Modelling and Analysis Tools for Biochemical Networks

Antonis Papachristodoulou

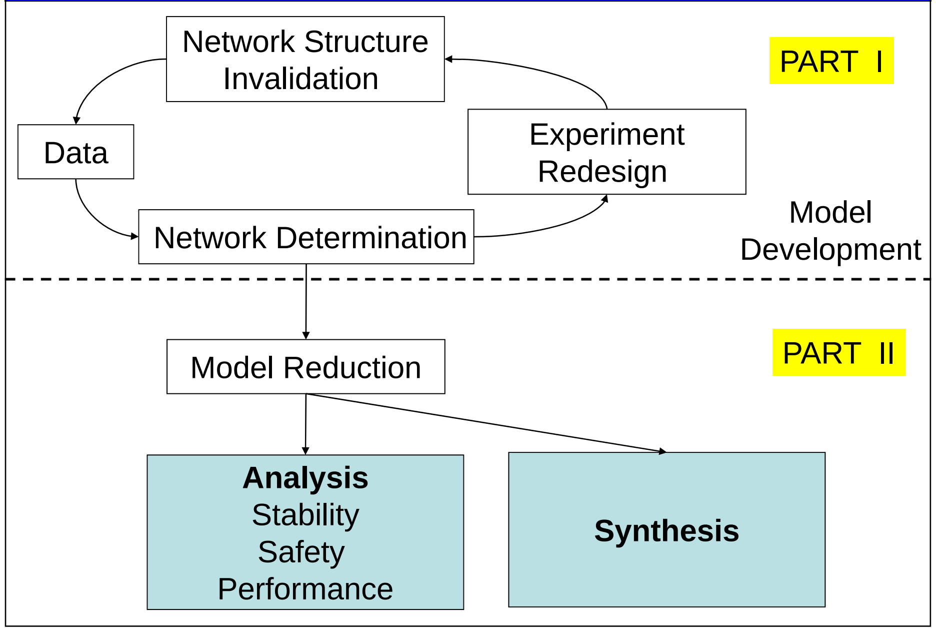
Department of Engineering Science
University of Oxford

www.eng.ox.ac.uk/control/sysos, antonis@eng.ox.ac.uk

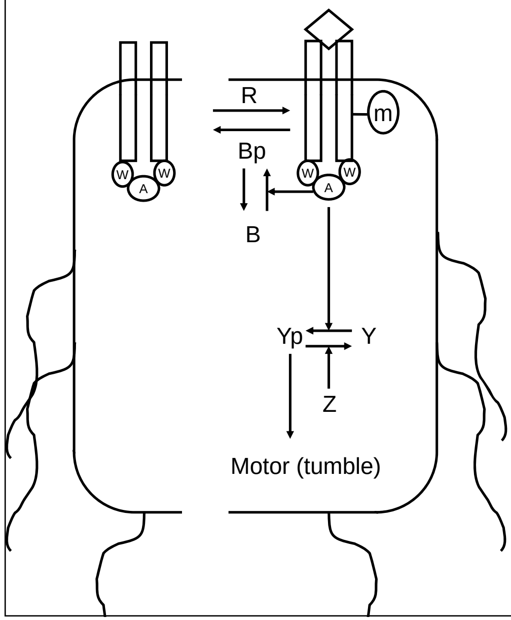
ICM 2010 Satellite Workshop on
Application of Control Theory & Optimization Techniques in Biochemical Pathways
16-18 August 2010, Hyderabad, India

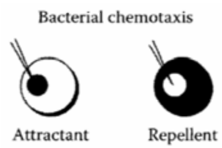
EPSRC Engineering and Physical Sciences Research Council | UNIVERSITY OF OXFORD | BBSRC bioscience for the future

