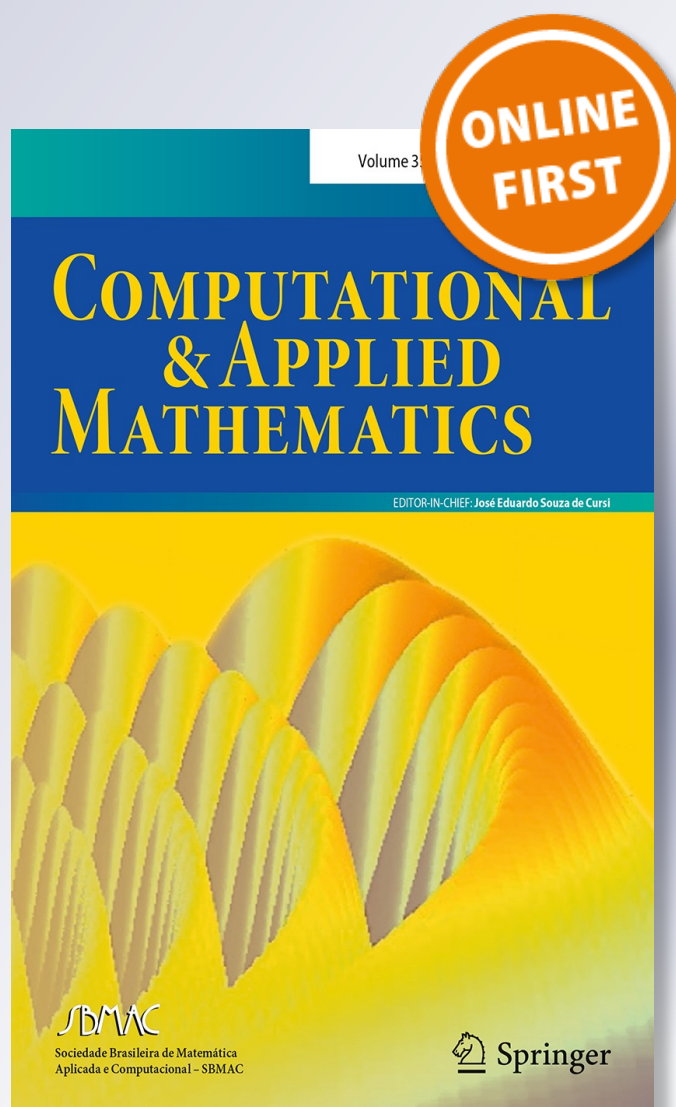# An efficient cost reduction procedure for bounded-control LQR problems

Vicente Costanza, Pablo S. Rivadeneira & John A. Gómez Múnera

Volume 3

ONLINE FIRST

COMPUTATIONAL & APPLIED MATHEMATICS

EDITOR-IN-CHIEF: José Eduardo Souza de Cursi

SBMAC
Sociedade Brasileira de Matemática
Aplicada e Computacional – SBMAC

Springer

Springer

CrossMark

An efficient cost reduction procedure for bounded-control LQR problems

Vicente Costanza[1] · **Pablo S. Rivadeneira**[1,2] ·
John A. Gómez Múnera[1]

Abstract A novel approach has been developed for approximating the solution to the constrained LQR problem, based on updating the final state and costate of a related regular problem, and on slightly shifting the switching times (the instants when the control meets the constraints). The main result is the expression of a suboptimal control in feedback form using the solution of some compatible Riccati equation. The gradient method is applied to reduce the cost via explicit algebraic formula for its partial derivatives with respect to the hidden final state/costate of the related regular problem and to the switching times. The numerical method is termed efficient because it does not involve integrations of states or cost trajectories, and reduces to its minimum the dimension of the unknown parameters at the final condition. All the relevant objects are calculated from a few auxiliary matrices, which are computed only once. The scheme is here applied to two case studies whose optimal solutions are known. The first example is a two-dimensional model of the 'cheapest stop of a train' problem. The second one refers to the temperature control of a metallic strip leaving a multi-stand rolling mill, a problem with a high-dimensional state.

✉ Vicente Costanza
   tsinoli@santafe-conicet.gov.ar

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

2   Universidad Nacional de Colombia - Sede Medellín, Facultad de Minas, Grupo "GITA",
    Cra. 80#65-223, Medellín, Colombia

Springer ✠BMAC

1 Introduction

The linear-quadratic regulator (LQR) problem is probably the most studied and used in the optimal control literature. Concurrently, the Hamiltonian formalism has also been at the core of the development of modern optimal control (Agrachev and Sachkov 2004; Athans and Falb 2006; Costanza and Rivadeneira 2014b; Pontryagin et al. 1964; Troutman 1996). In this paper, a Hamiltonian approach to the bounded control LQR problem will be pursued. When the n-dimensional finite-horizon problem for unbounded controls is regular, it leads to a set of $2n$ linear ordinary differential equations (ODEs) with two-point boundary-value conditions, known as the Hamilton Canonical Equations (HCEs). There are well-known methods (see for instance Costanza and Neuman 2009; Sontag 1998) to transform the boundary-value problem into an initial-value one. In the infinite-horizon, bilinear-quadratic regulator, and also in the change of set-point servo problems, there are also some attempts to find the missing initial condition for the costate variable from the data of each particular problem, which allows to integrate the Hamiltonian equations, on-line with the underlying control process (Costanza and Neuman 2006). For nonlinear systems, this line of work is still in its beginnings (Costanza and Neuman 2006; Costanza and Rivadeneira 2008; Costanza et al. 2009).

The bounded-control context may lead to non-regular optimal control problems, for whose solution there are not standard recipes (Athans and Falb 2006; Speyer and Jacobson 2010; Sontag 1998). Since the early 60s, the Pontryagin Maximum Principle (PMP) has been the standard theoretical setup to treat such non regular situations. In Kojima and Morari (2004) the infinite-horizon constrained LQR problem is studied by introducing singular-value decomposition applied to finite-time horizon linear subproblems, and then by joining all subproblems data into a big quadratic programming scheme leading to the optimal solution. The solution of the same problem is pursued through the convex duality approach in Goebel and Subbotin (2007), where a method of calculating the optimal stabilizing feedback without relying on discrete optimization is outlined. Recently, in Pannocchia et al. (2015), the same problem was tackled by approximating the input by a piecewise linear function on a finite-time discretization. The solution of this approximate problem results in a standard quadratic programming scheme. But notice that none of these approaches maintain the nonlinear character during a finite horizon.

In Sakizlis et al. (2005) the finite-horizon constrained LQR problem is analyzed and approximate solutions on a feedback form are suggested. The control law results nonlinear in the state and piecewise differentiable with respect to the switching times. However, the method is suitable only for low-dimensional systems, since for large systems the control needs to be calculated off-line, inconveniences that we try to avoid in what follows.

Our proposal takes advantage of the relationships between PMP and the classical Hamilton–Jacobi formalism. The theoretical result used here was initially phrased as follows: the optimal solution to a given bounded control LQR problem can be generated by saturating the solution of another unbounded LQR problem, having the 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 different matrix coefficient (Costanza and Rivadeneira 2013). Off-line and on-line schemes were developed to detect this new initial condition $\tilde{x}_0$ and final penalization matrix $\tilde{S}$ (Costanza and Rivadeneira 2014a; Costanza et al. 2013) under diverse circumstances. In Costanza and Rivadeneira (2014a) algebraic formula were developed to calculate the partial derivatives of the cost with respect to $\tilde{S}$ and $\tilde{x}_0$. Here this situation is considerably improved.

The main contribution of this article stems from a reformulation of the previous result in Costanza and Rivadeneira (2013). It is shown here that, instead of the final penalization $n \times n$ matrix $\tilde{S}$ and the hidden initial state $\tilde{x}_0$, only a $2n$-dimensional vector (ρ, μ) needs to be updated. Also, appropriate algebraic formula are devised to calculate the partial derivatives of the cost with respect to the new parameters, thus not involving ODEs numerical integrations. From the point of view of the computational effort, this reduction in the dimension of the parameters, along the algebraic calculation of partial derivatives may be determining when treating high-dimensional systems.

Since the control strategy is intended to work in feedback form when the control is between bounds, the precise knowledge of the initial condition of the underlying unrestricted LQR process is not substantial. In such a context it is enough to know the location of the time-instants where the control meets the constraints. Therefore, the proposed scheme updates ρ, μ, and the switching times τ_i referred above, while the total cost is being reduced via the gradient method (Bramanti et al. 1994; Fletcher and Reeves 1964; Pardalos and Pytlak 2008). In general the resulting control will at least be suboptimal, although in the two test cases illustrated here the optimal values were reached through the proposed procedure.

The numerical scheme takes advantage of the on-line availability of the Riccati matrices that correspond to a range of final penalty parameter values, generated from the solutions to a pair of first-order partial differential equations (Costanza 2007; Costanza and Neuman 2009).

The article has the following structure: after the Sect. 1, the regular LQR results and the main auxiliary matrices that will be used in the sequel are presented. Then the bounded-control version of the problem is described together with the theoretical results substantiating the proposed numerical procedure. Afterwards the algebraic formula to be employed in the numerical updating of the parameters are explicitly given, and the whole numerical scheme for cost reduction is elaborated. Two applications of the numerical scheme are then illustrated: one of them is the classical two-dimensional problem known as 'the cheapest stop of a train' (Agrachev and Sachkov 2004) , already revisited in Costanza and Rivadeneira (2013); and the second one is a linearized system coming from an industrial problem, initially stated through a first-order PDE, and here transformed in a high-dimensional ODE control system. Conclusions and perspectives are exposed at the end.

