Control structure and limitations of biochemical networks

F. López-Caamal[†*], D. A. Oyarzún[†], J. A. Moreno[°], D. Kalamatianos[†]

Abstract— **Biochemical networks typically exhibit intricate topologies that hinder their analysis with control-theoretic tools. In this work we present a systematic methodology for the identification of the control structure of a reaction network. The method is based on a bandwidth reduction technique applied to the incidence matrix of the network's graph. In addition, in the case of mass-action and stable networks we show that it is possible to identify linear algebraic dependencies between the time-domain integrals of some species' concentrations. We consider the extrinsic apoptosis pathway and an activation–inhibition mechanism to illustrate the application of our results.**

I. INTRODUCTION

The use of Systems and Control methods to analyze biochemical networks has gained recent attention in the field of Systems Biology [1], [2]. However, their intricate topology usually renders the use of control-theoretic principles difficult to apply [3]. An ubiquitous example of this difficulty is the lack of an explicit plant/controller separation, even for regulated processes in the biological context.

The identification of the modular structure of a complex biochemical network is an appealing approach, as it may allow to reveal how the interaction among the subsystems determines the behavior of the system as a whole [1]. This idea has been used, for example, in model analysis [4] and model estimation from time-series data [5]. In this paper, we tackle the problem of identifying possible subsystems within a given dynamical model. We present a method to perform a modular decomposition of a biochemical network. This is done by a bandwidth reduction technique applied to the incidence matrix of the network's graph, which reveals lower-dimensional subsystems that may be simpler to analyze.

We illustrate the method with the Extrinsic Apoptosis Pathway (EXAP) as presented in [6], which has been subject of several analyses. The bistability of this model has been studied e.g. in [7], [8] and [9], using Montecarlo approach, bifurcation analysis, and the notion of input-to-state stability, respectively. In [10] the model sensitivity was assessed using the Structured Singular Value (SSV) analysis [11] so as to identify the most robust and fragile reactions regarding

the stability of one of the fixed points. The application of our methodology to the EXAP model reveals that it can be represented as an equivalent decentralized control system.

At a second stage, we deal with the computation of the time-domain integrals of the species' concentrations in biochemical networks. The importance of this problem was stressed in Heinrich *et al* [12], whereby by comparing the integrals of certain signals the authors showed an input amplification phenomenon in the MAP-K pathway. In the same direction, Sontag and Chaves [13] tackled the more general problem of computing the $\mathcal{L}_2$ norm of a biochemical system, and identified conditions under which its value is identical to that of an associated linear system. In this work, we consider stable mass–action networks of binary reactions and show that one can find a linear combination of the state, which has an integral that can be explicitly computed in terms of the model parameters, initial conditions and the $\mathcal{L}_1$ norm of the input. The application of this result to one of the subsystems identified in the EXAP model reveals an algebraic dependency between the integrals of two species concentrations. As a second case study, we use the result in a model for an activation–inhibition mechanism and derive a closed–form formula for the integral of one species concentration.

II. BACKGROUND

A. Dynamical models of biochemical networks

A biochemical network composed of n species interacting via m reactions can be described by

$$\dot{\boldsymbol{c}} = \boldsymbol{N}\boldsymbol{v}(\boldsymbol{c}) + \boldsymbol{D}\boldsymbol{u}. \tag{1}$$

where $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{N} \in \mathbb{R}^{n \times m}$, $\boldsymbol{v}(\boldsymbol{c}) \in \mathbb{R}^m$, $\boldsymbol{D} \in \mathbb{R}^{n \times p}$, and $\boldsymbol{u} \in \mathbb{R}^p$. The state $\boldsymbol{c}$ is the vector of species concentrations. The matrix $\boldsymbol{N}$ describes the stoichiometry of the network, so that its $(i,j)^{\text{th}}$ entry, denoted as N_{ij}, is the stoichiometric coefficient of the ith compound in the jth reaction. If $N_{ij} > 0$ ($N_{ij} < 0$), the compound appears as a product (substrate) in the corresponding reaction. The vector $\boldsymbol{v}(\boldsymbol{c})$ contains the rates at which the reaction occur, whereas $\boldsymbol{u}$ acts as a control input and represents the rate at which (some of) the compounds are supplied to the network. This control input (and the matrix $\boldsymbol{D}$) allows to describe the interaction of the network with its surrounding. The reaction rates in $\boldsymbol{v}(\boldsymbol{c})$ are usually nonlinear functions of the state $\boldsymbol{c}$, the form of which depends on the specific chemical mechanisms enabling the reaction. A simple and commonly used model for these reactions is given by the Law of Mass Action (see [14], for example): under the assumption that the medium is homogeneous and the species appear in large

