Complex Structures and Dynamics

Applications to Chemistry, Biology, Social Science and Economics

Markus Kirkilionis

July 19, 2012

Abstract

The deterministic mass-action kinetics as a large particle number continuum approximation. Functional analysis related to the continuum limit. Polynomial differential equations. Some related algebraic geometry, in particular coordinate transformations. Stability of equilibria and extremal currents (often called elementary flux modes in the systems biology community), modularity of mass-action reaction systems. Bifurcation theory for mass-action systems. The interaction graph and the graphical representation of feedback loops. The Thomas-Soul theorem. Other graph concepts: The Feinberg school and the species-complex-linkage graph. The deficiency theorem. Outlook to generalisations of reaction kinetics including macro-molecules.

Let there be r reactions $\mathbf{R} = (R_1, ..., R_r)$, and different species $\mathbf{S} = \{S_1, ..., S_n\}$. A single *reaction* is defined by

$$\sum_{i=1}^{n} n_{ij}^{-} S_i \to \sum_{i=1}^{n} n_{ij}^{+} S_i,$$

with n_{ij}^{-} and n_{ij}^{+} being the stoichiometric coefficients of the forward and backward reactions. The coefficients

$$n_{ij} := n_{ij}^{+} - n_{ij}^{-}$$

define the stoichiometric matrix N. The law-of-mass-action induces the function

$$v : \mathbb{R}_{+}^{n} \times \mathbb{R}_{+}^{n} \to \mathbb{R}_{+}^{r}$$

which is the reaction velocity depending on concentrations $\mathbf{x} = (x_1, ...x_n)$ and the parameter $k = (k_1, ..., k_n)$ (containing the reaction constants). The triple $S = (\mathbf{X}, \mathbf{R}, v)$ is called the *reaction scheme*. Dynamics:

$$\dot{x} = Nv(x, p)$$

The molecularity of the species X_i in reaction R_j is encoded in the *kinetic exponent* κ_{ij}. All these exponents are assembed in the *kinetic matrix* κ. These exponents can be defined by

$$\kappa_{ij} = \frac{\partial \log(v_j(x_0, k_j))}{\partial \log(x_{0,i})}.$$

With this definition in most cases the n_{ij}^- equal the κ_{ij}. Each reaction rate can be written as

$$v_j(x, k_j) = k_j \prod_{i=1}^{m} x^{\kappa_{ij}}.$$

All these monomials form the reaction vector v we have already encountered.

An illustrative example we consider throughout this chapter is a well-known model describing the activity of the mitogen-activated protein kinase (MAPK). MAPKs play an important part in the signalling processes of eukaryotic cells by intervening with a multitude of proteins and phosphorylating them. They themselves undergo phosphorylation by a MAPK/ERK kinase (MEK) and dephosphorylation by a phosphatase. In this model we use a notation for species adapted to the biochemical interpretation. Let S_i, with $i = 0$, be the MAPK kinase, and with subscripts $i = 1$ and $i = 2$ the molecules with single and double phosphorylation, respectively. The symbols E and F represent MEK and the phosphatase. The reaction scheme for the model is:

$$E + S_0 \leftrightharpoons ES_0 \to E + S_1 \leftrightharpoons ES_1 \to E + S_2$$
$$F + S_2 \leftrightharpoons FS_2 \to F + S_1 \leftrightharpoons FS_1 \to F + S_0.$$

Therefore we have with the obvious notational identifications:

$$\mathcal{S} = \{S_1, \ldots, S_9\} = \{E, F, S_0, S_1, S_2, ES_0, ES_1, FS_1, FS_2\},$$

i.e. $s = 9$ in this example. Moreover, using the formal sum notation, we have

$$\begin{aligned}\mathcal{C} =& \{C_1, \ldots, C_{10}\} \\ =& \{E + S_0, ES_0, E + S_1, ES_1, E + S_2, F + S_2, FS_2, F + S_1, FS_1, F + S_0\},\end{aligned}$$

i.e. $c = 10$ in this example, and all stiochiometric coefficients are 1 and therefore do not appear. There are 12 reactions, i.e. $r = 12$. Note that we used the symbol " $\leftrightharpoons$ " to denote a *reversible reaction*. This means there are two reactions in opposite direction of each other, the complexes left and right of these arrows can be simultaneously target and source complexes.

We are now able to introduce a first graph associated with reaction networks, the directed *reaction graph* $\overrightarrow{G}_R = (V, \overrightarrow{E})$. We simply define the vertex set $V_{\overrightarrow{G}_R}$ as the set of complexes of the reaction network, i.e. $V(\overrightarrow{G}_R) = \mathcal{C}$. There is a directed edge between two complexes whenever there is a reaction, pointing from the source complex to the target complex. For our MAPK example reaction system this means

$$\overrightarrow{E}(\overrightarrow{G}_R) = \{(C_1, C_2), (C_2, C_1), (C_2, C_3), (C_3, C_4), (C_4, C_3), (C_4, C_5), \\ (C_6, C_7), (C_7, C_6), (C_7, C_8), (C_8, C_9), (C_9, C_8), (C_9, C_{10})\}.$$

