

A convexity result for the optimal control of a class of positive nonlinear systems^{*}

Franco Blanchini^{*}, Patrizio Colaneri^{**}
Richard H. Middleton^{***}

^{*} *Dipartimento di Matematica e Informatica, Università di Udine,
33100 Udine, Italy (e-mail blanchini@uniud.it)*

^{**} *DEI, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133
Milano, Italy (colaneri@elet.polimi.it)*

^{***} *School of Electrical and Computer Engineering, University of
Newcastle, Callaghan, New South Wales, 2308, Australia (email
richard.middleton@newcastle.edu.au)*

Abstract

In this paper a class of input-parametrized bilinear positive systems is considered. This class is characterized by the fact that the input variables affect only the diagonal entries of the dynamical matrix. The class of systems considered is relevant to a variety of dynamical models arising in system biology and compartmental systems. Given a final time t_f , it is proven that the any component of the state vector at t_f is a convex functional of the input variables. If a linear cost of the final state is considered this result has the important consequence that any Pontryagin solution of the associated optimal control problem is optimal and can be numerically computed by using standard gradient-type algorithms. A few extensions are given and an example is provided to illustrate the theory.

1. INTRODUCTION

This note deals with optimal control for a class of input-parametrized linear positive systems. The main motivation lies in the observation that in many biological models the input represents a scheduling function that acts on the system parameters and has to be designed to minimize a cost representing the infection over a certain interval of time. The study of optimal control problems for such systems may then help inform application areas such as infection treatment protocols. For instance, in the recent work [Hernandez-Vargas et al. \(2013\)](#), a certain class of simplified HIV treatment switched models are considered in order to prolong the viral rebound and mitigate the infection effects. This class of problems also arise in certain types of epidemiological models, for example in SI (susceptible/infection) models (e.g. [Rami et al. \(2013\)](#); [Moreno et al. \(2002\)](#)), at small initial infection stage.

The class of systems considered is related to that of switched/hybrid systems, for which the optimal control problem has been widely studied [Cassandras et al. \(2001\)](#), [Dmitruk and Kaganovich \(2011\)](#) and the variational approach has been developed by [Rapoport \(1996\)](#) and [Margaliot \(2006\)](#). The optimal control problem is inherently nonlinear and therefore non-convex in general. There is therefore no guarantee that solutions satisfying the Pontryagin maximum principle provide globally optimal solutions. Moreover, finding such solutions is in general

a formidable task, since it requires backward and forward iterations of the state-costate dynamics with no guarantee of convergence.

Convexity of the cost with respect to the control function is therefore a very strong property that ensures optimality of the Pontryagin solutions and allows the use of simple interior point algorithms to find an optimal solution.

In this note we prove that, for an important class of input-parametrized positive systems, the cost functional is indeed convex with respect to the control variables.

The paper is organized as follows. The class of systems and the associated optimal control problem are defined in Section 2. The main result on convexity is introduced in Section 3 and the proof of the main theorem is given in Section 4, where also some hints on the possible algorithms are given together with possible important extensions on the class of considered cost functions that preserve convexity. Section 5 includes an example taken from [Rami et al. \(2013\)](#) and dealing with a epidemiological model with small initial infection status.

We use the following notation. The semiring of nonnegative n -tuples of real numbers is \mathbb{R}_+^n . A square matrix $A = [a_{ij}]$ is *Metzler* if its off-diagonal entries are nonnegative, $a_{ij} \geq 0$ for every $i \neq j$. The symbol $\mathbf{1}$ denotes a vector with all entries equal to 1. For details on positive systems the reader is referred to [Farina and Rinaldi \(2000\)](#).

2. OPTIMAL CONTROL PROBLEM

Consider a nonlinear system described by

$$\dot{x} = A(u)x = (M + \Lambda(u))x, \quad x(0) = x_0 \quad (1)$$

^{*} Part of this paper was presented in the IFAC NOLCOS 2010. Supported by Science Foundation Ireland 07/IN/I1838. This work was performed in part whilst the authors were at the Hamilton Institute, The National University of Ireland, Maynooth. Corresponding author R. H. Middleton.

where $x(t)$ is the n -dimensional state vector, M is Metzler, $\Lambda(u)$ is an $n \times n$ diagonal matrix of convex and continuous functions of u and $u(t)$, for each t , is a m -dimensional control variable belonging to the polytope

$$\mathcal{U} \doteq \{u \in \mathbb{R}_+^m : \mathbf{1}'u = 1\}.$$

We assume further that u , as a function of time, belongs to the functional space \mathcal{U}_s of piecewise continuous functions taking values in \mathcal{U} , in the interval $[0, t_f]$ where $t_f > 0$ is a given final time instant.

Since $A(u)$ is Metzler for any u , \mathbb{R}_+^n is positively invariant for any $u \in \mathcal{U}_s$, i.e. the system is *positive*.