2 Equations for regular LQR optimal control problems

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

$$\mathcal{J}(u) = \int_0^{t_f} [x'(\tau)Qx(\tau) + u'(\tau)Ru(\tau)]\mathrm{d}\tau + x'(t_f)Sx(t_f), \qquad (1)$$

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

$$\dot{x} = Ax + Bu := f(x, u), \quad x(0) = x_0. \qquad (2)$$

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

As usual, the (real, time-constant) matrices in Eqs. (1) and (2) are 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. In this paper the input dimension will always be $m = 1$, the generalization to $m > 1$ requiring involved but no substantially different manipulations. The quadratic 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 for each pair (x, λ), and this occurs when u takes the explicit control value

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

(here independent of x), which is usually called 'the H-minimal control'. Finding the optimal control for a regular problem (see for instance Athans and Falb 2006; Sontag 1998) requires to solve the two-point boundary-value problem known as the 'Hamilton Canonical Equations' (HCEs),

$$\dot{x} = H_\lambda^0(x, \lambda); \quad x(0) = x_0, \tag{6}$$
$$\dot{\lambda} = -H_x^0(x, \lambda); \quad \lambda(t_f) = 2Sx(t_f), \tag{7}$$

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_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, i.e. Eqs. (6) and (7) for the LQR problem take the form

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

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

It is well known that the solution in feedback form to the unrestricted regular problem, as posed above, relies in turn on the solution $P(\cdot)$ to the 'Riccati Differential Equation' (RDE)

$$\dot{P} = PWP - PA - A'P - Q; \quad P(t_f) = S, \tag{10}$$

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

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

and, from 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 shown in the following sections.

The Hamiltonian matrix of the original problem,

$$\mathbf{H} := \begin{pmatrix} A & -W/2 \\ -2Q & -A' \end{pmatrix}, \tag{14}$$

and the associated fundamental matrix

$$\mathbf{U}(t) := e^{-\mathbf{H}t}, \tag{15}$$

will also be employed in devising algebraic formula for the partial derivatives of the cost. The matrix $\mathbf{U}(t)$ is $2n \times 2n$. For convenience, its $n \times n$ partition is denoted as

$$\begin{pmatrix} \mathbf{U}_1(t) & \mathbf{U}_2(t) \\ \mathbf{U}_3(t) & \mathbf{U}_4(t) \end{pmatrix} := \mathbf{U}(t). \tag{16}$$

Additional relevant objects from the LQR theory will be used in the sequel, for instance the matrices $\alpha(t_f, S), \beta(t_f, S)$ (see Bernhard 1972; Costanza et al. 2009; Sontag 1998) defined as:

$$\begin{pmatrix} \alpha(t_f, S) \\ \beta(t_f, S) \end{pmatrix} \triangleq \mathbf{U}(t_f) \begin{pmatrix} I \\ 2S \end{pmatrix} = \begin{pmatrix} \mathbf{U}_1(t_f) + 2\mathbf{U}_2(t_f)S \\ \mathbf{U}_3(t_f) + 2\mathbf{U}_4(t_f)S \end{pmatrix}. \tag{17}$$

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

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

and in such a case the matrices α, β are also related to the boundary conditions of the HCEs by the following relations:

$$x(t_f) = \alpha(t_f, S)^{-1}x(0), \quad \lambda(0) = \beta(t_f, S)x(t_f), \quad \lambda(t_f) = 2Sx(t_f). \tag{19}$$

3 The bounded-control case

The manipulated variable in most 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 can not 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 $\mathbb{R}$.

$$u(t) \in \mathbb{U} := [u_{\min}, u_{\max}]. \tag{20}$$

The qualitative features of optimal control solutions to bounded problems are significantly different from those of unbounded ones (Pontryagin et al. 1964). But questions about how much they actually differ, which classes of problems lead to bang–bang controls, and whether their solutions are just saturations of the optimal trajectories of unbounded problems, are still open.

The search for solutions to restricted problems most frequently falls in the domains of the Pontryagin Principle (PMP) (Pontryagin et al. 1964). However, even when solved, PMP is not flexible enough to treat state perturbations: no optimal feedback laws arise from the application of PMP equations, only at best open-loop control strategies.
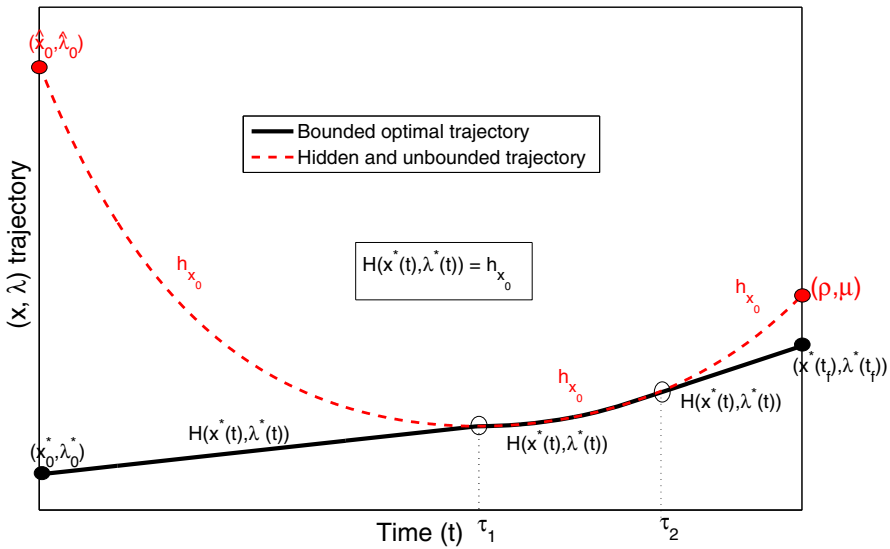
Fig. 1 Typical qualitative behavior of optimal phase trajectories for the original restricted problem $\{x_{x_0}^*, \lambda_{x_0}^*\}$, and for the related unrestricted process $\{\hat{x}, \hat{\lambda}\}$

In this paper the following result (Costanza and Rivadeneira 2013) will be exploited:

Let us assume that there exists a time instant $t \in (0, t_f)$ where $u_{x_0}^*(t) \in (u_{\min}, u_{\max})$. Then there exists a time interval $I \subset (0, t_f)$ containing t such that the optimal phase trajectory $\{x_{x_0}^*, \lambda_{x_0}^*\}$ of the original $(A, B, Q, R, S, t_f, \mathbb{U}, x_0)$-problem coincides with the optimal phase trajectory $\{\hat{x}, \hat{\lambda}\}$ corresponding to a $(A, B, Q, R, \hat{S}, t_f, \mathbb{R}, \hat{x}_0)$-problem.

The $(A, B, Q, R, \hat{S}, t_f, \mathbb{U}, \hat{x}_0)$-problem is not control-restricted ($\mathbb{U} \equiv \mathbb{R}$), and it is called the 'underlying', 'hidden', or, in short, the $\wedge$-problem. In what follows, it will be assumed that there exists just one maximal 'regular' interval $(\tau_1, \tau_2) \subset (0, t_f)$ where the control takes values in $(u_{\min}, u_{\max})$.

The result transforms the original problem with the 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 final conditions ρ, μ). A typical behavior of these objects is depicted in Fig. 1.

3.1 A novel approach to updating relevant parameters

Let us suppose that $(A, B, Q, R, \hat{S}, t_f, \mathbb{U}, \hat{x}_0)$ are the objects characterizing the underlying unrestricted problem, and that $\{\hat{x}, \hat{\lambda}\}$ is its optimal state/costate trajectory, and let us define $\rho := \hat{x}(t_f)$, $\mu := \hat{\lambda}(t_f)$,

$$\tilde{S} := \frac{1}{2} \frac{\mu \mu'}{\rho' \mu} \tag{21}$$

(whenever $\rho' \hat{S} \rho > 0$, otherwise $\tilde{S} := \hat{S}$). Notice that the $\frac{(n+1)n}{2}$ elements of $\tilde{S}$ are generated from the $2n$ elements of ρ and μ. The following theorem provides the basement for the new cost reduction procedure:

Theorem 1 *The $(A, B, Q, R, \tilde{S}, t_f, \mathbb{U}, \hat{x}_0)$-problem has the same optimal state/costate trajectory as the $(A, B, Q, R, \hat{S}, t_f, \mathbb{U}, \hat{x}_0)$-problem, namely $\{\hat{x}, \hat{\lambda}\}$.*

Proof To prove it, let us denote the trajectory $\{\tilde{x}, \tilde{\lambda}\}$ as the solution to the Hamiltonian equations of any of the problems (such ODEs do not depend on $\tilde{S}$, $\hat{S}$) with final conditions $\tilde{x}(t_f) = \rho$, $\tilde{\lambda}(t_f) = 2\tilde{S}\rho = \mu$.

