COMPUTATION OF AMPLIFICATION FOR SYSTEMS ARISING FROM CELLULAR SIGNALING PATHWAYS

Eduardo D. Sontag [*,1] **Madalena Chaves** [**,2]

[*] *Department of Mathematics, Rutgers University, New Brunswick, NJ 08903*
[**] *Department of Mathematics, Rutgers University, New Brunswick, NJ 08903*

Abstract: A commonly employed measure of the signal amplification properties of an input/output system is its induced $\mathcal{L}^2$ norm, sometimes also known as H_∞ gain. In general, however, it is extremely difficult to compute the numerical value for this norm, or even to check that it is finite, unless the system being studied is linear. This paper describes a class of systems for which it is possible to reduce this computation to that of finding the norm of an associated linear system. In contrast to linearization approaches, a precise value, not an estimate, is obtained for the full nonlinear model. The class of systems that we study arose from the modeling of certain biological intracellular signaling cascades, but the results should be of wider applicability.

Keywords: norms, gains, nonlinear systems, signaling pathways

1. INTRODUCTION

The analysis of signaling networks constitutes one of the central questions in systems biology. There is a pressing need for powerful mathematical tools to help understand and conceptualize their information processing and dynamic properties. One natural question is that of quantifying the amount of "signal amplification" in such a network, meaning in some sense the ratio between the size of a response or output and that of the input that gave rise to it. See for instance (Heinrich *et al.*, 2002) for a recent paper in this line of work.

In control theory, a routine way to quantify amplification is by means of the induced $\mathcal{L}^2$ norm or "H_∞ gain" of a system. A major difficulty when trying to apply these techniques to signaling networks is that such systems are usually highly nonlinear. Thus, typically, mathematical results are only given for small inputs or "weakly activated" systems, see for instance (Heinrich *et al.*, 2002; Chaves *et al.*, 2004). For large signals, that is, when analyzing the full nonlinear system, even deciding if the norm is finite or not is usually a very hard question.

In this paper, motivated by the particular systems studied in (Heinrich *et al.*, 2002; Chaves *et al.*, 2004), we introduce a class of nonlinear systems, which includes all these motivational examples as well as many others, and we show finiteness and how to obtain precise values for norms, by reducing the problem of norm estimation to the same problem for an associated linear system. This associated system is sometimes a linearization of the original system around an equilibrium point, though it need not be. In any case, the techniques are not at all related to linearization techniques, but instead borrow from comparison theorems, ISS-like estimates, and the theory of positive systems.

---

[1] Email: `sontag@math.rutgers.edu` . Supported in part by NIH Grant R01 GM46383 and NSF Grant CCR-0206789,
[2] Email: `madalena@math.rutgers.edu` . Supported in part by Sanofi-Aventis.

2. DEFINITIONS AND STATEMENTS OF RESULTS

We deal with systems of the following special form:

$$\dot{x}(t) = A(x(t))\,x(t) \;+\; B(x(t))\,u(t)\,,\; x(0) = 0 \quad (1)$$

(or just "$\dot{x} = A(x)x + B(x)u$"), where dot indicates time derivative, and states $x(t)$ as well as input values $u(t)$ are vectors with nonnegative components: $x(t) \in \mathbb{R}^n_{\geq 0}$ and $u(t) \in \mathbb{R}^m_{\geq 0}$ for all $t \geq 0$, for some positive integers n and m. We view A and B as matrix valued functions

$$A : \mathbb{R}^n_{\geq 0} \to \mathbb{R}^{n\times n}\,, \quad B : \mathbb{R}^n_{\geq 0} \to \mathbb{R}^{n\times m}$$

where $\mathbb{R}^k_{\geq 0} = (\mathbb{R}_{\geq 0})^k$, for any positive integer k, is the set of vectors $\xi \in \mathbb{R}^k$ in Euclidean k-space with all coordinates $\xi_i \geq 0$, $i = 1, \ldots, k$. Associated to these systems we also have an output or measurement

$$y(t) = h(x(t)) = C(x(t))x(t)$$

taking values $y(t) \in \mathbb{R}^p$, for some integer p, where $C : \mathbb{R}^n_{\geq 0} \to \mathbb{R}^{p\times n}$.

Assumptions We make several assumptions concerning the matrix functions A, B, and C, as follows.

Stability:

The matrix $A(0)$ is Hurwitz, that is, all eigenvalues of $A(0)$ have negative real parts.