[†] F. López-Caamal, D. A. Oyarzún and D. Kalamatianos are with the Hamilton Institute, National University of Ireland Maynooth, Co. Kildare, Ireland. Tel: + 353 (0)1 7086100. {fernando.lopezcaamal.2009; diego.oyarzun; dimitris.kalamatianos}@nuim.ie

[°] J.A. Moreno is with the Coordinación de Automatización, Instituto de Ingeniería, Universidad Nacional Autónoma de México (UNAM), Ciudad Universitaria, 04510 México, D.F., México. Tel: +52-55-56233600 ext. 8811. JMorenoP@ii.unam.mx

[*] Author for correspondence

molecule numbers, the rate of the ith reaction is proportional to the probability of the reactants to collide, which in turn is proportional to the product of their concentrations to the power of the molecularity. The mass–action reaction rate for a reaction with n_r reactants and n_p products is

$$v_i(\boldsymbol{c}) = k_i \prod_{i=1}^{n} [c_i]^{d_{ij}}, \tag{2}$$

where d_{ij} is the molecularity of reactant c_i in the jth reaction and $[c_i]$ denotes its concentration. In this paper we consider networks of binary reactions with molecularity one, i. e., reactions of the form

$$v_i(\boldsymbol{c}) = k_i f_i(\boldsymbol{c}), \tag{3}$$

where $f_i(\boldsymbol{c})$ is the product of two elements of $\boldsymbol{c}$. In the following a reversible reaction will be modeled as two independent reactions (one in the forward (backward) direction).

B. Bandwidth reduction for the incidence matrix of a graph

A graph is a pair of sets $G = (V, E)$ with $E \in V \times V$. The elements of V are called the *vertices* of the graph G and the elements of E are its *edges* [15]. Two vertices, z and q, are called *adjacent* if $(z, q) \in E$; the set of vertices adjacent to z is denoted as $T(z)$. The incidence matrix of G is defined as $\mathcal{S} = \{s_{ij}\}$ with $s_{ij} = 1$ if $(i, j) \in E$ and $s_{ij} = 0$ otherwise. Let $r(z)$ be a chosen labeling for vertex z, then bandwidth of z with respect to the labeling r is

$$\mathcal{B}_r(z) = \max\{|r(z) - r(q)|, \forall q \in T(z)\}.$$

The bandwidth of a graph G with respect to a labeling r is

$$\mathcal{B}_r(G) = \max\{B_r(z), \forall z \in V\}.$$

Let $\mathcal{B}(G)$ be the minimum $\mathcal{B}_r(G)$ over all possible labelings r, then the *bandwidth reduction problem* consists in finding a labeling r that minimizes $\mathcal{B}_r(G)$. In terms of the incidence matrix, the bandwidth reduction problem consists in finding a permutation of the rows and the columns of $\mathcal{S}$ that keeps all the non-zero elements of $\mathcal{S}$ in a band that is as close as possible to the main diagonal. Several algorithms for solving this problem have been proposed since the late 1960's, including the Reverse Cuthill-McKee procedure, GPS, and Tabu search based algorithms (see [16] and the references therein).

III. Control structure identification

The control analysis of biochemical networks is usually impeded by the complexity of their topologies. In this section we present a method for splitting a biochemical network into subsystems that may be simpler to analyze independently.

A. Modular decomposition of a biochemical network

Consider the biochemical network in (1) and its Jacobian matrix $\boldsymbol{J}(\boldsymbol{c}) \equiv \{J_{ij}(\boldsymbol{c})\}$ given by

$$\boldsymbol{J}(\boldsymbol{c}) = \boldsymbol{N} \frac{\partial v(\boldsymbol{c})}{\partial \boldsymbol{c}}. \tag{4}$$

The incidence matrix associated to the (graph of) network (1) is then defined as $\mathcal{S} \equiv \{s_{ij}\}$ with

$$s_{ij} = \begin{cases} 1 & , J_{ij} \neq 0 \text{ for } c \neq 0 \\ 0 & , J_{ij} = 0 \text{ for } c \neq 0 \end{cases},$$

The ith row of $\mathcal{S}$ shows whether or not there exists a dependence of the ith differential equation on the jth chemical species. Therefore, the block entries in the diagonal will appear when a set of variables is interacting with each other, thus defining a subsystem. The structure of the network can then be inferred from the sparsity structure in $\mathcal{S}$. The block matrices appearing in the diagonal of $\mathcal{S}$ can be interpreted as subsystems, whereas the off-diagonal entries represent the interconnection among them (See Figure 2 in Section III-B).

