

# Compositional Analysis of Autocatalytic Networks in Biology

Gentian Buzi, Ufuk Topcu, and John C. Doyle

**Abstract**—Autocatalytic pathways are a necessary part of core metabolism. Every cell consumes external food/resources to create components and energy, but does so using processes that also require those same components and energy. Here, we study effects of parameter variations to the stability properties autocatalytic pathway models and the extent of the regions of attraction around nominal operating conditions. Motivated by the computational complexity of optimization-based methods for estimating regions of attraction for large pathways, we take a compositional approach and exploit a natural decomposition of the system, induced by the underlying biological structure, into a feedback interconnection of two input-output subsystems: a small subsystem with complicating nonlinearities and a large subsystem with simple dynamics. This decomposition simplifies the analysis of large pathways by assembling region of attraction certificates based on the input-output properties of the subsystems. It enables us to numerically construct block-diagonal Lyapunov functions for families of pathways that are not amenable to direct analysis. Furthermore, it leads to analytical construction of Lyapunov functions for a large family of autocatalytic pathways.

## I. INTRODUCTION

Autocatalytic pathways are a necessary part of core metabolism. Every cell consumes external food/resources to create components and energy, but does so using processes that also require those same components and energy. Indeed all biological networks viewed at a large enough scope are massively autocatalytic. In this paper, we study the metabolic network with the topology of Fig. 1. This network is composed of a chain of reactions that convert one metabolite to another, ultimately producing multiple copies of the final product of the pathway. The pathway also requires the consumption of copies of its final product to convert its input into the first metabolite. All the reactions are catalyzed by enzymes. Additionally, the product of the pathway inhibits the enzyme that catalyzes the autocatalytic (first) reaction. An example of such a pathway is the glycolysis pathway [1], [2], [3].

Biochemical networks with different topologies and different reaction rate constraints have been a subject of study for many decades. General results regarding the number of the possible steady states and the convergence properties for certain networks with mass-action kinetics (the so-called networks of deficiency zero and one) have been established [4], [5], [6]. The theory of monotone dynamical systems [7] has proven a powerful tool for understanding the behavior of biological systems. Using this theory, certain network topologies have been shown to have global convergence

to steady states for quite general reaction rates [8]. Cyclic interconnection networks are closer to the topology shown in Fig. 1 and for these networks a sufficient stability criterion for cyclic interconnection with inhibition of the first reaction by the end product of the network has been established [9], [10]. These results are extended to prove global asymptotic stability using diagonal stability and passivity of subsystems if certain criteria are satisfied by the reaction rates [11], [12].

In most cases, the difficulty associated with analyzing autocatalytic networks, with the topology and feedback structure of Fig. 1, is that they do not exhibit such global properties. These networks can have multiple equilibrium points, each with large regions of attraction (RoA). The nonlinear nature of the autocatalytic reaction rate makes analysis of behavior due to non-infinitesimal perturbations from nominal operating conditions, such as estimating regions of attraction, difficult. This nonlinearity is caused by the coupling of positive and negative feedback, which induces many of the interesting properties in these networks. Depending on the feedback gain, the (stable) dynamics near the equilibrium points can be dominated by either positive or negative feedback. In these autocatalytic networks, changes in the concentration of metabolites and catalyzing enzymes during the lifetime of the cell can perturb the system from the nominal operating point of the pathway. Can the controller restore the system to normal operating conditions from these perturbations? We investigate effects of such perturbations through the estimation of invariant subsets of the RoA around nominal operating conditions (i.e., a measure of the set of perturbations from which the cell recovers). In previous work [3], we used a numerical procedure, composed of system theoretic characterizations and optimization-based formulations, for estimating the RoA in such networks with applications to the glycolytic pathway. This procedure is effective on relatively small pathways, however it becomes impractical for large ones (due to computational complexity). However, the underlying biological structure offers a natural decomposition of the system into a feedback interconnection of two input-output subsystems: a small subsystem with complicating nonlinearities and a large subsystem with simple dynamics. This decomposition simplifies the analysis of large pathways through a compositional framework for RoA analysis based on the input-output properties of the subsystems. It enables us to numerically construct block-diagonal Lyapunov functions (that provide algebraic descriptions of invariant subsets of the RoA) for families of pathways that are not amenable to direct analysis. Furthermore, it leads to analytical construction of Lyapunov functions for a large family of autocatalytic pathways.

The authors are with the Control and Dynamical Systems, California Institute of Technology, Pasadena, CA, 91125, USA `genti`, `utopcu`, `doyle@cds.caltech.edu`

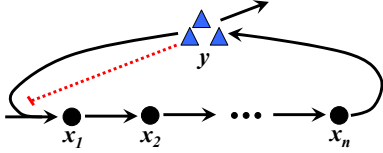


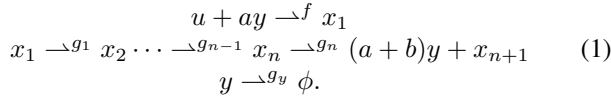
Fig. 1. The autocatalytic network is composed of a chain of reactions (black arrows) that convert one metabolite (black circles) to another. The pathway requires the consumption of the final product (blue triangles) to convert the input of the pathway into the first metabolite. Additionally, the product of the pathway inhibits (dotted red line) the enzyme that catalyzes the autocatalytic reaction.

In this paper, we omit the proofs of the results due to space considerations. See [13], [14] for more details on the proofs and for more general network topologies that allow for the inclusion of the degradation of intermediate metabolites as well as the reversibility of the chemical reactions.

The rest of the paper is organized as follows: We introduce an autocatalytic pathway model in section II followed by linear stability analysis in section III. A compositional analysis framework is introduced in section IV. This framework is used to analytically construct Lyapunov functions for a family of autocatalytic pathways in sections IV-A and IV-B. The scope and limitations of this framework are discussed in section IV-C and its generalization is discussed in section V.