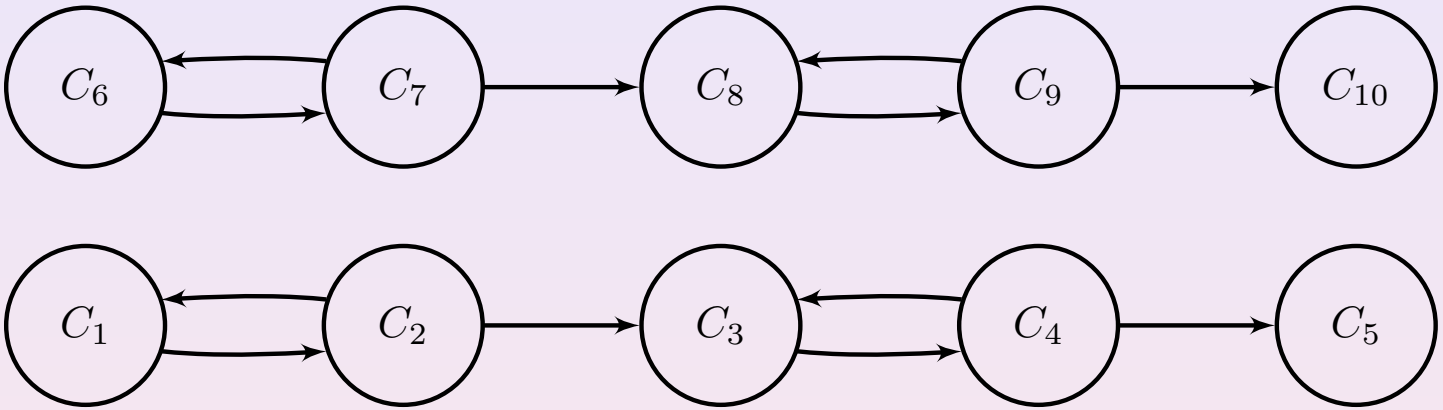


Figure: The directed reaction graph $\overrightarrow{G}_R$: of the MAPK reaction network example.

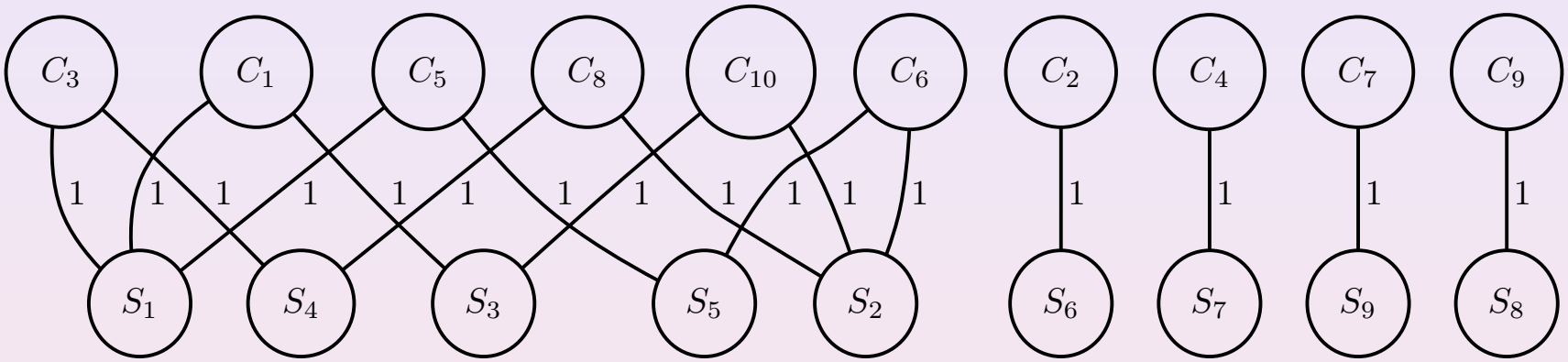


Figure: The undirected bipartite complex-species graph G_{CS} of the MAPK reaction network example.

For the directed graph the respective adjacency and incidence matrices are of importance. The first one, I_a contains the information whether the complex is the initial (entry -1) or the end vertex of an edge (entry 1). This entry distinguishes reactant complexes from product complexes. The second one, I_k contains nonzero entries only for initial vertices, i.e. for reactant complexes. The entries are the weight of the corresponding edge, which is the rate constant k_l.

The reaction scheme $\dot{x} = Nv$ can now be rewritten in the form

$$\dot{x} = YI_aI_k\Psi(x),$$

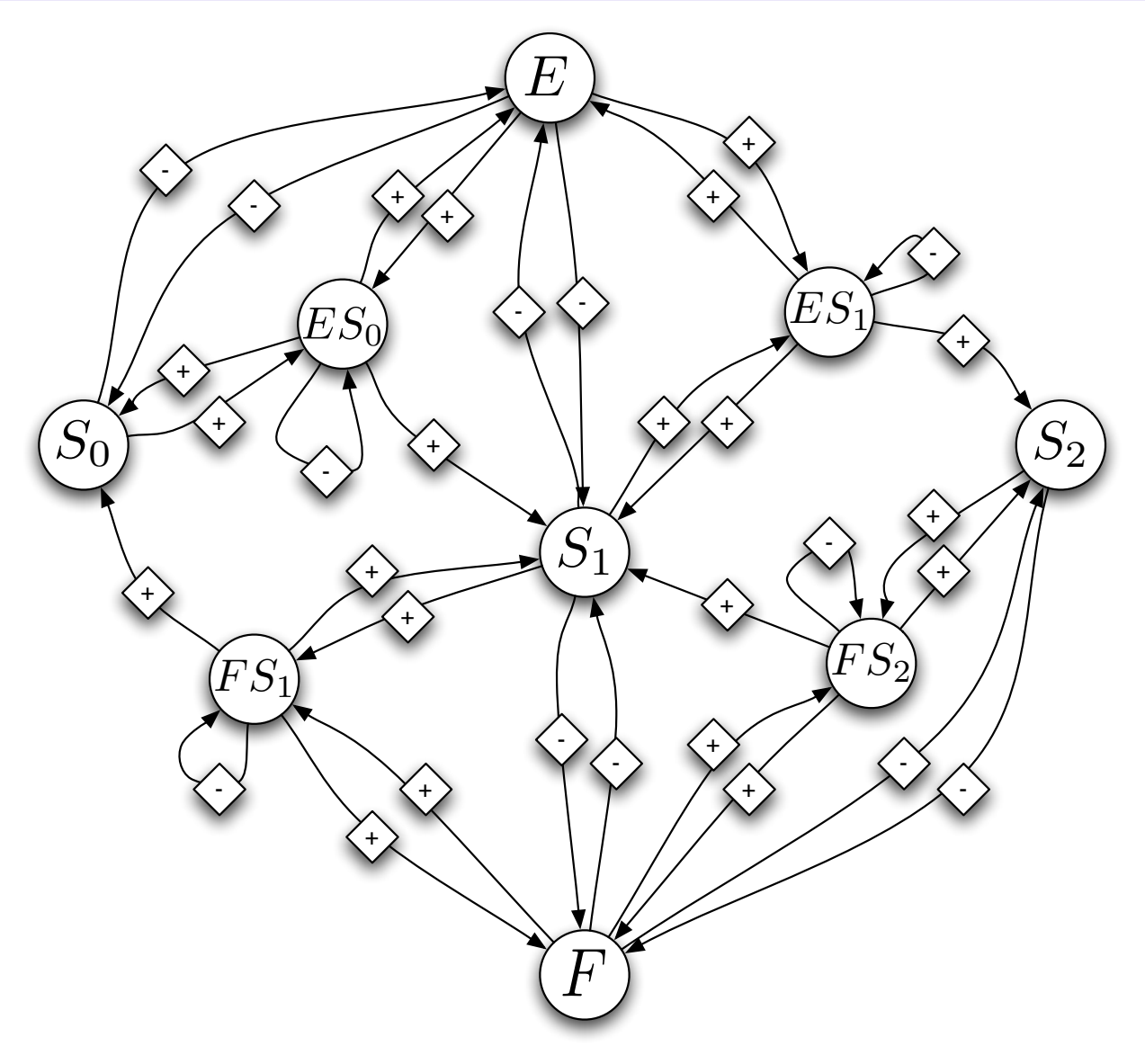
so $N = YI_a$ and $v = I_k\Psi(x)$. In other words the nonlinearities of the dynamical system defining the reaction scheme can be investigated in terms of an incidence and two adjacencies matrices, and a vector of monomials.