From the above discussion we see that one way of rearranging the structure of the network is by reducing the bandwidth of $\mathcal{S}$. The resulting labeling provides a means of identifying subsystems in the network, together with their corresponding interconnections. The proposed method can be summarized as follows:

S.1 Obtain the incidence matrix $\mathcal{S}$
S.2 Reorder the label of the vetices in order to reduce the bandwidth
S.3 Reorder the labeling so biochemical meaning is preserved (heuristic)
S.4 Identify the subsystems present as the blocks in the main diagonal and the interconnection signals as the off-diagonal terms (biologically motivated)

In the next section we illustrate the proposed decomposition technique for a model of the extrinsic apoptosis pathway (EXAP) [6].

B. Case study: EXAP model

Apoptosis is a programmed cell death mechanism based upon the activation of caspases. The caspases are enzymes capable of dismantling proteins. Initiator caspases cleave inactive pro-forms of effector caspases, thereby activating them. In turn, effector caspases trigger the apoptotic process by cleaving essential proteins for the cell survival. The initiation of this reaction network is regulated by caspase inhibitors, which avoid caspases to further react. The function of this cellular process is the removal of diseased and older cells for recycling and is essential for maintaining homeostasis. In general, two different pathways are considered: the intrinsic and extrinsic pathway. Further details on the modeling and systems analysis of the apoptotic process can be found in [17].

In the EXAP the initiator caspase is Caspase 8 and the effector caspase is Caspase 3. CARP (Cellular Apoptosis Regulatory Protein) represents the inhibitor of C_{8a}; and IAP (Inhibitor of Apoptosis Protein), the inhibitor of C_{3a}. The subindex 'a' denotes the active form of the caspase and the absence of the subindex, the pro-form version. C_{8a} can either be degraded *via* its inhibitor or activate the effector caspase C_3. In turn, the C_{3a} can be either degraded by its inhibitor or can activate the initiator caspase C_8. Once a considerable
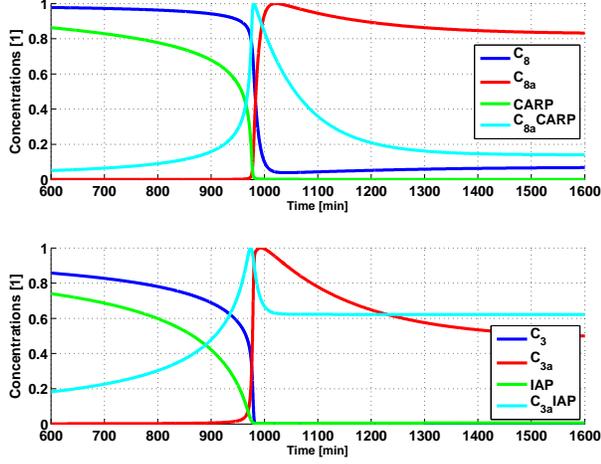
Fig. 1. Normalized trajectories of the states of the EXAP with nominal parameters and $c_0 = (130000, 1000, 21000, 0, 40000, 0, 40000, 0)$.

Fig. 2. Incidence matrix of the system (5). The green squares represent the subsystems in which the whole systems can be decomposed.

pool of C_{3a} is present in the cytosol, the cell undergoes the apoptotic process. The reactions of this pathway are:

$$C_{8a} + C_3 \xrightarrow{k_1} C_{3a} + C_{8a} \qquad C_{3a} + IAP \xrightarrow{k_4} C_{3a}$$

$$C_{3a} + IAP \underset{k_{m3}}{\overset{k_3}{\rightleftarrows}} IAP : C_{3a} \qquad C_{3a} \xrightarrow{k_6} 0$$

$$C_{8a} \xrightarrow{k_5} 0 \qquad IAP \underset{k_{m8}}{\overset{k_8}{\rightleftarrows}} 0$$

$$IAP : C_{3a} \xrightarrow{k_7} 0 \qquad C_3 \underset{k_{m10}}{\overset{k_{10}}{\rightleftarrows}} 0$$

$$C_8 \underset{k_{m9}}{\overset{k_9}{\rightleftarrows}} 0 \qquad CARP \underset{k_{m12}}{\overset{k_{12}}{\rightleftarrows}} 0$$

$$C_{8a} + CARP \underset{k_{m11}}{\overset{k_{11}}{\rightleftarrows}} CARP : C_{8a} \qquad CARP : C_{8a} \xrightarrow{k_{13}} 0$$

$$C_{3a} + C_8 \xrightarrow{k_2} C_{8a} + C_{3a}$$

Let $c = ([C_8] \ [C_{8a}] \ [C_3] \ [C_{3a}] \ [IAP] \ [IAP : C_{3a}] \ [CARP] \ [CARP : C_{8a}])^T$. Assuming a mass-action reaction mechanism the mathematical model is