## II. A MODEL FOR AUTOCATALYTIC PATHWAYS

Consider the autocatalytic metabolic pathway with multiple intermediate metabolite reactions (see Fig. 1)



Here,  $u$  is some precursor and source of energy for the pathway with no dynamics associated,  $y$  denotes the product of the pathway,  $x_i$  are intermediate metabolites,  $\phi$  is a null state,  $a$  is the number of  $y$  molecules that are invested in the pathway, and  $a+b$  is the number of  $y$  molecules produced.  $A \xrightarrow{f} B$  denotes a chemical reaction that converts the chemical species  $A$  to the chemical species  $B$  at rate  $f$ . The corresponding ordinary differential equations in terms of the stoichiometry matrix and reaction fluxes are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \\ \dot{y} \end{bmatrix} = S \begin{bmatrix} f(y) \\ g_1(x_1) \\ \vdots \\ g_{n-1}(x_{n-1}) \\ g_n(x_n) \\ g_y(y) \end{bmatrix} \quad (2)$$

for  $x_i \geq 0$ ,  $y \geq 0$ , where

$$S := \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & \ddots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & -1 \\ -a & 0 & \cdots & 0 & b+a & -1 \end{bmatrix}$$

In (2),  $g_1, \dots, g_n$ , and  $g_y$  are continuous, monotone, increasing functions that vanish at 0. These assumptions on  $g_i$  and  $g_y$  are consistent with large classes of chemical kinetics models such as mass-action and Michaelis-Menten. The rate  $f(y)$  of the autocatalytic reaction captures the negative feedback of the output ( $y$ ) via the inhibition of the catalyzing enzyme of the first reaction. We choose  $f(y) = \frac{Vy^q}{1+\gamma y^h}$  that is consistent with biological intuition and experimental data in the case of the glycolysis pathway [15], [1]. In this parameterization  $V > 0$  depends on the concentration of the input ( $u$ ),  $q > 0$  captures the strength of autocatalysis and  $\gamma > 0$  and  $h > 0$  capture the strength of inhibition (negative feedback). Note that  $f$  is not monotone and captures the coupling between the autocatalysis and inhibition. For the rest of the paper, we take  $a = b = 1$  and note that the generalization of the results for  $a > 0$  and  $b > 0$  is straightforward.

The system (2) can have different number of points in the positive orthant, but we will examine the behavior of the system at or around the fixed point of interest  $\bar{\xi} = (\bar{x}_1, \dots, \bar{x}_n, \bar{y}) \succ \mathbf{0}$ . Without loss of generality, we use the normalization  $\bar{y} = 1$  which implies that  $V = g_y(\bar{y})(1 + \gamma)$  and perform a change of coordinates to move the fixed point  $\bar{\xi}$  to the origin. Then, with some abuse of notation, the vector field can be written as

$$\begin{aligned}
 \dot{x}_1 &= f(y) - g_1(x_1) \\
 \dot{x}_2 &= g_1(x_1) - g_2(x_2) \\
 &\vdots \\
 \dot{x}_n &= g_{n-1}(x_{n-1}) - g_n(x_n) \\
 \dot{y} &= 2g_n(x_n) - f(y) - g_y(y)
 \end{aligned} \quad (3)$$

in  $\mathcal{D} := \{\xi \in \mathbb{R}^{n+1} : \xi \succeq -\bar{\xi}\}$  where  $f(y) = \frac{g_y(\bar{y})(1+\gamma)(y+1)^q}{1+\gamma(y+1)^h} - g_y(\bar{y})$ . Note that  $g_1, \dots, g_n$ , and  $g_y$  in (3) are continuous, monotone, increasing and to vanish at 0 (since they are shifted and translated versions of the original reaction rates).

## III. LINEAR STABILITY ANALYSIS

Define  $\hat{h} := \frac{\gamma}{1+\gamma}h$ ,  $k_i := \frac{\partial g_i}{\partial x_i}|_0 > 0$ ,  $k_y := \frac{\partial g_y}{\partial y}|_0 > 0$ ,  $r_y := g_y(0) > 0$ . The linearization of (3) around the origin is given by  $[\dot{\tilde{x}}^T \dot{\tilde{y}}^T]^T = J[\tilde{x}^T \tilde{y}^T]^T$  where  $J$  is

$$J = \begin{bmatrix} -k_1 & 0 & \cdots & 0 & 0 & r_y(q - \hat{h}) \\ k_1 & -k_2 & & 0 & 0 & 0 \\ & \ddots & \ddots & & \vdots & \vdots \\ 0 & 0 & & -k_{n-1} & 0 & 0 \\ 0 & 0 & & k_{n-1} & -k_n & 0 \\ 0 & 0 & \cdots & 0 & 2k_n & -k_y - r_y(q - \hat{h}) \end{bmatrix}.$$

*Remark 1:* Define  $J_0 := J|_{\hat{h}=0}$  and note that the linearization of (3) can be viewed as the closed-loop dynamics of

$$\begin{aligned}
 \dot{\tilde{\xi}} &= J_0 \tilde{\xi} + r_y [1 \ 0 \ \cdots \ 0 \ -1]^T u \\
 \eta &= [0 \ \cdots \ 0 \ 1]^T \tilde{\xi}.
 \end{aligned}$$

using a proportional (negative) feedback controller of gain  $\hat{h}$  (i.e.,  $u = -\hat{h}\eta$ ).

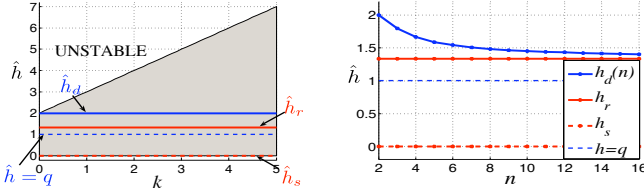


Fig. 2. **Left:** Illustration of the stability bounds ( $\hat{h}_d$ ,  $\hat{h}_s$ , and  $\hat{h}_r$ ). The grey area marks the set of parameters that define stable systems. The area between the blue lines corresponds to systems that satisfy the secant condition. **Right:** As the pathway size increases, the upper bound on the feedback strength  $\hat{h}$  given by the secant condition gets smaller.

