Control of convex-monotone systems

Anders Rantzer and Bo Bernhardsson

***Abstract*— We define the notion of convex-monotone system and prove that for such systems the state trajectory $x(\cdot)$ is a convex function of the initial state $x(0)$ and the input trajectory $u(\cdot)$. This observation gives a useful class of nonlinear dynamical systems for which control design can be performed by convex optimization. Applications to evolutionary dynamics of diseases and voltage stability in power networks are presented.**

I. Introduction

Control of monotone systems [1] has attracted increasing attention over the past decade. One reason is the variety of applications, such as chemical reaction networks, power systems, network flow models and populations models [13], [18]. Another reason is that monotone systems, and positive systems as a special case, lend themselves to analysis and synthesis methods with very good scalability properties [15], [16], [17].

This paper introduces yet another interesting feature of monotone systems: By exploiting monotonicity, we show that optimal control problems for certain nonlinear dynamical systems, with right-hand sides described by convex functions, can be stated in terms of convex optimization. The main result, Theorem 2, is introduced in Section II. After this, two different applications are described: optimal design of drug therapies using the convex-monotone properties of evolutionary dynamics of diseases in Section III and voltage stabilization for electrical networks in Section IV. Finally, some proofs are included in an appendix.

II. Main result

This paper is concerned with systems of the form

$$\dot{x}(t) = f(x(t), u(t)), \qquad x(0) = a, \qquad (1)$$

with $x(t) \in X \subset \mathbb{R}^n$, $u(t) \in U \subset \mathbb{R}^m$, where X and U are convex and (1) has a unique solution $x(t) = \phi_t(a,u)$. The system is said to be *monotone*, see [1], if the solution is a monotone function of the initial state a and the input trajectory u, i.e. if

$$(a_0, u_0) \le (a_1, u_1) \quad \Longrightarrow \quad \phi_t(a_0,u_0) \le \phi_t(a_1,u_1),$$

where inequalities are interpreted element-wise.

Before we state the main observation we remind the reader about the following fact which follows from more general results in [1].

Both authors are affiliated with Automatic Control LTH and the LCCC Linnaeus center, Lund University, Box 118, SE-221 00 Lund, Sweden, `rantzer at control.lth.se`.

Proposition 1: If $f \in C^1$ (i.e. continuously differentiable), the following statements are equivalent

- (i) The system (1) is monotone.
- (ii) The inequalities $\frac{\partial f_i}{\partial x_j} \ge 0$, $\frac{\partial f_i}{\partial u_k} \ge 0$ hold for all i, j, k with $i \ne j$.
- (iii) The solution to $\dot{x}(t) = f(x(t), u(t)) + v$, $x(0) = a$ is a monotone function of a, u, v.

A proof of Proposition 1 is included in the appendix to make the presentation self-contained.

If (1) is a monotone system and every row of f is also convex, the system will be said to be *convex-monotone*. The main observation in this paper is the following:

Theorem 2: If $f \in C^1$ and (1) is a convex-monotone system having the unique solution $x(t) = \phi_t(a,u)$, then each component of $\phi_t(a,u)$ is a convex function of (a,u).

Proof. Let $x_0(t) = \phi_t(a_0,u_0)$ and $x_1(t) = \phi_t(a_1,u_1)$. For $\lambda \in [0,1]$ introduce $x_\lambda = (1-\lambda)x_0 + \lambda x_1$ and

$$\begin{aligned} a_\lambda &= (1-\lambda)a_0 + \lambda a_1 \\ u_\lambda &= (1-\lambda)u_0 + \lambda u_1 \\ v_\lambda &= (1-\lambda)f(x_0,u_0) + \lambda f(x_1,u_1) - f(x_\lambda,u_\lambda) \end{aligned}$$

