problem, as posed above, relies in turn on the solution $P(\cdot)$ to the Riccati differential equation (DRE)

$$\dot{\pi} = \pi W \pi - \pi A - A'\pi - Q; \quad \pi(T) = S, \tag{10}$$

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

$$\lambda^*(t) = 2P(t)x^*(t), \tag{11}$$

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

$$u^*(t) = u^0(x^*(t), \quad \lambda^*(t)) = -R^{-1}B'P(t)x^*(t), \tag{12}$$

or equivalently to the optimal feedback law

$$u_f(t, x) = -R^{-1}B'P(t)x. \tag{13}$$

When the control values are restricted, the global regularity of the Hamiltonian can not be assured, and therefore the search for the optimal control strategy becomes more involved, as may be observed in the following Sections. Additional relevant objects from the LQR theory will be used in the sequel, for instance the matrices $\alpha(T, S)$, $\beta(T, S)$, solutions to the following pair of first-order, quasilinear partial differential equations (see Costanza 2008; Costanza and Rivadeneira 2008a, 2011a,b; Costanza and Neuman 2009 for details):

$$\alpha_T - \alpha_S M = -\alpha N, \quad \alpha(0, S) = I; \tag{14}$$

$$\beta_T - \beta_S M = -\beta N, \quad \beta(0, S) = 2S; \tag{15}$$

where the matrix coefficients are

$$M := A'S + SA + Q - SWS, \tag{16}$$

$$N := A - WS. \tag{17}$$

These matrices allow us to calculate, for any unbounded LQR problem, the solution $P(\cdot, T, S)$ to its DRE through the formula

$$P(t, T, S) = \frac{1}{2}\beta(T - t, S)\,[\alpha(T - t, S)]^{-1}\,\forall t \in [0, T], \tag{18}$$

and in such a case the matrices α, β are also related to the boundary conditions by the following relations (Bernhard 1972; Sontag 1998; Costanza et al. 2011):

$$x(0) = \alpha(T, S)x(T), \quad \lambda(0) = \beta(T, S)x(T). \tag{19}$$

The manipulated variable in most of the 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 (Pontryagin et al. 1962; Athans and Falb 1966). 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 following section is an attempt to typify the number and behavior of switching points in the constrained-control LQR case.

3 Characterizing optimal phase-trajectories

In this section, the main result is proved and its consequences are discussed. It is found that, if the optimal strategy $u^*_{x_0}$ for a bounded-control $(A, B, Q, R, S, T, \mathbb{U}, x_0)$-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 ($\mathbb{U} = \mathbb{R}$)$(A, B, Q, R, \hat{S}, T, \mathbb{R}, \hat{x}_0)$-problem.

Theorem 3.1 *Let us assume that there exists a time $\tau \in (0, T)$ where $u^*_{x_0}(\tau) \in (a, b)$. Then there exists a unique time interval $I \subset (0, T)$ containing τ such that the optimal phase trajectory $\left\{x^*_{x_0}, \lambda^*_{x_0}\right\}$ of the original $(A, B, Q, R, S, T, \mathbb{U}, x_0)$-problem coincides with the optimal phase trajectory $\{\hat{x}, \hat{\lambda}\}$ corresponding to a $(A, B, Q, R, \hat{S}, T, \mathbb{R}, \hat{x}_0)$-problem.*

Proof The PMP standard formulation for the original problem indicates [Agrachev and Sachkov 2004; Pontryagin et al. 1962; Troutman 1996, and Eq. (9)] 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:

$$\dot{\lambda} = -2Qx^*_{x_0} - A'\lambda; \quad \lambda(T) = 2Sx^*_{x_0}(T), \tag{20}$$

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

$$\dot{x} = Ax + Bu^*_{x_0}; \quad x(0) = x_0. \tag{21}$$

For the Hamiltonian of this problem, namely

$$H(x, \lambda, u) := x'Qx + Ru^2 + \lambda'(Ax + Bu), \tag{22}$$

the related functions

$$h_t(u) := H\left(x^*_{x_0}(t), \lambda^*_{x_0}(t), u\right) \tag{23}$$

can be constructed for all $t \in [0, T]$. Then the PMP also asserts that each h_t should take its minimal value at $u = u^*_{x_0}(t)$ and, for the class of autonomous problems at hand, that

$$h_t(u^*_{x_0}(t)) \equiv \bar{h}_{x_0}, \tag{24}$$

a constant in the whole optimization interval $[0, T]$. But from standard results (Kalman et al. 1969; Sontag 1998), it is immediately deduced that, for each t, the control trajectory denoted by

$$\tilde{u}(t) := -\frac{1}{2}R^{-1}B'\lambda^*_{x_0}(t), \tag{25}$$

allows to construct the optimal strategy in the following way:

$$u^*_{x_0}(t) = \begin{cases} a & \text{if } \tilde{u}(t) < a \\ \tilde{u}(t) & \text{if } a \le \tilde{u}(t) \le b \\ b & \text{if } \tilde{u}(t) > b. \end{cases} \tag{26}$$

It is clear from the assumptions that $\tilde{u}(\tau) \in (a, b)$. Then by continuity of regular controls, this situation should extend to a maximal nontrivial interval $I := [\tau_1, \tau_2] \subset [0, T]$, where the optimal state and costate variables $\{x^*_{x_0}, \lambda^*_{x_0}\}$ verify the following identities

(i) from Eq. (20):

$$\dot{\lambda} = -2Qx^*_{x_0} - A'\lambda; \quad \lambda(\tau) = \lambda^*_{x_0}(\tau); \tag{27}$$

(ii) and by replacing $\tilde{u}$ from Eq. (25) into the dynamics:

$$\dot{x} = Ax - \frac{1}{2}W\lambda^*_{x_0}; \quad x(\tau) = x^*_{x_0}(\tau). \tag{28}$$