Maximization at $\xi = 0$:

For each $\xi \in \mathbb{R}^n_{\geq 0}$, $A(\xi) \leq A(0)$, $B(\xi) \leq B(0)$, and $C(\xi) \leq C(0)$, meaning that $A(\xi)_{ij} \leq A(0)_{ij}$ for each $i, j \in \{1, \ldots, n\}$, $B(\xi)_{ij} \leq B(0)_{ij}$ for each $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$. and $C(\xi)_{ij} \leq C(0)_{ij}$ for each $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, m\}$.

Positivity of system:

For each $\xi \in \mathbb{R}^n_{\geq 0}$ and each $i \in \{1, \ldots, n\}$ such that $\xi_i = 0$, it holds that: $A(\xi)_{ij} \geq 0$ for all $j \neq i$ and $B(\xi)_{ij} \geq 0$ for all j. Also, for every $\xi \in \mathbb{R}^n_{\geq 0}$, $C_{ij}(\xi) \geq 0$ for all i, j.

Local Lipschitz assumption:

The matrix functions $A(\xi)$, $B(\xi)$, and $C(\xi)$ are locally Lipschitz in ξ.

Remarks about the form of the system The special form assumed for the system is in itself not very restrictive, since every (affine in controls) system $\dot{x} = F(x) + B(x)u$ may be written in this fashion, provided only that F be a continuously differentiable vector field and $F(0) = 0$, for instance by taking $A(\xi) = \int_0^1 F'(\lambda\xi)\,d\lambda$, where F' indicates the Jacobian of F. This reduction to a "state dependent linear form" $\dot{x} = A(x)x + B(x)u$ is often useful in control theory, where it appears for instance in the context of "state-dependent Riccati equation" approaches to optimal control. Of course, the difficulty is in satisfying the above assumptions for A and B.

A special case in which these hypotheses are satisfied is that of models of cell signaling cascades as in (Heinrich *et al.*, 2002; Chaves *et al.*, 2004). These are systems whose equations can be written as follows (with n arbitrary and $m = 1$):

$$\begin{aligned}
\dot{x}_1 &= \alpha_1 u(c_1 - x_1) - \beta_1 x_1 \\
\dot{x}_i &= \alpha_i x_{i-1}(c_i - x_i) - \beta_i x_i\,, \quad i = 2, \ldots, n
\end{aligned}$$

and output $y = x_n$, and the α_i's, β_i's, and c_i's are all positive constants. We represent this system in the above form using: $A(\xi)_{1,1} = -\beta_1$, $A(\xi)_{i,i-1} = \alpha_i c_i$ for $i = 2, \ldots, n$, $A(\xi)_{i,i} = -\alpha_i\xi_{i-1} - \beta_i$ for $i = 2, \ldots, n$, $B(\xi)_{1,1} = \alpha_1 c_1 - \alpha_1\xi_1$, and all other entries zero. Note that $A(\xi) \leq A(0)$ and $B(\xi) \leq B(0)$, for all $\xi \in \mathbb{R}^n_{\geq 0}$, because $-\alpha_i\xi_i \leq 0$ for all i. The matrix $A(0)$ is lower triangular with negative diagonals, and hence is Hurwitz. Positivity holds as well: if $i = 1$ and ξ is such that $\xi_1 = 0$, then $A(\xi)_{1j} = 0$ for all $j \neq 1$ and $B(\xi)_{11} = \alpha_1 c_1 > 0$; if instead $i > 1$ and ξ is such that $\xi_i = 0$, then $A(\xi)_{ij} = 0$ for all $j \notin \{i-1, i\}$, $A(\xi)_{i,i-1} = \alpha_i c_i > 0$, and $B(\xi)_{i1} = 0$. Finally, the functions $A(\cdot)$ and $B(\cdot)$ are linear, and hence Lipschitz. The matrix $C(\xi) = (0, 0, \ldots, 0, 1)^{\mathrm{T}}$ is constant and nonnegative. Thus all properties hold for this example.

A linear one-dimensional system $\dot{x}_{n+1} = x_n - \ell x_{n+1}$ may be cascaded at the end, as in (Chaves *et al.*, 2004), and the output is in that case redefined as $y = x_{n+1}$; this may be again modeled in the same way, and the assumptions still hold.