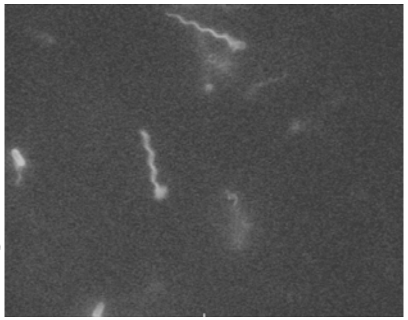
Motivation and Outline



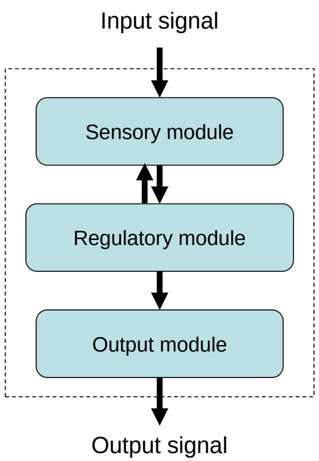
Bacterial Chemotaxis

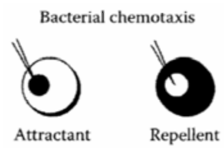


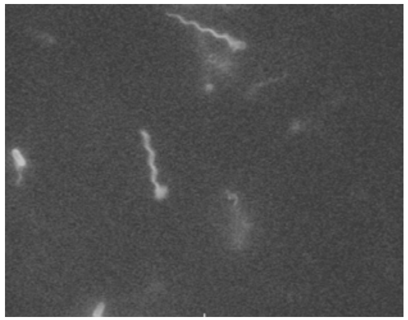




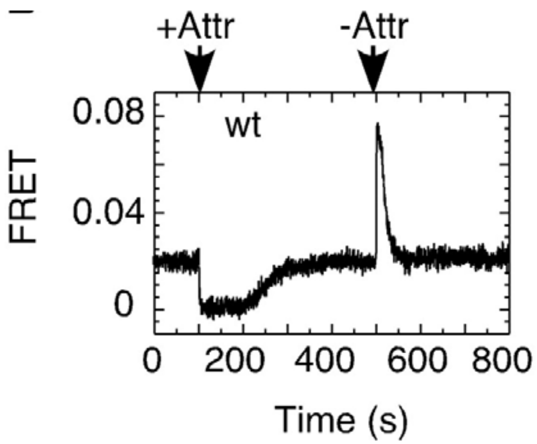
Bacterial Chemotaxis







Perfect Adaptation and Integral Feedback in *E. coli*

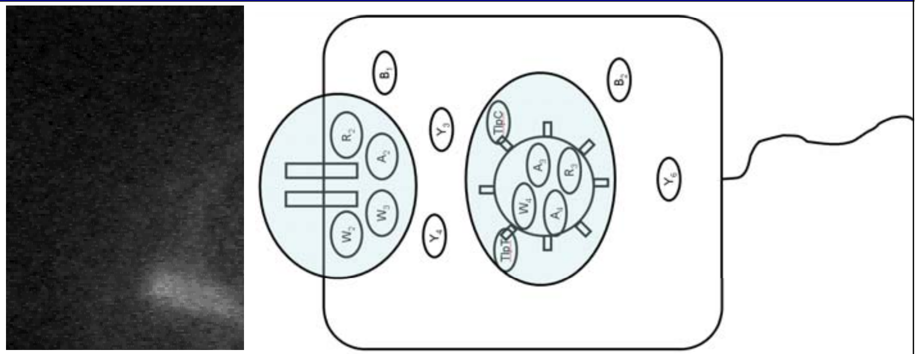


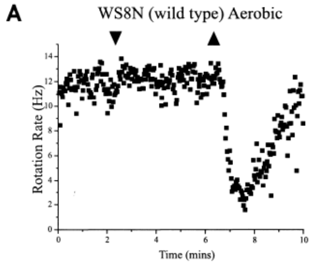
Sourjik et al., PNAS

- Perfect adaptation
- Hence Integral Feedback

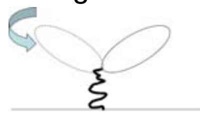
Yi et al., PNAS

Rhodobacter Sphaeroides





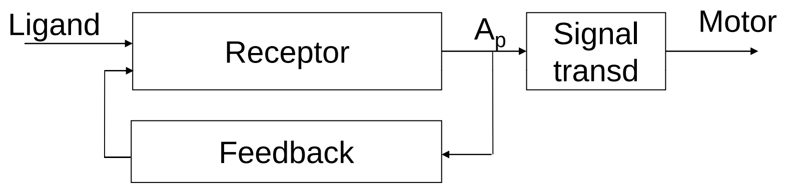
- Also perfect adaptation.
- Hence Integral feedback.