For  $\hat{h} > q$ , stability of the linearized dynamics can be studied using the so-called secant condition [9], [10]. To this end, for a given  $n > 1$  define  $\theta(n) := (\sec(\pi/(n+1)))^{n+1}$ ,  $\hat{h}_d(n) := q + \frac{k_y}{r_y} \frac{\theta(n)}{2+\theta(n)}$ , and  $\hat{h}_d(1) := q + \frac{k_y}{r_y}$ . Then,  $J$  is Hurwitz for  $q < \hat{h} < \hat{h}_d(n)$ . Furthermore, if  $k_1 = k_2 = \dots = k_n = k_y + r_y(q - \hat{h})$ , then  $J$  is Hurwitz only if  $\hat{h} < \hat{h}_d(n)$  [9], [10]. The upper bound  $\hat{h}_d(n)$  depends on the pathway size  $n$ , consumption rates  $k_y$  and  $r_y$  of the pathway output  $y$ , and the autocatalysis parameter  $q$ , but not on the rate  $k_i$  of the intermediate reactions. Note that as  $n$  increases, the maximum allowable stabilizing feedback gain decreases. For general feedback gain  $\hat{h}$  (i.e., removing the restriction  $\hat{h} > q$ ), we establish a different set of bounds.

**Proposition 1:** Let  $\hat{h}_r := q + \frac{1}{3} \frac{k_y}{r_y}$  and  $\hat{h}_s := q - \frac{k_y}{r_y}$ . If  $\hat{h}_s < \hat{h} \leq \hat{h}_r$ , then  $J$  is Hurwitz for arbitrary size  $n$  and arbitrary values of intermediate reaction rates  $k_i$ . The bounds above are tight, i.e., for any value of gain  $\hat{h} \notin (\hat{h}_s, \hat{h}_r]$ , one can construct an unstable pathway of appropriate size.  $\triangleleft$

For general stoichiometry of autocatalysis, i.e.,  $a, b > 0$  (instead of  $a = b = 1$ ), Proposition 1 holds for  $\hat{h}_r = q + \frac{b}{2a+b} \frac{k_y}{r_y}$  and  $\hat{h}_s = q - \frac{k_y}{r_y}$ . Note that  $\hat{h}_r = \lim_{n \rightarrow \infty} \hat{h}_d(n)$  (see Fig. 2 (right)). Both the secant condition bound  $\hat{h}_d$  and  $\hat{h}_r$  can be conservative for specific pathways, especially for those with high intermediate reaction rates (as illustrated in Fig. 2 (left) for  $n = 1, q = 1, k_y = 1, r_y = 1$ ). Later in this paper, we will revisit these bounds.

#### IV. A SIMPLE DECOMPOSITION

Having established conditions on the linear stability of system (3), we now study the RoA properties of these dynamics. To this end, we decompose the system (3) into 2 subsystems  $S_1$  and  $S_2$  governed by

$$\begin{aligned} \dot{x}_1 &= z - g_1(x_1) \\ \dot{x}_2 &= g_1(x_1) - g_2(x_2) \\ &\vdots \\ \dot{x}_n &= g_{n-1}(x_{n-1}) - g_n(x_n) \\ w &= g_n(x_n) \end{aligned} \quad (4)$$

and

$$\begin{aligned} \dot{y} &= 2w - f(y) - g_y(y) \\ z &= f(y) \end{aligned} \quad (5)$$

and interconnected as shown in Fig. 3.  $S_1$  and  $S_2$  are single-input-single-output systems with  $n$  and 1 states, respectively. Note that  $S_1$  is composed of a chain of chemical reactions and the reactions rates  $g_i$  are monotone functions. Because

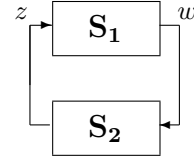


Fig. 3. The system (3) is represented as feedback interconnection of 2 subsystems  $S_1$  and  $S_2$ , given by (4) and (5).

of its special structure, it is relatively easy to establish its input-output properties for even large values of  $n$ . On the other hand,  $S_2$  captures the most dominant nonlinearity: the dynamics of  $y$  and its involvement in autocatalysis and feedback regulation  $f$ . Using this decomposition, we have separated the complexity of high dimensional states (in  $S_1$ ) from the complexity of the important nonlinearity of autocatalysis control (in  $S_2$ ).

**Definition 1:** A continuously differentiable function  $U : \mathbb{R}^m \rightarrow \mathbb{R}$  is positive definite in a neighborhood  $\mathcal{N}$  of the origin if  $U(0) = 0$  and for all nonzero  $x \in \mathcal{N}$ ,  $U(x) > 0$ . For such  $U$  and  $\beta \in \mathbb{R}$ , define the set  $\Omega_{U,\beta}$  to be the connected component of  $\{x \in \mathbb{R}^m : U(x) \leq \beta\}$  containing the origin.

Reference [16] proposes a method for computing invariant subsets of the RoA of the asymptotically stable equilibrium points of systems with unmodeled dynamics, where the unmodeled dynamics are required to satisfy certain gain relations or dissipation inequalities. We apply this idea to the feedback interconnection of Fig. 3 to compute invariant subsets of the RoA for (3). For notational simplicity, let us rewrite the dynamics in (4) and (5), respectively, as

$$\dot{x} = F_1(x, z), \quad w = G_1(x), \quad (6)$$

$$\dot{y} = F_2(y, w), \quad z = G_2(y). \quad (7)$$

**Proposition 2:** Let  $\delta, \kappa \in \mathbb{R}$ . Let  $U_1 : \mathbb{R}^n \rightarrow \mathbb{R}$  be a continuously differentiable, positive definite function such that

$$\nabla U_1(x) \cdot F_1(x, z) \leq z^2 + 2\delta wz - \kappa w^2, \quad \forall (x, z) \quad (8)$$

and  $U_2 : \mathbb{R} \rightarrow \mathbb{R}$  be a continuously differentiable, positive definite in some neighborhood  $\mathcal{N}_0$  of the origin such that

$$\nabla U_2(y) \cdot F_2(y, w) \leq \kappa w^2 - 2\delta wz - z^2, \quad \forall w \text{ and } y \in \mathcal{N}_0. \quad (9)$$

Define  $U(x, y) := U_1(x) + U_2(y)$  and let  $\beta$  be such that

$$\Omega_{U,\beta} \subset \mathbb{R}^n \times \mathcal{N}_0 \quad (10)$$

and  $\Omega_{U,\beta}$  is bounded. Then, the following statements hold:  $\Omega_{U,\beta}$  is an invariant subset under the flow of (3). If one of the inequalities (8) and (9) is strict (i.e., equality is satisfied only at  $(x, z) = (0, 0)$  and  $(y, w) = (0, 0)$  respectively), then  $\Omega_{U,\beta}$  is a subset of the RoA of the origin for (3).  $\triangleleft$