$$
\begin{aligned}
\dot{c}_1 &= -k_2 c_4 c_1 - k_9 c_1 + k_{m9} \\
\dot{c}_2 &= k_2 c_4 c_1 - k_5 c_2 - k_{11} c_2 c_7 + k_{m11} c_8 \\
\dot{c}_3 &= -k_1 c_2 c_3 - k_{10} c_3 + k_{m10} \\
\dot{c}_4 &= k_1 c_2 c_3 - k_6 c_4 - k_3 c_4 c_5 + k_{m3} c_6 \qquad (5)\\
\dot{c}_5 &= k_{m3} c_6 - k_8 c_5 + k_{m8} - (k_3 + k_4) c_4 c_5 \\
\dot{c}_6 &= -(k_{m3} + k_7) c_6 + k_3 c_4 c_5 \\
\dot{c}_7 &= k_{m11} c_8 - k_{12} c_7 + k_{m12} - k_{11} c_2 c_7 \\
\dot{c}_8 &= -(k_{m11} + k_{13}) c_8 + k_{11} c_2 c_7
\end{aligned}
$$

Under nominal conditions and parameters [18], the system's dynamic is shown in Figure 1, where the response is due only to initial concentrations. The phase portrait of this model with nominal parameters shows three fixed points, two of them are stable nodes and one is a saddle [18]. This model was built to reproduce a switch behavior, which resembles the initiation of the apoptosis process, by a high level of C_{3a}. By applying the procedure of previous section we can identify the subsystems shown in Figure 2. In addition, in each subsystem we can identify blocks that can be interpreted as plant and controller. This reveals a decentralized control structure for the EXAP model, as shown in Figure 3. The plant and controller blocks are defined as

$$
\mathcal{P}_1 : \begin{cases}
\dot{c}_1 &= -k_2 c_4 c_1 - k_9 c_1 + k_{m9} \\
\dot{c}_2 &= k_2 c_4 c_1 - k_5 c_2 \underbrace{-k_{11} c_2 c_7 + k_{m11} c_8}_{u_1}
\end{cases}
$$

$$
\mathcal{K}_1 : \begin{cases}
\dot{c}_7 &= k_{m11} c_8 - k_{12} c_7 + k_{m12} - k_{11} \overbrace{c_2}^{v_1 = y_1} c_7 \\
\dot{c}_8 &= -(k_{m11} + k_{13}) c_8 + k_{11} c_2 c_7
\end{cases}
$$

$$
\mathcal{P}_2 : \begin{cases}
\dot{c}_3 &= -k_1 c_2 c_3 - k_{10} c_3 + k_{m10} \\
\dot{c}_4 &= k_1 c_2 c_3 - k_6 c_4 \underbrace{-k_3 c_4 c_5 + k_{m3} c_6}_{u_2}
\end{cases}
$$

$$
\mathcal{K}_2 : \begin{cases}
\dot{c}_5 &= k_{m3} c_6 - k_8 c_5 + k_{m8} - (k_3 + k_4) \overbrace{c_4}^{v_2 = y_2} c_5 \\
\dot{c}_6 &= -(k_{m3} + k_7) c_6 + k_3 c_4 c_5
\end{cases}
$$

whereas the interconnection signals are

$$
\begin{aligned}
y_1 &= C_{8a}, & v_1 &= C_{8a}, & w_1 &= C_{3a}, \\
y_2 &= C_{3a}, & v_2 &= C_{3a}, & w_2 &= C_{8a}.
\end{aligned}
$$

In this example, the modular decomposition proposed is used to show the structure of the system. This interconnection scheme may be used to identify input-output properties of the subsystems and hence infer properties of the whole system dynamics.

IV. TIME-DOMAIN LIMITATIONS

The structure of the network in (1) may impose algebraic dependencies to the state variables. For example, when rank $N < n$ some species satisfy conservation laws that
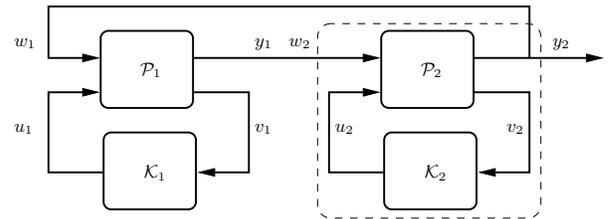
Fig. 3. Decentralized controller structure of system (5). The dotted line encircles the system to which the further analysis is performed.

constraint their possible trajectories [14]. Although such limitations can be recognized by inspecting the stoichiometric matrix, there may be other ones that are not as easy to detect; in this section we present a method to detect algebraic dependencies that may appear between the time-domain integrals of the state variables.