Then $v_\lambda \ge 0$ due to convexity of f. If we let $y_\lambda(t)$ be the solution to

$$\dot{y}(t) = f(y(t), u_\lambda(t)) + v_\lambda(t), \quad y(0) = a_\lambda \qquad (2)$$

then from Proposition 1 (iii) we get

$$\phi_t(a_\lambda, u_\lambda) \le y_\lambda(t).$$

However, x_λ solves (2), so $y_\lambda \equiv x_\lambda$ and therefore

$$\phi_t(a_\lambda,u_\lambda) \le y_\lambda(t) = x_\lambda(t) = (1-\lambda)\phi_t(a_0,u_0) + \lambda\phi_t(a_1,u_1).$$

This completes the proof. □

The observation in Theorem 2 is believed to be new. Using Theorem 2 convex-monotone dynamical systems can be combined with convex constraints and convex objective functions into convex optimization problems, which can be solved efficiently, even if the dynamics are nonlinear. We will illustrate this with two applications. Even though we only describe small examples in this paper, the methods are well suited for large-scale systems.

III. Evolutionary Dynamics of Diseases

Design of combination therapies for diseases such as HIV and cancer has recently been studied using control theory [7], [8], [9], [10], [14]. The basis is a model of the form

$$\dot{x}(t) = \left(A + \sum_i u_i D^i\right) x(t)$$

where each state x_k is the concentration of a mutant and each input u_i is a drug doses. Here A is a Metzler matrix[1] that describes the mutation dynamics without drugs, while $D^1, \dots, D^m$ are diagonal matrices modelling the effects of drugs. A basic problem is to determine $u_1, \dots, u_m \geq 0$, possibly subject to additional constraints, such that the state vector x decays as fast as possible.

We will now show that the system is actually convex-monotone after a transformation to logarithmic variables. This means that the choice of optimal time-varying drug doses $u_1(t), \dots, u_m(t)$ can be done by convex optimization using an application of Theorem 2:

Corollary 3: Given a Metzler matrix A, let $x(t)$ be the solution of

$$\dot{x}(t) = \left(A + \sum_{i=1}^m u_i(t) D^i\right) x(t) \qquad x(0) = a > 0$$

where $D^1, \dots, D^m$ are diagonal matrices. Then $\log x_k(t)$ is a convex function of (a, u).

Proof. Let a_{kj} be the entries of A and let D^i_k be the kth diagonal element of D^i. Define $z_k = \log x_k$. Then

$$\dot{z}_k(t) = \frac{\dot{x}_k(t)}{x_k(t)} = \sum_{kj} a_{kj} \frac{x_j(t)}{x_k(t)} + \sum_i u_i(t) D^i_k$$

$$\dot{z}_k(t) = \sum_{kj} a_{kj} \exp(z_j - z_k) + \sum_i u_i(t) D^i_k$$

which is a convex monotone system since $a_{kj} \geq 0$ when $k \neq j$ and A is Metzler. Hence the claim follows from Theorem 2. □

Remark 1. Restrictions enforcing a piecewise constant control signal can be added without destroying convexity.

With constant u_k and if the initial state is not taken into account, an alternative problem focusing on asymptotic growth rate, can be formulated as selection of u_i that minimize the Perron-Frobenius eigenvalue λ_{PF} of the matrix $A + \sum_i u_i D^i$. This can be done by convex optimization in the following way:

Proposition 4: Consider a Metzler matrix A and diagonal matrices $D^i = \operatorname{diag}\{D^i_1, \dots, D^i_n\}$, $i = 1, \dots, m$. Given $u \in \mathbb{R}^m$, let $\lambda_{\text{PF}}(A + \sum_i u_i D^i)$ be the Perron-Frobenius eigenvalue of $A + \sum_i u_i D^i$. Then

$$\min_{u \in U} \lambda_{\text{PF}}(A + \sum_i u_i D^i)$$

$$= \min_{u \in U, z_1, \dots, z_n} \max_k \left(\sum_{j=1}^n a_{kj} \exp(z_j - z_k) + \sum_{i=1}^m u_i D^i_k\right)$$

Remark 2. The expression on the right hand side is a convex function of $u_1, \dots, u_m, z_1, \dots, z_n$ and hence describes a convex optimization problem if U is convex. There is no sign-restriction on the elements of D.