Proposition 2 constructs RoA certificates for the overall system (3) based on the certificates for the input-output properties of the subsystems  $S_1$  and  $S_2$  established using the dissipation inequalities in (8) and (9), respectively. The quadratic function  $z^2 + 2\delta wz - \kappa w^2$  in (8) is called a *supply rate* for  $S_1$  (similarly its negative is used as a supply rate in (9) for  $S_2$ ). One diversion from the classical dissipation

inequalities literature [17] is that the inequality in (9) is local, i.e., it is supposed to hold in certain bounded subsets (for example in certain sublevel sets of associated storage functions) of the state-space but not necessarily the whole state space [16]. Using special forms of the inequalities (8) and (9), we show that for specific ranges of feedback gain  $\hat{h}$ , we can analytically find a diagonal function  $U$  that satisfy the conditions in Proposition 2 (consequently,  $\Omega_{U,\beta}$  is an invariant subset of the RoA).

#### A. Region of attraction estimation via a local small-gain type condition

Consider the case where  $\delta = 0$  and  $\kappa = 1$  in (8) and (9). That is, we search for a positive definite function  $U_1 : \mathbb{R}^n \rightarrow \mathbb{R}$  such that

$$\nabla U_1(x) \cdot F_1(x, z) \leq z^2 - w^2, \quad \forall (x, z) \quad (11)$$

and another function  $U_2 : \mathbb{R} \rightarrow \mathbb{R}$ ,  $U_2(y)$  positive definite in some neighborhood  $\mathcal{N}_0$  of the origin such that

$$\nabla U_2(y) \cdot F_2(y, w) < w^2 - z^2, \quad \forall w \text{ and } \forall y \in \mathcal{N}_0 \setminus \{0\}. \quad (12)$$

*Lemma 1:* The function  $U_1 : \mathbb{R}^n \rightarrow \mathbb{R}$

$$U_1(x) := \sum_{i=1}^n 2 \int_0^{x_i} g_i(\xi) d\xi, \quad (13)$$

is continuously differentiable, positive definite and satisfies (11).  $\triangleleft$

*Lemma 2:* If  $\hat{h}_s < \hat{h} < \hat{h}_r$ , then there exists a nonempty neighborhood  $\mathcal{N}_0$  of the origin such that the function  $U_2 : \mathbb{R} \rightarrow \mathbb{R}$

$$U_2(y) := \frac{1}{2} \int_0^y (f(\xi) + g_y(\xi)) d\xi \quad (14)$$

is continuously differentiable, positive definite in  $\mathcal{N}_0$  and satisfies (12).  $\triangleleft$

Define

$$D_y := \left\{ y : (g_y(y) - f(y)) \left( f(y) + \frac{1}{3}g_y(y) \right) \geq 0 \right\},$$

the set of points  $y$  for which the graph of  $f$  lies between the graph of  $g_y$  and  $-\frac{1}{3}g_y$  (Fig. 4 (left)). The set  $\mathcal{N}_0$  in Lemma 2 is constructed as the largest sublevel set of  $U_2$  that is a subset of  $D_y$ .

*Proposition 3:* For  $\hat{h}_s < \hat{h} < \hat{h}_r$ , let  $U_1$  and  $U_2$  be given by (13) and (14), respectively, and  $\mathcal{N}_0$  be as in Lemma 2. Let  $U(x, y) = U_1(x) + U_2(y)$ . Then, for any  $\beta$  such that  $\Omega_{U,\beta} \subset \mathbb{R}^n \times \mathcal{N}_0$  and  $\Omega_{U,\beta}$  is bounded,  $\Omega_{U,\beta}$  is an invariant subset of the RoA of the origin for system (3).  $\triangleleft$

Proposition 3 follows from Proposition 2, Lemmas 1 and 2.

*Example 1:* Let  $g_i(x_i) = k_i x_i$ ,  $g_y(y) = y$ ,  $\gamma = \frac{3}{2}$ ,  $q = 1$ , and  $h = 2$ . Then,  $\hat{h} = \frac{6}{5}$ ,  $\hat{h}_r = \frac{4}{3}$ ,  $U_1(x) = k_1 x_1^2 + \dots + k_n x_n^2$ ,  $U_2(y) = \frac{5}{12} \log(5 + 6y + 3y^2) + \frac{1}{4}y^2 - \frac{1}{2}y - \frac{5}{12} \log 5$ ,  $U(x, y) = U_1(x) + U_2(y)$ , and  $\mathcal{N}_0 = \{y \mid y > -1\}$  (Fig. 4 (left)). The maximum sublevel set  $\beta$  of  $U$  that satisfies (10) is  $\beta = 0.3682$ . Fig. 4 (right) shows a slice of  $\Omega_{U,\beta}$  in the  $y - x_i$  subspace.

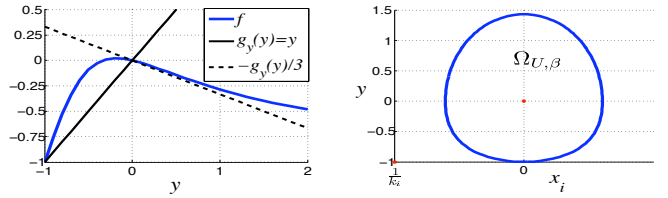


Fig. 4. **Left:**  $\mathcal{N}_0$  can be constructed as the set of points  $y$  where the graph of  $f$  lies between  $g_y$  and  $-\frac{1}{3}g_y$ . **Right:** The  $y - x_i$  slice of the invariant subset of RoA  $\Omega_{U,\beta}$  ( $x_j = 0$ ,  $j \neq i$ ).