Therefore, by uniqueness of ODE solutions, in $I := [\tau_1, \tau_2]$ the optimal state/costate trajectories $\{x^*_{x_0}, \lambda^*_{x_0}\}$ coincide with a solution $\{\hat{x}, \hat{\lambda}\}$ to the Hamiltonian flow, i.e. with a trajectory of the system

$$\frac{d}{\mathrm{d}t}\begin{pmatrix} x \\ \lambda \end{pmatrix} = \mathbf{H}\begin{pmatrix} x \\ \lambda \end{pmatrix}, \tag{29}$$

where $\mathbf{H}$ is the Hamiltonian matrix for the linear-quadratic problem:

$$\mathbf{H} := \begin{pmatrix} A & -\frac{1}{2}W \\ -2Q & -A' \end{pmatrix}, \tag{30}$$

which is clearly the optimal solution of the unbounded $(A, B, Q, R, \hat{S}, T, \mathbb{R}, \hat{x}_0)$-problem. This proves the existence of the interval I alluded to in the statement.

Now, in such a 'regular' interval I, the following identity holds:

$$H(\hat{x}, \hat{\lambda}, \tilde{u}) = H^0(\hat{x}, \hat{\lambda}) = H(x^*_{x_0}, \lambda^*_{x_0}, u^*_{x_0}) = \bar{h}_{x_0}. \tag{31}$$

But if the optimal control $u^*_{x_0}$ saturates itself (i.e., if there exist time subintervals I_a with $u^*_{x_0}(t) \equiv a$ or $I_b u^*_{x_0}(t) \equiv b$), then the solution trajectory $\{\hat{x}, \hat{\lambda}\}$, existing and being unique in the whole interval $[0, T]$, forces $\{x^*_{x_0}, \lambda^*_{x_0}\}$ to differ from it in the interior of any subinterval I_a, I_b. Were another regular interval possible afterwards, the phase trajectories could not coincide any more, as well as their Hamiltonians. This proves that I is unique in $[0, T]$. □

The following assertions have practical significance in designing off-line control algorithms:

Corollary 3.2 *There exists a unique matrix trajectory* $\hat{P}(\cdot)$ *defined in* $[0, T]$, *solution to the same DRE of the original problem, namely,*

$$\dot{\pi} = \pi W \pi - \pi A - A'\pi - Q, \tag{32}$$

subject to some symmetric positive semidefinite final condition

$$\pi(T) = \hat{S} \tag{33}$$

(*where, in general,* $\hat{S} \neq S$). *The Riccati matrix* $\hat{P}(\cdot)$ *interrelates the state and costate trajectories* $\{\hat{x}, \hat{\lambda}\}$ *in the form*

$$\hat{\lambda}(t) = 2\hat{P}(t)\hat{x}(t)\forall t \in [0, T]. \tag{34}$$

Proof It follows from Jurdjievic (2006) that the Hamiltonian flow corresponding to Eq. (29) must admit the internal relationship expressed in Eq. (34) via some differentiable symmetric positive semidefinite matrix function $\hat{P}(\cdot)$. But by replacing this relation into ODEs (29), it follows that $\hat{P}(\cdot)$ must be a solution to the DRE (32). The final condition (33) characterizes the appropriate Riccati matrix for the subjacent unbounded problem. □

From the Theorem 3.1 and its Corollary 3.2, it follows immediately that $\hat{S} = S$ if and only if the control strategies $u^*_{x_0}$ and $\tilde{u}$ coincide in the whole interval $[0, T]$.

Corollary 3.3 *The optimal control of the restricted* $(A, B, Q, R, S, T, \mathbb{U}, x_0)$*-problem can be expressed in feedback form by saturating the optimal control of the related* $(A, B, Q, R, \hat{S}, T, \mathbb{R}, \hat{x}_0)$*-problem.*

Proof From Eqs. (25, 34), it follows that the optimal control $\hat{u}$ for the $(A, B, Q, R, \hat{S}, T, \mathbb{R}, \hat{x}_0)$-problem can be expressed as a feedback law, namely

$$\hat{u}(t) = -R^{-1}B'\hat{P}(t)\hat{x}(t); \tag{35}$$

and then, in the 'regular' subinterval $[\tau_1, \tau_2]$,

$$\tilde{u}(t) = -\frac{1}{2}R^{-1}B'\lambda^*_{x_0}(t) = -\frac{1}{2}R^{-1}B'\hat{\lambda}(t) = \hat{u}(t). \tag{36}$$

In particular, $\hat{u}(\tau_i) = \tilde{u}(\tau_i)$ assume the bound values a or b for $i = 1, 2$; and since both $\hat{u}$ and $\hat{u}$ are differentiable under parameter variations, then their values must remain in the exterior of (a, b) for some nontrivial time intervals before τ_1 and after τ_2 (or either reach the endpoints of $[0, T]$). The Hamiltonian Canonical Equations and the Pontryagin Maximum Principle, valid along the $\{\hat{x}, \hat{\lambda}\}$ and $\{x^*_{x_0}, \lambda^*_{x_0}\}$ trajectories, force $\hat{u}(t)$ and $\tilde{u}(t)$ to remain either both out from (a, b), or either both inside $[a, b]$ (and in the last case they must simultaneously coincide).