Remark 3. The fact in Proposition 4 that the spectral radius of a non-negative matrix depends convexly on the diagonal elements has been known since [4]. Even closer to our result is [5]. Minimization of the Perron-Frobenius eigenvalue was treated for slightly different parameter dependence in [3, pp. 165-167].

Proof. A Metzler matrix M has all eigenvalues in the left half plane if and only if there exists a vector $x > 0$ with $Mx < 0$. Hence $\lambda_{\text{PF}}(A + \sum_i u_i D^i) < \gamma$ if and only if there exists $x > 0$ with

$$(A + \sum_i u_i D^i - \gamma I)x < 0$$

With $z_k = \log x_k$ for $k = 1, \dots, n$ this can be written

$$\sum_{j=1}^n a_{kj} x_j + \sum_i u_i D^i_k x_k - \gamma x_k < 0$$

$$\sum_{j=1}^n a_{kj} \frac{x_j}{x_k} + \sum_i u_i D^i_k - \gamma < 0$$

$$\sum_{j=1}^n a_{kj} \exp(z_j - z_k) + \sum_i u_i D^i_k - \gamma < 0$$

Minimization of γ gives the desired result. □

For a numeric example, we study a small toy model inspired by [7], [8], with $n = 4$ viral genotypes with viral populations x_i, $i = 1 \dots n$, and $m = 3$ different possible drug therapies that can be combined with the weight u_i, $i = 1, \dots m$.

The behaviour is given by

$$\dot{x} = \left(A + \sum_{i=1}^m u_i D^i\right) x, \qquad \text{with } A = \delta I + \mu M \quad (3)$$

with the clearance rate $\delta = -0.24$, viral mutation rate $\mu = 10^{-4}$, and mutation matrix

$$M = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

Replication rates for viral variants and therapies used in the analysis are described by Table 1. These rates do not correspond to efficiency of existing treatments but

[1] A matrix is Metzler iff it has nonnegative elements outside its main diagonal

Virus variant	Therapy 1	Therapy 2	Therapy 3
Type 1 (x_1)	$D_1^1 = 0.05$	$D_1^2 = 0.10$	$D_1^3 = 0.30$
Type 2 (x_2)	$D_2^1 = 0.25$	$D_2^2 = 0.05$	$D_2^3 = 0.30$
Type 3 (x_3)	$D_3^1 = 0.10$	$D_3^2 = 0.30$	$D_3^3 = 0.30$
Type 4 (x_4)	$D_4^1 = 0.30$	$D_4^2 = 0.30$	$D_4^3 = 0.15$

Fig. 1. Replication rates for viral variants and therapies.

were chosen to present an illustrative small version of a therapy scheduling problem. To describe a treatment scenario the weights must lie in the unit simplex

$$u_i(t) \geq 0 \text{ with } \sum_i u_i(t) = 1. \quad (4)$$

Remark 4. The use of a treatment $u(t)$ which is not a vertex of the simplex corresponds to fast switching between existing therapies. One should also note that $u = 0$ does in these parameters not correspond to the dynamics without virus treatment and is not a viable alternative.

We solve the two different optimization problems and start by finding the treatment combination (constant u_i) that minimizes the Perron-Frobenius eigenvalue of $A + \sum_i u_i D^i$. In CVX, a package for specifying and solving convex optimization problems [6], this can be formulated as

```
variables z(n,1) u(nrd,1) lambda
minimize lambda
subject to
u >= 0
sum(u) == 1
for k=1:n
    A(k,:)*exp(z-z(k)) + D(k,:)*u  <= lambda
end
```

where $D \in \mathbb{R}^{n\times m}$ contains the elements of D^i in its ith column. The result for the parameters in Table 1 is

$$\lambda_{\text{PF}} = -0.008,$$

which means that all virus variants are decaying. Without having treatment 3 available, $u_3 = 0$, the result is

$$\lambda_{\text{PF}} = 0.06$$

i.e. positive growth rate.