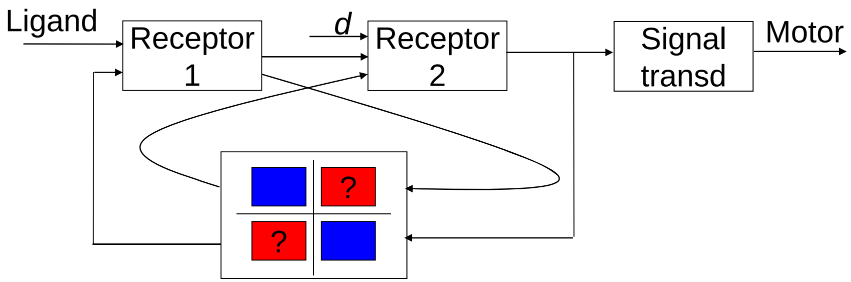
Martin et al, *J. Bact.* 183(24) pp.7135-7144

Block Diagram Representation

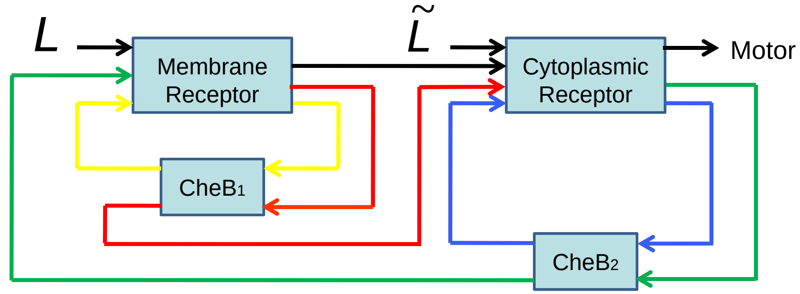
E. coli



R. sphaeroides



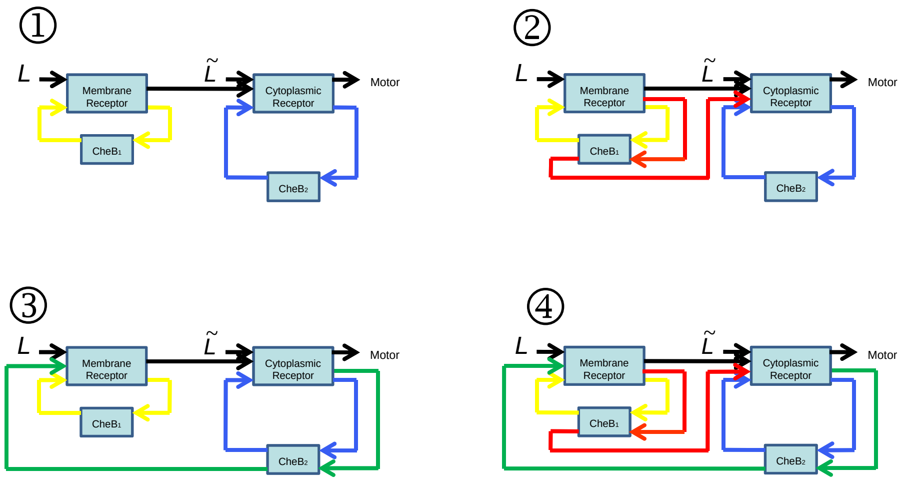
Feedback Interconnection



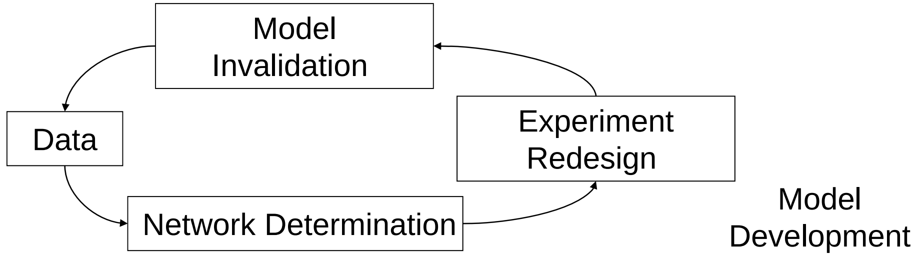
- Which feedback configuration is correct?
- Created four models with different feedback connectivities, each reproducing wild type data, then assessed their performance.
- Undertaken experiments to invalidate models.