Since $\{\hat{x}, \hat{\lambda}\}$ is the solution to Eq. (29), it verifies

$$\begin{pmatrix} \hat{x}(0) \\ \hat{\lambda}(0) \end{pmatrix} = \mathrm{e}^{-\mathbf{H}t} \begin{pmatrix} \hat{x}(t) \\ \hat{\lambda}(t) \end{pmatrix} \quad \forall t \in [0, T], \tag{37}$$

with $\hat{x}_0 := \hat{x}(0)$ differing in general from x_0. Known $\hat{x}_0$ and $\hat{S}$, the control $\hat{u}(\cdot)$ can be generated as in any regular LQR problem, and therefore $u^*_{x_0}(\cdot)$ can be expressed as a feedback law from Eqs. (26, 36, 35). □

From now on, a control strategy that has the form of $u^*_{x_0}$ in Eq. (26) will be denoted with the superscript sat standing for 'saturation', which permits us to write the outcome simply as

$$u^*_{x_0} = \hat{u}^{\mathrm{sat}}. \tag{38}$$

These results transform the original problem with bounded controls (whose solution must be looked for in the infinite-dimensional space of admissible control trajectories) into a finite-dimensional search (for the hidden initial condition $\hat{x}_0$ and final penalization matrix $\hat{S}$).

4 Off-line analytical and numerical approaches

4.1 A one-dimensional example: the exponential function

It is well known (Troutman 1996) that the exponential function $x(t) = \mathrm{e}^t$ is the optimal state trajectory, and the optimal control strategy at the same time, of the following fixed-endpoint unbounded-controls problem

$$\dot{x} = u, x(0) = 1, x(1) = e, \; u(t) \in \mathbb{R}, \tag{39}$$

$$\mathcal{J}(u) = \int_0^1 \left[x^2(t) + u^2(t)\right] \mathrm{d}t. \tag{40}$$

As posed, this is not an LQR problem since the desired target is not the state $\bar{x} = 0$. But it is interesting to see that it can be solved through the Euler–Lagrange equation of Variational Calculus. The solution will then be the same if the set of control values is restricted to $\mathbb{U} = [a, b]$, with $a \leq 1$and $b \geq e$. But things change when the bounds are in the interior of $[a, b]$ and/or the final state is free. Here, after the change of variables

$$z := x - e, \tag{41}$$

the following related LQR problem will be treated:

$$\dot{z} = u, \quad z(0) = z_0 = 1 - e, u(t) \in [a, b], \tag{42}$$

$$\mathcal{J}(u) = \int_0^1 \left[z^2(t) + u^2(t)\right] \mathrm{d}t + [z(T)]^2, \tag{43}$$

$$A = 0, \quad B = 1, \quad Q = 1, \quad R = 1.$$

After some numerical explorations, the following values have been chosen for the control bounds and final penalization coefficient:

$$a = 1.44, \quad b = 2, \quad S = 13. \tag{44}$$

The problem is reduced to solve several nonlinear algebraic equations dictated by the PMP, which can be handled by standard numerical software. The answer found included two generalized switching points and an optimal final state

$$\tau_1 = 0.0925862, \quad \tau_2 = 0.61704, \quad z(1) = -0.0997093. \tag{45}$$

The optimal control and some linked objects result in

$$u^*(t) = \begin{cases} 2 & \forall t \in [0, \tau_1], \\ \hat{u}(t) & \forall t \in [\tau_1, \tau_2], \\ 1.44 & \forall t \in [\tau_2, 0], \end{cases} \tag{46}$$

$$\hat{u}(t) = -\frac{\hat{\lambda}(t)}{2} = d_1 \mathrm{e}^t - d_2 \mathrm{e}^{-t}, \tag{47}$$

$$\hat{z}(t) = d_1 \mathrm{e}^t + d_2 \mathrm{e}^{-t},$$

$$d_1 = 0.212802, \quad d_2 = -1.93792, \tag{48}$$

$$\hat{S} = 9.6, \quad \bar{h}_{z_0} = -1.64955, \quad \mathcal{J}(u^*) = 3.7309, \tag{49}$$

which are illustrated in Fig. 1.

4.2 A two-dimensional example: the cheapest stop of a train

Here another classical case-study, known in the literature as ‘the cheapest stop of a train’, will be revisited [see for instance Agrachev and Sachkov (2004) for the control–energy–cost, fixed-endpoints, unbounded controls version; Howlett et al. (2009) and the references therein for the numerical determination of switching times in more involved cases].

The dynamics of such a problem in its simplest form is linear:

$$\dot{x}_1 = x_2; \quad \dot{x}_2 = u, \tag{50}$$

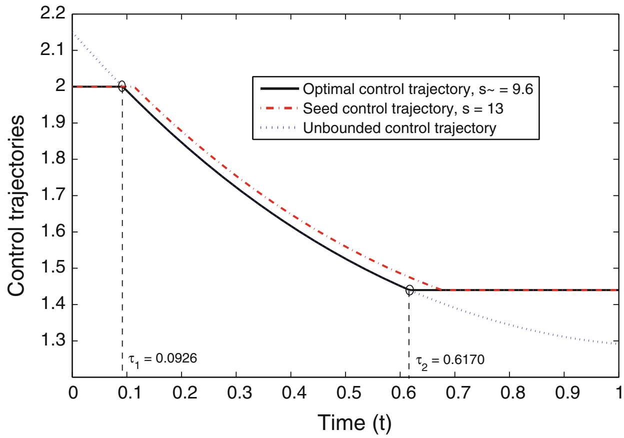

Fig. 1 Control trajectories for related optimal control problems in one dimension

or, in matrix notation,

$$\dot{x} = f(x, u) = Ax + Bu, \tag{51}$$

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{52}$$

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. The pair (A, B) is controllable. Unless indicated, the nominal initial conditions $x(0) = x_0$ chosen for illustration will be kept fixed at

$$x_1(0) = 1; \quad x_2(0) = -1. \tag{53}$$

The 'flexible endpoint' problem is associated with a 'quadratic final penalty' $K(x(T)) = x'(T)Sx(T)$ in the cost objective function, as announced in Eq. (1). The values adopted for the cost parameters are $Q = 10I$, $R = 0.5$, $S = 100I$.