We now study the benefit of allowing time varying functions $u_i(t)$ satisfying (4). Assuming the initial value $x(0) = [1000, 0.1, 0.1, 0.001]$ known we solve the convex optimization problem of minimizing total virus population at time $T = 200$ days, i.e.

$$\min_u \sum_k x_k(T)$$

subject to (3) and (4). The result is given in figures 2-4. The final value on total virus population is improved from 0.08 to 0.03 compared to solving the same optimization problem with constant u_i. The final

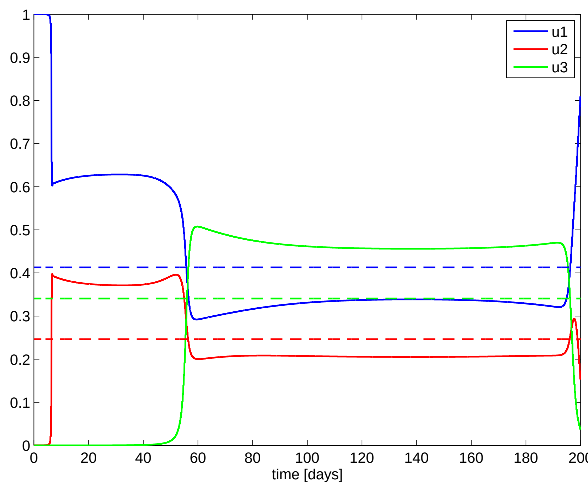

Fig. 2. Optimal time varying virus treatment u_1, u_2, u_3 (solid) vs optimal constant virus treatment (dashed).

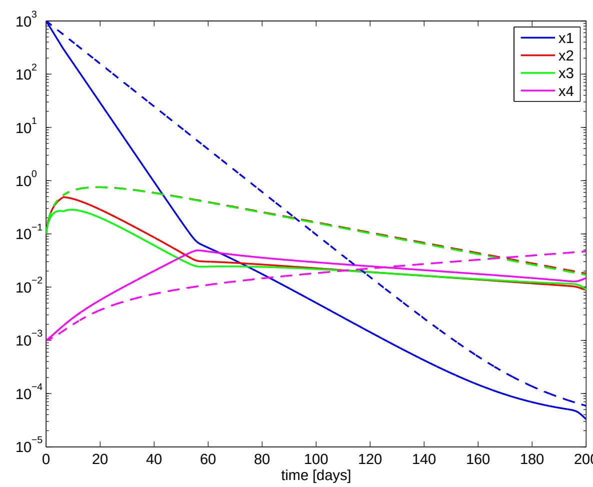

Fig. 3. Virus population x_1, x_2, x_3, x_4 for time varying treatment (solid) vs constant treatment (dashed). The final value on total virus population is improved from 0.08 to 0.03.

value on total virus population achieved with constant treatment for 200 days, is obtained already after 100 days with time-varying treatment.

IV. Voltage stability in a DC network

Our next example is from the study of power systems. The fundamental role of nonnegative matrices in this context has recently been exploited in the stationary context of power flow optimization [11], [12]. Equipped with the results of section II, we are now also ready to treat voltage stability and power flow under non-stationary conditions.