The two graphs defining the system of kinetic equations are on the basis of a series of graphs, each of them emerging from its predecessor by a certain formation law, which leads to a weighted directed pseudo-graph. This directed pseudograph encodes basic information about the stability of the stationary solutions of the kinetic equations. Its adjacency matrix is the part of the Jacobian matrix $\widetilde{Jac}(j)$. This originates from the special form of the convex Jacobian matrix. Actually all occurring factor matrices of

$$\widetilde{Jac}(j) = N\ diag(\sum_i j_i E_i)\ \kappa^t = Y\ I_a\ diag(\sum_i j_i E_i)\ \kappa^t$$

can be considered as adjacency matrices and incidence matrices of various graphs. The importance of the directed pseudo-graph results from the fact that instabilities of the stationary solutions of the kinetic equations can be recognized by its cycles and loops, which are reflected as matrix structures in the according adjacency matrix. These structures are called *feedback-loops*.

The Interaction Graph



The so-called Thomas's first conjecture was proven by Soulé. Here we restate it in reverse form:

Theorem (Thomas-Soulé)

If a system has no positive cycles in $G_{int}(x)$ for any x, then it cannot exhibit multi-stationarity.

Aside from the relationship between cycles and Jacobian entries, which lead to conditions about positivity of the determinant of $-J$ and its minors, Soulé made use of Gale-Nikaidô theory to show that the RHS of the dynamical system is injective and hence cannot exhibit multistationarity.

We say that complexes y and y' belong to the same linkage class if there exists an undirected path in the reaction diagram connecting the two complexes. The deficiency of a reaction network (denoted by the symbol δ) is defined by the following formula,

$$\delta = c - l - r,$$

where c is the number of complexes, l is the number of linkage classes and r is the rank of the stoichiometric matrix N. It holds that the deficiency index is always nonnegative. The stochiometric subspace for a reaction network is the span of the reaction vectors, namely $Im(N)$. Two vectors y and y' are stoichiometrically compatible if $y' - y \in Im(N)$. Stoichiometric compatibility is an equivalence relation that induces a partition of the space $\mathbb{R}^n_+$ into equivalence classes. Each positive stoichiometric compatibility class is a space of the form $\{x_0 + Im(N)\} \cap \mathbb{R}^N_+$, where x_0 is some positive initial concentration.

We state the following version of a deficiency theorem:

Theorem (Deficiency-Zero Theorem)

Consider a mass-action reaction network of deficiency zero. Assume that the network is weakly reversible. Then the following alternatives hold for any arbitrary parameter set:

1. *The system admits neither a positive equilibrium, nor a positive periodic orbit.*
2. *The system has the following properties: each positive stoichiometric compatibility class contains precisely one equilibrium, this equilibrium is asyptotically stable, and there is no nontrivial periodic orbit.*

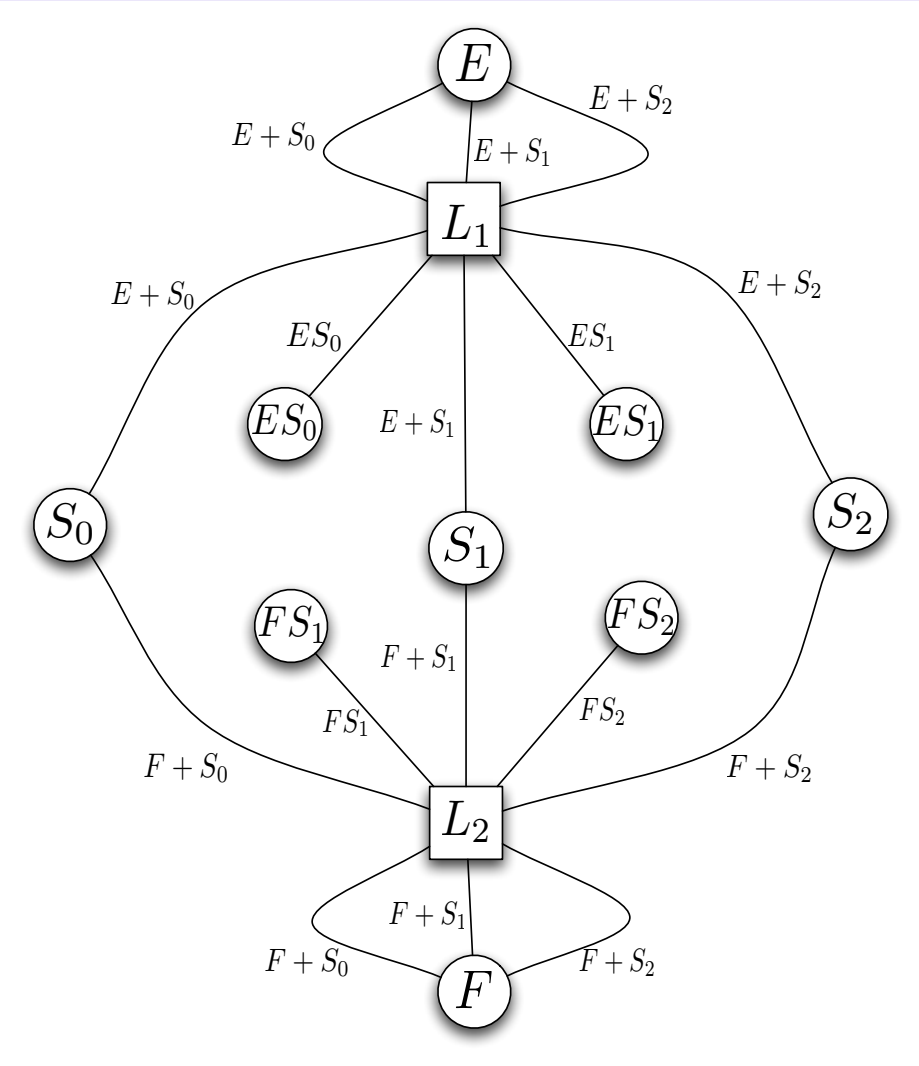
The power of the deficiency-zero theorem is that by the definition of the deficiency index we can identify a class of networks which cannot have multiple positive steady steady.