A. Computation of time-domain integrals

Assume that for a constant control $\boldsymbol{u} = \bar{\boldsymbol{u}}$ the dynamics of (1) are stable and let $\bar{\boldsymbol{c}} \in \mathbb{R}_{\geq 0}^n$ be the corresponding equilibrium. Define the deviation coordinates as $\boldsymbol{e} = \boldsymbol{c} - \bar{\boldsymbol{c}}$ and $\boldsymbol{u_e} = \boldsymbol{u} - \bar{\boldsymbol{u}}$. We further assume that the control satisfies

$$\boldsymbol{u_T} = \int_0^\infty \boldsymbol{u_e} dt < \infty. \tag{6}$$

A Taylor expansion of (1) around the equilibrium $(\bar{\boldsymbol{c}}, \bar{\boldsymbol{u}})$ yields

$$\dot{\boldsymbol{e}} = \boldsymbol{J}(\bar{\boldsymbol{c}})\boldsymbol{e} + \frac{1}{2} \begin{pmatrix} \boldsymbol{e}^T \boldsymbol{H} (N_1 \boldsymbol{v}(\boldsymbol{c}))|_{\boldsymbol{c}=\bar{\boldsymbol{c}}} \boldsymbol{e} \\ \vdots \\ \boldsymbol{e}^T \boldsymbol{H} (N_m \boldsymbol{v}(\boldsymbol{c}))|_{\boldsymbol{c}=\bar{\boldsymbol{c}}} \boldsymbol{e} \end{pmatrix} + \boldsymbol{D}\boldsymbol{u_e}, \tag{7}$$

Where N_i is the ith row of the the stoichiometric matrix and $\boldsymbol{H}$ denotes the Hessian operator. Note that the expression in (7) is exact, since for binary reactions the higher order derivatives of $\boldsymbol{v}(\boldsymbol{c})$ vanish. Equation (7) can be rewritten as

$$\dot{\boldsymbol{e}} = \boldsymbol{J}(\bar{\boldsymbol{c}})\boldsymbol{e} + \boldsymbol{N}\boldsymbol{B}\boldsymbol{g}(\boldsymbol{e}) + \boldsymbol{D}\boldsymbol{u_e}, \tag{8}$$

where $\boldsymbol{B} = \operatorname{diag}\{b_i\}$ with

$$b_i = \begin{cases} k_i & \text{, if the } i^{\text{th}} \text{ reaction is binary} \\ 0 & \text{, otherwise.} \end{cases}$$

and the entries of $\boldsymbol{g}(\boldsymbol{e}) \in \mathbb{R}^m$ defined as

$$g_i(\boldsymbol{e}) = \begin{cases} f_i(\boldsymbol{e}) & \text{, if the } i^{\text{th}} \text{ reaction is binary} \\ 0 & \text{, otherwise.} \end{cases}$$

Integration of (8) yields

$$\boldsymbol{e}(\infty) - \boldsymbol{e}(0) = \boldsymbol{A} \int_0^\infty \boldsymbol{e} \, dt + \boldsymbol{N}\boldsymbol{B} \int_0^\infty \boldsymbol{g}(\boldsymbol{e}) \, dt + \boldsymbol{D} \int_0^\infty \boldsymbol{u_e} \, dt,$$

where we have defined $\boldsymbol{A} = \boldsymbol{J}(\bar{\boldsymbol{c}})$. The stability of (1) together with (6) imply that $\boldsymbol{e}(\infty) = 0$ and hence

$$\int_0^\infty \boldsymbol{e} dt = -\boldsymbol{A}^{-1}\boldsymbol{e}(0) - \boldsymbol{A}^{-1}\boldsymbol{N}\boldsymbol{B} \int_0^\infty \boldsymbol{g}(\boldsymbol{e}) \, dt - \boldsymbol{A}^{-1}\boldsymbol{D}\boldsymbol{u_T}. \tag{9}$$

From this equation the forthcoming proposition is straightforward.