Let us first explain the issue of voltage stability in the simplest possible case; a single resistive transmission line from generator to load, see Figure 5. Given generator voltage u_1 and line current i the load voltage u_2 can be computed by Ohm's law $u_2 = u_1 - Ri$ where R

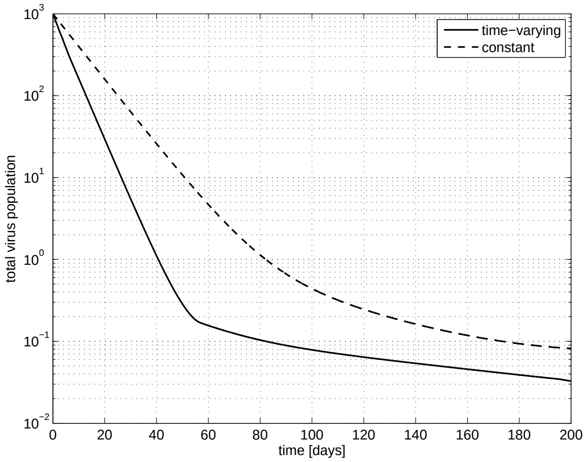

Fig. 4. Total virus population $\sum x_k(t)$ for time varying treatment (solid) vs constant treatment (dashed). Note that the final value on total virus population (0.08) achieved with constant treatment for 200 days, is obtained already after 100 days with time-varying treatment.

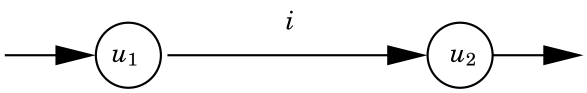

Fig. 5. A single transmission line from generator to load

is the line resistance. In particular, the power $p = iu_2$ delivered to the load is upper bounded by

$$p = i(u_1 - Ri) \leq \frac{u_1^2}{4R}.$$

When the voltage drops in a power network, an active load could try to counteract the power loss by extracting more current. Such a behaviour is described by the following load model:

$$\frac{di}{dt} = \frac{\widehat{p}}{u_1 - Ri} - i.$$

where $\widehat{p}$ is the power needed by the load. Notice that if u_1 is constant and the initial current is bigger than $u_1/(2R)$, then current will increase further and the load voltage $u_1 - Ri$ will eventually go down to zero. This is called a voltage collapse.

Now consider an arbitrary network of generators and loads connected by transmission lines. The voltages at the generators and loads are given by vectors $u^G \in \mathbb{R}^m$ and $u^L \in \mathbb{R}^n$ respectively. The voltages are mapped into vectors of external currents $i^G \in \mathbb{R}^m$ and $i^L \in \mathbb{R}^n$ according to the equation

$$\begin{bmatrix} -i^G(t) \\ i^L(t) \end{bmatrix} = \underbrace{\begin{bmatrix} Y^{GG} & Y^{GL} \\ Y^{LG} & Y^{LL} \end{bmatrix}}_{Y} \begin{bmatrix} u^G(t) \\ u^L(t) \end{bmatrix} \qquad (5)$$

Voltages are always assumed to be positive. The sign convention for currents is that when power is produced by a generator, the corresponding entry of $i^G(t)$ is positive and when power is extracted by a load, the corresponding entry of $i^L(t)$ is positive. The matrix Y is a symmetric Metzler matrix. Each off-diagonal element is given by the admittance (inverse of resistance) of the corresponding transmission line. The diagonal elements are negative and such that the all row sums of Y are equal to zero. The vanishing row sums of Y correspond to the fact that no currents will flow through the network when all voltages are equal, regardless of their value. In particular, Y is not invertible and the voltages cannot be uniquely determined from the currents.

To study voltage stability, we need a model for the load dynamics. Resistive loads are modelled by the equation

$$u_k^L = R_k i_k^L$$

Such "feedback loops" can be closed without changing the structure of Y, so resistive loads can be ignored without loss of generality. Instead, we will focus on loads that adjust their current to compensate for voltage deviations:

$$\frac{di_k^L}{dt}(t) = \frac{\widehat{p}_k}{u_k^L(t)} - i_k^L(t)$$

In particular, the load current i_k^L increases if the delivered power $u_k^L i_k^L$ is too small and decreases if it is too large. Writing this load model on vector form with the expression for load voltage from (5) gives

$$\frac{di^L}{dt}(t) = \widehat{p}./[(Y^{LL})^{-1}(i^L - Y^{LG}u^G)] - i^L(t)$$

where the notation "./" denotes entry-wise division. The fact that Y^{LL} is Metzler and Hurwitz implies that $(Y^{LL})^{-1} \leq 0$ [2, page 137]. Hence, *the system is convex-monotone with state i^L and input $-u^G$.*