Induced gains Assume given a system (1). We consider the operator T that assigns the solution function x to each input u. To be more precise, we consider inputs $u \in \mathcal{L}^2([0, \infty), \mathbb{R}^m_{\geq 0})$, and define $x = Tu$ as the unique solution of the initial value problem (1). In principle, this solution is only defined on some maximal interval $[0, \mathcal{T})$, where $\mathcal{T} > 0$ depends on u; however, we will show below that $\mathcal{T} = +\infty$, and that x is again square integrable (and nonnegative), so we may view x as an element of $\mathcal{L}^2([0, \infty), \mathbb{R}^n_{\geq 0})$ and T as an (nonlinear) operator

$$T \;:\; \mathcal{L}^2([0, \infty), \mathbb{R}^m_{\geq 0}) \to \mathcal{L}^2([0, \infty), \mathbb{R}^n_{\geq 0})\,.$$

We will write $|\cdot|$ for Euclidean norm, and use $\|\cdot\|$ to denote $\mathcal{L}^2$ norm: $\|u\|^2 = \int_0^\infty |u|^2\,dt$. For the operator T, we consider the usual induced operator norm:

$$\|T\| \;:=\; \sup_{u\neq 0} \frac{\|Tu\|}{\|u\|}\,.$$

We will show that $\|T\| < \infty$ for the systems that we are considering. In order to see this, we first consider the linear system

$$\dot{z} \;=\; A(0)z + B(0)u\,, \quad z(0) = 0 \qquad (2)$$

with output $v = \ell(z) = C(0)z$, and its associated operator

$$L : \mathcal{L}^2([0,\infty), \mathbb{R}^m_{\geq 0}) \to \mathcal{L}^2([0,\infty), \mathbb{R}^n_{\geq 0}) : u \mapsto z\,.$$

Since $A(0)$ is a Hurwitz matrix, $z(t)$ is defined for all $t \geq 0$, and L indeed maps $\mathcal{L}^2$ into $\mathcal{L}^2$. Furthermore, its induced norm $\|L\|$, the "H_∞ gain" of the system with output $y = z$, is finite; see for instance (Doyle *et al.*, 1992). (The H_∞ gain is defined for arbitrary-valued inputs $u \in \mathcal{L}^2([0,\infty), \mathbb{R}^m)$; we will remark below, cf. Section 5, that the same norm is obtained when only nonnegative inputs are used in the maximization.) Moreover, the $\mathcal{L}^2 \to \mathcal{L}^\infty$ (or "H_2") induced gain is also finite. Therefore, using $\|\cdot\|_\infty$ to denote supremum norm $\|z\|_\infty = \sup_{t\geq 0} |z(t)|$, we can pick a common constant $c \geq 0$ such that

$$\|Lu\| \leq c\,\|u\| \;\text{ and }\; \|Lu\|_\infty \leq c\,\|u\| \qquad (3)$$
$$\text{for all } \; u \in \mathcal{L}^2([0,\infty), \mathbb{R}^m_{\geq 0})$$

where c upper bounds both $\|L\|$ and $\|L\|_\infty$ (we use $\|L\|_\infty$ for operators to denote induced $\mathcal{L}^2 \to \mathcal{L}^\infty$ norm).

Our object of study are the compositions with the output maps, i.e. the input/output operators:

$$T_o : \mathcal{L}^2([0,\infty), \mathbb{R}^m_{\geq 0}) \to \mathcal{L}^2([0,\infty), \mathbb{R}^p_{\geq 0})$$
$$: u \mapsto y = C(x)x = C(Tu)Tu$$

and

$$L_o : \mathcal{L}^2([0,\infty), \mathbb{R}^m_{\geq 0}) \to \mathcal{L}^2([0,\infty), \mathbb{R}^p_{\geq 0})$$
$$: u \mapsto v = C(0)z = C(0)Lu$$

and their corresponding induced norms. Our main result is as follows:

Theorem 1. The norm of T_o is finite, and $\|T_o\| = \|L_o\|$.

3. PRELIMINARY RESULTS

We start our proof by remarking that the solutions of (1) remain in $\mathbb{R}^n_{\geq 0}$. To see this, we need to verify the following property (this is a standard invariance fact; see for instance (Angeli and Sontag, 2003) for a discussion in a related context):

for each $i = 1, \ldots, n$, each $\xi \in \mathbb{R}^n_{\geq 0}$ such that $\xi_i = 0$, and each $\mu \in \mathbb{R}^m_{\geq 0}$,

$$(A(\xi)\xi + B(\xi)\mu)_i \;\geq\; 0\,.$$

Since $\xi_i = 0$, we need to prove that $\sum_{j\neq i} A(\xi)_{ij}\xi_j + \sum_j B(\xi)_{ij}\mu_j$ is nonnegative, but this is implied by the positivity assumption.