Proposition 4.1: Define $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{e}$ with $\boldsymbol{C} \in \mathbb{R}^{q \times n}$ and such that

$$\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{N}\boldsymbol{B} = \boldsymbol{0}. \tag{10}$$

Then,

$$\int_0^\infty \boldsymbol{y} \, dt = -\boldsymbol{C}\boldsymbol{A}^{-1} \left(\boldsymbol{e}(0) + \boldsymbol{D}\boldsymbol{u_T} \right). \tag{11}$$

Condition (10) indicates that $\boldsymbol{C}$ has to be chosen so that its rows form a basis for the kernel of $\boldsymbol{B}\boldsymbol{N}^T\boldsymbol{A}^{-T}$. We foresee two possible applications of Proposition 4.1: firstly, since y is a linear combination of the species' concentrations (in deviation coordinates), the formula in (11) allows the identification of algebraic constraints between the integral of the different species' concentrations. Secondly, if elements of $\boldsymbol{y}$ are also elements of $\boldsymbol{e}$, then (11) gives a closed-form formula for their integral in terms of the initial conditions, the $\mathcal{L}_1$ norm of the input, and the parameters of the network. Next we present an example for each of these two scenarios.

B. Case study: EXAP system

As a follow-up to the case study in Section III-B, here we analyze time-domain limitations that appear in the dynamics of subsystems $\mathcal{P}_2$ and $\mathcal{K}_2$ in Figure 3. To isolate these subsystems form the rest of the network, $w_2 = c_2$ is treated as a constant input to $\mathcal{P}_2$. As a simplification of the network, the degradation and synthesis of $[IAP]$ are replaced with the control $\boldsymbol{u_e}$. With this considerations, the model can be rewritten as in (8) with:

$$\boldsymbol{N} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & -1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\boldsymbol{B} = \operatorname{diag}\{0, k_3, 0, k_4, 0, 0, 0, 0, 0, 0\}$$

$$\boldsymbol{v}(c_2) = \begin{pmatrix} k_1 w_2 c_3 \\ k_3 c_4 c_5 \\ k_{m3} c_6 \\ k_4 c_4 c_5 \\ k_6 c_4 \\ k_7 c_6 \\ k_8 c_5 \\ k_{m8} \\ k_{10} c_3 \\ k_{m10} \end{pmatrix}, \quad \boldsymbol{g}(\boldsymbol{e}) = \begin{pmatrix} 0 \\ e_2 e_3 \\ 0 \\ e_2 e_3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{D} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Here $\boldsymbol{e} = \boldsymbol{c}_2 - \bar{\boldsymbol{c}}_2$, $\boldsymbol{c}_2 = (c_3 \quad c_4 \quad c_5 \quad c_6)^T$ and

$$\bar{\boldsymbol{c}}_2 = \left(\frac{k_{m10}}{k_{10} + k_a w_2} \quad \frac{k_{m10} k_1 w_2}{k_{10} + k_1 w_2} \quad 0 \quad 0 \right)^T.$$

This system is stable for constant inputs and the matrix

$$\boldsymbol{C} = \begin{pmatrix} 0 & 1 & 0 & \frac{k_7}{k_6} \end{pmatrix} \tag{12}$$

satisfies (10). Thus we can use expression (11) to compute the integral of $y = \boldsymbol{C}\boldsymbol{e}$.

Considering nil initial concentrations and the inputs in Figure 4, we get

$$\int_0^\infty y \, dt = 0,$$

which implies

$$\int_0^\infty e_2 \, dt = -\frac{k_7}{k_6} \int_0^\infty e_4 \, dt.$$

This means that the integral of e_2 is proportional to the one of e_4. This is a consequence of a trade-off in the dynamic
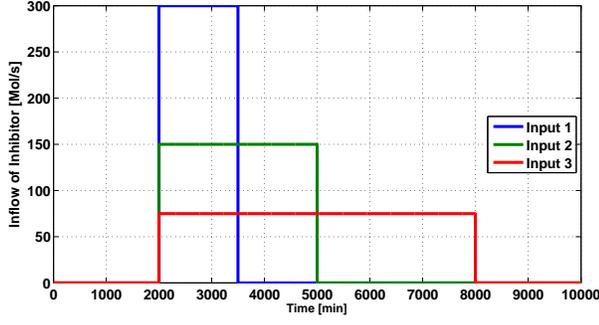
Fig. 4. Inputs to the network. Note that the area of all the signals is the same.

behavior. In the biological context, this fact shows that $IAP : C_{3a}$ and C_{3a} will mutually degrade. It is a well known fact that the role of IAP is a regulator of the triggering of the apoptosis mechanism. Here this fact is shown by the dynamic constraint among the integrals of C_{3a} and $IAP : C_{3a}$.