Four Possible Feedback Interconnections

Models are initially fitted to wild-type ligand response data



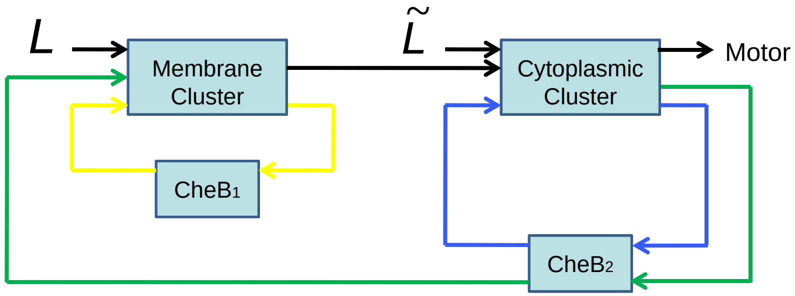
Methodology



- Iterate to invalidate models and create new hypotheses.
- New models reproduce all available data.
- New experiments are designed and undertaken in the lab.
- Hope is that the new data can be used to invalidate some models.
- We have actually used this method in the laboratory.

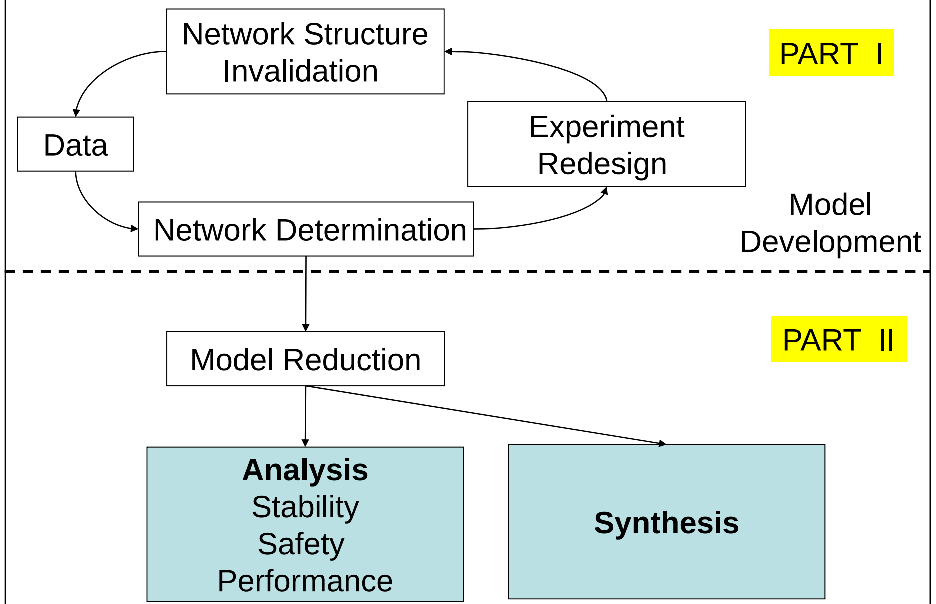
BMC Systems Biology 2009, 3:25
BMC Bioinformatics 2009, 10:132
BMC Systems Biology 2009, 3:103
BMC Systems Biology 2010, 4:38

From model invalidation procedure, only one model survives…



- 16 nonlinear ODEs are used to describe this system
- Uncertainties in most parameters.
- What can we now say about this system?
- How can we evaluate its stability, performance or understand its properties?

Motivation and Outline



Mathematical Setting

- Dynamical system model:

$$\dot{x} = f(x,u,\theta) \qquad \text{or} \qquad x_{k+1} = f(x_k,u_k,\theta)$$
$$y = h(x) \qquad\qquad\qquad y_k = h(x_k)$$

- State-space/Input Space/Parameter Space:

$$x \in X \subseteq \mathbb{R}^{n_x}, \quad u \in U \subseteq \mathbb{R}^{n_u}, \quad \theta \in \Theta \subseteq \mathbb{R}^{n_\theta}, \quad y \in \mathbb{R}^{n_y}$$

- All functions are assumed to be sufficiently smooth.
- Model has been proposed *a - priori* but can be uncertain or under the influence of disturbances/inputs.
- Interested in analysis questions.

Questions of Interest

Is the system *robustly performing well*?