#### B. Compositional estimation of the RoA using another dissipation inequality

Consider the case  $\kappa = 0$  in (8) and (9). That is, we search for a continuously differentiable, positive definite function  $U_1 : \mathbb{R}^n \rightarrow \mathbb{R}$  such that

$$\nabla U_1(x) \cdot F_1(x, z) \leq z^2 + 2\delta w z, \quad \forall (x, z) \quad (15)$$

and  $U_2 : \mathbb{R} \rightarrow \mathbb{R}$ , such that  $U_2$  is continuously differentiable, positive definite in some neighborhood  $\mathcal{N}_0$  of the origin and

$$\nabla U_2(y) \cdot F_2(y, w) < -z^2 - 2\delta w z, \quad \forall w \text{ and } \forall y \in \mathcal{N}_0 \setminus \{0\}. \quad (16)$$

Next we find  $U_1$  and  $U_2$  satisfying (15) and (16), respectively, for  $q < \hat{h} < \hat{h}_d(n)$  and a specific  $\delta$ . We first show that a diagonal function  $U_1$  exists for the linearization of (4) and  $0 < \delta \leq \frac{1}{2}\theta(n)$ , and then show that it exists for  $S_1$ . We treat the special case of the linearization separately since it outlines a procedure of finding  $U_1$  by solving a simple linear matrix inequality for the case when  $g_i(x_i) = k_i x_i$  (i.e., mass-action kinetics). The linearization of system (4) is given by

$$\begin{aligned} \dot{x} &= A_1 x + B_1 z \\ w &= C_1 x \end{aligned} \quad (17)$$

where  $C_1 = [0 \ \dots \ 0 \ k_n]$ ,

$$A_1 = \begin{bmatrix} -k_1 & 0 & \dots & 0 & 0 \\ k_1 & -k_2 & & 0 & 0 \\ \vdots & & \dots & & \vdots \\ 0 & 0 & & -k_{n-1} & 0 \\ 0 & 0 & \dots & k_{n-1} & -k_n \end{bmatrix}. \quad (18)$$

*Lemma 3:* For  $0 < \delta \leq \frac{1}{2}\theta(n)$ , there exists a positive, quadratic function  $U_1 : \mathbb{R}^n \rightarrow \mathbb{R}$  of the form  $U_1(x) = x^T P x$ , where  $P$  is a diagonal matrix satisfying

$$\nabla U_1(x)(A_1 x + B_1 z) \leq z^2 + 2\delta w z \quad \forall (x, z). \quad (19)$$

$\triangleleft$

*Remark 2:* Note that the dissipation inequality in (19) can be equivalently written as a linear matrix inequality.

*Lemma 4:* For  $0 < \delta \leq \frac{1}{2}\theta(n)$ , there exist positive real numbers  $d_1, \dots, d_n$  such that the function

$$U_1(x) := \sum_{i=1}^n d_i \int_0^{x_i} g_i(\xi) d\xi \quad (20)$$

satisfies (15).  $\triangleleft$

The proof of Lemma 4 builds on a result in [12, Corollary 3 and its proof].

*Lemma 5:* Let  $n \geq 1$  be a given integer,  $\delta = \theta(n)/2$ , and  $\hat{h}$  satisfy  $q \leq \hat{h} \leq \hat{h}_d(n)$ . Then, there exists a nonempty neighborhood  $\mathcal{N}_0$  of the origin such that the function  $U_2 : \mathbb{R} \rightarrow \mathbb{R}$

$$U_2(y) := -\delta \int_0^y f(\xi) d\xi \quad (21)$$

is continuously differentiable, positive definite in  $\mathcal{N}_0$ , and satisfies (16).  $\triangleleft$

In Lemma 5, the set  $\mathcal{N}_0$  can be constructed as the largest sublevel set of  $U_2$  that is a subset of

$$\left\{ y : -f(y) \left( f(y) + \frac{\delta}{1+\delta} g_y(y) \right) \geq 0 \right\},$$

the set of points  $y$  for which the graph of  $f$  lies in the areas between the graphs of  $\frac{\delta}{1+\delta} g_y$  and the map  $y \mapsto 0$ .

*Proposition 4:* Let  $n \geq 1$  be a given positive integer,  $U(x, y) := U_1(x) + U_2(y)$ , for  $U_1$  and  $U_2$  as defined in (20) and (21),  $\delta = \frac{1}{2}\theta(n)$ , and  $\mathcal{N}_0$  be as constructed above. Then, for  $q < \hat{h} < \hat{h}_d(n)$  and  $\beta$  such that  $\Omega_{U,\beta} \subset \mathbb{R}^n \times \mathcal{N}_0$  and  $\Omega_{U,\beta}$  is bounded,  $\Omega_{U,\beta}$  is an invariant subset of the RoA of the origin for system (3).  $\triangleleft$

Proposition 4 is a direct consequence of Proposition 2, Lemma 4, and Lemma 5.

Proposition 1 and the secant condition establish bounds  $\hat{h}_r$  and  $\hat{h}_d(n)$  on  $\hat{h}$  that guarantee stability of the pathway for arbitrary rates (and, in the case of  $\hat{h}_r$ , number) of intermediate reactions. Propositions 3 and 4 imply that systems with gains that obey those bounds are simple to analyze, i.e., estimates of RoA for such systems can be constructed through compositional analysis based on local dissipation inequalities. On the other hand, they do not provide any guarantees on how large these estimates are. In general the size of these subsets will depend on the properties of  $f$  and  $g_y$ . Given  $U$  as constructed above, one can search for its largest sublevel set that is contained in the set where  $U$  is decreasing along the vector field of (3). Another approach is to combine invariant subsets of the RoA, since if  $\Omega_{U,\beta_1}$  and  $\Omega_{W,\beta_2}$  are invariant subsets of RoA of the fixed point, then so is  $\Omega_{U,\beta_1} \cup \Omega_{W,\beta_2}$ . For example, if  $q \leq \hat{h} \leq h_r$ , then we can use  $\Omega_{U,\beta}$  and the Lyapunov functions that we get from Proposition 3 and 4 to improve the estimate.