Let us notice that

$$\tilde{\lambda}(t_f) = 2\tilde{S}\rho = \mu = 2\hat{S}\rho , \tag{22}$$

and therefore $\{\tilde{x}, \tilde{\lambda}\}$ is the optimal trajectory of the $(A, B, Q, R, \hat{S}, t_f, \mathbb{U}, \hat{x}_0)$-problem, i.e.

$$\{\tilde{x}, \tilde{\lambda}\} \equiv \{\hat{x}, \hat{\lambda}\}. \tag{23}$$

But since both pairs are solutions of the Hamiltonian equations, they must also satisfy $\tilde{\lambda} = 2\tilde{P}\tilde{x}$, $\hat{\lambda} = 2\hat{P}\hat{x}$, with $\tilde{P}$ and $\hat{P}$ two solutions to the RDE with final conditions $\tilde{S}$ and $\hat{S}$, respectively. Therefore the solution x to

$$\dot{x} = (A - W\tilde{P})x, \quad x(0) = \hat{x}_0 \tag{24}$$

is optimal for the $(A, B, Q, R, \tilde{S}, t_f, \mathbb{U}, \hat{x}_0)$ -problem. Now, by defining $\delta := x - \hat{x}$, it follows that

$$\begin{aligned}\dot{\delta} &= A\delta - W(\tilde{P}x - \hat{P}\hat{x}) = A\delta - W(\tilde{P}x - \tilde{P}\tilde{x}) \\ &= A\delta - W(\tilde{P}x - \tilde{P}\hat{x}) = (A - W\tilde{P})\delta. \end{aligned} \tag{25}$$

Since $\delta(0) = 0$, this means that $x \equiv \hat{x} \equiv \tilde{x}$. □

It follows immediately that:

Corollary 2 $\tilde{P}\hat{x} = \tilde{P}\tilde{x} = \hat{P}\hat{x}$, *and therefore the optimal feedback* (13) *is the same for the $\wedge$ -and the $\sim$-problems.*

Remark 3 As a useful consequence of these new results, it can be observed that the unknown parameters $(\hat{S}, \hat{x}_0)$ can be replaced by (ρ, μ) since: (i) the trajectory $(\hat{x}, \hat{\lambda})$ characterized by $(\hat{S}, \hat{x}_0)$ is also uniquely determined by (ρ, μ), and (ii) the Riccati matrices $\tilde{P}$, $\hat{P}$ arising from their final values $\hat{S}, \tilde{S}$, respectively, have the same effect in constructing the optimal feedback, and therefore they are equally effective for control purposes. Also, it is to be noticed that $\hat{S}$ has $\frac{(n+1)n}{2}$ unknown entries (since it is symmetric), and $\hat{x}_0$ another n unknowns. In total $\frac{(n+3)n}{2}$ unknowns. In the present setup, this number is reduced to $2n$ (for ρ, μ). For $n = 100$, this amounts to a reduction from 5150 to 200 unknowns parameters.

The numerical procedure will attempt then to reach the hidden (ρ, μ), starting from an appropriate $(\rho_{\text{seed}}, \mu_{\text{seed}})$, and by repeated application of the gradient method force the cost to decrease until convergence.

3.2 Algebraic formulas used in the procedure

3.2.1 Auxiliary objects

The following type of feedback control laws will be frequently used in the sequel

$$\tilde{u}(t) := \begin{cases} u_{\min}, & \forall t \in [0, \tau_1) \\ -R^{-1}B'P(t, \rho, \mu)x(t), & \forall t \in [\tau_1, \tau_2), \\ u_{\max}, & \forall t \in [\tau_2, t_f] \end{cases} \tag{26}$$

where $\tilde{u}(t)$ is a short notation for $\tilde{u}_{\rho,\mu,\tau_1,\tau_2}(t)$, which will be used to indicate that the feedback law is associated to the parameters $(\rho, \mu, \tau_1, \tau_2)$. The new strategy proposes variations of the variables ρ and μ in such a way that the matrix $\tilde{S}$ results completely determined by the final conditions. More precisely, the procedure should take into account the following considerations:

(i) The initial matrix $\tilde{S}_0$ will be constructed from the optimal state and the costate trajectories of the unconstrained problem (1), (2) by applying the feedback (13). The costate verifies at t_f that $\lambda(t_f) = 2P(t_f)x(t_f) = 2Sx(t_f)$, and then by denoting $\rho_{seed} := x(t_f)$, $\mu_{seed} := \lambda(t_f)$, the seed matrix $\tilde{S}_0$ is defined as

$$\tilde{S}_0 := \frac{1}{2} \frac{\mu_{seed}\mu'_{seed}}{\rho'_{seed}\mu_{seed}}. \tag{27}$$

(ii) A 'seed' strategy is adopted to start the iterative method below, namely

$$u_{seed}(t) := \begin{cases} u_{min} \text{ if } -R^{-1}B'P(t, \tilde{S}_0)x(t) & \leq u_{min} \\ u_{max} \text{ if } -R^{-1}B'P(t, \tilde{S}_0)x(t) & \geq u_{max} \\ -R^{-1}B'P(t, \tilde{S}_0)x(t) & \text{otherwise} \end{cases} . \tag{28}$$

The state trajectory corresponding to the control u_{seed} and starting at x_0, i.e. $x_{u_{seed}}$, will be denoted as x_{seed}. Notice that from the seed control and state trajectories, simulated for the new final matrix $\tilde{S}_0$, the first values for the saturation times, denoted $\tau_{1,0} \leq \tau_{2,0}$, can be detected if they exist.

(iii) In what follows $P(t, \rho, \mu)$ will denote the solution to the RDE (10), with final condition $P(t_f) = \tilde{S}$. When the value of ρ, μ are clear from the text, the notation may simplify to $P(t)$. The following identity will also be used ($\eta := \rho$ or μ) from (18),

$$\frac{\partial P(t, \rho, \mu)}{\partial \eta} = \frac{\partial \left[\frac{1}{2}\beta(t_f - t, \rho, \mu) \left[\alpha(t_f - t, \rho, \mu) \right]^{-1} \right]}{\partial \eta}$$

$$= \frac{1}{2} \left[\beta_\eta \alpha^{-1} - \beta \alpha^{-1} \alpha_\eta \alpha^{-1} \right] (t_f - t, \rho, \mu)$$

$$= \frac{1}{2} \left[\beta_\eta - 2P(t, \rho, \mu) \alpha_\eta \right] \alpha^{-1}, \tag{29}$$

where $\alpha, \alpha_\eta := \frac{\partial \alpha}{\partial \eta}, \beta_\eta := \frac{\partial \beta}{\partial \eta}$ should be evaluated at $(t_f - t, \rho, \mu)$, and from Eqs. (16), (17), the partial derivatives α_η, β_η result

$$\alpha_\eta = 2\mathbf{U}_2 \frac{\partial S}{\partial \eta}, \quad \beta_\eta = 2\mathbf{U}_4 \frac{\partial S}{\partial \eta}. \tag{30}$$

(iv) The 'saturated' fundamental matrix:

$$\Psi(t, \tau) := \int_\tau^t e^{A(t-\sigma)}d\sigma = e^{At} \int_\tau^t e^{-A\sigma}d\sigma = e^{A(t-\tau)} \int_0^{t-\tau} e^{-As}ds, \tag{31}$$

and the related matrices

$$\hat{\Psi}(t, \tau) := \int_\tau^t \Psi'(\sigma, \tau)Qe^{A(\sigma-\tau)}d\sigma = \int_0^{t-\tau} \Psi'(s+\tau, \tau)Qe^{As}ds, \tag{32}$$

$$\check{\Psi}(t, \tau) := \int_\tau^t e^{A'(\sigma-\tau)}Qe^{A(\sigma-\tau)}d\sigma = \int_0^{t-\tau} e^{A's}Qe^{As}ds. \tag{33}$$

will also be needed in the sequel. By defining the matrix functions $F_1(\Delta)$, $F_2(\Delta)$, $G(\Delta)$, $H(\Delta)$, $K(\Delta)$, $L_2(\Delta)$, $L_3(\Delta)$, $M_1(\Delta)$, and $M_2(\Delta)$, dependent on the scalar Δ, according with the following matrix partitions:

$$\begin{pmatrix} F_1(\Delta) & G(\Delta) \\ 0 & F_2(\Delta) \end{pmatrix} := \exp\left(\begin{bmatrix} -A' & Q \\ 0 & A \end{bmatrix}\Delta\right), \tag{34}$$

$$\begin{pmatrix} I & H(\Delta) \\ 0 & e^{-A\Delta} \end{pmatrix} := \exp\left(\begin{bmatrix} 0 & I \\ 0 & -A \end{bmatrix}\Delta\right), \tag{35}$$

$$\begin{pmatrix} I & M_1(\Delta) & K(\Delta) \\ 0 & L_2(\Delta) & M_2(\Delta) \\ 0 & 0 & L_3(\Delta) \end{pmatrix} := \exp\left(\begin{bmatrix} 0 & I & 0 \\ 0 & -A' & Q \\ 0 & 0 & A \end{bmatrix}\Delta\right); \tag{36}$$

then by employing appropriate formula for some special integrals of Hamiltonian matrix exponentials (Van Loan 1978), and resorting to integration by parts and similar algebraic manipulations; finally the integrals in Eqs. (31)–(33) can be expressed as follows:

$$\check{\Psi}(t, \tau) = F_2'(t - \tau) G(t - \tau), \tag{37}$$

$$\Psi(t, \tau) = e^{A(t-\tau)} H(t - \tau), \tag{38}$$

$$\hat{\Psi}(t, \tau) = \Psi'(t, \tau) e^{-A'(t-\tau)} \check{\Psi}(t, \tau) - K(t - \tau); \tag{39}$$

showing that all of them can be calculated without resorting to numerical integrations.