From (5) we get that the generator currents can be expressed in terms of the inputs and states states

$$i^G = \left[Y^{GL}(Y^{LL})^{-1}Y^{LG} - Y^{GG}\right]u^G - Y^{GL}(Y^{LL})^{-1}i^L$$

The first term on the right hand side depends linearly on u^G and the second is a convex function multiplied by a non-negative matrix, so i^G is a convex function of u^G, just like $-u^L$, i^L and di^L/dt. Similarly, the differentiated relationship

$$\frac{di^G}{dt} = \left[Y^{GL}(Y^{LL})^{-1}Y^{LG} - Y^{GG}\right]\frac{du^G}{dt} - Y^{GL}(Y^{LL})^{-1}\frac{di^L}{dt}$$

shows that the same is true for di^G/dt. In particular, it is possible to use convex optimization to stabilize the system dynamics in spite of bounds on voltages, currents and their derivatives. It follows that the set of initial states that can be saved from a voltage collapse is convex.

Finally, consider the network in Figure 6. The two transmission lines have admittances $y_1, y_2 \geq 0$. Writing Kirchhoff's current law for the nodes in Figure 6

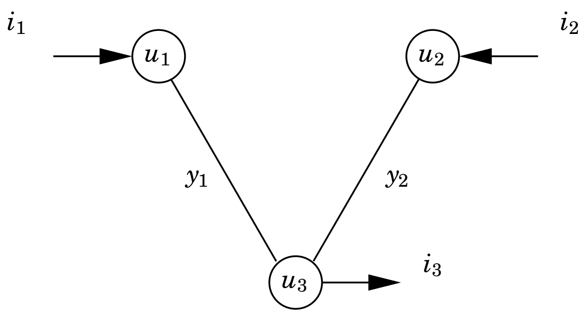

Fig. 6. A transmission network with two generators and one load.

gives

$$\begin{bmatrix} -i_1 \\ -i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} -y_1 & 0 & y_1 \\ 0 & -y_2 & y_2 \\ y_1 & y_2 & -y_1 - y_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Node 3 is an active load with

$$\frac{di_3}{dt} = \frac{\widehat{p}(y_1 + y_2)}{y_1 u_1 + y_2 u_2 - i_3} - i_3$$

For constant generator voltages u_1 and u_2, the load voltage $u_3 = y_1 u_1 + y_2 u_2 - i_3$ could shrink to zero in finite time, which means voltage collapse. To avoid this, it is clear from monotonicity that the voltages u_1 and u_2 should be raised as quickly as possible. However, the rate of increase is limited by the flexibility of the generators to increase power production. Limited flexibility in generator 2 can be modelled as a bound on the rate of change in the generator current

$$i_2 = \frac{y_2}{y_1 + y_2}(i_3 + y_1 u_2 - y_1 u_1).$$

Such an upper bound becomes another convex constraint on the voltage trajectories u_1, u_2.

Altogether, our analysis shows that the problem to bring an electrical network into equilibrium, subject to active loads and constraints on rates and amplitudes, can be addressed using trajectory optimization in a convex monotone system.