SCL Graph

Definition

The species-complex-linkage graph $G_{SCL} = (V, G)$ is an undirected bipartite graph which has two types of vertices: The species (represented by V_1) and the complex linkage classes (represented by V_2). We draw an edge from a species vertex to a linkage vertex if the linkage class has a complex which contains this particular species.

SCL Graph


E
$E + S_2$
$E + S_0$
$E + S_1$
L_1
$E + S_0$
$E + S_2$
ES_0
ES_1
ES_0
$E + S_1$
ES_1
S_0
S_2
S_1
FS_1
FS_2
$F + S_1$
FS_1
FS_2
$F + S_0$
$F + S_2$
L_2
$F + S_1$
$F + S_0$
$F + S_2$
F

Modularity

Is there a 'dissection' or 'modularity' of the nonlinear network behaviour? This must be a central idea of any complexity theory based on networks, especially in systems biology (Savageau and others).
Reaction schemes: consider the stationary states $Nv = 0$.
Transformation into reaction coordinates yields the following expression for the Jacobian J:

$$J = N\ diag(v)\kappa^t diag(h_0)$$

with κ being the kinetic matrix and $h_0 = x_0^{-1}$ is the inverse of stationary species concentrations.

The intersection of the set of stationary reaction rates with the kernel of N induces a convex polyhedric cone.

$$K_v = \{v \in \mathbb{R}^r | Nv = 0\ ,\ v \geq 0\} = (ker(N)) \cap \mathbb{R}^r_+$$

$$= \sum_{i=1}^{t} j_i E_i\ ,\ j_i > 0\ \forall i\}$$

The minimal number of generating vectors E_i are called *extremal currents* (Clarke), in the bioinformatics literature also called "elementary flux modes". All equilibrium reactions can be written in the form $v(j) = \sum_{i=1}^{t} j_i E_i$.

Substitution of the extremal currents into the Jacobian yields

$$\tilde{J}(v) = N\ diag(\sum_{i=1}^{t} j_i E_i)\kappa^t diag(h_0).$$

With the implicit definition $J(v) = \tilde{J}(j)diag(h_0)$ the local asymptotic stability of every equilibrium can be obtained by summing up the stability properties of the different extremal currents:

$$\tilde{J}(j)diag(h_0) = j_1\tilde{J}(E_1)diag(h_0) + \ldots + j_t\tilde{J}(E_t)diag(h_0).$$

Extremal currents can be stable, unstable, or mixing stable (Clarke, definition based on local Lyapunovfunctions). A stable and mixing stable extremal current is called positive loop. Unstable or not mixing stable extremal currents are called stoichiometric generators. They lead to complex dynamical behaviour.

Toric Variety (1)

In bifurcation theory, the behaviour of the system depends on the parameters k_i and c_i. Applying bifurcation theory to stoichiometric network analysis requires some additional restrictions on the convex cone. Here every ray in the convex cone, i.e., $[z_1, ..., z_r]$ with $z \in ker(N) \cap \mathbb{R}^r_+$, corresponds to some positive solution of

$$Nv(x_0, k) = 0 \tag{8}$$

for some values of k. To find such positive solution x, it is sufficient to solve the system

$$z = v(x, k) = x_0 v(x, k)$$

where constant x_0 is introduced because of the ambiguous length of z. Then, the interior of the convex cone corresponds to all positive solutions of (8) for any value of k.

Toric Variety (2)

However, for fixed values of k and a given $z > 0$, a solution for the system $z = v(x,k) = x_0 v(x,k)$ is not guaranteed. In fact, z needs to satisfy additional conditions that have been derived by Karin. A positive solution exists if, and only if, $z \in V(I_{tor}^{def})$ where $V(I_{tor}^{def})$ is an affine deformed toric variety of the deformed toric ideal,

$$I_{tor}^{def} = \{f \in \mathbb{Q}[z] | f(v(x,k)) = 0\} \subseteq \mathbb{Q}(k)[z].$$

Since each ray of the cone is parametrized by a different choice of convex parameters j, there are also restrictions on j. A positive solution will exist if $j \in V(J) \cap \mathbb{R}_+^M$ where $V(J)$ is a variety of the new ideal $J \subseteq \mathbb{Q}(k)[j]$ created by substituting $z = \sum_{i=1}^{M} j_i E_i$ into I_{tor}^{def}. For our purposes, the preferred basis of I_{tor}^{def} in all examples will be a Gröbner basis with lexicographic ordering.