Similarly, solutions of (1) remain in $\mathbb{R}^n_{\geq 0}$, as also $(A(0)\xi + B(0)\mu)_i \geq 0$ if $\xi_i = 0$.

The next observation is a key one:

Lemma 3.1. Every solution of (1), with $u \in \mathcal{L}^2$, is defined for all $t \geq 0$. Moreover, for any two solutions x of (1) and (2) with the same input u, it holds that $0 \leq x_i(t) \leq z_i(t)$ for each coordinate $i = 1, \ldots, n$ and each $t \geq 0$.

Proof. We use the following comparison principle for differential equations. Suppose that $f(t,\xi)$ and $g(t,\xi)$ are such that $f_i(t,\xi) \leq g_i(t,\xi)$ for all $i = 1, \ldots, n$ and all $\xi \in \mathbb{R}^n_{\geq 0}$, and that we consider the solutions of $\dot{x} = f(t,x)$ and $\dot{z} = g(t,z)$ with the same initial condition (or, more generally, initial conditions $x(0) \leq z(0)$). Then, provided that g is quasi-monotone (and suitable regularity conditions hold, as here), we may conclude that $x(t) \leq z(t)$ (componentwise) for all $t \geq 0$ for which both solutions are defined. See for instance (Smith, 1995; Lakshmikantham and Leela, 1969). Quasi-monotonicity means that $\partial g_i/\partial \xi_j \geq 0$ for all $i \neq j$.

Let us now take any fixed control and let $f(t,\xi) = A(\xi)\xi + B(\xi)u(t)$, $g(t,\xi) = A(0)\xi + B(0)u(t)$. We have that $f(t,\xi) \leq g(t,\xi)$ coordinatewise, because $A(\xi) \leq A(0)$ and $B(\xi) \leq B(0)$ by assumption. To see that g is quasi-monotone, one needs to verify that $A(0)_{ij} \geq 0$ for all $i \neq j$. but this follows from the positivity assumption on (A, B). Thus the comparison principle tells us that $x(t) \leq z(t)$ for all $t \geq 0$ for which the solution x is defined (the solution z is defined for all t, since (2) is linear and $A(0)$ is a Hurwitz matrix). We already observed that x is bounded below by zero; thus, the maximal solution x is bounded on any finite interval, and hence it is indeed defined for all t, and the Lemma follows. ■

Corollary 3.2. For each $u \in \mathcal{L}^2$, the solution Tu of (1) is in $\mathcal{L}^2$, and the operator T has finite norm. Moreover,

$$\|Tu\| \leq \|Lu\| \leq c\,\|u\|$$

and

$$\|Tu\|_\infty \leq \|Lu\|_\infty \leq c\,\|u\|$$

where c is any constant as in (3), so in particular $\|T\| \leq \|L\| \leq c$ and $\|T\|_\infty \leq \|L\|_\infty \leq c$. Similarly, the i/o operator T_o also has finite norm, $\|T_o u\| \leq \|L_o u\|$ and $\|T_o u\|_\infty \leq \|L_o u\|_\infty$ for all $u \in \mathcal{L}^2$, and $\|T_o\| \leq \|L_o\|$, $\|T_o\|_\infty \leq \|L_o\|_\infty$.

Proof. Pick any u, and let $x = Tu$ and $z = Lu$. By the Lemma, $0 \le x_i(t) \le z_i(t)$ for all t, so

$$\|x\|^2 = \int_0^\infty \sum_{i=1}^n x_i(s)^2\, ds \le \int_0^\infty \sum_{i=1}^n z_i(s)^2\, ds = \|z\|^2 .$$

So $\|Tu\| \le \|Lu\| \le c\,\|u\|$, and since u was arbitrary it follows that $\|T\| \le \|L\|$. Similarly,

$$\|x\|_\infty = \sup_{t\ge 0} |x(t)| \;\le\; \sup_{t\ge 0} |z(t)| = \|z\|_\infty$$

leads to $\|Tu\|_\infty \le \|Lu\|_\infty$ and $\|T\|_\infty \le \|L\|_\infty$.