In the literature (see Alt 2003; Bryson and Ho 1969 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, which are difficult to solve).

It can be shown (Costanza and Rivadeneira 2012) that the problem has at the most one nontrivial regular interval $[\tau_1, \tau_2] \subset [0, T]$,and that $u^*(t) \equiv a$ in $[0, \tau_1]$, $u^*(t) \equiv b$ in $[\tau_2, T]$. By denoting $\bar{x} = x^*(T)$, $\bar{\lambda} = \lambda^*(0)$, the exact solution procedure to determine the unknowns $\{\tau_1, \tau_2, \bar{x}_1, \bar{x}_2, \bar{\lambda}_1, \bar{\lambda}_2\}$, amounts to solve the following:

(i) In the interval $[\tau_2, T]$ the state obeys

$$\dot{x} = Ax + Bb; x(T) = \bar{x}; \tag{54}$$

which can be analytically solved, with $\bar{x}$ as an explicit unknown. By inserting the state solution into the corresponding dynamics for the costate, as in Eqs. (9), subject to $\lambda(T) = 2S\bar{x}$ as required by the PMP, the analytical integration for the costate can be performed.

(ii) In the regular interval $[\tau_1, \tau_2]$ both the state and costate follow the flow of the Hamilton Canonical Equations, i.e.

$$\begin{pmatrix} x(t) \\ \lambda(t) \end{pmatrix} = \mathrm{e}^{\mathbf{H}(t-\tau_2)} \begin{pmatrix} x(\tau_2) \\ \lambda(\tau_2) \end{pmatrix}, \tag{55}$$

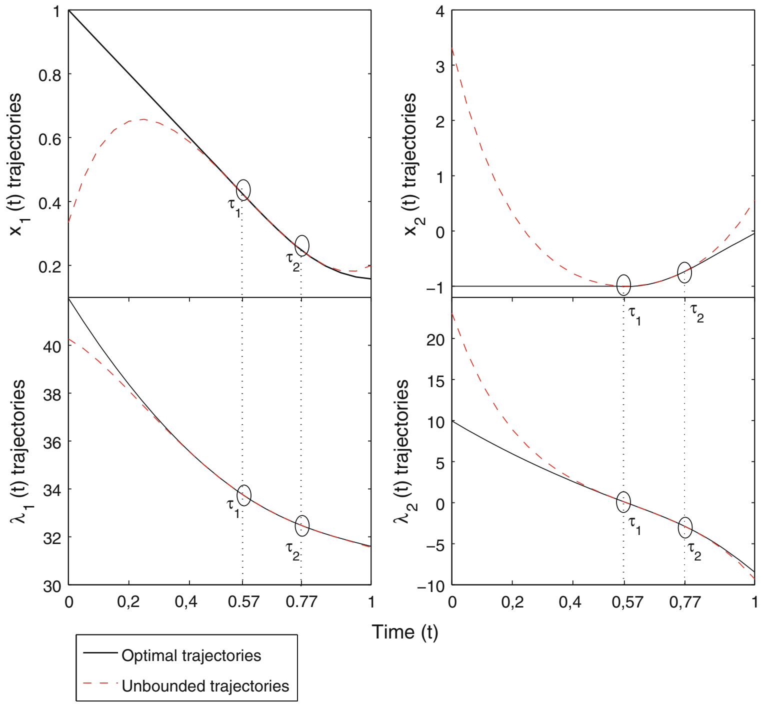

Fig. 2 State and costate trajectories $\{x^*, \lambda^*\}$ for the original constrained optimal control problem, and $\{\hat{x}, \hat{\lambda}\}$ for the 'hidden' unconstrained problem. Both trajectory pairs have the same value of the Hamiltonian $\bar{h}_{x_0} = -21.962$

(iii) In the interval $[0, \tau_1]$ the equations are

$$\dot{x} = Ax + Ba; \quad x(0) = x_0; \tag{56}$$

$$\dot{\lambda} = -2Qx - A'\lambda; \quad \lambda(0) = \bar{\lambda}. \tag{57}$$

These equations are subject to a number of matching conditions, expressed as equalities whose solutions must meet, for instance

$$\begin{aligned} x_{(r)}(\tau_2) &= x_{(\ell)}(\tau_2), \quad x_{(r)}(\tau_1) = x_{(\ell)}(\tau_1), \\ \lambda_{(r)}(\tau_2) &= \lambda_{(\ell)}(\tau_2), \quad \lambda_{(r)}(\tau_1) = \lambda_{(\ell)}(\tau_1), \end{aligned} \tag{58}$$

where the superscript (r) denotes 'approaching from the right' and (ℓ) 'from the left'. The number of equations (58) is greater than needed to solve for the unknowns. Choosing an array of independent equations and solving them through standard mathematical software, the unknowns are found (reported below) and the optimal trajectories are generated (illustrated in Fig. 2):

$$\begin{aligned} \tau_1 &= 0.575342, \quad \tau_2 = 0.776409, \\ \bar{x}_1 &= 0.158018, \quad \bar{x}_2 = -0.0423203 \\ \bar{\lambda}_1 &= 41.962, \quad \bar{\lambda}_2 = 9.9603. \end{aligned} \tag{59}$$