Along with the matrix $\mathbf{U}(t)$, all these auxiliary matrix functions can be evaluated once, interpolated off-line, and kept in memory to be recalled as needed during the optimization procedure. It will then be assumed in the sequel that they will be available as functions of their two variables (t, τ), in the range $[0, t_f] \times [0, t_f]$.

When $u(t) \equiv u_{min}$ in $[0, \tau_1]$, the state at τ_1 results

$$x(\tau_1) = e^{At} x_0 + \Psi(\tau_1, 0) B u_{min}. \tag{40}$$

(v) Here, it is important to notice that the state at τ_2 may be calculated from the Hamiltonian flow. Since for each $t \in [\tau_1, \tau_2)$ the control is $-R^{-1}B'P(t, \rho, \mu)x(t)$ and the costate (corresponding to this piece of a regular trajectory), denoted $\tilde{\lambda}(t)$, is $\tilde{\lambda}(t) = 2P(t, \rho, \mu)x(t)$, and

$$\begin{pmatrix} x(t) \\ \tilde{\lambda}(t) \end{pmatrix} = \mathbf{U}(\tau_1 - t)\begin{pmatrix} x(\tau_1) \\ \tilde{\lambda}(\tau_1) \end{pmatrix}, \tag{41}$$

this finally results in

$$\begin{aligned} x(\tau_2) &= \mathbf{U}_1(\tau_1 - \tau_2)x(\tau_1) + \mathbf{U}_2(\tau_1 - \tau_2)\tilde{\lambda}(\tau_1) \\ &= (\mathbf{U}_1(\tau_1 - \tau_2) + 2\mathbf{U}_2(\tau_1 - \tau_2)P(\tau_1, \eta))x(\tau_1). \end{aligned}$$

In what follows, the argument t of $\mathbf{U}(t)$ will also be omitted when the context is clear. Also, when $u(t) \equiv u_{max}$ in $[\tau_2, t_f]$, then the state at the final time t_f is

$$x(t_f) = e^{A(t_f - \tau_2)} x(\tau_2) + \Psi(t_f, \tau_2) B u_{max}. \tag{42}$$

3.2.2 The partial derivatives of the cost

Previous results in Dhamo and Tröltzsch (2011) indicate that the optimal control $u^*(\cdot)$ for the original $(A, B, Q, R, S, t_f, \mathbb{U}, x_0)$ -problem can be obtained by saturating the optimal control $\hat{u}(\cdot)$ corresponding to the unrestricted $\wedge$-problem. But actually it is not necessary to determine the whole $\hat{u}(\cdot)$ trajectory since we know it saturates outside $\mathcal{I} = [\tau_1, \tau_2]$, and also we know the values that $\hat{u}(\cdot)$ assumes for $t \notin [\tau_1, \tau_2]$. So it is enough to calculate τ_1, τ_2 and $\hat{P}(t)$, for $t \in \mathcal{I}$ to define the whole $u^*(\cdot)$ strategy. Then, we will search for $\rho, \mu, \tau_1, \tau_2$ to avoid calculating the open loop $\hat{u}(\cdot)$ and saturating it afterwards. This amounts to end in a number of $2(n + 1)$ parameters to be updated via the gradient method.

It is known (Dhamo and Tröltzsch 2011) that the total cost $\mathcal{J}(\tilde{u})$ is differentiable as a function of the variables $(\rho, \mu, \tau_1, \tau_2)$. It is time-partitioned here for convenience:

$$J(\rho, \mu, \tau_1, \tau_2) := \mathcal{J}(\tilde{u}) = J_1 + J_2 + J_3 + J_4, \tag{43}$$

where J_1 accounts for the trajectory cost associated with the saturation period $[0, \tau_1]$, J_2 with the regular interval $[\tau_1, \tau_2]$, J_3 with $[\tau_2, t_f]$, and J_4 measures the final penalty. Denoting by u_1 the saturated value of the control in $[0, \tau_1]$, and by u_2 the value assumed in $[\tau_2, t_f]$, the partial costs may be expressed as follows:

$$J_1 = Ru_1^2\tau_1 + \int_0^{\tau_1} x'(t) \, Qx(t) \, \mathrm{d}t, \tag{44}$$

$$J_2 = x'(\tau_1)P(\tau_1) x(\tau_1) - x'(\tau_2) P(\tau_2) x(\tau_2), \tag{45}$$

$$J_3 = Ru_2^2(t_f - \tau_2) + \int_{\tau_2}^{t_f} x'(t) \, Qx(t) \, \mathrm{d}t, \tag{46}$$

$$J_4 = x'(t_f) Sx(t_f). \tag{47}$$

The derivatives D_η with respect to the variables ρ and μ (denoted indistinctly by the variable η) result:

$$D_\eta J_1 = 0, \; D_\eta J_2 = x'(\tau_1) \left[\frac{\partial P(\tau_1)}{\partial \eta} - 4(\mathbf{U}_1 + 2\mathbf{U}_2 P(\tau_1))' P(\tau_2)\mathbf{U}_2 \frac{\partial P(\tau_1)}{\partial \eta} \cdots, \right.$$
$$\left. -(\mathbf{U}_1 + 2\mathbf{U}_2 P(\tau_1))' \frac{\partial P(\tau_2)}{\partial \eta} (\mathbf{U}_1 + 2\mathbf{U}_2 P(\tau_1)) \right] x(\tau_1), \tag{48}$$

where the partial derivatives of P are calculated from Eqs. (29), (30), and from

$$\frac{\partial \widetilde{S}_{kl}}{\partial \rho_i} = \frac{1}{2} (\mu_k \mu_l) \left(-\frac{\mu_i}{(\rho'\mu)^2} \right), \tag{49}$$

$$\frac{\partial \widetilde{S}_{kl}}{\partial \mu_j} = \frac{1}{2} \left[\frac{Z_{j(kl)}}{(\rho'\mu)} - \frac{(\mu_k \mu_l)\,\rho_j}{(\rho'\mu)^2} \right], \tag{50}$$

where

$$Z_{j(kl)} := \frac{\partial \mu_k \mu_l}{\partial \mu_j} = \begin{cases} 2\mu_j & \text{if } k = l = j \\ \mu_k & \text{if } l = j \text{ and } k \neq j \\ \mu_l & \text{if } k = j \text{ and } l \neq j \\ 0 & \text{if } k \neq j \text{ and } l \neq j \end{cases}. \tag{51}$$

$$D_\eta J_3 = \int_{\tau_2}^{t_f} 2x'(t)Q\frac{\partial x(t)}{\partial \eta}dt = 4\left[x'(\tau_2)\check{\Psi}(t_f, \tau_2) + \cdots \right.$$

$$\left. + u_2 B'\hat{\Psi}(t_f, \tau_2)\right]\mathbf{U}_2(\tau_1 - \tau_2)\frac{\partial P(\tau_1, \eta)}{\partial \eta}x(\tau_1), \quad (52)$$

after replacing $x'(t) = x'(\tau_2) e^{A'(t-\tau_2)} + B'\Psi'(t, \tau_2)u_2$, and expanding

$$\frac{\partial x(t)}{\partial \eta} = e^{A(t-\tau_2)}\frac{\partial x(\tau_2)}{\partial \eta} = e^{A(t-\tau_2)}2\mathbf{U}_2(\tau_1 - \tau_2)\frac{\partial P(\tau_1, \rho, \mu)}{\partial \eta}x(\tau_1). \quad (53)$$

Finally,

$$D_\eta J_4 = 2x'(t_f) S \frac{\partial x(t_f)}{\partial \eta} = 4x'(t_f) S e^{A(t_f-\tau_2)}\mathbf{U}_2(\tau_1 - \tau_2)\frac{\partial P(\tau_1, \eta)}{\partial \eta}x(\tau_1). \quad (54)$$

The derivatives with respect to the switching times τ_1 and τ_2 are

$$D_{\tau_1} J_1 = Ru_1^2 + x'(\tau_1)Qx(\tau_1), \quad (55)$$

$$D_{\tau_1} J_2 = 2x'(\tau_1)P(\tau_1)Bu_1 + x'(\tau_1)P(\tau_1)WP(\tau_1)x(\tau_1)\ldots$$

$$-x'(\tau_1)Qx(\tau_1) - 2x'(\tau_2)P(\tau_2)\frac{\partial x(\tau_2)}{\partial \tau_1}, \quad (56)$$

$$D_{\tau_1} J_3 = \int_{\tau_2}^{t_f} 2x'(t)Qe^{A(t-\tau_2)}dt\frac{\partial x(\tau_2)}{\partial \tau_1}$$

$$= \int_{\tau_2}^{t_f}\left(e^{A(t-\tau_2)}x(\tau_2) + \Psi(t_f, \tau_2)Bu_2\right)' Qe^{A(t-\tau_2)}dt\frac{\partial x(\tau_2)}{\partial \tau_1}$$

$$= x'(\tau_2)\int_{\tau_2}^{t_f} e^{A'(t-\tau_2)}Qe^{A(t-\tau_2)}dt\frac{\partial x(\tau_2)}{\partial \tau_1}\cdots$$

$$+ B'u_2\int_{\tau_2}^{t_f} \Psi'(t_f, \tau_2)Qe^{A(t-\tau_2)}dt\frac{\partial x(\tau_2)}{\partial \tau_1}$$

$$= (x'(\tau_2)\check{\Psi}(t_f, \tau_2) + u_2 B'\hat{\Psi}(t_f, \tau_2))\frac{\partial x(\tau_2)}{\partial \tau_1}. \quad (57)$$