The positivity and the maximization properties for C imply that, for each coordinate i of the outputs $y(t) = C(x(t))x(t)$ and $v(t) = C(0)z(t)$, we have $0 \le y_i(t) = \sum_{j=1}^n C_{ij}(x(t))x_j(t) \le \sum_{j=1}^n C_{ij}(0)z_j(t) = v_i(t)$, so the inequalities for T_o and L_o follow by an analogous reasoning. ∎

Note that the inequality $\|T_o\| \le \|L_o\|$ gives the finiteness statement as well as one-half of the equality in the main theorem.

For any matrix Q, we denote by $|Q|$ its induced operator norm as an operator in Euclidean space, that is, the smallest constant d such that $|Q\xi| \le d\,|\xi|$ for all ξ.

Lemma 3.3. There is a nondecreasing and continuous function $M : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ such that:

$$|A(\xi) - A(0)| \le M(|\xi|)\,|\xi|$$

$$|B(\xi) - B(0)| \le M(|\xi|)\,|\xi|$$

$$|C(\xi) - C(0)| \le M(|\xi|)\,|\xi|$$

for all $\xi \in \mathbb{R}^n_{\ge 0}$.

Proof. This is a simple consequence of the local Lipschitz property. On each ball $\mathcal{B}(R) = \{\xi \mid |\xi| \le R\}$, we pick the smallest common Lipschitz constant $M_0(R)$ for $A(\cdot)$, $B(\cdot)$, and $C(\cdot)$. The function M_0 is nondecreasing, and hence can be majorized by a continuous and nondecreasing function M. Since $\xi \in \mathcal{B}(|\xi|)$, we have that $|A(\xi) - A(0)| \le M(|\xi|)\,|\xi|$, and similarly for B and C. ∎

Corollary 3.4. For each function $x \in \mathcal{L}^2 \bigcap \mathcal{L}^\infty$:

$$\|A(x(\cdot)) - A(0)\| \le M(\|x\|_\infty)\,\|x\|$$

$$\|B(x(\cdot)) - B(0)\| \le M(\|x\|_\infty)\,\|x\|$$

$$\|C(x(\cdot)) - B(0)\| \le M(\|x\|_\infty)\,\|x\|$$

where M is as in Lemma 3.3.

Proof. We have:

$$\begin{aligned}
\|A(x(\cdot)) - A(0)\|^2 &= \int_0^\infty |A(x(s)) - A(0)|^2\, ds \\
&\le \int_0^\infty M(|x(s)|)^2\,|x(s)|^2\, ds \\
&\le \int_0^\infty M(\|x\|_\infty)^2\,|x(s)|^2\, ds \\
&= M(\|x\|_\infty)^2 \int_0^\infty |x(s)|^2\, ds \\
&= M(\|x\|_\infty)^2\,\|x\|^2
\end{aligned}$$

and similarly for B and C. ∎

4. PROOF OF THE MAIN RESULT

Pick any input $u \in \mathcal{L}^2$ and consider once again the respective solutions $x = Tu$ and $z = Lu$. By Corollary 3.2, we know that both $\|x\| \le c\,\|u\|$ and $\|x\|_\infty \le c\,\|u\|$. Therefore, using Corollary 3.4, we also have that:

$$\|A(x(\cdot)) - A(0)\| \le cM(c\,\|u\|)\,\|u\|$$

$$\|B(x(\cdot)) - B(0)\| \le cM(c\,\|u\|)\,\|u\|$$

$$\|C(x(\cdot)) - C(0)\| \le cM(c\,\|u\|)\,\|u\|$$

where M is as in Lemma 3.3. Let $\varphi : \mathbb{R}_{\ge 0} \to \mathbb{R}^n_{\ge 0}$ be the function $\varphi(t) :=$

$$(A(0) - A(x(t)))\,x(t) \;+\; (B(0) - B(x(t)))\,u(t)\,.$$

By the Cauchy-Schwartz inequality,

$$\begin{aligned}
\|(A(x(\cdot)) - A(0))\,x(\cdot)\| &\le \|A(x(\cdot)) - A(0)\|\,\|x\| \\
&\le c^2 M(c\,\|u\|)\,\|u\|^2
\end{aligned}$$

and

$$\begin{aligned}
\|(B(x(\cdot)) - B(0))\,u(\cdot)\| &\le \|B(x(\cdot)) - B(0)\|\,\|u\| \\
&\le cM(c\,\|u\|)\,\|u\|^2
\end{aligned}$$

from which we conclude that

$$\|\varphi\| \le \gamma(\|u\|)\,\|u\|$$

with $\gamma(r) = (c^2 + c)M(cr)r$, and γ is a function of class $\mathcal{K}$, i.e. continuous, strictly increasing, and with $\gamma(0) = 0$.