Additional relevant quantities may be calculated, for instance

$$\bar{h}_{x_0} = -21.962, \quad \mathcal{J}(u^*_{x_0}) = 15.1632, \tag{60}$$

$$x^*(\tau_1) = \begin{pmatrix} 0.424658 \\ -1 \end{pmatrix}, \quad x^*(\tau_2) = \begin{pmatrix} 0.242469 \\ -0.713092 \end{pmatrix}, \tag{61}$$

$$\hat{x}(0) = \begin{pmatrix} 0.334096 \\ 3.31771 \end{pmatrix}, \quad \hat{x}(T) = \begin{pmatrix} 0.199617 \\ 0.560481 \end{pmatrix}, \tag{62}$$

$$\hat{\lambda}(0) = \begin{pmatrix} 40.267 \\ 23.0974 \end{pmatrix}, \quad \hat{\lambda}(T) = \begin{pmatrix} 31.5591 \\ -9.2941 \end{pmatrix}, \tag{63}$$

$$\hat{S} = \begin{pmatrix} 528.3 & -160 \\ -160 & 48.7 \end{pmatrix} = \hat{P}(T), \quad \hat{P}(0) = \begin{pmatrix} 20.3 & 4.02 \\ 4.02 & 3.08 \end{pmatrix}. \tag{64}$$

5 The numerical on-line scheme

The purely off-line procedure implied by the PMP, as illustrated in the previous subsection, is equivalent to find both the hidden state $\hat{x}_0$ and the matrix $\hat{S}$, from which the generalized switching times τ_1, τ_2 and the optimal control can be calculated. Besides being a difficult nonlinear-programming problem, often avoided or overlooked in Engineering practice, the procedure may be criticized for generating an open-loop recipe, non adaptable under perturbations. To cope with both inconveniences, an on-line scheme is devised below, more in the line of standard numerical software for Engineering applications. The resulting control will normally be suboptimal, though always better than the 'seed' control, i.e. the straightforward saturation of the solution to the original problem in its unconstrained version.

It will be assumed that the time horizon is partitioned by two saturation times of the form $0 < \tau_1 < \tau_2 < T$; the optimal control resulting saturated in $[0, \tau_1] \cup [\tau_2, T]$ and regular in $[\tau_1, \tau_2]$.

The only off-line calculations needed are:

(i) the matrices $\alpha(t, S), \beta(t, S), t \in [0, T]$, solutions to Eqs. (14–15), in their analytic or interpolated approximate forms. The Riccati matrices can be obtained from α, β by using Eq. (18),

(ii) the matrix-valued function

$$\Psi(t, \tau) := \int_\tau^t \mathrm{e}^{A(t-\sigma)} d\sigma = \mathrm{e}^{At} \int_\tau^t \mathrm{e}^{-A\sigma} d\sigma, \tag{65}$$

needed for calculating the state-transition map in the saturation periods, i.e. for piecewise-constant control trajectories. Notice that Ψ depends only on the matrix A, and when A is invertible $\Psi(t, \tau) = A^{-1}\left(\mathrm{e}^{A(t-\tau)} - I\right)$, so it is relatively simple to have it solved in closed terms, or to have it numerically stored in memory. For instance, for the case-study in Sect. 4.2, A is not invertible but its exponential is just $\mathrm{e}^A = I + A$ and Ψ follows immediately:

$$\Psi(t, \tau) = \begin{pmatrix} t - \tau & \frac{(t-\tau)^2}{2} \\ 0 & t - \tau \end{pmatrix}, \tag{66}$$

(iii) a simulation of the dynamics applying the 'seed' control, i.e. the feedback form

$$u_{\text{seed}}(t, x) = \left[-R^{-1} B' P(t, T, S) x\right]^{\text{sat}}, \tag{67}$$

and of its resulting state trajectory x_{seed}, i.e. the solution to

$$\dot{x} = \begin{cases} Ax + Ba & \text{if } -R^{-1}B'P(t,T,S)x \le a, \\ Ax + Bb & \text{if } -R^{-1}B'P(t,T,S)x \ge b, \\ [A - WP(t,T,S)]\,x & \text{otherwise} \end{cases} \tag{68}$$

with $x(0) = x_0$. Then, a first approximation $\tau_{1,0} \le \tau_{2,0}$ to the optimal saturation points τ_1, τ_2will become available. In the same line of reasoning, the initial estimates are adopted for the hidden matrix $\hat{S}$, i.e. $\hat{S}_0 := S$. Also, a subdivision of the time-horizon of the form $t_0 = 0 < t_1 < t_2 < \cdots < t_N = T$ is adopted to make possible intermediate calculations, updating parameters, and deciding changes in the control strategy. Then, the on-line scheme proceeds through the following steps:

(i) For $t \in [0, t_1]$ the control is set to

$$u_1 \equiv u_{\text{seed}}(0, x_0), \tag{69}$$

as an approximation to $\left[-R^{-1}B'P(t,T,\hat{S}_0)\hat{x}(t)\right]^{\text{sat}}$ during the initial sampling period.

(ii) In the meantime, the relevant parameters are updated via some version of the gradient method (Pardalos and Pytlak 2008). In treating the example announced in Sect. 4.2, the following prescriptions are adopted, starting with $k = 0$:

$$\hat{S}_{k,j} := \hat{S}_{k,j-1} - \gamma_S \frac{\partial J}{\partial S}(\hat{S}_{k.j-1}, \tau_{1,k}, \tau_{2,k}),\ j = 1, 2, \ldots, \tag{70}$$

$$\hat{S}_{k+1} \approx \lim_j \hat{S}_{k,j}, \tag{71}$$

and on the same lines,

$$\tau_{i,k+1} \approx \lim_j \left(\tau_{i,k} - \gamma_{\tau_i} \frac{\partial J}{\partial \tau_i}\right)_j ;\ i = 1, 2, \tag{72}$$

where $\gamma_S, \gamma_{\tau_1}, \gamma_{\tau_2}$ are appropriate constants and J denotes a new object, representative of the total cost $\mathcal{J}(u)$, constructed from:

$$J(\hat{S}, \tau_1, \tau_2) := J_{[0,\tau_1]} + J_{[\tau_1,\tau_2]} + J_{[\tau_2,T]} + J_T, \tag{73}$$