C. Case study: activation–inhibition mechanism

Consider a simple activation inhibition mechanism represented by the forthcoming reactions:

$$0 \xrightarrow{k_1} x_1 \xrightarrow{k_4} 0$$
$$x_1 + w \underset{k_{2b}}{\overset{k_{2f}}{\rightleftharpoons}} x_2 + w$$
$$x_2 \xrightarrow{k_5} 0$$
$$x_2 + x_3 \underset{k_{3b}}{\overset{k_{3f}}{\rightleftharpoons}} x_4 \xrightarrow{k_6} 0$$
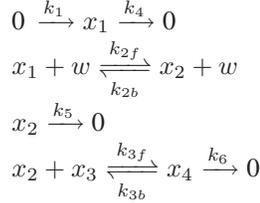
Here $[w]$ is assumed to be constant and x_1 can be interpreted as an inactive form of a species of interest, x_2 as the activated version, x_3 as the inhibitor and x_4 as the complex formed by the inhibitor and x_2. The dynamical model can be written

in the form of (8), where:

$$N^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & 1 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad v(x) = \begin{pmatrix} k_1 \\ k_{2f} x_1 w \\ k_{2b} x_2 w \\ k_{3f} x_2 x_3 \\ k_{3b} x_4 \\ k_4 x_1 \\ k_5 x_2 \\ k_6 x_4 \end{pmatrix}$$

$$B = \text{diag} \{0, 0, 0, k_{3f}, 0, 0, 0, 0\}$$
$$g(e) = (0 \ 0 \ 0 \ e_2 e_3 \ 0 \ 0 \ 0 \ 0)^T$$
$$D = (0 \ 1 \ 0 \ 0)^T$$

Here $x = ([x_1] \ [x_2] \ [x_3] \ [x_4])^T$, $e \equiv x - \bar{x}$ where $\bar{x}$ is the unique positive fixed point:

$$\bar{x} = \left(\frac{k_1 (k_{2b} + k_5)}{k_{2b} k_4 + k_5 (k_{2f} w + k_4)} \quad \frac{k_1 k_{2f} w}{k_{2b} k_4 + k_5 (k_{2f} w + k_4)} \quad 0 \quad 0 \right)^T$$

In this case the matrix $A^{-1} N B$ has only one nonzero entry:

$$A^{-1} N B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence the area of any linear combination of the first, second and fourth states can be calculated in terms of the structure, parameters and exogenous inputs to the system. Choose, for example $C = (0 \ 1 \ 0 \ 0)$. Note this vector lies in the left kernel of $A^{-1} N B$. Then its integral is given by the following closed–form expression:

$$\int_0^\infty y \, dt = \alpha(k_{2f} w \quad k_{2f} w + k_4 \quad -(k_{2f} w + k_4) \quad 0) \cdot$$
$$\cdot [e_0 + D u_T],$$

where $\alpha = (k_5 k_4 + k_{2b} k_4 w + k_5 k_{2f} w)^{-1}$. For the initial concentrations and parameters $\int_0^\infty e_2 \, dt = 106.0377 \times 10^6$. Figure 6 shows the response of the system due to initial

Fig. 5. Upper: Trajectories of $e_2 + \frac{k_7}{k_6} e_4$. Lower: $\int_0^\infty (e_2 + \frac{k_7}{k_6} e_4) dt$. Note that the final value in all cases is equal, despite the input applied. The inputs are shown in Figure 4.
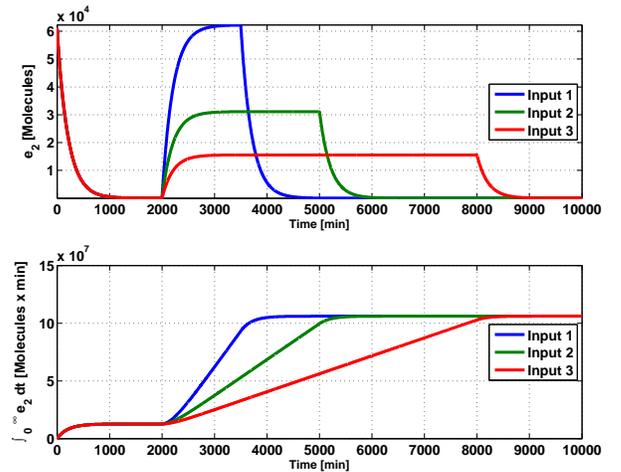
Fig. 6. Upper: Trajectories of e_2 with the initial conditions $e_0 = (2100 \ 60000 \ 1000 \ 1000)^T$ and $\{k_1, k_{2f}, k_{2b}, k_{3f}, k_{3b}, k_4, k_5, k_6\} = \{1, 0.1, 0.02/w, 0.1, 1, 1, 0.003, 1\}$. Lower: The integral of e_2. Note that despite the form of the inputs is different in all cases, the final value of the integral is the same.

concentrations and three different inputs which affect the concentration of e_2. These inputs are shown in Figure 4. Note that all of them have the same area. This is consistent with the result proposed in Proposition 4.1, since the value of the integral only depends on the integral of the input, rather than the shape. We also note that, although the dynamics depend on velocity constants k_1, k_{3f}, k_{3b}, k_6, they do not affect the integral's value.