### C. Existence of block diagonal Lyapunov functions

The decomposition of (3) into  $S_1$  and  $S_2$  provides a convenient way of searching for block diagonal Lyapunov functions, i.e we search for  $U(x, y) = U_1(x) + U_2(y)$ ,  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}$ . The next proposition examines the linearizations of  $S_1$  and  $S_2$  and shows that, if there is such a Lyapunov function for the linearization of system (3), then we can construct it by finding  $U_1$  and  $U_2$  satisfying (11) and (12), respectively. The linearization of system (4) is given by (17), and let the linearization of  $S_2$  be written as

$$\begin{aligned} \dot{y} &= a_2 y + 2w \\ z &= -\sigma y \end{aligned} \quad (22)$$

where  $\sigma = -\frac{\partial f}{\partial y}|_{y=0} = r_y(\hat{h} - q)$ ,  $a_2 = -k_y + \sigma$ . We rewrite the linearization of (3) as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = J \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A_1 & -\sigma B_1 \\ 2C_1 & a_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (23)$$

*Proposition 5:* If there exist positive definite matrices  $P_1 \in \mathbb{R}^{n \times n}$  and  $p_2 \in \mathbb{R}$  such that  $J^T P + P J$  is negative definite, where

$$P := \begin{bmatrix} P_1 & 0 \\ 0 & p_2 \end{bmatrix},$$

then the quadratic functions  $U_1(x) := -\frac{\sigma^2}{2a_2 p_2} x^T P_1 x$ ,  $U_2(y) := p_2 y^2$ , are continuously differentiable, positive definite and along with the constant  $\delta := \frac{\sigma}{k_y - \sigma}$  satisfy

$$\begin{aligned} \nabla U_1(x) \cdot (A_1 x + B_1 z) &< z^2 + 2\delta w z, \quad \forall x, z \\ \nabla U_2(y) \cdot (a_2 y + 2w) &\leq -z^2 - 2\delta w z, \quad \forall y, w. \end{aligned}$$

$\triangleleft$

Next, we state a partial converse result.

*Proposition 6:* For  $g_1, \dots, g_n$  such that  $\frac{\partial}{\partial x_i} g_i(x_i)|_0 = k$ , for  $i = 1, \dots, n$ , and for  $\hat{h} > \hat{h}_d(n)$  there is no quadratic block diagonal Lyapunov function of the form  $U(x, y) = U_1(x) + U_2(y)$  for the linearized dynamics in (23).  $\triangleleft$

Proposition 6 demonstrates a limitation of the analysis based on the simple decomposition of (3) given in (4) and (5). We next investigate how such limitations may be partially alleviated using more general decompositions of (3).

## V. GENERAL DECOMPOSITION

Consider that the system in (3) is decomposed into two input-output (sub)systems  $S_1$  and  $S_2$  with  $n_1$  and  $n_2$  states, respectively, as in Fig. 3, where  $n_1 + n_2 = n + 1$ . Let  $S_1$  and  $S_2$ , respectively, be governed by

$$\begin{aligned} \dot{\zeta} &= \tilde{F}_1(\zeta) + \tilde{B}_1 z \\ w &= \tilde{G}_1(\zeta) = g_n(x_n), \end{aligned} \quad (24)$$

$$\begin{aligned} \dot{\psi} &= \tilde{F}_2(\psi) + \tilde{B}_2 w \\ z &= \tilde{G}_2(\psi) = g_{n_2-1}(x_{n_2-1}), \end{aligned} \quad (25)$$

where  $\tilde{B}_1 = [1 \ 0 \ \dots \ 0]^T$ ,  $\tilde{B}_2 = [2 \ 0 \ \dots \ 0]^T$ , and the correspondence between the original states  $x_1, \dots, x_n, y$  and the new state variables (in (24) and (25)) is  $\psi = (y, x_1, \dots, x_{n_2-1})$  and  $\zeta = (x_{n_2}, \dots, x_n)$ . We call any such feedback interconnection of subsystems with  $n_1$  and  $n_2$  states a  $(n_1, n_2)$ -decomposition of (3) (with the convention that the decomposition in (4) and (5) being a  $(n, 1)$  decomposition). Fig. 5 illustrates the  $(n_1, n_2)$  decomposition where  $S_1$  corresponds to part of the pathway as marked and  $S_2$  represents the dynamics of the rest of the pathway.

Similarly to the previous decomposition, we search for a neighborhood  $\mathcal{N}_2 \subset \mathbb{R}^{n_2}$  of the origin, and a continuously differentiable, positive definite function  $U_1 : \mathbb{R}^{n_1} \rightarrow \mathbb{R}$  in  $\mathbb{R}^{n_1}$  and another continuously differentiable, positive definite function  $U_2 : \mathbb{R}^{n_2} \rightarrow \mathbb{R}$  in  $\mathcal{N}_2$  such that

$$\nabla U_1(\zeta) \cdot \dot{\zeta} \leq \rho(w, z), \quad \forall (\zeta, z) \in \mathbb{R}^{n_1} \times \mathbb{R} \quad (26)$$

$$\nabla U_2(\psi) \cdot \dot{\psi} \leq -\rho(w, z), \quad \forall (\psi, w) \in \mathcal{N}_2 \times \mathbb{R}. \quad (27)$$

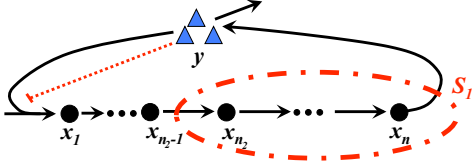


Fig. 5. Illustration of the  $(n_1, n_2)$  decomposition of the autocatalytic pathway model in Fig. 1.

where  $\rho : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a supply rate parametrized by  $\delta$  and  $\kappa$

$$\rho(w, z) := z^2 + 2\delta wz - \kappa w^2.$$

Define

$$U(\zeta, \psi) := U_1(\zeta) + U_2(\psi), \quad (28)$$

and  $\beta$  be such that

$$\Omega_{U,\beta} \subset \mathbb{R}^{n_1} \times \mathcal{N}_0$$

and  $\Omega_{U,\beta}$  is bounded.

*Proposition 7:* Let  $U$  and  $\beta$  be defined as above. Then,  $\Omega_{U,\beta}$  is invariant under the flow of (3), and if one of the inequalities (8) and (9) is strict, then  $\Omega_{U,\beta}$  is a subset of the RoA of the origin of (3).  $\triangleleft$

We now focus on a special case where  $\frac{\partial}{\partial x_i} g_i(x_i)|_0 = k$  for  $i = 1, \dots, n$ . For this case, we showed, in the previous section, that for  $\hat{h}_s < \hat{h} < \hat{h}_d(n)$ , we can find block diagonal Lyapunov functions that give an estimate of RoA. On the other hand, Proposition 6 establishes that, for  $\hat{h} > \hat{h}_d(n)$ , the  $(n, 1)$ -decomposition cannot produce a (quadratic) block diagonal Lyapunov function of the form  $U(x, y) = U_1(x) + U_2(y)$  for the linearization of (3). The following example demonstrates that Lyapunov functions for systems with  $\hat{h} > \hat{h}_d(n)$  can be constructed by increasing the number of states of  $S_2$ .