$$J_{[0,\tau_1]} := \left[\widetilde{x'Qx} + Ru_1^2\right]\tau_1, \tag{74}$$

where u_1 denotes the saturated (equal to a or b) constant value of the control applied during $[0, \tau_1]$,

$$J_{[\tau_1,\tau_2]} := x'(\tau_1)\,P(\tau_1)\,x(\tau_1) - x'(\tau_2)\,P(\tau_2)\,x(\tau_2)\,, \tag{75}$$

where $P(\tau_i), i = 1, 2$ are short notations for $P(\tau_i, T, \hat{S})$, since for the unsaturated periods the increments in cost can be calculated from the value function V, known to be in this case

$$V(t,x) = x'P(t)x. \tag{76}$$

For each updating of $(\hat{S}, \tau_1, \tau_2)$, the special points $x(\tau_i)\,, i = 1, 2$, and $x(T)$ can be approximated by using Eqs. (19, 65) in the following form

$$x(\tau_1) = \mathrm{e}^{A\tau_1}x_0 + \Psi(\tau_1, 0)Bu_1, \tag{77}$$

$$x(\tau_2) = \alpha^{-1}(\tau_2 - \tau_1, \hat{S})x(\tau_1)\,, \tag{78}$$

$$x(T) = \mathrm{e}^{A(T-\tau_2)}x(\tau_2) + \Psi(T, \tau_2)Bu_2, \tag{79}$$

where u_2 denotes the suboptimal saturated constant value of the control to be applied during the interval $[\tau_2, T]$, i.e.

$$u_2 \equiv -R^{-1}B'P(\tau_{2,0}, T, \hat{S}_0)x_{\text{seed}}\left(\tau_{2,0}\right) \approx \left[-R^{-1}B'P(\tau_{2,0}, T, \hat{S}_0)\hat{x}\left(\tau_{2,0}\right)\right]^{\text{sat}}.$$

The remaining terms in Eq. (73) are calculated from

$$J_{[\tau_2,T]} := \left[\widetilde{x'Qx} + Ru_2^2\right](T-\tau_2), \quad J_T := x'(T)\,Sx(T), \tag{80}$$

where S is the original final penalization matrix. In Eqs. (74, 80) the term $\int x'Qx\mathrm{d}t$ of the original cost $\mathcal{J}(u)$ has been replaced by rectangular approximations

$$\int_t^{t+\Delta t} x'Qx\mathrm{d}t \approx \widetilde{x'Qx}\cdot\Delta t, \tag{81}$$

and in the numerical trials, the simplest mean-value approximation was used successfully, namely

$$\widetilde{x'Qx} = \frac{1}{2}\left[x'(t)Qx(t) + x'(t+\Delta t)Qx(t+\Delta t)\right]. \tag{82}$$

It can be proved that the partial derivatives $\frac{\partial J}{\partial S}$ and $\frac{\partial J}{\partial \tau_i}$ exist and are continuous (Jurdjievic 2006; Dhamo and Tröltzsch 2011).

(iii) The duration of each sampling period $[t_k, t_{k+1}]$ should be balanced between two conflicting objectives: first, $t_{k+1} - t_k$ should be small enough so as to obtain a reliable updating of the parameters $\hat{S}, \tau_1, \tau_2$ through simple approximations of the partial derivatives of J; but also the time extent must be compatible with the computational effort needed to reach convergence in Eqs. (71, 72). The procedure of point (ii) is repeated for the successive $k = 1, 2, \ldots,$ as far as $t_{k+1} \leq \tau_{1,k}$. The appropriate control should remain saturated and the prescription should be, simply $u_{(k)}(t) \equiv u_1$.
Notice that the Riccati matrix P corresponds now to a final condition $\pi(T) = \hat{S}_k$, and that P can be easily recovered from the matrices α, β, available online.
This scheme is continued until reaching the first updated saturation time, i.e. until for some k, $t_k \leq \tau_{1,k} < t_{k+1}$.
When that happens, the sampling time t_{k+1} is set equal to $\tau_{1,k}$, and the period $[t_k, t_{k+1}]$ is treated as before in points (ii–iii).

(iv) Now, after reaching the last updated saturation point $\tau_{1,k}$, ideally the feedback laws should coincide

$$-R^{-1}B'P(t, T, \hat{S}_k)x(t) = -R^{-1}B'P(t, T, \hat{S}_k)x^*(t) \tag{83}$$

(here $x(t)$ denotes the actual value of the state of the system at time t). However, since perturbations are taken into account, then the following feedback, eventually suboptimal but robust, is adopted:

$$u_{(k)}(t) = \left[-R^{-1}B'P(t, T, \hat{S}_k)x(t)\right]^{\text{sat}}. \tag{84}$$