V. Appendix: Proof of Proposition 1

Lemma 5 (Variational Formula): Let $x(t,\alpha)$ be the solution to

$$\dot{x} = f(x, u(t) + \alpha v(t)), \quad x(0,\alpha) = x_0 + \alpha z_0$$

where $f \in C^1$, then $\Delta_\alpha(t) := \frac{\partial x(t,\alpha)}{\partial \alpha}$ satisfies

$$\frac{d}{dt}\Delta_\alpha(t) = A(t)\Delta_\alpha(t) + B(t)v(t), \quad \Delta_\alpha(0) = z_0,$$

where $A(t) = \frac{\partial f}{\partial x}(x^*, u^*)$ and $B(t) = \frac{\partial f}{\partial u}(x^*, u^*)$ are evaluated at $x^* = x(t,\alpha), u^* = u(t) + \alpha v(t)$. The solution is given by

$$\Delta_\alpha(t) = \Phi_A(t,0)z_0 + \int_0^t \Phi_A(t,s)B(s)v(s)ds,$$

where $\Phi_A(t,s)$ is the solution to

$$\frac{d}{dt}\Phi_A(t,s) = A(t)\Phi_A(t,s)$$
$$\Phi_A(s,s) = I.$$

Proof. Classical. □

Lemma 6 (Positivity of Fundamental Matrix): Let $\Phi_A(t,s)$ be the fundamental matrix to the system

$$\dot{x}(t) = A(t)x(t),$$

where $A(t)$ is locally bounded and Metzler, i.e. $A_{ij} \geq 0$ for $i \neq j$, then

$$\Phi_A(t,s) \geq 0, \quad t \geq s.$$

Proof. For any compact interval I one can find a constant c such that $B(t) := A(t) + cI \geq 0$ for $t \in I$. We have $\Phi_A(t,s) = e^{-c(t-s)}\Phi_B(t,s)$ where

$$\frac{d}{dt}\Phi_B(t,s) = B(t)\Phi_B(t,s)$$
$$\Phi_B(s,s) = I$$

from which follows that $\Phi_B(t,s) \geq 0$ if $t \geq s$ and t and s belong to I. This proves the result. □

Proof of Proposition 1.

$(i) \Rightarrow (ii):$ Let $x(t,\varepsilon)$ denote the solution to (1) with $x(0) = a + \varepsilon e_j$ where e_j denotes the j:th unit vector. For the ith component we have $x_i(0) = x_i(0,\varepsilon)$ if $i \neq j$ and

$$\dot{x}_i(0,\varepsilon) - \dot{x}_i(0) = f_i(x(0) + \varepsilon e_j, u(0)) - f_i(x(0), u(0))$$
$$= \varepsilon \frac{\partial f_i}{\partial x_j} + o(\varepsilon)$$

so by monotonicity $\frac{\partial f_i}{\partial x_j} \geq 0$. Similarly, using perturbations of the form $u(t,\varepsilon) = u(t) + \varepsilon w_j(t)$ with $w_j(t)$ being a step function in coordinate j, we get $\frac{\partial f_i}{\partial u_j} \geq 0, \forall i, j$.

$(ii) \Rightarrow (i):$ For $\alpha \in [0,1]$ let $x(t,\alpha)$ be the solution to

$$\dot{x}(t,\alpha) = f(x(t,\alpha), u(t) + \alpha\Delta_u(t)), \quad x(0,\alpha) = a + \alpha\Delta_a.$$

We need to prove that

$$\Delta_u(t) \geq 0, \Delta_a \geq 0 \quad \Rightarrow \quad x(t,1) \geq x(t,0).$$

From the fundamental theorem of calculus and the variational formula

$$x(t,1) - x(t,0) = \int_0^1 \frac{\partial x(t,\alpha)}{\partial \alpha} d\alpha$$
$$= \int_0^1 \left(\Phi_A(t,0)\Delta_a + \int_{s=0}^t \Phi_A(t,s)B(s,\alpha)\Delta_u(s)\,ds \right) d\alpha$$

where $\Phi_A(t,s)$ is the fundamental matrix to $\dot{z}(t) = A(t,\alpha)z(t)$ with $A(t,\alpha) = \frac{\partial f}{\partial x}(x(t,\alpha), u(t,\alpha))$, and where $B(t,\alpha) = \frac{\partial f}{\partial u}(x(t,\alpha), u(t,\alpha))$. From Lemma 6 it follows that $\Phi_A(t,s)$ and $B(t,\alpha)$ are nonnegative. Hence the integral is nonnegative and the result follows.