*Example 2:* Consider the  $(1, 2)$ -decomposition of a pathway with  $n = 2$  and  $g_i(x_i) = kx_i$ . For  $\kappa > 0$ ,  $\delta = \frac{\kappa-1}{2}$ , and  $U_1(x_1) = \frac{1}{2}(1 + \kappa)x_1^2$ , it follows that  $\nabla U_1(x_1) \cdot (z - kx_1) - z^2 - (\kappa - 1)wz + \kappa w^2 = -(z - kx_1)^2 \leq 0$  for all  $(x, z)$ . The inequality in (27) boils down to  $\nabla U_2(\psi) \cdot (\bar{F}_2(\zeta) + \bar{B}_2 w) \leq \kappa w^2 - (\kappa - 1)wz - z^2$ . Fig. 6 shows the set of parameters for which this decomposition yields a block-diagonal Lyapunov function for  $q = 1$  and  $\frac{\partial}{\partial y} g_y(y)|_0 = 1$ . Similar analysis is repeated for a pathway with  $n = 6$  and the effect of increasing values of  $n_2$  in these decompositions is shown in Fig. 7.

Fig. 8 shows that, as the size of the pathway increases, the size of the stability region (in the parameter space) decreases and range of feedback gains becomes limited. For a fixed pathway size, as feedback gains increase the corresponding systems approach the stability boundary (in the parameter space) and we need to use decompositions with larger number of states in  $S_2$ .

In the case  $g_1, \dots, g_n, g_y$  are polynomial (e.g. mass-action kinetics) and  $f$  is rational, the search for the polynomials  $U_1$  and  $U_2$  that satisfy the conditions in Proposition 7 can be formulated as a sum-of-squares (SOS) programming problem [18]. Under these assumptions, the SOS relaxation for (26) leads to a (relatively) low degree polynomial in  $n_1 + 1$

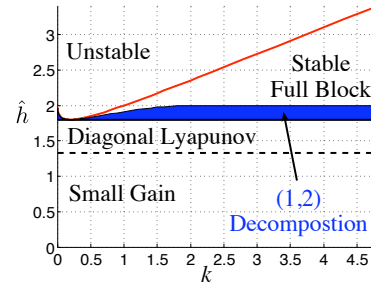


Fig. 6. Stability diagram for the three-state system with  $g_i(x_i) = kx_i$ ,  $i = 1, 2, q = 1$ ,  $\frac{\partial}{\partial y} g_y(y)|_0 = 1$ . The  $(2, 1)$ -decomposition can only be used to construct Lyapunov functions for the parameter sets labeled Diagonal Lyapunov and Small Gain. In addition to these sets, the  $(1, 2)$ -decomposition can be used to construct Lyapunov functions for the set in blue also.

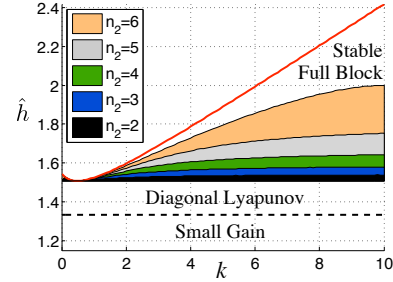


Fig. 7. For longer pathways different  $(n_1, n_2)$ -decompositions are used to produce Lyapunov functions that verify subsets of RoA of the origin. For a fixed  $k$  as the feedback gain  $\hat{h}$  approaches the stability boundary, it requires larger systems  $S_2$  (increase in  $n_2$ ), i.e., the computational complexity of finding a Lyapunov function increases. Figure shows the seven-state system ( $n = 6$ ) with  $q = 1$ ,  $\frac{\partial}{\partial y} g_y(y)|_0 = 1$ ,  $g_i(x_i) = kx_i$ ,  $i = 1, \dots, 6$ .

(indeterminate) variables while the SOS relaxation for (27) leads to a (relatively) high degree polynomial in  $n_2 + 1$  (indeterminate) variables. By solving the resulting SOS program, instead of the SOS programs used in [3] for “direct” RoA computations for (3), we obtain a significant reduction in computational complexity. This reduction stems from the fact that we replace the original SOS program in  $n + 1 = n_1 + n_2$  variables by another with two constraints (in addition to other smaller constraints): one in low number of (indeterminate) variables ( $n_2 + 1$ ) and the other in comparable (to the original SOS program) number of (indeterminate) variables ( $n_1 + 1$ ) but lower degree of polynomials. It should also be noted that for increasing values of  $n_2$ , the computational complexity of the SOS program for the conditions in Proposition 7 will be

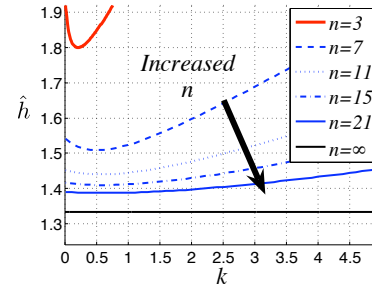


Fig. 8. As the size of the pathway increases, the size of the stability region (in the parameter space) decreases. Figure shows the stability bounds for the feedback gain  $\hat{h}$  as a function of  $k$ , for  $n$ -state systems with  $q = 1$ ,  $\frac{\partial}{\partial y} g_y(y)|_0 = 1$ ,  $g_i(x_i) = kx_i$ ,  $\forall i$ .

mainly determined by that of the SOS constraints for (27). Consequently, the decomposition in (24) and (25) and the compositional analysis strategy in Proposition 7 lead to two conflicting trends: increasing values of  $n_2$  (the state dimension of  $S_2$ ) render the analysis potentially less conservative at an expense of higher computational complexity.