Hopf Bifurcation

Assume that for a system

$$\dot{x} = f(x, k) \quad x \in \mathbb{R}^n, k \in \mathbb{R}$$

there exists a smooth curve of equilibria $(x(k), k)$ with $x(k_0) = k_0$. This system has a simple Hopf bifurcation if the following conditions on the coefficients of the characteristic polynomial of the Jacobian $D_x f(x_0, k_0)$ hold:

(CH1) $a_0(k_0) > 0, H_1(k_0) > 0, \ldots, H_{n-2}(k_0) > 0, H_{n-1}(k_0) = 0;$

(CH2) $\frac{d}{dk} H_{n-1}(k_0) \neq 0.$

Hopf Bifurcation

Let

$$\dot{x} = Nv(x,k) \quad x \in \mathbb{R}^n, k \in \mathbb{R}^p_+$$

be a system consisting of M currents with corresponding convex parameters $j_1, ..., j_M$. Let H_i denote the $i \times i$ Hurwitz determinants of the characteristic polynomial

$$det(\lambda I - Jac(j,h)) = \lambda^n + \alpha_{n-1}(j,h)\lambda^{n-1} + \cdots + \alpha_1(j,h)\lambda + \alpha_0(j,h). \quad (9)$$

For a set of (j,h) satisfying conditions:

(HB1) $j \in V(J) \cap \mathbb{R}^M_+$ and $h \in \left\{1/x : Nv(x,k) = \sum_{i=1}^M j_i E_i\right\} \cap \mathbb{R}^n_+$

(HB2) $\alpha_0(j,h) > 0$, $H_1(j,h), ..., H_{n-2}(j,h) > 0$, $H_{n-1}(j,h) = 0$; and

(HB3) $\sum_{s=1}^M \frac{\partial H_{n-1}}{\partial j_s} \frac{\partial j_s}{\partial k_i} \neq 0$ where k_i is any one of the reaction constants

then there exists a constellation $k \in \mathbb{R}^p_+$ at which the network undergoes a simple Hopf bifurcation.

We apply the conditions to a modified Selkov model of glycolytic oscillations described by Eiswirth et al.. The model describes two species interacting through five reactions, where S_1 denotes the product Fructokinase-1,6-biphosphate (F1,6BP) S_2 denotes adenosine triphosphate (ATP):

$$\begin{aligned} 2S_1 + S_2 &\xrightarrow{k_1} 3S_1 \\ \emptyset &\underset{k_3}{\overset{k_2}{\rightleftarrows}} S_1 \\ \emptyset &\underset{k_5}{\overset{k_4}{\rightleftarrows}} S_2 \end{aligned}$$

Their dynamics take the form,

$$\begin{aligned} \dot{x}_1 &= k_1 x_1^2 x_2 + k_2 - k_3 x_1 \\ \dot{x}_2 &= -k_1 x_1^2 x_2 + k_4 - k_5 x_2 \end{aligned}$$

The stoichiometric matrix and vector of reaction rates are,

$$N = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 1 & -1 \end{bmatrix} v(x;k) = \begin{bmatrix} k_1 x_1^2 x_2 \\ k_2 \\ k_3 x_1 \\ k_4 \\ k_5 x_2 \end{bmatrix}$$

The model consists of three extreme currents:

$$E_1 = (0,1,1,0,0) \quad E_2 = (0,0,0,1,1) \quad \text{and} \quad E_3 = \quad (1,0,1,1,0).$$

The first two extreme currents describe subnetworks of inflow and outflow of F1,6BP and ATP, respectively. The third current combines the autocatalytic formation of F1,6BP via ATP with outflow of F1,6BP and inflow of ATP.

Convex parameters j that lie in the variety of the toric ideal $V(J)$ must satisfy the equations,

$$\begin{aligned} k_5 k_3^2 j_3 - k_1 (j_1 + j_3)^2 j_2 &= 0 \\ j_1 - k_2 &= 0 \\ j_2 + j_3 - k_4 &= 0 \end{aligned} \tag{10}$$

Via Hermite Normal Form we can calculate $x_1 = \frac{j_1 + j_3}{k_3}$, $x_2 = \frac{j_2}{k_5}$ and,

$$\begin{aligned} h_1 &= \frac{k_3}{j_1 + j_3} \\ h_2 &= \frac{k_5}{j_2} \end{aligned} \tag{11}$$

The convex Jacobian is,

$$Jac(j) = \begin{bmatrix} (j_3 - j_1) h_1 & j_3 h_2 \\ -2 j_3 h_1 & -(j_2 + j_3) h_2 \end{bmatrix}$$

The characteristic polynomial is,

$$\lambda^2+(h_2j_2+j_3h_2+h_1j_1-j_3h_1)\lambda+j_3^2h_2h_1-j_3h_1h_2j_2+h_1j_1j_3h_2+h_1h_2j_1j_2$$

so, the Hurwitz determinants follow as,

$$\begin{aligned} a_0 &= j_3^2h_2h_1 - j_3h_1h_2j_2 + h_1j_1j_3h_2 + h_1h_2j_1j_2 \\ H_1 &= h_2j_2 + j_3h_2 + h_1j_1 - j_3h_1. \end{aligned}$$