Remark 1. The class of systems described by (1) includes linear systems in a polytope when

$$\Lambda(u) = \sum_{i=1}^M \Lambda_i u_i \quad (2)$$

with $u \in \mathcal{U}_s$. This system can be considered to be the embedding of a linear switched system. In this case $\Lambda(u) = \Lambda_\sigma$ where the switching signal σ is a left continuous piecewise constant signal taking values in the finite set $\mathcal{M} = \{1, 2, \dots, M\}$ so that $u_i(t) = 1$ if $\sigma(t) = i$ and $u_i(t) = 0$ otherwise.

For instance, such a system can represent a simplified model of treatment of HIV infection dynamics (see for example Hernandez-Vargas et al. (2013)), where x represents the concentrations of various viral mutants in a patient, and σ represents the selection of a suitable therapy. Other examples can be found in the widespread SIR (or SI, or SIRS) models of epidemiology over a network, in the initial infection phase (see for example Rami et al. (2013); Moreno et al. (2002)). The embedded system is important in the context of optimal control, since it is capable of capturing possible singular arcs corresponding to sliding trajectories of the system state.

We now considered an optimal control problem for system (1), with terminal cost functional:

$$J(u, x_0) := c'x(t_f) \quad (3)$$

where c is a nonnegative vector. The problem is to find $u^\circ \in \mathcal{U}_s$ that minimizes the cost, i.e. $\min_{u \in \mathcal{U}_s} J(u, x_0)$.

In the literature on optimal control, direct sufficient conditions (for instance associated with Hamilton-Jacobi-Bellman equations) are often impractical. Therefore, many results give necessary conditions for optimality using the Pontryagin principle, as a starting point for seeking optimal solutions. We therefore define a Pontryagin solution as a candidate optimal control solution satisfying the necessary conditions. For further details see e.g. Bressan and Piccoli (2007).

Definition 1. A triple $u^\circ(t) : [0, t_f] \times \mathcal{U}$, $x^\circ(t)$, $\pi^\circ(t)$, that satisfies (for almost all t) the system of equations:

$$\dot{x}^\circ(t) = A(u^\circ(t))x^\circ(t) \quad (4)$$

$$-\dot{\pi}^\circ(t) = A(u^\circ(t))'\pi^\circ(t) \quad (5)$$

$$u^\circ(t) \in \operatorname{argmin}_{u \in \mathcal{U}} \{\pi^{\circ\prime}(t)A(u)x^\circ(t)\} \quad (6)$$

with boundary conditions $x^\circ(0) = x_0$, $\pi^\circ(t_f) = c$, is called a *Pontryagin solution* for the optimal control problem.

3. CONVEXITY

We are now ready to state the main result of the paper, namely the convexity of the map $u \in \mathcal{U}_s \rightarrow J(u, x_0) \in \mathbb{R}_+$.

Theorem 1. Consider system (1) and assume that M is a Metzler matrix, $\Lambda(u)$ is a diagonal matrix composed by convex functions of $u \in \mathcal{U}$, and consider the cost function (3), where c is a nonnegative vector. Then, the functional $u \rightarrow c'x(t_f)$ from \mathcal{U}_s to \mathbb{R}_+ is convex.

The proof of the above result will be given next. However, to stress its importance, the following theorem stating the optimality of a Pontryagin solution associated with the optimal control problem is provided.

Theorem 2. Consider (1) and assume that M Metzler, $\Lambda(u)$ is a diagonal matrix composed by convex functions of $u \in \mathcal{U}$, and consider the cost function (3), where c is a nonnegative vector. Then the optimal control problem admits at least one Pontryagin solution $(u^\circ, x^\circ, \pi^\circ)$ and u° is a global optimal control signal relative to x_0 . Moreover, the value of the optimal cost functional is $\pi^{\circ\prime}(0)x_0$.

Proof Note first that the optimal control for system (1) and cost (3) always exists. Indeed, see e.g. Theorem 5.1.1 in Bressan and Piccoli (2007), a sufficient condition is that the sets of velocities $F(x, u) := \{A(u)x; u \in \mathcal{U}\}$ are convex and that the vector field is bounded by an affine function of the norm of the state variable, i.e. $\|(M + \Lambda(u))x\| \leq \alpha(1 + \|x\|)$ for some positive scalar α , and for all $x \in \mathbb{R}_+^n$ and $u \in \mathcal{U}$. These conditions are of course satisfied in our case. Let an optimal triple be $x^\circ, u^\circ, \pi^\circ$. This triple is a Pontryagin solution, as defined in Definition 1. Indeed, the Hamiltonian function associated with system (1) and the linear cost (3) is

$$H(x, u, \pi) = \pi(t)'A(u)x(t)$$

and $\dot{\pi}(t) = -(\frac{\partial H}{\partial x})' = -A(u)'\pi(t)$, $\dot{x}(t) = (\frac{\partial H}{\partial \pi})' = A(u)x(t)$, with $\pi(t_f) = c$ and $x(0) = x_0$. The transversal conditions are satisfied and for all $u \in \mathcal{U}$:

$$H(x^\circ, u^\circ, \pi^\circ) \leq H(x^\circ, u, \pi^\circ).$$

In view of the Pontryagin principle, the triple $(x^\circ, \pi^\circ, u^\circ)$ satisfies the necessary conditions for optimality. Theorem 1 states the convexity of the cost with respect to $u \in \mathcal{U}_s$. This fact is sufficient, see (Bressan and Piccoli, 2007, Theorem 5.1.1), to guarantee optimality. \square