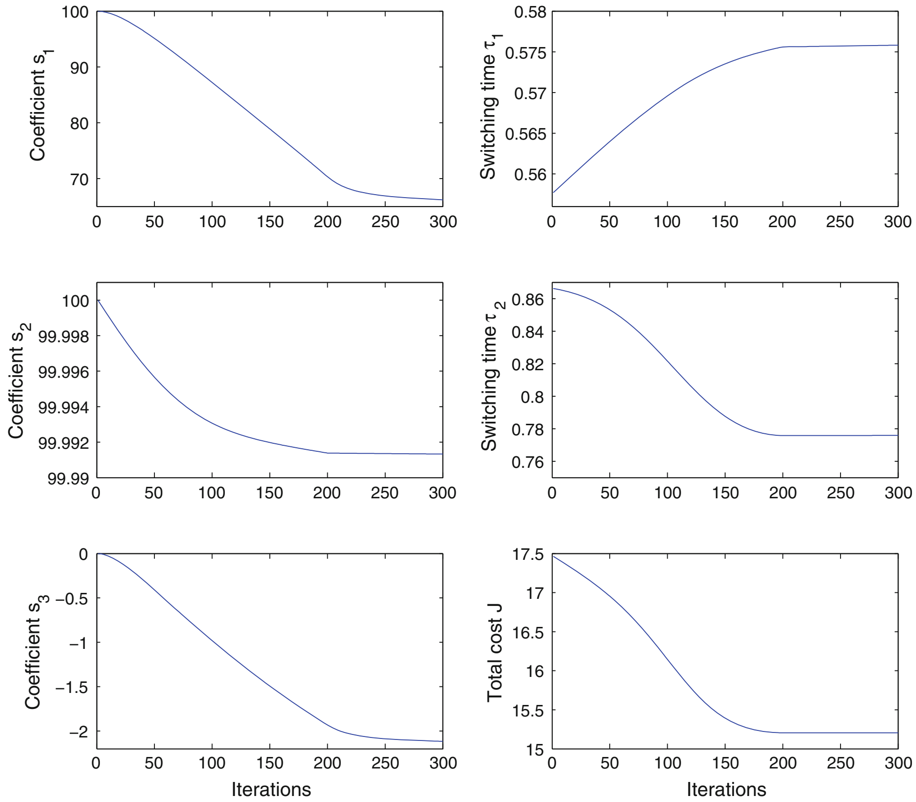

Fig. 3 Results of gradient-method-iterations for unknowns $\hat{S}$, τ_1, τ_2. The evolution of the modified total cost $J(\hat{S}, \tau_1, \tau_2)$ is also shown. The tuned values for iteration parameters were $\gamma_S = 2.0 \times 10^2$, $\gamma_{\tau_1} = 3.0 \times 10^{-6}$, $\gamma_{\tau_2} = 5.5 \times 10^{-4}$

Parameters $\hat{S}$ and τ_2 will keep on being updated as before, τ_1 remaining equal to its last corrected value.

(v) After the last updating for τ_2, the control should be constant again and equal to u_2 till the end.

In Figs. 3, 4, and 5 and Table 1, the results of applying the proposed online numerical scheme to the 'cheapest stop of the train' example, with same parameters as used in Sect. 4.2, are illustrated. The gradient-based iteration on the parameters sends the seed $\hat{S}_0 = S = 100$ approximately to the limiting matrix

$$\hat{S}_{350} = \begin{pmatrix} 64.94 & -2.11 \\ -2.1 & 99.99 \end{pmatrix}, \tag{85}$$

which is a local minimum but differs from the (global) optimal $\hat{S}$ calculated offline and reported in Eq. (64). The simple gradient method seems unable to lead the seed $\hat{S}_0$ out from its local basin, but still the total cost $\hat{J} = 15.22$ is much closer to the optimal $J^* = 15.16$ than the one obtained by saturating the unconstrained solution ($J_{\text{seed}} = 17.42$).

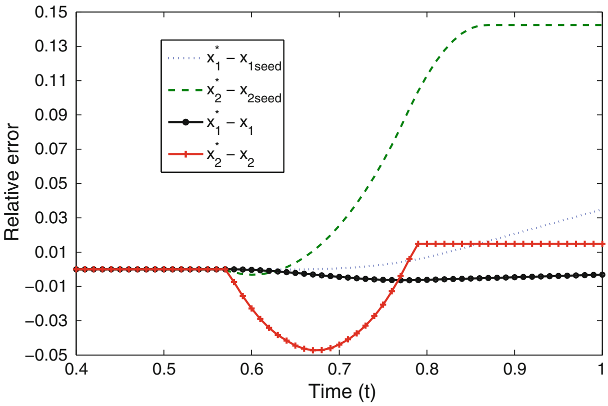

Fig. 4 State performance comparison between the online $x(.)$ and the seed $x_{\text{seed}}(.)$ trajectories with respect to the optimal solution $x^*(.)$

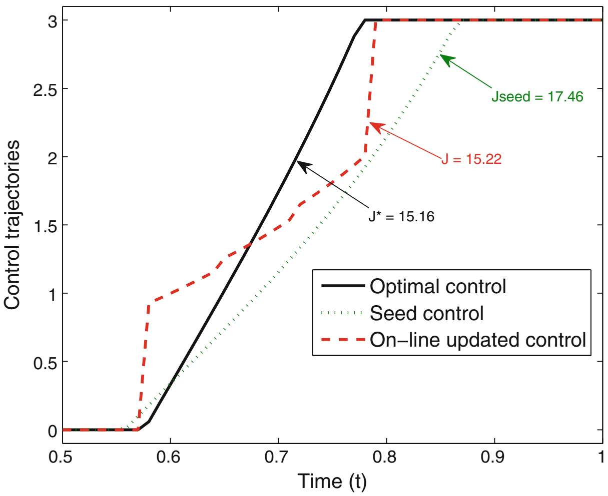

Fig. 5 Optimal and suboptimal control strategies

Table 1 Values of final state norms, switching times, and total costs of the control trajectories

	Optimal trajectory	Seed trajectory	On-line updated trajectory
$\|x(T)\|_2$	0.1636	0.2223	0.1710
τ_1	0.5753	0.5576	0.5716
τ_2	0.7764	0.8662	0.7892
J	15.16	17.42	15.22