Condition (HB2) is satisfied if $a_0 > 0$ and $H_1 = 0$, namely,

$$2j_3h_2j_2 + j_3^2h_2 - 2j_3j_2h_1 + h_2j_2^2 \quad > \quad 0\,. \tag{12}$$

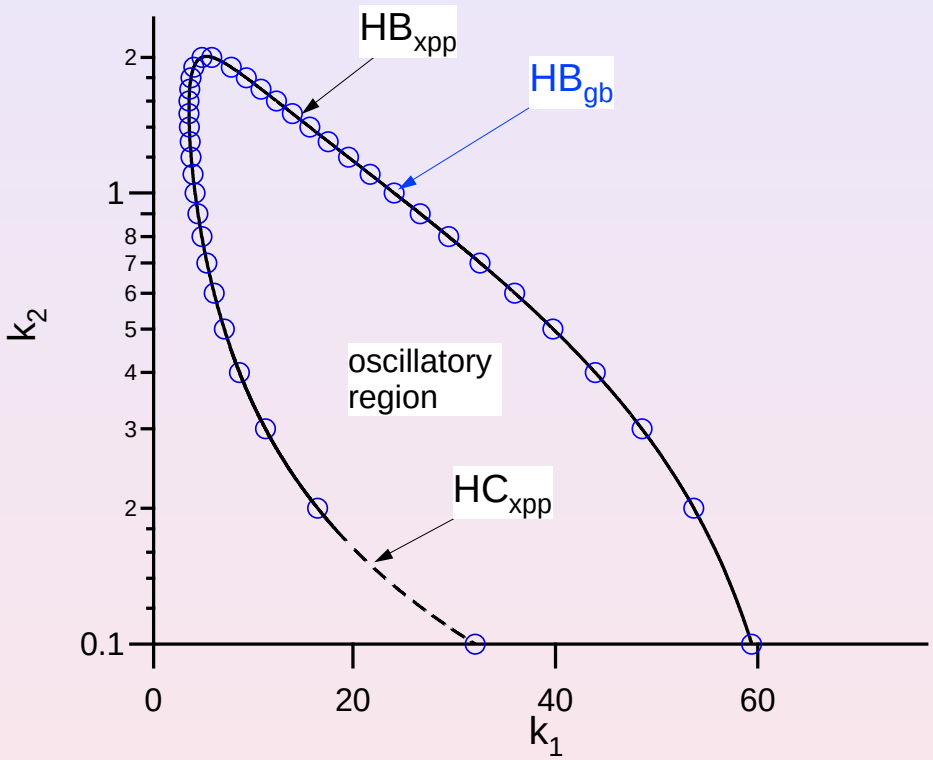
Since $\frac{\partial H_1}{\partial j_1} = h_1$ and $\frac{\partial H_1}{\partial j_2} = h_2$, (HB3) reduces checking that $\frac{\partial H_1}{\partial j_3} \neq 0$, or equivalently,

$$h_2 - h_1 \neq 0. \tag{13}$$

Bifurcation Diagramms

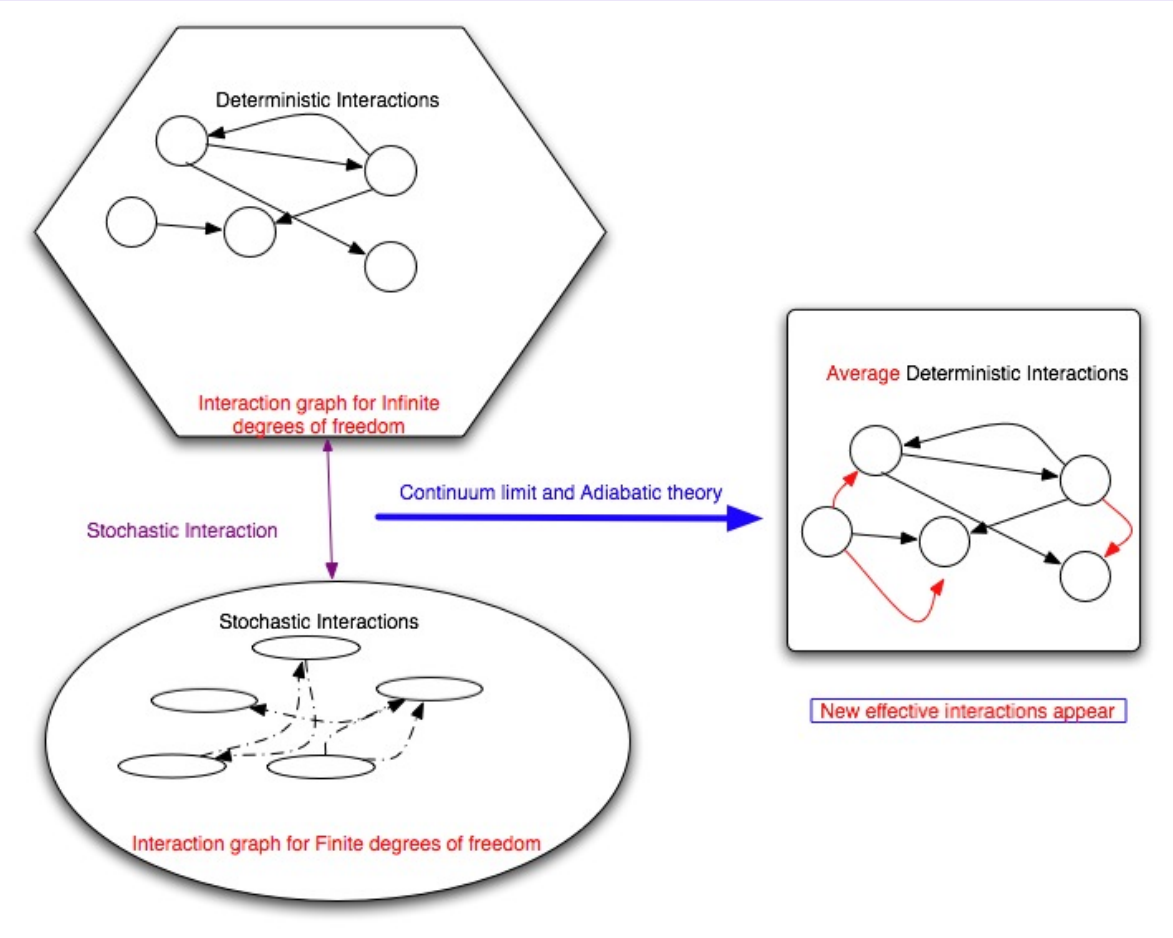
Equations (10) where each $j_t > 0$ for $t = 1, .., 3$ and (11) - (13), gives a set of conditions on (j, h) pair for which there exists a parameter set k such that the system undergoes a simple Hopf bifurcation. In the next figure we compare Hopf bifurcation calculated via our method and a Hopf bifurcation calculated by XPPaut. There is good agreement between the two curves of bifurcations.