## VI. SUMMARY

Many autocatalytic metabolic networks, such as the glycolytic pathway, have the topology of Fig. 1. In this paper we explicitly account for the intermediate reactions in the model given by (2) and demonstrate the effect of the pathway size  $n$  on the stability of the pathway and computational complexity of verifying RoA properties around nominal operating conditions (i.e., fixed points). A direct application of a secant condition [9], [10], [12] gives an upper bound  $\hat{h}_d(n)$  on the feedback gains  $\hat{h}$ , that result in stable pathways of fixed size  $n$  and arbitrary (fixed) intermediate reaction rates. We also establish lower and upper bounds, respectively  $\hat{h}_s$  and  $\hat{h}_r$ , on feedback gains that result in arbitrarily large, stable pathways. These bounds depend on the stoichiometry and strength of autocatalysis, and the demand of the pathway product, but not on the other parameters. They are tight in the sense that unstable pathways can be constructed for gains that lie outside the ranges established by them. The bound  $\hat{h}_d(n)$  yields that the range of possible stable gains gets smaller as the pathway size ( $n$ ) increases.

A simple representation of the system (2), as the feedback interconnection of a  $n$ -state system  $S_1$  given by (4) and a single state system  $S_2$  given by (5), allows us to analytically construct diagonal Lyapunov functions via the use of local dissipation or small-gain type inequalities for systems with feedback gain less than  $\hat{h}_d(n)$ . This simple decomposition allows us to construct block diagonal Lyapunov functions (i.e., functions of the type  $U(x, y) = U_1(x) + U_2(y)$ ,  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}$ ) for (2). However this decomposition is not as useful for gains larger than  $\hat{h}_d(n)$ . In fact, we show that no quadratic block-diagonal Lyapunov function exists for the linearization of (2) if  $\hat{h} > \hat{h}_d$  and the intermediate reaction rates are identical.

To construct Lyapunov functions for gain  $\hat{h} > \hat{h}_d$ , we use a more general  $(n_1, n_2)$ -decomposition scheme where the system in (2) is represented as a feedback interconnection of an  $n_1$ -state system  $S_1$  given by (24) and an  $n_2$ -state system  $S_2$  given by (25), with generally  $n_1$  (much) bigger than  $n_2$ . If the vector field of system (2) is rational, we can construct Lyapunov functions by using local dissipation inequalities and solving the corresponding SOS programs. As the size  $n_2$  of system  $S_2$  increases, so does the computational complexity of these optimization problems. On the other hand,  $(n_1, n_2)$ -decomposition with larger  $n_2$  can be used to construct Lyapunov functions for (2) with higher gains. This reinforces the notion that realizations that are more (computationally) “complex” are those with high feedback gains (i.e., gains that are close to the stability boundary, and therefore fragile to perturbations in the parameter space). This framework also shows that, from the computational

point of view, the increase in pathway size has two adverse effects: it increases the computational cost of constructing Lyapunov functions and limits the range of feedback gains that lead to stable pathways (therefore requires  $(n_1, n_2)$ -decomposition with large  $n_2$ .)

## REFERENCES

- [1] S. Dano, M. F. Madsen, H. Schmidt, and G. Cedersund, “Reduction of a biochemical model with preservation of its basic dynamic properties,” *FEBS Journal*, vol. 273, no. 21, 2006.
- [2] F. A. Chandra, G. Buzi, and J. C. Doyle, “Linear control analysis of the autocatalytic glycolysis system,” in *Proc of the American Control Conference*, 2009, pp. 319–324.
- [3] G. Buzi, U. Topcu, and J. C. Doyle, “Quantitative nonlinear analysis of autocatalytic networks with applications to glycolysis,” in *Proc. of the American Control Conference*, 2010.
- [4] M. Feinberg, “The existence and uniqueness of steady states for a class of chemical reaction networks,” *Arch. Rational Mech. Anal.*, no. 132, pp. 311–370, 1995.
- [5] —, “Chemical reaction network structure and the stability of complex isothermal reactors: II Multiple steady states for networks of deficiency one,” *Chem. Eng. Sci.*, no. 43, pp. 1–25, 1988.
- [6] F. Horn and R. Jackson, “General mass action kinetics,” *Arch. Rational Mech. Anal.*, vol. 47, no. 2, pp. 81–116, 1972.
- [7] E. D. Sontag, “Monotone and near-monotone biochemical networks,” *Syst Synth Biol*, vol. 1, no. 2, pp. 59–87, 2007.
- [8] P. de Leenheer, D. Angeli, and E. D. Sontag, “Monotone chemical reaction networks,” *Journal of Mathematical Chemistry*, vol. 41, no. 3, pp. 295–314, 2007.
- [9] C. D. Thron, “The secant condition for instability in biochemical feedback control – Parts I and II,” *Bulletin of Mathematical Biology*, vol. 53, pp. 383–424, 1991.
- [10] J. J. Tyson and H. G. Othmer, “The dynamics of feedback control circuits in biochemical pathways,” in *Progress in theoretical biology*, R. Rosen and F. M. Snell, Eds. New York, NY, USA: Academic Press, 1978, vol. 5, pp. 1–62.
- [11] M. Arcak and E. D. Sontag, “A passivity-based stability criterion for a class of biochemical reaction networks,” *Mathematical Biosciences and Engineering*, vol. 5, no. 1, pp. 1–19, 2008.
- [12] —, “Diagonal stability of a class of cyclic systems and its connection with the secant criterion,” *Automatica*, vol. 42, no. 9, pp. 1531–1537, 2006.
- [13] G. Buzi, “Control theoretic analysis of autocatalytic networks in biology with applications in glycolysis,” Ph.D. Dissertation, California Institute of Technology, 2010.
- [14] G. Buzi, U. Topcu, and J. Doyle, “Analysis of autocatalytic networks in biology,” *Automatica*, 2010, under review.
- [15] M. Banuelos, C. Gancedo, and J. M. Gancedo, “Activation by phosphate of yeast phosphofructokinase,” *J Biol Chem*, vol. 252, no. 18, pp. 6394–6398, 1977.
- [16] U. Topcu, A. Packard, P. Seiler, and G. Balas, “Stability region estimation for systems with unmodeled dynamics,” in *Proc. European Control Conf.*, 2009.
- [17] J. C. Willems, “Dissipative dynamical systems I: General theory,” *Archive for Rational Mechanics and Analysis*, vol. 45, pp. 321–343, 1972.
- [18] P. Parrilo, “Semidefinite programming relaxations for semialgebraic problems,” *Mathematical Programming Series B*, vol. 96, no. 2, pp. 293–320, 2003.