Consider the difference $w(t) = z(t) - x(t)$. Note that $w(0) = 0$. Evaluating $\dot{w} = [A(0)z + B(0)u] - [A(x)x + B(x)u]$ and rearranging terms,

$$\dot{w}(t) = A(0)w(t) + \varphi(t)\,.$$

Using once again that $A(0)$ is a Hurwitz matrix, we know that, for some constant $d \ge 0$ which depends only on $A(0)$ and not on the particular input u being used, $\|w\| \leq d\,\|\varphi\|$. Therefore, $\|w\| \leq \gamma(\|u\|)\,\|u\|$, after redefining $\gamma(r) := d\gamma(r)$.

In terms of the outputs $y = T_o u = C(x)x$ and $v = L_o u = C(0)z$,

$$
\begin{aligned}
\|v - y\| &= \|C(0)z - C(x(\cdot))x\| \\
&\leq \|C(0)(z-x)\| + \|(C(0) - C(x(\cdot))x\| \\
&\leq |C(0)|\,\|z - x\| + \|C(0) - C(x(\cdot))\|\,\|x\| \\
&\leq |C(0)|\,\gamma(\|u\|)\,\|u\| + c^2 M(c\,\|u\|)\,\|u\|^2
\end{aligned}
$$

and we can again write the last term as $\gamma(\|u\|)\,\|u\|$ if we redefine $\gamma(r) := |C(0)|\,\gamma(r) + c^2 M(cr)r$.

The triangle inequality gives us that $\|Lu\| - \|Tu\| \leq \|Lu - Tu\|$ and $\|L_o u\| - \|T_o u\| \leq \|L_o u - T_o u\|$, and Corollary 3.2 gives $\|Tu\| \leq \|Lu\|$ and $\|T_o u\| \leq \|L_o u\|$, so we may summarize as follows:

Proposition 4.1. There is a function $\gamma \in \mathcal{K}$ such that

$$0 \leq \|Lu\| - \|Tu\| \leq \gamma(\|u\|)\,\|u\|$$

and

$$0 \leq \|L_o u\| - \|T_o u\| \leq \gamma(\|u\|)\,\|u\|$$

for any input $u \in \mathcal{L}^2$. □

To conclude the proof of Theorem 1, we must show that $\|T_o\| \geq \|L_o\|$. Let $g = \|L_o\|$, and pick a minimizing sequence u_n, $n = 1, 2, \ldots$ of nonzero inputs in $\mathcal{L}^2$, that is,

$$\lim_{n\to\infty} \frac{\|L_o u_n\|}{\|u_n\|} = g\,.$$

Pick a sequence of real numbers $\varepsilon_n > 0$ such that $v_n := \varepsilon_n u_n \to 0$ (for example, $\varepsilon_n = (n\,\|u_n\|)^{-1}$). Since L_o is a linear operator, $\|L_o v_n\| = \varepsilon_n\,\|L_o u_n\|$, and since $\|v_n\| = \varepsilon_n\,\|u_n\|$, also $\|L_o v_n\|/\|v_n\| = \|L_o u_n\|/\|u_n\|$. Applying the second inequality in Proposition 4.1:

$$0 \leq \frac{\|L_o v_n\|}{\|v_n\|} - \frac{\|T_o v_n\|}{\|v_n\|} \leq \gamma(\|v_n\|) \to 0$$

which gives that $\frac{\|T_o v_n\|}{\|v_n\|} \to g$, and therefore $\|T_o\| \geq g$, as desired. □

5. POSITIVE VS. ARBITRARY INPUTS

We have shown that the norm of the nonlinear system (1) can be exactly computed by finding the norm of the associated linear system (2). The computation of induced $\mathcal{L}^2$ norms for linear systems is a classical area of study, and amounts to the maximization, over the imaginary axis, of the largest singular value of the transfer matrix of the system (the Laplace transform of the impulse response), the H_∞ norm; see for instance (Doyle *et al.*, 1992). There is, however, a potential gap in the application of this theory to our problem, namely, the usual definition of H_∞ norm corresponds to maximization over *arbitrary* inputs $u \in \mathcal{L}^2([0,\infty), \mathbb{R}^m)$, not necessarily inputs with values in $\mathbb{R}^m_{\geq 0}$ as considered in this paper. We close this gap now, by showing that the same result is obtained, for systems (2), whether one optimizes over arbitrary or over nonnegative inputs. We give two proofs, one elementary and the other one less trivial but leading to a stronger conclusion.