For the final penalization, from Eq. (42),

$$D_{\tau_1} J_4 = 2x'(t_f)Se^{A(t_f-\tau_2)}\frac{\partial x(\tau_2)}{\partial \tau_1}, \quad (58)$$

where the last term can be expressed in terms of the data as

$$\frac{\partial x(\tau_2)}{\partial \tau_1} = \mathbf{U}_1(\tau_1 - \tau_2)(Ax(\tau_1) + Bu_1) - (A - WP(\tau_2))x(\tau_2) + \cdots$$

$$2\mathbf{U}_2(\tau_1 - \tau_2)[P(\tau_1)(WP(\tau_1)x(\tau_1) + Bu_1) - (A'P + Q)x(\tau_1)]. \quad (59)$$

Similarly, the derivatives with respect to τ_2 are

$$D_{\tau_2} J_1 = 0, \quad (60)$$

$$D_{\tau_2} J_2 = x'(\tau_2)(P(\tau_2)WP(\tau_2) + Q)x(\tau_2), \tag{61}$$

$$D_{\tau_2} J_3 = -(Ru_2^2 + x'(\tau_2)Qx(\tau_2)), \tag{62}$$

$$D_{\tau_2} J_4 = -2x'(t_f)Se^{A(t_f - \tau_2)}(WP(\tau_2)x(\tau_2) + Bu_2). \tag{63}$$

3.2.3 Updating the parameters

First approximations $\tau_{1,0}$, $\tau_{2,0}$ to the optimal saturation points τ_1, τ_2 become available after simulating the state trajectory x_{seed} as described in Sect. 3.2.1. For $t \in [0, t_f]$ the control is set to

$$\tilde{u}_0(t) \equiv u_{\text{seed}}(t). \tag{64}$$

The parameters $(\rho, \mu, \tau_1, \tau_2)$ are then updated to construct successive control strategies $\tilde{u}_j$; $j = 1, 2, \ldots$ that decrease the value of the total cost,

$$\mathcal{J}(\tilde{u}_{j+1}) \leq \mathcal{J}(\tilde{u}_j) \leq \cdots \mathcal{J}(u_{\text{seed}}); \quad j = 1, 2, \ldots \tag{65}$$

according to the prescriptions of the gradient method:

$$\begin{cases} \rho_j = \rho_{j-1} - \gamma_j \frac{\partial J}{\partial \rho}(\rho_{j-1}, \mu_j, \tau_{1,j}, \tau_{2,j}) \\ \mu_j = \mu_{j-1} - \gamma_j \frac{\partial J}{\partial \mu}(\rho_j, \mu_{j-1}, \tau_{1,j}, \tau_{2,j}) \end{cases}, \tag{66}$$

$$\begin{cases} \tau_{1,j} = \tau_{1,j-1} - \gamma_j \frac{\partial J}{\partial \tau_1}(\rho_j, \mu_j, \tau_{1,j-1}, \tau_{2,j}) \\ \tau_{2,j} = \tau_{2,j-1} - \gamma_j \frac{\partial J}{\partial \tau_2}(\rho_j, \mu_j, \tau_{1,j}, \tau_{2,j-1}) \end{cases}, \tag{67}$$

until convergence or practical decision to stop, where γ_j is a positive, small real number measuring the portion of the gradient vector to be applied in each iteration, chosen and tuned by the user.

Some more sophisticated versions of the gradient method can be explored in this context. For instance, in the second example of this paper the method of Fletcher and Reeves (Bramanti et al. 1994; Fletcher and Reeves 1964) was essayed, producing a moderate improvement in the number of iterations needed to reach the optimum. The conjugate search directions s^j are calculated sequentially starting from

$$s^0 := -\nabla J(p^0), \tag{68}$$

where p denotes the parameters involved $(\rho, \mu, \tau_1, \tau_2)$, and p^0 contains their seed values. The value s^0 is the initiating search direction and ∇ is the (row) gradient operator. The updating equation for the search direction (prescribed by this technique) is

$$s^{j+1} = -\nabla J(p^{j+1}) + s^j \frac{\nabla J(p^{j+1})\nabla' J(p^{j+1})}{\nabla J(p^j)\nabla' J(p^j)}, \quad j = 1, 2, \ldots, \tag{69}$$

and the updating of parameters is conducted as follows:

$$(\rho, \mu, \tau_1, \tau_2)_{j+1} = (\rho, \mu, \tau_1, \tau_2)_j + \gamma_j s^j. \tag{70}$$

3.2.4 Schematic description of the numerical procedure

A practical instrumentation of the results in previous sections would require the following basic steps:

(i) Calculate the auxiliary matrices as described in Eq. (2) and Sect. 3.2.1 and store them in memory. Obtain first approximations of the state and control trajectories, x_{seed} and u_{seed} respectively (for instance, those corresponding to the saturation of the optimal control of the subjacent unrestricted regular problem). These trajectories provide initial tentative values for the parameters $(\rho, \mu, \tau_1, \tau_2)$.

(ii) Update the parameters through the chosen version of the gradient method and of the scalar γ, as proposed in Sect. 3.2.3.

(iii) Confirm by numerical integration that the real cost is actually decreasing, and in such a case return to step (ii). This step can be executed every several iterations of the gradient method, according to the programmer intuition and the particularities of the application.

(iv) If the cost does not decrease, reduce the speed γ until it does, or

(v) Adopt the last results as best approximation to the optimal control attainable by this method, and stop.

4 Applications and numerical results

4.1 A low-dimensional example

The first example of this paper refers to the 'cheapest stop of a train' (Agrachev and Sachkov 2004; Costanza and Rivadeneira 2013, 2014a, b; Costanza et al. 2013) and its LQR problem defined by the objects:

$$
\begin{aligned}
\dot{x}_1(t) &= x_2(t); & x_1(0) &= 1, \\
\dot{x}_2(t) &= u(t); & x_2(0) &= -1, \\
u(t) &\in [0, 3]; & 0 &\le t \le t_f = 1,
\end{aligned}
\tag{71}
$$

$Q = 10I$, $R = 0.5$, and $S = 100I$. The optimal solution was already presented in Costanza and Rivadeneira (2013). The result of updating parameters from the gradient method applied to $\tau_1, \tau_2, \rho, \mu$, are depicted in Figs. 2, 3, 4, 5 and 6. In Fig. 3, the seed and the optimal trajectories are illustrated. The Fig. 4 shows the reduction of the total cost until obtaining the minimal value, i.e., its optimal value $J^* = 15.1632$.

The updating of the switching times may lead to different situations, as depicted in Fig. 2. The updated values, after an application of the gradient as in Eq. (67), are denoted in Fig. 2 by τ_1, τ_2. The updates of ρ, μ may lead to different types of regular intervals, for instance $[\tau_1^{\text{sat},1}, \tau_2^{\text{sat},1}]$ or $[\tau_1^{\text{sat},2}, \tau_2^{\text{sat},2}]$. In case 1, the control is constructed as

$$
u^1(t) := \begin{cases} u_{\min} & \text{if } t \in [0, \tau_1] \\ \tilde{u}(t) & \text{if } t \in [\tau_1, \tau_2] \\ u_{\max} & \text{if } t \in [\tau_2, t_f] \end{cases},
\tag{72}
$$

i.e., the saturated values τ_1, τ_2 are respected, forcing the control to jump due to the values of $\tilde{u}(t)$ in the regular interval. In the case 2,
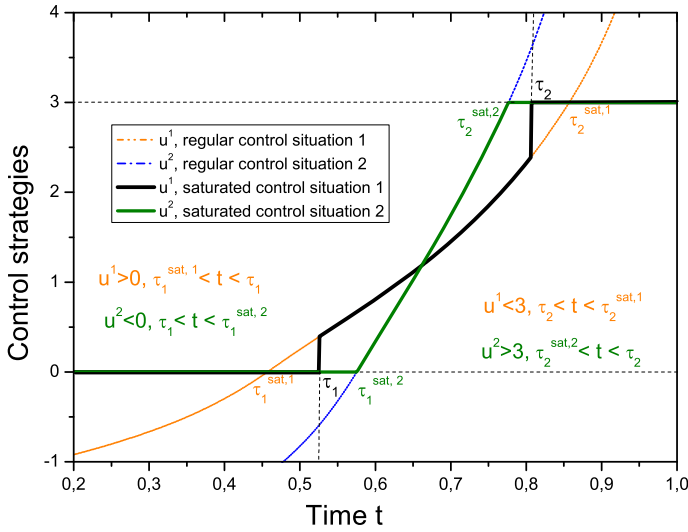
Fig. 2 How to update the switching times under different situations when the gradient method is applied
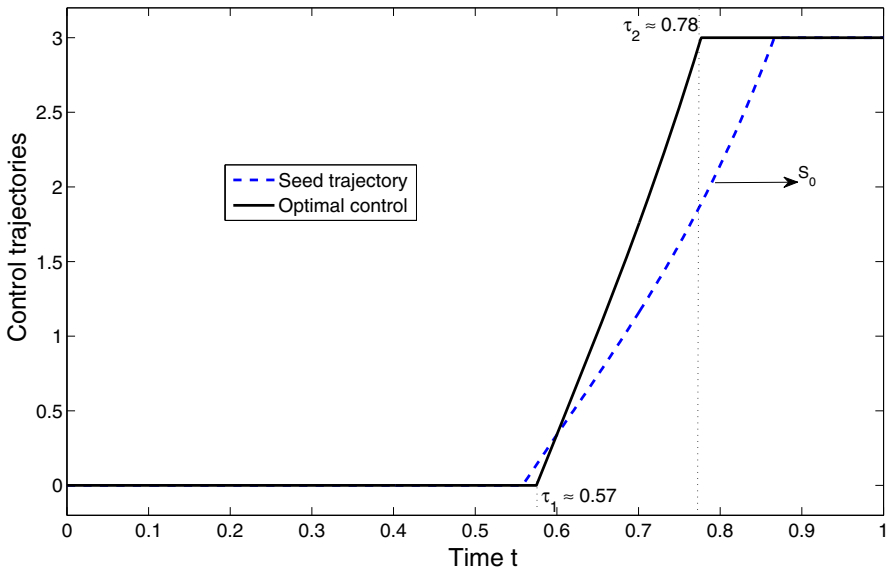