4. PROOF OF THEOREM 1

Here we prove Theorem 1. Given $u \in \mathcal{U}_s$, system (1) become a linear time-varying positive system described by

$$\dot{x}(t) = (M + \Lambda(u(t)))x(t) \quad (7)$$

Let $\Phi(u, t, t_0)$ be the transition matrix of $M + \Lambda(u(t))$, i.e.

$$\frac{d}{dt}\Phi(u, t, t_0) = (M + \Lambda(u(t)))\Phi(u, t, t_0), \quad \Phi(u, t_0, t_0) = I$$

Given $t_f > 0$, and a positive vector c rewrite the cost as

$$J(u, x_0) = c'\Phi(u, t_f, t_0)x_0$$

We now prove that

- (i) the cost function $J(u, x_0)$ from \mathcal{U} into \mathbb{R}_+ is convex
- (ii) the cost functional $u \rightarrow J(u, x_0)$ from $\bar{\mathcal{U}}_s$ into \mathbb{R}_+ is convex, where $\bar{\mathcal{U}}_s = \{\Lambda : [0, t_f] \rightarrow \mathcal{L}\}$ is the set of piecewise constant functions in \mathcal{U} .

(iii) the cost functional $u \rightarrow J(u, x_0)$ from \mathcal{U}_s into \mathbb{R}_+ is convex (Theorem 1).

Let us start by stating two technical lemmas.

Lemma 1. Consider a scalar function $f : \mathcal{U} \rightarrow \mathbb{R}_+$ defined as

$$f(u) = e^{w'z(u)}$$

where $w \in \mathbb{R}_+^n$ and assume that the entries of vector $z(u)$ are twice differentiable convex functions of $u \in \mathcal{U}$. Then the function $f(u)$ is convex.

Proof. The gradient of $f(u)$ is $\frac{\partial f(u)}{\partial u} = f(u)w'Z_u$ and the Hessian is $\frac{\partial^2 f(u)}{\partial^2 u} = f(u)Z_u'w w'Z_u + \sum_{i=1}^n w_i \frac{\partial^2 z_i(u)}{\partial^2 u}$, where Z_u' is the gradient of $z(u)$. The vector w is by assumption non-negative i.e. $w_i \geq 0, i = 1, 2, \dots, n$. Moreover, the Hessian matrices $\frac{\partial^2 z_i(u)}{\partial^2 u}, i = 1, 2, \dots, n$ are positive semidefinite, since we assume $z(u)$ is a convex function of u . Therefore $\frac{\partial^2 f(u)}{\partial^2 u}$ is positive semidefinite, and therefore $f(u)$ is convex. \square

Lemma 2. Let $f_k(u), k = 1, 2, \dots, p$ be a sequence of convex functions on a convex domain, and assume that the sequence point-wise converges to a function $f(u)$. Then $f(u)$ is convex.

Proof. Assume by contradiction that $f(u)$ is not convex, that is, there exist two points u_1 and u_2 and $0 < \alpha < 1$ such that, denoting by $u = \alpha u_1 + (1 - \alpha)u_2$, we have

$$f(u) > \alpha f(u_1) + (1 - \alpha)f(u_2) \quad (8)$$

On the other hand, from the convexity assumption on f_k ,

$$-f_k(u) + \alpha f_k(u_1) + (1 - \alpha)f_k(u_2) \geq 0 \quad (9)$$

for all k . Taking the limit, we have

$$f(u) \leq \alpha f(u_1) + (1 - \alpha)f(u_2)$$

which contradicts (8).

4.1 (i): Convexity of the cost in \mathcal{U}

We first consider the case of a constant u , i.e. point (i) above. Note first that for a constant u the transition matrix is $\Phi(u, t_f, 0) = e^{(M+\Lambda(u))t_f}$, so that the cost is $J(u, x_0) = c'e^{(M+\Lambda(u))t_f}x_0$.

Lemma 3. For any non-negative c and x_0 , the function $f(u) : \mathcal{U} \rightarrow \mathbb{R}_+$ defined by

$$f(u) = J(u, x_0) = c'e^{(M+\Lambda(u))t_f}x_0 \quad (10)$$

is convex in u .