V. CONCLUSIONS

In this work, two studies which can be useful for the analysis of biochemical networks are presented.

Firstly the identification of subsystems and its interconnection topology are achieved *via* the bandwidth reduction of the incidence matrix associated to the dynamical model. The effectiveness of the methodology depends on the sparsity of this matrix, since in highly coupled systems the subsystems identification and isolation might be difficult. The case study is the extrinsic apoptosis pathway, in which an eight–states model was decomposed into four subsystems of dimension two. The structure in which these subsystems are interconnected resembles a decentralized controller scheme.

The second analysis focuses on the identification of a linear combination of the state whose time–domain integral can be computed in terms of the initial concentrations, model parameters and the $\mathcal{L}_1$ norm of the inputs. Here such a subspace is determined for mass–action reaction networks. The use of this linear transformation is exemplified with two case studies. In the first one a subsystem previously identified on the EXAP is analyzed, whence a dynamic trade–off between two states arises. The later case study shows how a closed–form expression for computing one state's time–domain integral can be obtained. This formula is function of *some* parameters solely, showing that the integral will be insensitive to variations of the rest.

REFERENCES

[1] M. E. Csete and J. C. Doyle, "Reverse engineering of biological complexity," *Science*, vol. 295, no. 5560, pp. 1664–1669, 2002.

[2] P. A. Iglesias and B. P. Ingalls (editors), *Control Theory and Systems Biology*. MIT Press, 2009.

[3] E. Sontag, "Molecular systems biology and control," *European Journal of Control*, vol. 11, pp. 1 – 40, 2005.

[4] J. J. Tyson, K. C. Chen, and B. Novak, "Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell," *Current Opinion in Cell Biology*, vol. 15, no. 2, pp. 221 – 231, 2003.

[5] E. August and A. Papachristodoulou, "Efficient, sparse biological network determination," *BMC Systems Biology*, vol. 3, no. 1, p. 25, 2009.

[6] T. Eissing, H. Conzelmann, E. D. Gilles, F. Allgower, E. Bullinger, and P. Scheurich, "Bistability analyses of a caspase activation model for receptor-induced apoptosis," *J. Biol. Chem.*, vol. 279, pp. 36892–36897, August 2004.

[7] T. Eissing, F. Allgower, and E. Bullinger, "Robustness properties of apoptosis models with respect to parameter variations and intrinsic noise," *Systems Biology, IEE Proceedings*, vol. 152, pp. 221–228, Dec. 2005.

[8] L. Corotenuto, V. Pace, D. Bellizzi, and G. De Benedictis, "Equilibrium, stability and dynamical response in a model of the extrinsic apoptosis pathway," *Journal of Biological Systems*, vol. 15, no. 3, pp. 261–285, 2007.

[9] M. Chaves, T. Eissing, and F. Allgower, "Bistable biological systems: A characterization through local compact input-to-state stability," *Automatic Control, IEEE Transactions on*, vol. 53, pp. 87–100, Jan. 2008.

[10] J. Shoemaker and F. Doyle, "Identifying fragilities in biochemical networks: Robust performance analysis of fas signaling-induced apoptosis," *Biophysical Journal*, 2008.

[11] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice Hall, 1996.

[12] R. Heinrich, B. G. Neel, and T. A. Rapoport, "Mathematical models of protein kinase signal transduction.," *Mol Cell*, vol. 9, pp. 957–970, May 2002.

[13] E. Sontag and M. Chaves, "Exact computation of amplification for a class of nonlinear systems arising from cellular signaling pathways," *Automatica*, vol. 42, no. 11, pp. 1987 – 1992, 2006.

[14] E. Klipp, A. Kowald, C. Wierling, and H. Lehrach, *Systems Biology in Practice: Concepts, Implementation and Application*. Wiley-VCH, May 2005.

[15] R. Diestel, *Graph Theory*. Springer-Verlag, 2000.

[16] R. Marti, M. Laguna, F. Glover, and V. Campos, "Reducing the bandwidth of a sparse matrix with tabu search," *European Journal of Operational Research*, vol. 135, pp. 450–459, December 2001.

[17] I. N. Lavrik, R. Eils, N. Fricker, C. Pforr, and P. H. Krammer, "Understanding apoptosis by systems biology approaches.," *Molecular bioSystems*, vol. 5, no. 10, pp. 1105–1111, 2009.

[18] E. Bullinger, "System analysis of a programmed cell death model," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pp. 7994–7999, Dec. 2005.