$(ii) \Rightarrow (iii)$: The Jacobian of the right-hand side with respect to the pair (u,v) is given by the nonnegative matrix $\begin{bmatrix} \frac{\partial f}{\partial u} & I \end{bmatrix} \geq 0$. The result hence follows from the implication $(ii) \Rightarrow (i)$, just proved, applied to the system with extended input $u := (u,v)$

$(iii) \Rightarrow (i)$: Trivial. □

VI. Acknowledgement

The work has been supported by the Swedish Research Council through the LCCC Linnaeus Center and the eLLIIT Excellence Center at Lund University.

References

[1] David Angeli and Eduardo D. Sontag. Monotone control systems. *IEEE Transactions on Automatic Control*, 48(10):1684–1698, October 2003.

[2] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*. Classics in Applied Mathematics. SIAM, 1994.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[4] Joel E.Cohen. Convexity of the dominant eigenvalue of an essentially nonnegative matrix. *Proc. Amer. Math. Soc.*, 81:657–658, 1981.

[5] L. Elsner. On convexity properties of the spectral radius of nonnegative matrices. *Linear Algebra and its Applications*, 61:31âĂŞ35, September 1984.

[6] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.0 beta. http://cvxr.com/cvx, September 2013.

[7] Esteban Hernandez-Vargas, Patrizio Colaneri, Richard Middleton1, and Franco Blanchini. Discrete-time control for switched positive systems with application to mitigating viral escape. *International journal of robust and nonlinear control*, 21:1093–1111, 2011.

[8] Esteban A. Hernandez-Vargas, Patrizio Colaneri, and Richard H. Middleton. Optimal therapy scheduling for a simplified hiv infection model. *Automatica*, pages 2874–2880, 2013.

[9] Vanessa Jonsson, Nikolai Matni, and Richard M. Murray. Reverse engineering combination therapies for evolutionary dynamics of disease: An h-infinity approach. In *Proceedings of 52nd IEEE Conference on Decision and Control*, Florence, December 2013.

[10] Vanessa Jonsson, Anders Rantzer, and Richard M. Murray. A scalable formulation for engineering combination therapies for evolutionary dynamics of disease. In *Proceedings of American Control Conference*, Portland, June 2014.

[11] J. Lavaei and S.H. Low. Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1):92–107, 2012.

[12] Javad Lavaei, Anders Rantzer, and Steven Low. Power flow optimization using positive quadratic programming. In *18th IFAC World Congress*, Milano, Italy, August 2011.

[13] Patrick De Leenheer, David Angeli, and Eduardo D. Sontag. Monotone chemical reaction networks. *Journal of Mathematical Chemistry*, 41(3):295–314, April 2007.

[14] Rutao Luo, Michael J. Piovoso, Javier Martinez-Picado, and Ryan Zurakowski. Optimal antiviral switching to minimize resistance risk in hiv therapy. *PLoS ONE*, 6(11):e27047, 11 2011.

[15] Luc Moreau. Stability of continuous-time distributed consensus algorithms. In *Proceedings of 43rd IEEE Conference on Decision and Control*, Bahamas, 2004.

[16] Anders Rantzer. Distributed control of positive systems. In *50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, Florida, USA, December 2011.

[17] Anders Rantzer, Björn S. Rüffer, and Gunther Dirr. Separable Lyapunov functions for monotone systems. In *Proceedings of IEEE Conference of Decision and Control*, Florence, December 2013.

[18] Eduardo D. Sontag. Monotone and near-monotone biochemical networks. *Syst Synth Biol*, 1:59–87, 2007.