6 Concluding remarks

A novel theoretical result and derived numerical procedures to solve bounded-control flexible-endpoint LQR problems have been substantiated and illustrated through two classical examples. The optimal control solution can, in principle, be generated off-line after detecting the initial condition and the final penalization matrix of a related-unbounded problem, which actually contains all the relevant information. But, even when this is possible, the result is really an open-loop prescription, since a hidden state variable has to be generated at all times. A closed-loop control is in general suboptimal, although it may be preferred in practice, when perturbations are expected to appear. With this objective, an efficient on-line algorithm is devised and applied to the two-dimensional case-study. The resulting strategies, either in the open-loop as in the feedback contexts, are quite different from the saturated form of the optimal control corresponding to the unrestricted problem with same parameters and initial condition, called here a 'seed' strategy. The seed scheme, often naively adopted in Engineering practice during the whole optimization period, is used here just to initiate the on-line numerical procedure. It should be acknowledged that, almost always, the new proposed procedure will be suboptimal. This is because the application of PMP principles to obtain the optimal solution is essentially an off-line calculation, and if any deviation from the optimal solution occurs (by mistake or by ignorance), then optimality will immediately be lost, no matter the subsequent effort. However, when the PMP solution was not previously found, or when only the 'seed' strategy is available, or when state perturbations appear in a real process-control situation; then no more than a suboptimal performance can be expected. The on-line updating of the parameter ($\hat{S}$) while the total cost is reduced (via the gradient method) should be regarded as a means to improve the seed control strategy as time evolves. This new scheme will result in the optimal strategy only when: (i) the right $\hat{S}$ is reached before the Riccati gain $P(\hat{S})$ has to be applied, and (ii) no state perturbations occur. As a consequence, the stability of the method is guaranteed since the cost is not allowed to increase and is bounded from below. Some positive features of the on-line proposed strategy are:

- The method is based on theoretical results ensuring that the hidden final penalization $\hat{S}$ and the appropriate (two at the most) saturation times τ_1, τ_2 are the critical objects to be ascertained.
- It takes advantage of the availability of α, β as functions of $(T - t, S)$, and consequently on the possibility of generating Riccati matrices $P(t, T, \hat{S})$ online by simple algebraic manipulations, as $\hat{S}$ is updated; i.e. the DRE does not need to be solved for any value of $\hat{S}$, not even offline.
- The control in Eq. (84) is given in feedback form, and therefore, the algorithm is unaffected by state perturbations due to fluctuations in environmental conditions.
- The updating of parameters $(\hat{S}, \tau_1, \tau_2)$ is performed via the gradient of the cost of the process, and this cost is calculated by simple algebraic formula instead of by predicting state and control trajectories and integrating the Lagrangian, as in most 'model predictive control' (MPC) techniques. This reduces the computational effort and allows for updating in shorter sampling times.
- Another conceptual difference with currently available MPC approaches is that here there exists a unique matrix $\hat{S}$ to look for in each LQR problem. This allows for further reduction on the computing effort, since there is no need for updating Riccati equations through receding horizon schemes.
- It is under exploration the online generation of the matrices α, β involved in the calculation of the optimal feedback gain at each sampling time. This step would improve the applicability of the algorithm to large-dimensional processes, especially to those governed by partial differential equations.

References

Agrachev A, Sachkov Y (2004) Control theory from the geometric viewpoint. Springer, Berlin

Alt W (2003) Approximation of optimal control problems with bound constraints by control parameterization. Control Cybern 32:451–472

Athans M, Falb PL (1966) Optimal control. An introduction to the theory and its applications. Mc-Graw-Hill, New York

Bernhard P (1972) Introducción a la Teoría de Control Óptimo. Cuaderno Nro. 4, Instituto de Matemática "Beppo Levi", Rosario, Argentina

Bryson AE, Ho YC (1969) Applied optimal control. Blaisdell

Costanza V (2008) Finding initial costates in finite-horizon nonlinear-quadratic optimal control problems. Opt Control Appl Methods 29:225–242

Costanza V, Neuman CE (2006) Optimal control of nonlinear chemical reactors via an initial-value Hamiltonian problem. Opt Control Appl Methods 27:41–60

Costanza V, Neuman CE (2009) Partial differential equations for missing boundary conditions in the linear-quadratic optimal control problem. Latin Am Appl Res 39:207–212

Costanza V, Rivadeneira PS (2008a) Feedback Óptimo del Problema Lineal-Cuadrático con Condiciones Flexibles. Paper A001VP. Proceedings XXI^{o} Congreso Argentino de AADECA, AADECA, Buenos Aires

Costanza V, Rivadeneira PS (2008b) Finite-horizon dynamic optimization of nonlinear systems in real time. Automatica 44:2427–2434

Costanza V, Rivadeneira PS (2011) Initial values for Riccati ODEs from variational PDEs. Comp Appl Math 30:331–347

Costanza V, Rivadeneira PS (2011) Approximating the solution to LQR problems with bounded controls. Latin Am Appl Res 41:339–351

Costanza V, Rivadeneira PS (2012) Minimizing control-energy in a class of bounded-control LQR problems. Opt Control Appl Methods (in press)

Costanza V, Rivadeneira PS, Spies RD (2011) Equations for the missing boundary values in the Hamiltonian formulation of optimal control problems. J Optim Theory Appl 149:26–46

Dhamo V, Tröltzsch F (2011) Some aspects of reachability for parabolic boundary control problems with control constraints. Comput Optim Appl 50:75–110

Howlett PG, Pudney PJ, Vu X (2009) Local energy minimization in optimal train control. Automatica 45:2692–2698

Jurdjievic V (2006) Geometric control theory. Cambridge University Press (digital printing of 1997 edition)

Kalman RE, Falb PL, Arbib MA (1969) Topics in mathematical system theory. McGraw-Hill, New York

Pardalos P, Pytlak R (2008) Conjugate gradient algorithms in nonconvex optimization. Springer, New York

Pontryagin LS et al (1962) The mathematical theory of optimal processes. Wiley, New York

Sontag ED (1998) Mathematical control theory. Springer, New York

Speyer JL, Jacobson DH (2010) Primer on optimal control theory. SIAM Books, Philadelphia

Troutman JL (1996) Variational calculus and optimal control. Springer, New York