Fig. 3 Control strategy resulting from application of the gradient method to the bidimensional example

$$
u^2(t) := \begin{cases} u_{\min} & \text{if } t \in [0, \tau_1^{\text{sat},2}] \\ \tilde{u}(t) & \text{if } t \in [\tau_1^{\text{sat},2}, \tau_2^{\text{sat},2}] \\ u_{\max} & \text{if } t \in [\tau_2^{\text{sat},2}, t_f] \end{cases}, \tag{73}
$$

since at τ_1 and τ_2 the regular control $\tilde{u}(t)$ could take values out of the allowed domain $\mathbb{U} = [u_{\min}, u_{\max}]$. All other situations are handled accordingly; i.e.: τ_1 and τ_2 are
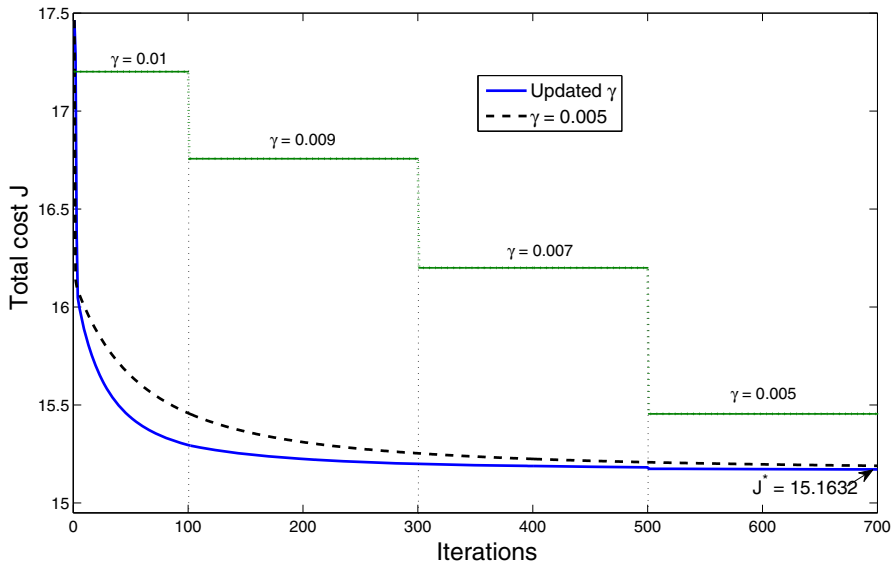
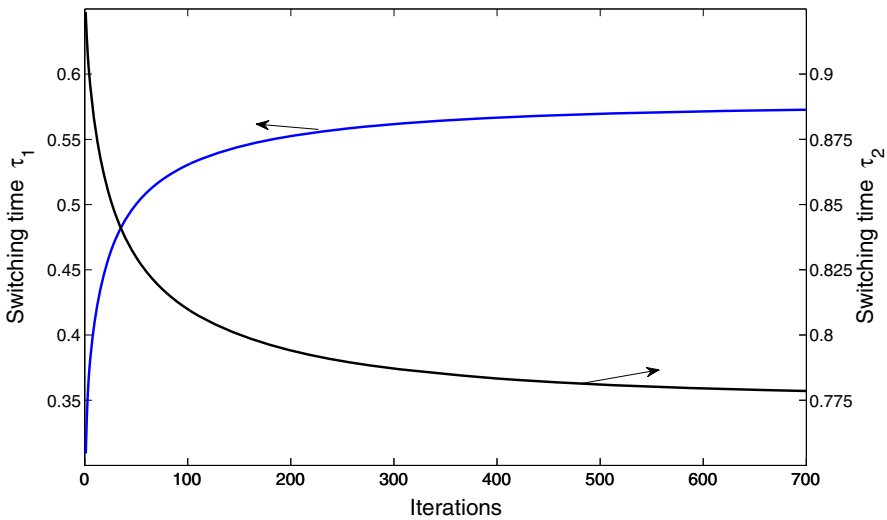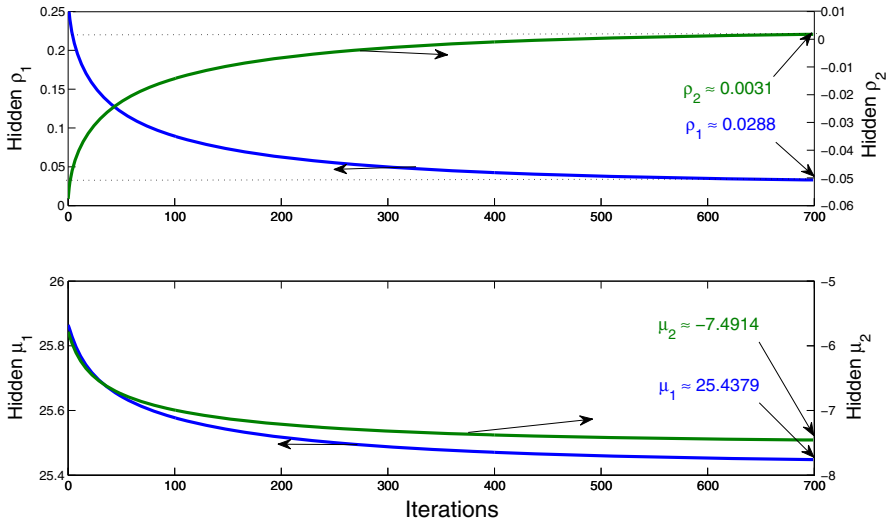**Fig. 4** Evolution of the total cost resulting from the gradient method

Fig. 5 Evolution of the τ_s resulting from the gradient method

respected as possible, but $\tau_i^{\mathrm{sat},j}$ are used when the values of $\tilde{u}(t)$ fall in the complement of $\mathbb{U}$.

Two series of iterations were performed to determine the best strategy for the updating of the parameter gamma: in the first one, the value of the parameter γ was maintained constant, $\gamma = 0.005$ (black dashed line in Fig. 4), and in the second one, it was updated whenever felt needed. It is clear that at the beginning of the method γ could be chosen moderately high, provided that the method decreases the cost. However, at some point this high γ is useless since big gradients may 'jump' over eventual local minima, so disrupting

Fig. 6 Evolution of the ρ_s and μ_s resulting from the gradient method

the decreasing of the cost. At this point γ must be reduced. In this case study, the first value was settled in $\gamma = 0.01$, and the last reached value was $\gamma = 0.005$ (blue solid line in Fig. 4). This way of changing the parameter γ showed a better performance than when it was kept constant. Figures 5 and 6 illustrate the evolution of τ_1, τ_2; and of ρ, μ, respectively. After 700 iterations the parameters converge to their optimal values. The final values are: $\tau_1 = 0.5753$, $\tau_2 = 0.7764$, and $\tilde{S}_{11} = 456.613$, $\tilde{S}_{21} = -134.475$, $\tilde{S}_{22} = 39.6015$, calculated from Eq. (21) and the values of ρ, μ are shown in Fig. 6.

4.2 A typical linearized model situation: the rolling mill

The second case-study models a rolling mill as described in Costanza and Rivadeneira (2014a) and Hearns and Grimble (2010), whose (infinite-dimensional) dynamics, coming from a standard energy balance, obeys the following first-order PDE

$$\frac{\partial \theta}{\partial t} = -V \frac{\partial \theta}{\partial z} + a(\theta_a - \theta) + b(\theta_a^4 - \theta^4), \tag{74}$$

where $\theta(t, z)$ is the temperature of the metallic strip at time t and location z in the trend, $V(t)$ is the linear speed of the strip, and θ_a is the ambient temperature (assumed constant in this set-up). The coefficients a, b weigh the rate of heating due to conduction and radiation, respectively. The system is simplified by neglecting radiation (small b), and by supposing that the temperature will stay around the equilibrium profile

$$\theta_{SS}(z) = \theta_a + (\theta_0 - \theta_a) \exp(-az/V_0), \tag{75}$$

which is the solution to Eq. (74) with $b = 0$, $\partial \theta / \partial t = 0$, $V(t) \equiv V_0$, and $\theta_0 := \theta_{SS}(0)$, some appropriate constant characterizing each physical set-up. The following definitions

$$\Delta \theta(t, z) := \theta(t, z) - \theta_{SS}(z); \quad u(t) := V(t) - V_0, \tag{76}$$

allow to approximately express the dynamics of the fluctuations through the 'linearized' version of Eq. (74), namely

$$\frac{\partial \Delta \theta}{\partial t} = -V_0 \frac{\partial \Delta \theta}{\partial z} - a \Delta \theta + \left[\frac{a}{V_0}(\theta_0 - \theta_a) \exp(-az/V_0) \right] u, \tag{77}$$

after neglecting the term $u \partial \Delta \theta / \partial z$, on the argument that it is the product of two 'small' quantities.