Proof. We first note that $f(u)$ in (10) is a well defined continuous function of u . Then recall a useful formula for the exponential of the sum of two matrices (Cohen (1981)):

$$e^{(M+\Lambda(u))t_f} = \lim_{k \rightarrow \infty} \left(e^{\frac{Mt_f}{k}} e^{\frac{\Lambda(u)t_f}{k}} \right)^k$$

Therefore, defining the functions

$$f_k(u) = c' \left(e^{\frac{Mt_f}{k}} e^{\frac{\Lambda(u)t_f}{k}} \right)^k x_0$$

we have

$$f_k(u) \rightarrow f(u)$$

Let us consider the generic function $f_k(u)$. Since $e^{\frac{Mt_f}{k}}$ is a nonnegative matrix and $e^{\frac{\Lambda(u)t_f}{k}}$ is a diagonal matrix with

elements $\xi_i \doteq e^{(\lambda_i(u)t_f/k)}$ we have that $f_k(u)$ is a positive polynomial in the variables ξ_i . Formally

$$f_k(u) = \sum_{k_1 + \dots + k_N = k} \left(\alpha_{k_1, k_2, \dots, k_N} \xi_1^{k_1} \xi_2^{k_2} \dots \xi_N^{k_N} \right)$$

with $\alpha_{\bullet} \geq 0$ formed from the appropriate sums of products of elements of $c, e^{\frac{Mt_f}{k}}$, and x_0 . On the other hand if we replace $\xi_i = e^{(\lambda_i(u)t_f/k)}$ we get that each monomial satisfies

$$\xi_1^{k_1} \xi_2^{k_2} \dots \xi_N^{k_N} = e^{\sum_{i=1}^n \lambda_i(u)k_i t_f/k} = e^{w'z(u)}$$

where $w' = [k_1 t_f/k \ k_2 t_f/k \ \dots \ k_N t_f/k]$ and $z(u)' = [\lambda_1(u) \ \lambda_2(u) \ \dots \ \lambda_n(u)]$. Since w is nonnegative and $\lambda_i(u)$ are convex functions of u in view of Lemma 1 we conclude that $f_k(u)$ is convex for all k . The proof then follows from the fact that $f_k(u) \rightarrow f(u) = J(u, x_0)$, and hence $J(u, x_0)$ is a convex function of u . \square

4.2 (ii): Convexity of the cost in $\bar{\mathcal{U}}_s$

Lemma 4. Let $u(t)$ be a piecewise constant function of t

$$u(t) = u^{[i]}, \quad t_{i-1} \leq t < t_i = t_{i-1} + T_i$$

with $i = 1, 2, \dots, K$. Then

$$J(u, x_0) = c^\top \prod_i e^{(M+\Lambda(u^{[i]}))T_i} x_0 \quad (11)$$

is convex in the values $u^{[i]}$.

Proof. The proof follows similar lines of argument to the proof of Lemma 3. We approximate each exponential

$$e^{(M+\Lambda(u^{[i]}))T_i} \approx \left(e^{\frac{MT_i}{k}} e^{\frac{\Lambda(u^{[i]})T_i}{k}} \right)^k$$

as before and we notice that we get a polynomial with positive coefficients in the unknowns $\xi_{i,j} = e^{\lambda_{i,j} T_i/k}$, where $\lambda_{i,j}, j = 1, 2, \dots, n$ are the elements on the diagonal of $\Lambda(u^{[i]})$. This polynomial is convex and hence the limit function is convex as well. \square

Remark 2. The convexity results have been presented in the case where the intervals $[t_{k-1}, t_k]$ are common to all functions. This is without loss of generality, since if each function of $u(t)$ has its own interval partition, we consider the “intersection” of these intervals.

4.3 Point (iii): Convexity of the cost in $\bar{\mathcal{U}}_s$

We now are in a position to prove the main result, Theorem 1. Given any piecewise continuous function of time $f(t)$ in the interval $[0, t_f]$, there exists a sequence of piecewise-constant functions f_k which converge to f in the integral norm

$$\int_0^{t_f} |f_k(t) - f(t)| dt \rightarrow 0$$

As a consequence $c^\top \Phi(u, t_f, t_0)x_0$ can be achieved as the limit of functions of the form (11). Repeating the convexity and limit argument, by contradiction assume that

$$u(t) = \alpha u^{\{1\}}(t) + (1 - \alpha)u^{\{2\}}(t)$$

and that

$$J(u, x_0) > \alpha J(u^{\{1\}}, x_0) + (1 - \alpha)J(u^{\{2\}}, x_0)$$

for some $0 < \alpha < 1$. On the other hand, the function u can be seen as the limit of a sequence of piecewise constant functions $\bar{u}(k)$, and the same for $u^{\{1\}}$ and $u^{\{2\}}$. Thanks to convexity with respect to \bar{U}_s , the opposite inequality is satisfied, i.e.

$$J(\bar{u}(k), x_0) \leq \alpha c^\top J(u^{\{1\}}(k), x_0) + (1 - \alpha)J(\bar{u}^{\{2\}}(k), x_0)$$

This leads to a contradiction.