The positivity assumptions imply that the operator L_o is a nonnegative convolution operator:

$$(L_o u)(t) = \int_0^t W(t-s)u(s)\,ds\,, \qquad (4)$$

$$W(t) \in (\mathbb{R}_{\geq 0})^{p\times m}\ \forall\, t \geq 0\,. \qquad (5)$$

Here $W(t) = C(0)e^{tA(0)}B(0)$, and its nonnegativity follows from the fact that e^{tF} has all entries nonnegative, provided that $F_{ij} \geq 0$ for all $i \neq j$. (This last fact is well-known: it is clear for small t from the expansion $e^{tF} = I + tF + o(t)$, and for large t by then writing e^{tF} as a product of matrices $e^{(t/k)F}$ with the positive integer k large enough.) We next show that any operator as in (4-5) has the same norm whether viewed as an operator on $\mathcal{L}^2([0,\infty), \mathbb{R}^m)$ or on $\mathcal{L}^2([0,\infty), \mathbb{R}^m_{\geq 0})$. Since the norm as an operator on nonnegative inputs is, obviously, upper bounded by the norm on arbitrary inputs, it will be enough to show that, for each $w \in \mathcal{L}^2([0,\infty), \mathbb{R}^m)$, there is another input $\tilde{w} \in \mathcal{L}^2([0,\infty), \mathbb{R}^m_{\geq 0})$ with $\|w\| = \|\tilde{w}\|$ and $\|L_o w\| \leq \|L_o \tilde{w}\|$.

Given such a w, we start by writing $w = u - v$, where u and v are picked in $\mathcal{L}^2([0,\infty), \mathbb{R}^m_{\geq 0})$ and orthogonal. (Such a decomposition is always possible. We define coordinatewise, for each $i = 1, \ldots, m$, $u_i := \max\{w_i, 0\}$ and $v_i := \max\{-w_i, 0\}$; clearly, $w = u - v$. The supports of u_i and v_i are disjoint, so $\langle u_i, v_i\rangle = \int_0^\infty u_i(t)v_i(t)\,dt = 0$ for each i, and also then $\langle u, v\rangle = \sum_{i=1}^m \langle u_i, v_i\rangle = 0$.) We now let $\tilde{w} := u + v$. Since u and v (or $-v$) are orthogonal, $\|w\|^2 = \|u\|^2 + \|-v\|^2 = \|u\|^2 + \|v\|^2 = \|\tilde{w}\|^2$, so $\|w\| = \|\tilde{w}\|$. Because L_o is nonnegative, both $x = L_o u$ and $y = L_o v$ are nonnegative. To finish the proof, we only need to see that $\|x - y\| \leq \|x + y\|$:

$$
\begin{aligned}
\|x - y\|^2 &= \int_0^\infty \textstyle\sum_{i=1}^p (x_i(t) - y_i(t))^2\,dt \\
&= \int_0^\infty \textstyle\sum_{i=1}^p (x_i(t)^2 + y_i(t)^2 - 2x_i(t)y_i(t))\,dt
\end{aligned}
$$

$$\leq \int_0^\infty \sum_{i=1}^p (x_i(t)^2 + y_i(t)^2 + 2x_i(t)y_i(t))\, dt$$
$$= \int_0^\infty \sum_{i=1}^p (x_i(t) + y_i(t))^2\, dt$$
$$= \|x + y\|^2 \ .$$

A different proof, which in fact also implies that the supremum in the definition of norm is achieved as a maximum, is as follows. We consider the adjoint L_o^* of L_o (seen as an operator on the Hilbert space $\mathcal{L}^2([0,\infty),\mathbb{R}^m)$), and the composition $M = L_o^* L_o : \mathcal{L}^2([0,\infty),\mathbb{R}^m) \to \mathcal{L}^2([0,\infty),\mathbb{R}^m)$. The operator M is self-adjoint and (since L_o is a convolution operator with an $\mathcal{L}^2$ kernel) compact. Its spectrum consists of real and nonnegative eigenvalues, and its largest eigenvalue λ is such that $\mu = \sqrt{\lambda}$ is the largest singular value of L_o, and equals the norm of L_o as an operator $\mathcal{L}^2([0,\infty),\mathbb{R}^m) \to \mathcal{L}^2([0,\infty),\mathbb{R}^p)$. Take any eigenvector u corresponding to λ, so $Mu = \lambda u$. It follows that $\|L_o u\|^2 = \langle L_o u, L_o u\rangle = \langle u, Mu\rangle = \langle u, \lambda u\rangle = \mu^2 \|u\|^2$, so u is a maximizing vector for L_o. Moreover, for a compact positive operator M on a Hilbert space, the Krein-Rutman Theorem says that, provided that there is a nonzero eigenvalue (which there is in this case, since M is self-adjoint and we may assume without loss of generality that $M \neq 0$), then the maximal eigenvalue λ admits a nonnegative eigenvector u. Thus $\|L_o u\|$ is maximized at this $u \in \mathcal{L}^2([0,\infty),\mathbb{R}^m_{\geq 0})$.