From the control theory perspective the state x in Eq. (77) may be identified, at each time t, with the whole z-profile of the temperature

$$x(t) := \{\Delta \theta(t, z), \quad z \in [0, L]\}. \tag{78}$$

This, in principle, makes the dynamics infinite-dimensional, of the form

$$\dot{x} = F(x, u), \tag{79}$$

F denoting the operator associated to the right-hand-side of Eq. (77), so placing its rigorous treatment out of the scope of this paper. An n-dimensional approximation has then been constructed by discretizing the z-variable in the form

$$h := L/n; \quad z_i := (i-1)h; \quad i = 1, \ldots, n; \tag{80}$$

next by defining n state variables x_i (or equivalently a vector state variable $x(\cdot)$ with values $x(t)$ in $\mathbb{R}^n$),

$$x_i(t) := \Delta \theta(t, z_i); \quad i = 1, \ldots, n; \tag{81}$$

$$x(t) := (x_1(t), x_2(t), \ldots, x_n(t))'; \tag{82}$$

and finally by approximating the z-partial derivative by some appropriate linear combination of the function $\Delta \theta(t, \cdot)$ evaluated at the discretized values z_i, for instance,

$$\frac{\partial \Delta \theta}{\partial z}(t, z_i) \approx \frac{x_{i+1}(t) - x_i(t)}{h}; \quad i = 1, \ldots, n-1; \tag{83}$$

$$\frac{\partial \Delta \theta}{\partial z}(t, z_n) \approx \frac{x_n(t) - x_{n-1}(t)}{h}. \tag{84}$$

After such manipulations the following structure of a linear control system is obtained

$$\dot{x} = Ax + Bu, \tag{85}$$

where the $n \times n$ matrix A and the column n-vector B take the form

$$A = (a_{ij}): \begin{cases} a_{ii} = V_0/h - a; & a_{i,i+1} = -V_0/h; \quad i = 1, \ldots, n-1 \\ a_{n,n-1} = V_0/h; & a_{nn} = -(a + V_0/h) \\ \text{all remaining elements equal to 0} \end{cases}, \tag{86}$$

$$B = (b_i) = \left[\frac{a}{V_0}(\theta_0 - \theta_a) \exp(-az_i/V_0); \quad i = 1, \ldots, n \right]. \tag{87}$$

The eigenvalues of the matrix A are dominated by the relation between the heat gained at each position by convection versus the heat extracted at that point by the environment, implicit in the term $V_0/h - a$ which appears in the main diagonal, except in its last element. The free evolution has to be stable to keep any physical meaning in the equations (the temperature can not grow forever). But, if control has to be relevant to increase stability, it is appropriate

to explore those situations near where the system might lose stability (for instance, due to environmental perturbations). With this contradictory objectives in mind, the following values for the parameters were investigated

$$V_0 = h = 1; \quad a = 1.001. \tag{88}$$

The discretized, ODE version (85) of Eq. (77) was numerically confirmed to be an acceptable approximation (Costanza and Rivadeneira 2014a). The initial state $x_0 = x(0)$ used for simulation of the system defined by Eqs. (85)–(87) was

$$x_i(0) = 100; \quad i = 1, \ldots, n, \tag{89}$$
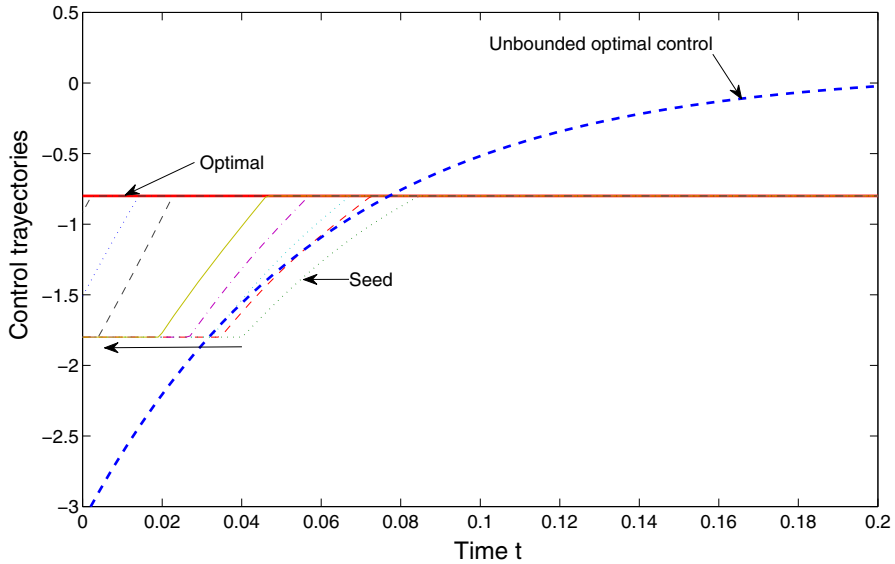
Fig. 7 Evolution of the control trajectories with the iterations until reaching the optimal control

Table 1 Seed and final values for parameters

Component	ρ_{seed}	ρ^*	μ_{seed}	μ^*
1	$-234,223$	$245,168$	$-1402,842$	$-631,500$
2	$317,068$	$-127,917$	$634,734$	$940,896$
3	$51,967$	$-265,032$	$1,383,567$	$1,518,771$
4	$594,135$	$-315,422$	$1,658,772$	$1,731,146$
5	$621,500$	$-333,940$	$1,759,913$	$1,809,198$
6	$631,557$	$-340,743$	$1,797,081$	$1,837,874$
7	$635,253$	$-343,274$	$1,810,772$	$1,848,498$
8	$636,609$	$-343,959$	$1,815,561$	$1,851,761$
9	$637,122$	$-34,545$	$1,818,557$	$1,856,167$
10	$637,254$	$-343,012$	$1,816,606$	$1,850,087$

with the following values for the reference temperatures (in °C):

$$\theta_a = 20, \quad \theta_0 = 700. \tag{90}$$

The cost objective of the LQR problem was assigned the following parameters

$$t_f = 0.5, \quad Q = 0.05 I_n, \quad R = 100, \quad S = 15 I_n, \tag{91}$$

and $n = 10$. The bounds imposed on control values were chosen here as

$$[u_{\min}, u_{\max}] = [-1.8, -0.8], \tag{92}$$

just to illustrate a case where the costs of the unrestricted problem and the first approximation to the restricted one (the 'seed') are different enough to justify looking for better approximations and to obtain a significant cost reduction through the implementation of the method. Indeed, the seed trajectory was calculated by applying Eq. (28), from which the initial switching times were detected: $\tau_{1,0} = 0.0413$, $\tau_{2,0} = 0.0865$, and the total cost resulted $J_{\text{seed}} = 1.1316 \times 10^6$, while the unbounded cost was $J_{\text{unbounded}} = 4.3338 \times 10^5$.

In Fig. 7, the seed control trajectory (dotted line) and its evolution towards the optimal control (solid line) are illustrated. In this example, the optimal control resulted $u^*(t) \equiv -0.8 = u_{\max} \ \forall \ t$. The optimal cost was retrieved after 25 iterations and its value was $J^* = 9.1010 \times 10^5$. The relative reduction of the total cost J^* with respect to the cost of the seed strategy J_{seed} was 20%. The seed and final values of ρ, μ obtained by the method are shown in Table 1.