4.4 Algorithms

We end this section by providing a computational scheme for the optimal control problem. The convexity property allows using different types of algorithms to find the solution of

$$\min_{u \in \mathcal{U}_s} J(u), \quad J(u) = c'x(t_f)$$

Computations can be cast in discrete-time, by taking, as done in the previous section, a subdivision of the interval $[0, t_f]$ into N intervals T_1, T_2, \dots, T_N . The control variable may be approximated as piecewise constant, i.e. (with $T_0 = 0, T_N = t_f$):

$$u(t) = \bar{u}^{[k]}, \quad t \in \left[\sum_{i=1}^{k-1} T_i, \sum_{i=1}^k T_i + T_k \right)$$

The discretized control is denoted by \bar{u} taking values in $\bar{\mathcal{U}}$, the cartesian products of \mathcal{U} . Hence the problem is to find

$$\min_{\bar{u} \in \bar{\mathcal{U}}} J(\bar{u}, x_0) = \min_{\bar{u} \in \bar{\mathcal{U}}} c' \prod_{i=N}^1 e^{(M+\Lambda(\bar{u}^{[i]})T_i)x_0} \quad (12)$$

The constrained optimization problem can be solved using the standard Matlab function `fmincon.m` which is based on an interior point method. Notice indeed that $\bar{\mathcal{U}}$ is a convex set and that $J(\bar{u}, x_0)$ is a convex function of \bar{u} .

To further assist the convex minimisation, note that we can explicitly compute the gradient of the cost. The gradient of $J(\bar{u}, x_0)$ is a mN -dimensional row vector and the j^{th} m -dimensional row entry $g^{[j]}(k)$ can be computed from (12). Indeed a simple computation shows that the generic r -th entry $g_r^{[j]}(k)$ of $g^{[j]}$ is

$$g_r^{[j]}(k) = c' \Psi_r^{[j]} \left(\Phi_r^{[j]}|_{\bar{u}(k)} \right) \Gamma_r^{[j]} x_0$$

where

$$\Psi_r^{[j]} = \prod_{i=N}^{j+1} e^{(M+\Lambda(\bar{u}^{[i]})T_i), \quad \Gamma_r^{[j]} = \prod_{i=j-1}^1 e^{(M+\Lambda(\bar{u}^{[i]})T_i}$$

and

$$\begin{aligned} \Phi_r^{[j]} &= \frac{\partial e^{(M+\Lambda(\bar{u}^{[j]})T_j}}{\partial \bar{u}_r^{[j]}} \\ &= \int_0^{T_j} e^{(M+\Lambda(\bar{u}^{[j]}))(t-\tau)} \frac{\partial \Lambda(\bar{u}^{[j]})}{\partial \bar{u}_r^{[j]}} e^{(M+\Lambda(\bar{u}^{[j]})\tau} d\tau \\ &= (0 \ I) e \begin{pmatrix} M + \Lambda(\bar{u}^{[j]}) & 0 \\ \frac{\partial \Lambda(\bar{u}^{[j]})}{\partial \bar{u}_r^{[j]}} & M + \Lambda(\bar{u}^{[j]}) \end{pmatrix} T_j \begin{pmatrix} I \\ 0 \end{pmatrix} \end{aligned}$$

Notice that the algorithm can be further enhanced by explicit computation of the Hessian matrix.

As a last observation, notice that the optimal cost $J(u^o, x_0)$ is a concave function of x_0 , see for instance the recent papers [Hernandez-Vargas et al. \(2011\)](#) and [Blanchini et al. \(2012\)](#). Then taking $x_0 \in \mathcal{A}$ for some predefined set of initial states, it may also be of interest to find a saddle point solution of the min-max problem

$$\min_{u \in \mathcal{U}_s} \max_{x_0 \in \mathcal{A}} J(x_0, u)$$

i.e. a solution u^*, x_0^* such that $J(x_0, u^*) \leq J(x_0^*, u^*) \leq J(x_0^*, u)$ for any $x_0 \in \mathcal{A}$, where $\mathcal{A} = \{x_0 \in \mathbb{R}_+^n : \mathbf{1}_n x_0 = 1\}$, and any $u \in \mathcal{U}_s$. In this regard, taking again the above discretization of u , we are able to write the computational scheme

$$\bar{u}^{[k+1]} = \text{Proj}_{\bar{\mathcal{U}}} \left(\bar{u}^{[k]} - \alpha g^{[k]} \right) \quad (13)$$

$$\bar{x}_0^{[k+1]} = \text{Proj}_{\mathcal{A}} \left(\bar{x}_0^{[k]} + \alpha h^{[k]} \right) \quad (14)$$

where $\text{Proj}_{\mathcal{A}}$ is the projection on \mathcal{A} , $h^{[k]}$ is the gradient of $J(x_0, u)$ with respect to x_0 at the k -th iteration. The vector h can be easily computed (thanks to the linearity of $J(x_0, u)$ with respect to x_0 , as $h = c' \prod_{i=N}^1 e^{(M+\Lambda(\bar{u}^{[i]})T_i}$.