6. CASCADES

Signaling systems are often built by cascading subsystems, so it is interesting to verify that a cascade of any number of systems which satisfy our properties again has the same form. It is enough, by induction, to show this for two cascaded systems

$$\dot{x} = A_1(x)x + B_1(x)u \quad v = C_1(x)x$$
$$\dot{z} = A_2(z)z + B_2(z)\tilde{u} \quad y = C_2(z)z$$

each of which satisfies our assumptions, under the series connection obtained by setting $\tilde{u} = v$. The composite system can be represented in terms of the following $A(\xi,\zeta)$ and $B(\xi,\zeta)$ matrices:

$$A = \begin{pmatrix} A_1(\xi) & 0 \\ B_2(\zeta)C_1(\xi) & A_2(\zeta) \end{pmatrix}, \ B = \begin{pmatrix} B_1(\xi) \\ 0 \end{pmatrix}$$

and output y.

It is easy to verify all the necessary properties. For example, the only nontrivial part of the maximization property amounts to checking that $B_2(\zeta)C_1(\xi) \leq B_2(0)C_1(0)$, which follows from $B_2(\zeta)C_1(\xi) \leq B_2(0)C_1(\xi)$ (using the maximization property for B_2 and the positivity of C_1) and $B_2(0)C_1(\xi) \leq B_2(0)C_1(0)$ (using maximization for C_1 and positivity of $B_2(0)$). Similarly, the only nontrivial part of the positivity property involves checking that $(B_2(\zeta)C_1(\xi))_{ij} \geq 0$ provided that $\zeta_i = 0$, for all j. But, for such a vector ζ, we know that $B_2(\zeta)_{ik} \geq 0$ for all k, so indeed $\sum_k B_2(\zeta)_{ik} C_1(\xi)_{kj} \geq 0$.

7. REMARKS AND CONCLUSIONS

We provided a way to compute, for systems of a special form, the induced $\mathcal{L}^2$ norm of the system. The special form includes a variety of cellular signaling cascade systems. An even wider class of systems can be included as well, provided that one extend our treatment to systems that are monotone with respect to orders other than that given by the first quadrant. Such orders have proven useful in analyzing, for example, MAPK cascades, see for example (Angeli and Sontag, 2003; Angeli *et al.*, 2004). The details of this extension will be provided elsewhere.

REFERENCES

D. Angeli, J.E. Ferrell, Jr., and E.D. Sontag, Detection of Multi-Stability, Bifurcations, and Hysteresis in a Large Class of Biological Positive-Feedback Systems, Proceedings of the National Academy of Sciences USA 101(2004) 1822-1827.

D. Angeli and E.D. Sontag, Monotone Control Systems, IEEE Trans. Autom. Control, 48(2003): 1684-1698.

M. Chaves, E.D. Sontag, and R. Dinerstein, Optimal Length and Signal Amplification in Weakly Activated Signal Transduction Cascades, J. Physical Chemistry, to appear.

J. Doyle, B. Francis and A. Tannenbaum, Feedback Control Systems, MacMillan Publishing Co, 1992. Also available on the web at: `http://www.control.utoronto.ca/-people/profs/francis/dft.html` .

R. Heinrich, B.G. Neel and T.A. Rapoport, Mathematical Models of Protein Kinase Signal Transduction, Molecular Cell, 9(2002) 957-970.

V. Lakshmikantham and S. Leela. Differential and Integral Inequalities, Academic Press, N.Y., 1969.

H.L. Smith, Monotone Dynamical Systems, AMS, Providence, 1995.

E.D. Sontag, Mathematical Control Theory: Deterministic Finite Dimensional Systems, (2^{nd} edition), Springer, NY, 1998.