The optimality of the obtained control trajectory was checked by calculating: (i) the solution to the Hamiltonian equations (9), backwardly from (ρ^*, μ^*), which reproduced the initial state condition, and (ii) the Hamiltonian along the trajectory, which resulted in $H(x^*(t), \lambda^*(t), u^*(t)) \equiv 2.2191 \times 10^6$, constant as expected.
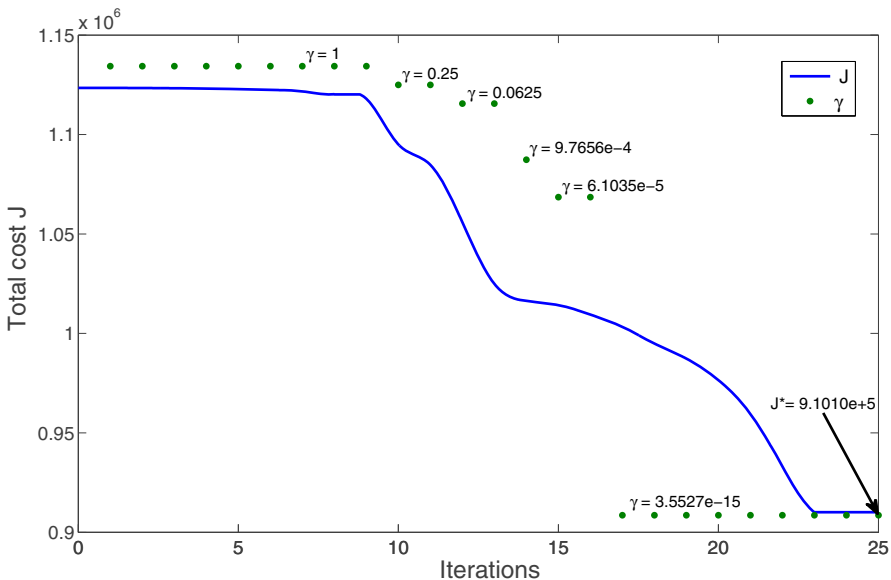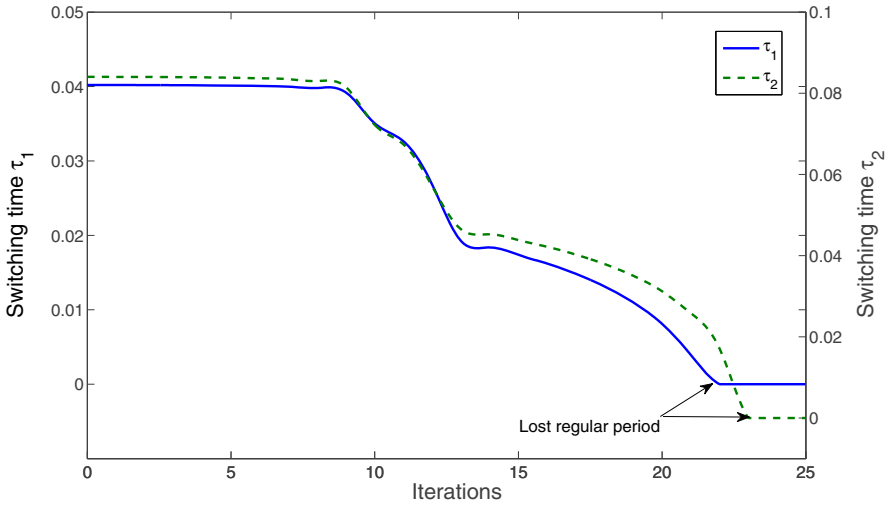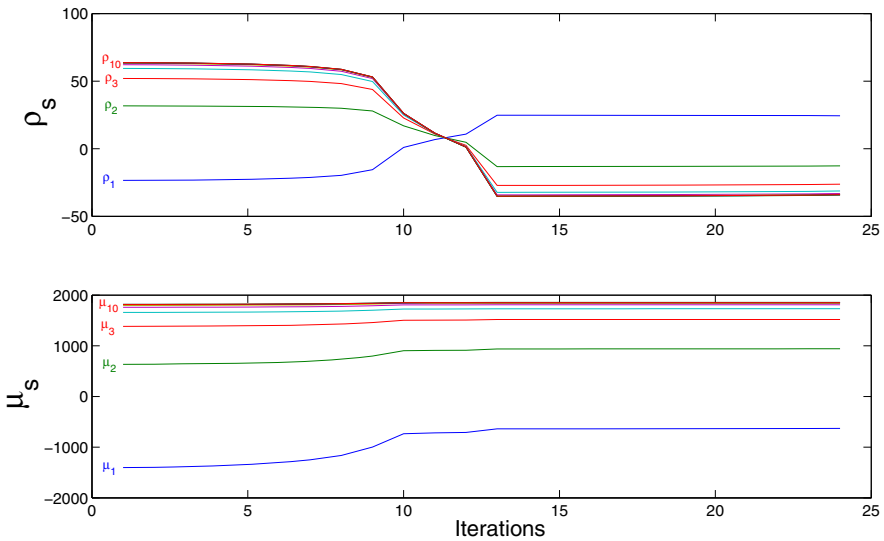

Fig. 8 Evolution of the total cost when applying the conjugate gradient method

Fig. 9 Evolution of the τ_s when applying the conjugate gradient method

Fig. 10 Evolution of the ρ and μ when applying the conjugate gradient method

For this problem of moderately high dimension, a comparison of CPU time was performed: the first try by updating the 55 coefficients of the matrix S, and the second one by updating the 20 coefficients of ρ, μ. As it was expected, the first experiment takes 458.95 s per iteration while the second one only 151.79 s. This means that the second method is 200% faster than the first one.

The parameter γ was changed to keep lowering the total cost. The initially value was $\gamma = 1$, and the final one was $\gamma = 1 \times 10^{-15}$, as it is marked in Fig. 8. The evolution of the variables τ_1, τ_2 are depicted in Fig. 9. All components of ρ, μ converge to their optimal values and their evolution can be seen in Fig. 10.

5 Conclusions and perspectives

Although feedback laws may be preferred in practice, when perturbations are expected to appear, a closed-loop control is in general suboptimal when there exist constraints in the manipulated variable. With these limitations in mind, an efficient algorithm was devised to approximate the open-loop optimal control via feedback, based on recent and new theoretical results. The resulting strategies are quite different from the saturated form of the optimal control corresponding to the unrestricted problem having the same parameters and initial condition. This saturated form is used here just as a first approximation, and so labeled as a 'seed' strategy.

In the paper it was assumed that there existed a unique regular interval $\mathcal{I} = (\tau_1, \tau_2) \subset (0, t_f)$. A systematic treatment of the case with a countable number of regular arcs is under development along similar lines. The next target would be to devise a numerical scheme able to be implemented on-line.

The updating of the parameters ρ, μ and the saturation times τ_i, as long as the total cost is reduced via algebraic formula, is the central idea. The stability of the method is guaranteed since (i) the cost is not allowed to increase, and (ii) it is bounded from below. Enhancing the speed is changing the value of the manipulated variable is probably the main issue of concern in the control-computational world, mainly due to real-time requirements. The increasing number of the variables to update in greater dimensions generates a need of calculation that can exceed current performance. In response to these necessities, the positive features of the new on-line proposed strategy are:

- The method is based on theoretical results ensuring that the hidden values ρ, μ and the appropriate saturation times τ_1, τ_2 are the critical objects to be ascertained.
- Calculating the final penalization matrix in terms of ρ and μ reduces the total number of components to update. With the method proposed in Costanza and Rivadeneira (2014a), the minimal number of variables to update is $\frac{n(n+3)}{2} + 2$, since $\tilde{S}$ is a symmetric matrix. By using the new strategy the total number of variables are reduced to $2n+2$, a significant reduction for high dimensional systems.
- The method takes advantage on the availability of α, β as functions of $(t_f - t, \tilde{S})$, and consequently on the possibility of generating Riccati matrices $P(t, t_f, \rho, \mu)$ on-line by simple algebraic manipulations, as ρ, μ are updated; i.e., the RDE does not need to be solved for any value of ρ, μ.
- The updating of parameters $(\rho, \mu, \tau_1, \tau_2)$ is performed via the gradient of the cost of the process, and this cost is calculated by simple algebraic formulas instead of by calculating state, control, and cost trajectories by ODE integrations.

References

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

Athans M, Falb PL (2006) Optimal control: an introduction to the theory and its applications. Dover, New York

Bernhard P (1972) Introducción a la Teoría de Control Óptimo. In: Cuaderno No. 4, Instituto de Matemática "Beppo Levi" , Rosario

Bramanti E, Bramanti M, Stiavetti P, Benedetti E (1994) A frequency deconvolution procedure using a conjugate gradient minimization method with suitable constraints. J Chemom 8:409–421

Costanza V (2007) Finding initial costates in finite-horizon nonlinear-quadratic optimal control problems. Optimal Control Appl Methods 29:225–242

Costanza V, Neuman CE (2006) Optimal control of nonlinear chemical reactors via an initial-value Hamiltonian problem. Optimal 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 (2008) Finite-horizon dynamic optimization of nonlinear systems in real time. Automatica 44:2427–2434

Costanza V, Rivadeneira PS (2013) Optimal saturated feedback laws for LQR problems with bounded controls. Comput Appl Math 32:355–371

Costanza V, Rivadeneira PS (2014) Online suboptimal control of linearized models. Syst Sci Control Eng 2:379–388

Costanza V, Rivadeneira PS (2014) Enfoque Hamiltoniano al control óptimo de sistemas dinámicos. OmniScriptum, Saarbrücken

Costanza V, Rivadeneira PS, González AH (2013) Minimizing control-energy in a class of bounded-control LQR problems. Optimal Control Appl Methods 35(3):361–382

Costanza V, Rivadeneira PS, Spies RD (2009) 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

Fletcher R, Reeves CM (1964) Function minimization for conjugate gradients. Comput J 7(2):149–154

Goebel R, Subbotin M (2007) Continuous time linear quadratic regulator with control constraints via convex duality. IEEE Trans Autom Control 52(5):886–892

Hearns G, Grimble MJ (2010) Temperature control in transport delay systems. In: The 2010 American control conference, Baltimore

Kojima A, Morari M (2004) LQ control of constrained continuous-time systems. Automatica 40:1143–1155

Pannocchia G, Rawlings JB, Mayne DQ, Mancuso G (2015) Whither discrete time model predictive control? IEEE Trans Autom Control 60(1):246–252

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

Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mishchenko EF (1964) The mathematical theory of optimal processes. Macmillan, New York

Sakizlis V, Perkins JD, Pistikopoulos EN (2005) Explicit solutions to optimal contol problems for contrained continuous-time linear systems. IEE Proc Control Theory Appl 152(4):443–452

Sontag ED (1998) Mathematical control theory, 2nd edn. 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

Van Loan CF (1978) Computing integrals involving the matrix exponential. IEEE Trans Autom Control 23(3):395–404