4.5 Extensions

The proof of convexity has been carried out by looking at a problem where the cost is a generic linear combination of the final state. However we know that any convex and nondecreasing function of a convex function is convex, and hence a similar analysis can be carried out considering such a cost. Linearity of the cost with respect to x_f and linearity of the system for any given control function also implies linearity with respect to the initial state, so implying concavity of the optimal cost with respect to x_0 . Notice that to preserve concavity it is possible to consider a cost that is a concave and nonincreasing function of the state.

Note also that the cost functional can be readily extended to an integral cost of the form

$$J(u, x_0) = c'x(t_f) + \int_0^{t_f} d'x(t)dt \quad (15)$$

where d is a nonnegative vector. Indeed the former problem (1) and cost (3) are recovered by augmenting the state as follows. Let $\xi = [x' \ \eta]'$, with $\eta = \int_0^t d'x(\tau)d\tau$ and $J(u, x_0) = [c' \ 1] \xi(t_f)$.

Following a similar rationale we can also establish an extension of the presented theory for systems affected by a constant input, i.e. $\dot{x} = A(u)x + b$, where b is a nonnegative vector. This can be achieved again by state augmentation, i.e. $\xi = [x' \ b]'$ so that $J(u, x_0) = [c' \ 0] \xi(t_f)$.

Finally, notice that the convexity property also holds when the cost is enriched by an additional term that depends on the control variable $u \in \mathcal{U}_s$, say $h(u)$, provided that h is a convex functional of u .

Therefore, our convexity properties holds when the cost is

$$J(x_0, u) = c'x(t_f) + \int_0^{t_f} d'x(t)dt + h(u)$$

under the assumption the the system structure $\dot{x} = (M + \Lambda(u))x$ is such that $\Lambda(u)$ is a diagonal matrix with convex

function of u , M is a nonnegative matrix and h is a convex functional of u .

5. EXAMPLE

The example is taken from [Rami et al. \(2013\)](#) and concerns the epidemiological model of a population divided into n groups. Each group is divided into two classes, i.e. $I_i(t)$ infectives and $S_i(t)$ susceptibles. Under the assumption that the total number $I_i(t) + S_i(t) = N_i$ is constant and letting $x_i(t) = I_i(t)/N_i$ one can write, for $i = 1, 2, \dots, n$:

$$\dot{x}_i(t) = (1 - x_i(t)) \sum_{j=1}^N \frac{\beta_{ij} N_j}{N_i} x_j(t) - (\gamma_i + \mu_i) x_i(t) \quad (16)$$

where β_{ij} is the rate at which susceptibles in group i are infected by infectives in group j , γ_i is the rate at which an infective individual in group i is cured and μ_i the birth and death rates in group i (assumed to be equal since the number of the total population in the same group is constant). Note that the set $0 \leq x_i \leq 1$, $i = 1, 2, \dots, 4$, is positively invariant.

We now assume that, for each group i , m different cures are possible, so that the rate γ_i depends on an additional index, say $\sigma = \{1, 2, \dots, m\}$, that represents the switching signal that orchestrates the different therapies for each group i . Therefore, we replace γ_i in (16) with $\gamma_{\sigma i}$. The addition of the therapy scheduling preserves the positive invariance property of the set $0 \leq x_i \leq 1$, $i = 1, 2, \dots, 4$. We also make the simplifying assumption that the change of therapies does not affect the rates β_{ij} . Finally, we linearize the system around the *disease free equilibrium* $x = 0$. The linearized system is then given by

$$\dot{x} = (M + \Lambda_\sigma)x \quad (17)$$

where M is a nonnegative matrix with entries $M_{ij} = \beta_{ij} N_j / N_i$ and $\Lambda_1, \Lambda_2, \dots, \Lambda_m$ are diagonal matrices. The entries of Λ_σ are denoted by $\lambda_{\sigma i}$ and are as follows: $\lambda_{\sigma i} = -\gamma_{\sigma i} - \mu_i$. We can define the associated *embedded* system as

$$\dot{x} = (M + \Lambda(u))x \quad (18)$$

where the vector $u = [u_1, u_2, \dots, u_m]'$ lives in the polytope $\sum_{i=1}^m u_i = 1$, $u_i \geq 0$ and $\Lambda(u) = \sum_{i=1}^m \Lambda_i u_i$. $\Lambda(u)$ is convex with respect to u , so that the result of Section 4.3 can be used to conclude that, given a final time t_f , the scalar cost functional $J(x_0, u) = f(x(t_f))$ is a convex functional of $u \in \mathcal{U}_s$ provided that f is a monotonically increasing function. In the following example we consider a finite time optimal control problem with integral cost, namely