Bifurcation Diagramms



Two parameter bifurcation diagram of the glycolytic model for $k_3 = 10$, $k_4 = 4$, $k_5 = 0.5$. k_1 and k_2 are bifurcation parameters. HB_{xpp} and HB_{gb} denote Hopf bifurcations detected by XPPaut and our method, respectively. XPPaut also detects a line of homoclinic bifurcations (HC).

A Hierachy of Limits



Extended Master Equation

Assuming some particles can have finitely many discrete states the general time evolution of P is now given by a generalised master equation (ME)

$$\frac{\partial P(t,\mathbf{n})}{\partial t} = (L_R^* + L_E^*) \circ P(t,\mathbf{n}) + \frac{1}{\epsilon} K^T(\mathbf{n})\, P(t,\mathbf{n}), \qquad (14)$$

with $\epsilon > 0$ a small parameter fixing the relative time scale on which the Markov chain evolves. Also P, L_R^*, L_E^* and K^T are sufficiently regular such that (14) has a unique solution for all times $t > 0$.

The Average Dynamics

Upon the assumption that the evolution of the MC is faster than the deterministic dynamics, the macroscopic evolution of the combined system is described by the average dynamics, which is given by

$$\begin{cases} \dfrac{dx_i(t)}{dt} = \displaystyle\sum_{\sigma=1}^{M} \mu_\sigma(\mathbf{x}(t))\, f_i^{(\sigma)}(\mathbf{x}(t)), \\ x_i(0) = x_{i,0}. \end{cases} \tag{15}$$

Here μ is an invariant measure of $\mathcal{K}(\mathbf{x})$, and $i = 1, ..., N$. The invariant measure μ can be a convex combination of invariant measures compatible with the initial values $\{p_i(0)\}_{i=1}^{M}$, namely the support of μ in Σ is equal to the support of $p(0)$.

We now define an extended reaction scheme for a genetic clock:

- A is the activator protein and M_A its corresponding mRNA,
- R is the repressor protein and M_R its corresponding mRNA,
- C is a complex formed by A and R.

Each gene can be either active or inactive, D_A, D_R denote the inactive states, and D'_A, D'_R the active states. We now collect all necessary reactions to model this situation:

Gene activation:

$$A + D_A \rightleftharpoons [\theta_A]\gamma_A D'_A, A + D_R \rightleftharpoons [\theta_R]\gamma_R D'_R$$

Transcription:

$$D_A \rightarrow^{\alpha_A} D_A + M_A, \qquad D'_A \rightarrow^{\alpha'_A} D'_A + M_A$$
$$D_R \rightarrow^{\alpha_R} D_R + M_R, \qquad D'_R \rightarrow^{\alpha'_R} D'_R + M_R$$

Translation:

$$M_A \rightarrow^{\beta_A} A, \qquad M_R \rightarrow^{\beta_R} R$$

Regulation and inhibition of protein A:

$$A + R \to^{\gamma_C} C, \qquad C \to^{\delta_A} R$$

Degradation:

$$M_A \to^{\delta_{MA}} \emptyset, \qquad M_R \to^{\delta_{MR}} \emptyset,$$

$$A \to^{\delta_A} \emptyset, \qquad R \to^{\delta_R} \emptyset.$$

We next need to introduce the following concentrations for smaller molecules in the system and their complexes:

$$a = [A], \quad c = [C], \quad r = [R], \quad m_A = [M_A], \quad m_R = [M_R].$$

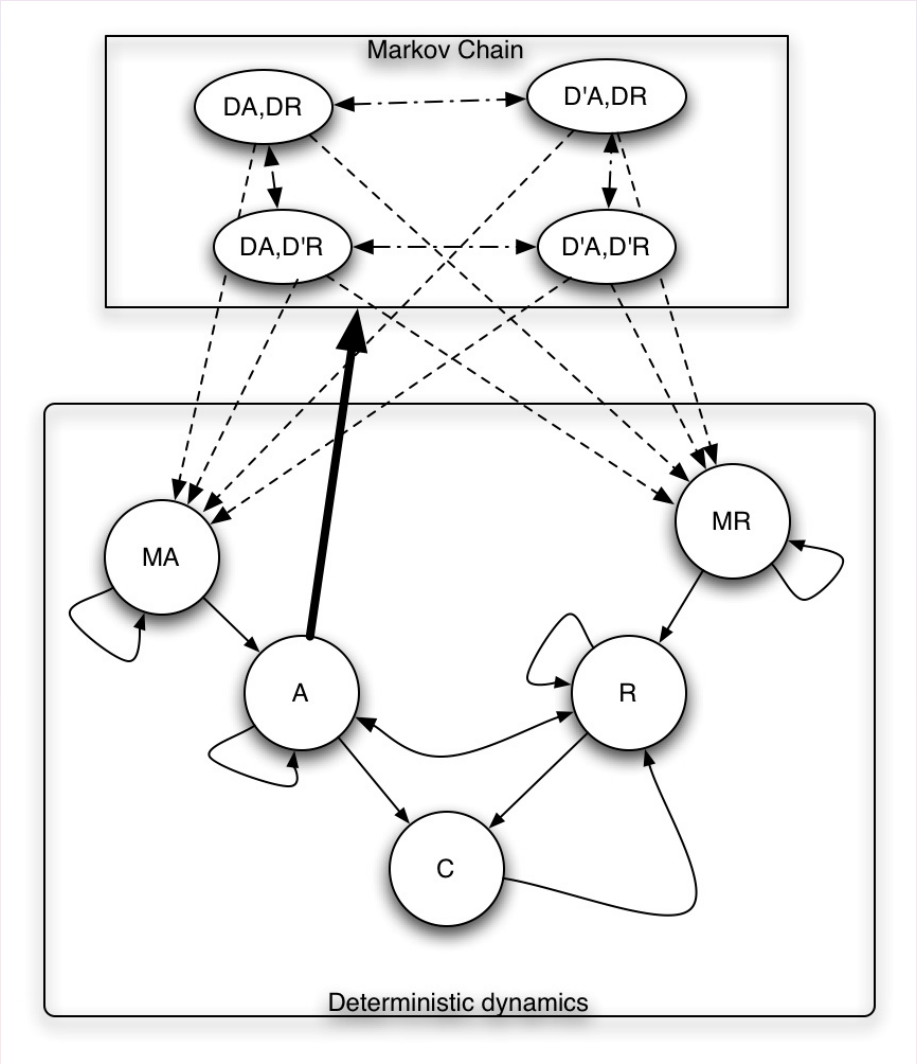


Figure: Combined interaction graph $\vec{\mathcal{I}}_C$ for the VKBL model.