$$J(x_0) = \int_0^{t_f} \bar{c}' x(t) dt \quad (19)$$

where \bar{c} is a nonnegative column vector. Notice that it is easy to cast the optimal control problem for (18), (19) in the formulation studied in the formulation studies in the previous section. Indeed it is enough to add an additional state variable z with initial value $z(0) = 0$ and equation $\dot{z} = \bar{c}' x$ so that $J(x_0) = c'[x(t_f)' z(t_f)]'$ with $c = [0 \dots 0 \ 1]'$. It goes without saying that if system (17) is stabilizable, it is possible to tackle the problem of finding the optimal control for the infinite horizon problem, i.e. the minimization of

$$\bar{J}(x_0) = \int_0^\infty \bar{c}' x(t) dt$$

The stabilization problem, even for positive systems, is a formidable task for dimensions greater than 2, see the recent paper [Blanchini et al. \(2012\)](#). A sufficient condition for stabilizability $\dot{x} = A_\sigma x$ is given by looking at the so-called co-positive Lyapunov-Metzler inequalities

$$p_i' A_i + \sum_{j \neq i} \mu_{ij} (p_j - p_i) + \bar{c}' < 0, \quad i = 1, 2, \dots, m \quad (20)$$

If there exist nonnegative vectors p_i , $i = 1, 2, \dots, m$, and nonnegative numbers μ_{ij} , $j \neq i$ satisfying (20), then the switching law $\sigma = \text{argmin}_i p_i' x$ is stabilizing and such that $J(x_0) < \min_i p_i' x_0$. If this is the case, one can optimize the upper bound with respect to the free parameters λ_{ij} .

5.1 Simulation results

Consider system (16) and its linearized version (17) with

$$M = \begin{bmatrix} 0.8147 & 0.6324 & 0.9575 & 0.9572 \\ 0.9058 & 0.0975 & 0.9649 & 0.4854 \\ 0.1270 & 0.2785 & 0.1576 & 0.8003 \\ 0.9134 & 0.5469 & 0.9706 & 0.1419 \end{bmatrix}$$

$$\Lambda_1 = \text{diag}\{-4.442, -2.5140, -6.0511, -2.4946\}$$

$$\Lambda_2 = \text{diag}\{-2.7808, -2.9818, -5.2238, -4.6278\}$$

The task is to find the optimal control that minimizes

$$J(x_0) = \int_0^\infty \sum_{i=1}^4 x_i(t) dt$$

with the initial state given by $x_0 = [0.05 \ 0.15 \ 0.25 \ 0.35]'$.

According to (19) we set $c = [0 \ 0 \ 0 \ 1]$. Moreover, letting

$$A(u) = \begin{bmatrix} M + \Lambda(u) & 0 \\ \bar{c}' & 0 \end{bmatrix}$$

and $\Lambda(u) = \Lambda_1 u_1 + \Lambda_2 (1 - u_1)$, we can tackle the problem of minimizing $J(x_0) = \lim_{t_f \rightarrow \infty} c' \xi(t_f)$ under the system dynamics given by $\dot{\xi} = A(u)\xi$.

First let us consider fixed u . Notice that both $A_1 = M + \Lambda_1$ and $A_2 = M + \Lambda_2$ are Hurwitz stable, so that the best value of the cost for constant u in the vertices can be easily found $J(x_0) = \min_i q_i' x_0$ where $q_i = -c'(\Lambda_i + M)^{-1}$, $i = 1, 2$. It turns out that $J(x_0) = 0.622$. Notice that by taking a constant u strictly inside the polytope, the best constant control is $u_1 = 0.26$, $u_2 = 0.74$, that corresponds to a sliding mode for the switching system. With this control the cost is $J(x_0) = 0.601$.

One can also optimize the values of the parameters μ_{ij} in order to minimize the upper bound of the cost given by the co-positive Lyapunov-Metzler inequalities (20). Taking $\mu_{12} = 0.74$, $\mu_{21} = 0.26$ one obtains $p_1 = [0.876 \ 0.933 \ 0.652 \ 1.083]$, $p_2 = [1.229 \ 0.833 \ 0.723 \ 0.725]$ and applying the associated sub-optimal switching strategy one obtains a cost equal to 0.582. The (state-feedback) control law generates a sort of periodic behaviour in that the control switches from $u_1 = 0$ to $u_1 = 1$ in a periodic fashion.

We have computed the optimal control numerically, with a time horizon of 20 seconds, and 0.1 second discretization, using an analytical gradient, as described in Section 4.4, and the interior point algorithm in the Matlab function, `fmincon`. Surprisingly, in this particular case, the opti-

mal cost for the linearized system, $J(x_0) = 0.582$, is equal to the cost obtained by the sub-optimal switching strategy. As for the concave-convex mixed strategy, see (13), (14), the results show that the worst initial state is $x_0 = \alpha[0.206 \ 0.1631, 0.0043 \ 0.126]'$, where α is any positive scalar. Notice that since the system is linear the cost is also linear with respect to x_0 and hence only the direction is important. The associated cost is $J(x_0) = 0.4321\alpha$.