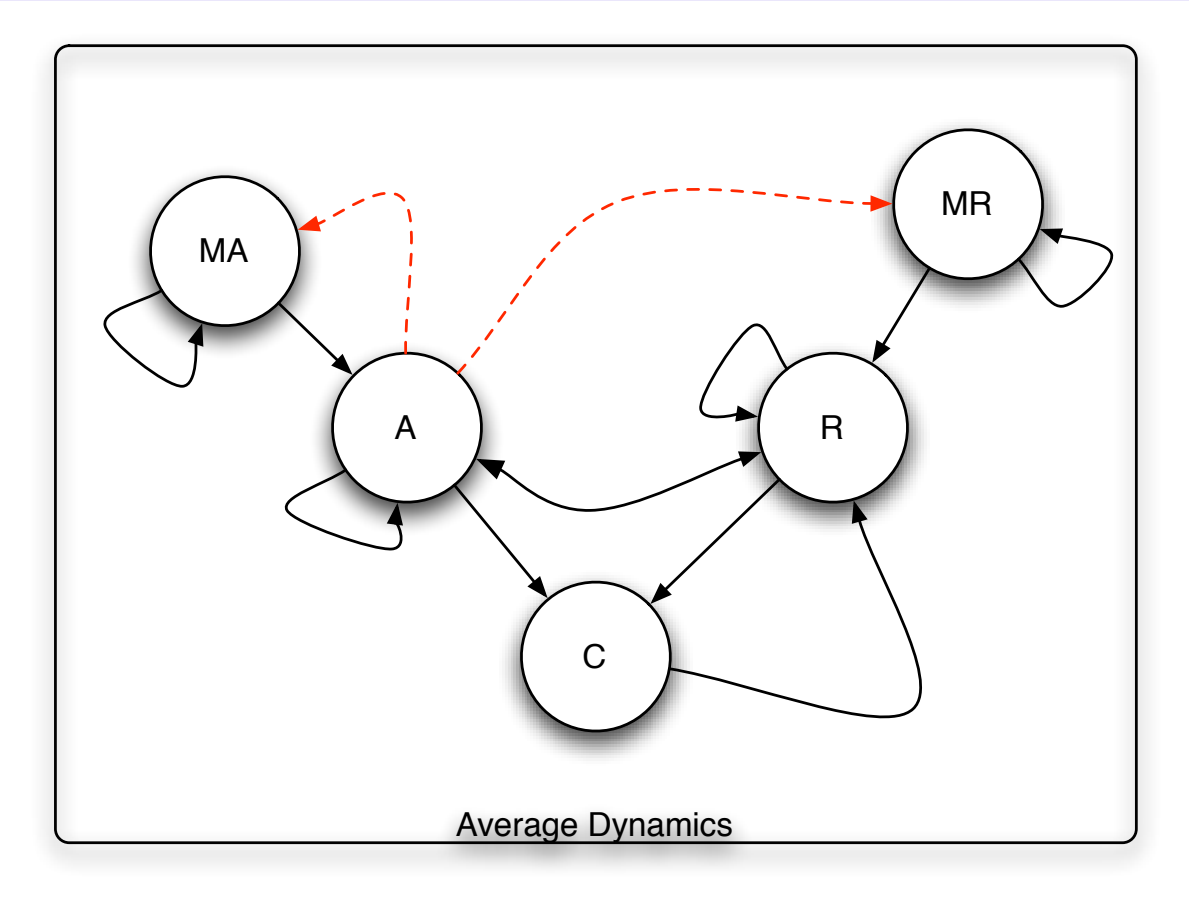


Figure: The averaged interaction graph $\bar{\vec{\mathcal{I}}}$ for the VKBL model.

$$\dot{a}(t) = -\delta_A\, a(t) - \gamma_C\, a(t)\, r(t) + \beta_A\, m_A(t),$$

$$\dot{c}(t) = -\delta_A\, c(t) + \gamma_C\, a(t)\, r(t),$$

$$\dot{r}(t) = -\delta_R\, r(t) - \gamma_C\, a(t)\, r(t) + \beta_R\, m_R(t) + \delta_A\, c(t),$$

$$\dot{m}_A = \frac{\alpha_A \theta_A}{\theta_A + \gamma_A a(t)} + \frac{\alpha'_A \gamma_A a(t)}{\theta_A + \gamma_A a(t)} - (\delta_{MA} + \beta_A)\, m_A(t),$$

$$\dot{m}_R = \frac{\alpha_R \theta_R}{\theta_R + a(t)\,\gamma_R} + \frac{\alpha'_R a(t)\,\gamma_R}{\theta_R + a(t)\,\gamma_R} - (\delta_{MR} + \beta_R)\, m_R(t). \tag{16}$$

The Average Dynamics

Upon the assumption that the evolution of the MC is faster than the deterministic dynamics, the macroscopic evolution of the combined system is described by the average dynamics, which is given by

$$\begin{cases} \dfrac{dx_i(t)}{dt} = \displaystyle\sum_{\sigma=1}^{M} \mu_\sigma(\mathbf{x}(t))\, f_i^{(\sigma)}(\mathbf{x}(t)), \\ x_i(0) = x_{i,0}. \end{cases} \tag{17}$$

Here μ is an invariant measure of $\mathcal{K}(\mathbf{x})$, and $i = 1, ..., N$. The invariant measure μ can be a convex combination of invariant measures compatible with the initial values $\{p_i(0)\}_{i=1}^{M}$, namely the support of μ in Σ is equal to the support of $p(0)$.