When applied to the nonlinear system, (16), the best constant control within the vertices is with $u_1 = 0$ and the associated cost is 0.411. The best constant control within the polytope is $u_1 = 0.26$, $u_2 = 0.74$ and the associated cost is 0.403. The switching strategy based on the LM inequalities provides a cost equal to 0.3922. If we compute the optimal control worked out for the linearized system, the cost is again equal (up to numerical errors) to the cost due to the switching strategy based on the co-positive Lyapunov-Metzler inequalities. Finally, the min-max optimal strategy with $x_0 = [0.206 \ 0.1631, 0.0043 \ 0.126]'$ has been applied to the nonlinear system, giving a cost equal to 0.407. The transient of the state-variables and the optimal input function are illustrated in Figs. 1, 2 respectively.

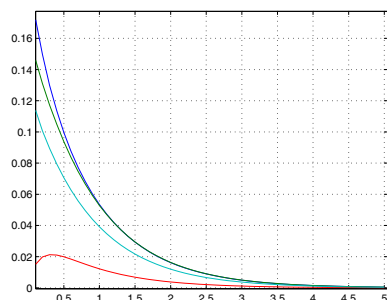


Figure 1. State variables of the nonlinear system (optimal mixed control for the linearized system)

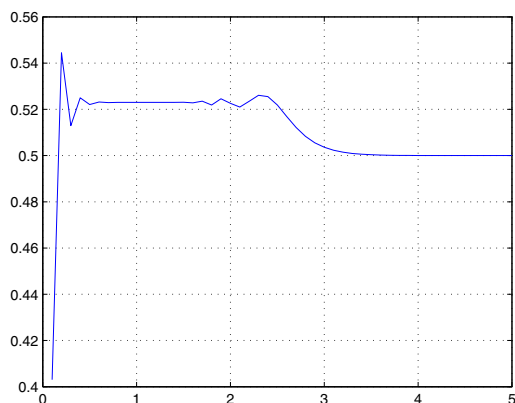


Figure 2. Optimal control variable associated with the worst initial state

6. CONCLUSIONS

In this paper we have examined the convexity of the cost functional with respect to the input function of a particular class of positive systems. Such systems are characterized

by the fact that the control variables affect the diagonal entries only, and these entries are convex functions of the control variables. The convexity property of the cost is important since is sufficient to guarantee the optimality of solutions satisfying the Pontragyn minimum principle in a optimal control context. This fact opens the way to the use efficient computational methods to find the optimal control. On the other hand, the class of systems considered is amenable to characterize important applications in the field of system biology. In particular, a simple example taken from a infective/susceptive epidemiological model is considered. Extensions of this work can consider questions such as: the search for different application models and the extension of the theory to a wider class of positive nonlinear dynamic systems.

REFERENCES

- Blanchini, F., Colaneri, P., and Valcher, M. (2012). Co-positive lyapunov functions for the stabilization of positive switched systems. *Automatic Control, IEEE Transactions on*, 57(12), 3038–3050. doi:10.1109/TAC.2012.2199169.
- Bressan, A. and Piccoli, B. (2007). *Introduction to Mathematical Control Theory*. AIMS Series on Applied Mathematics, Philadelphia, 2007.
- Cassandras, C., Pepyne, D., and Wardi, Y. (2001). Optimal control of a class of hybrid systems. *Automatic Control, IEEE Transactions on*, 46(3), 398–415.
- Cohen, J. (1981). Convexity of the dominant eigenvalue of an essentially nonnegative matrix. *Proceedings of the American Mathematical Society*, 81(4), 657–658.
- Dmitruk, A. and Kaganovich, A. (2011). Maximum principle for optimal control problems with intermediate constraints. *Computational Mathematics and Modeling*, 22(2), 180–215.
- Farina, L. and Rinaldi, S. (2000). *Positive linear systems: Theory and applications*. Wiley New York.
- Hernandez-Vargas, E., Colaneri, P., Middleton, R., and Blanchini, F. (2011). Discrete-time control for switched positive systems with application to mitigating viral escape. *International Journal of Robust and Nonlinear Control*, 21(10), 1093–1111.
- Hernandez-Vargas, E.A., Colaneri, P., and Middleton, R.H. (2013). Optimal therapy scheduling for a simplified hiv infection model. *Automatica*, 49(9), 2874–2880.
- Margaliot, M. (2006). Stability analysis of switched systems using variational principles: An introduction. *Automatica*, 42(12), 2059–2077.
- Moreno, Y., Pastor-Satorras, R., and Vespignani, A. (2002). Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 26(4), 521–529.
- Rami, M., Bokharaie, V., Mason, O., and Wirth, F. (2013). Stability criteria for sis epidemiological models under switching policies. *arXiv preprint arXiv:1306.0135*.
- Rapoport, L. (1996). Asymptotic stability and periodic motions of selector-linear differential inclusions. *Robust control via variable structure and Lyapunov techniques*, 269–285.