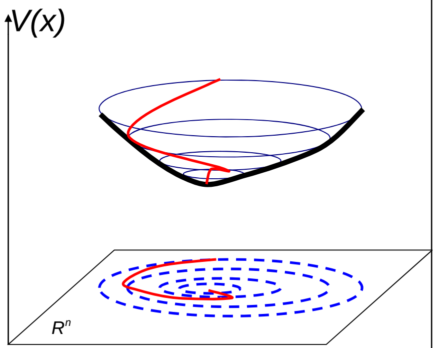
- Stability/Robust Stability of the equilibrium.
  Assume: $u = 0,\ f(x^*,0,\theta) = 0,$ is x^* stable/as.stable?
- Performance/Robust Performance.
  Is some performance metric guaranteed?
- Safety/Robust Safety
  Does dynamical system ever evolve into unsafe regions?
- Eventuality/Robust Eventuality
  Does it visit a particular set of states?

Analysis in Dynamical Systems

Lyapunov:
$\dot{x} = f(x),\ f(0) = 0,\ f$ Lipschitz,
$x \in X \subseteq \mathbb{R}^{n_x},\ 0 \in X$
If $\exists\ V : X \to R$ cont. differentiable s.t.

$V(0) = 0$, and $V(x) > 0$ in $X \setminus \{0\}$
$$-\dot{V}(x) = -\frac{\partial V(x)}{\partial x} f(x) > 0 \text{ in } X \setminus \{0\}$$

then $x = 0$ is asymptotically stable.



Checking if $p(x) \geq 0$ is NP-hard when $\deg(p) \geq 4$, $p \in \mathbb{R}[x]$.

Positive Polynomials and Sum of Squares

Shor: $p(x)$ is Sum of Squares $\Rightarrow$ $p(x)$ is positive semi-definite

$$p(x) = \sum_{i=1}^{m} f_i^2(x) \underset{\not\Leftarrow}{\Rightarrow} p(x) \geq 0$$

Worst-case polynomial time complexity | NP-hard when deg(p) ≥ 4

Can be reduced to a semidefinite programme (SDP) (Parrilo), which can be setup and solved using SOSTOOLS:
www.eng.ox.ac.uk/control/sostools

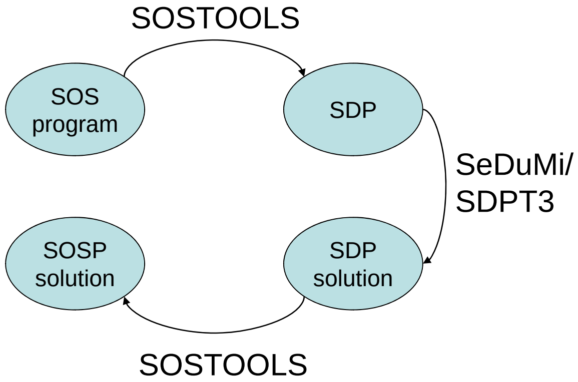
Can also search for unknown coefficients of $p(x)$ so that $p(x)$ is SOS

Prajna, Papachristodoulou, Parrilo CDC02

SOSTOOLS

Formulates and solves the equivalent semidefinite programme (SDP)

www.eng.ox.ac.uk/control/sostools



How to Construct Polynomial Lyapunov Functions

Lyapunov: For $\dot{x} = f(x)$, with $f(0) = 0$, find V such that

$$V(x) > 0$$

$$-\dot{V}(x) = -\frac{\partial V}{\partial x} f(x) > 0$$

Then $x = 0$ is asymptotically stable.

For $\dot{x} = f(x)$, with $f(0) = 0$, find V, $\varphi(x) > 0$, $\psi(x) > 0$ such that

$$V(x) - \varphi(x) \text{ is SOS}$$

$$-\frac{\partial V}{\partial x} f(x) - \psi(x) \text{ is SOS}$$

Then $x = 0$ is asymptotically stable.

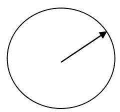
$$\varphi(x) = \sum_{i=1}^{n} \sum_{j=1}^{\deg(\varphi(x))/2} c_{ij} x_i^{2j}, \quad c_{ij} \geq 0, \quad \sum_{j=1}^{\deg(\varphi(x))/2} c_{ij} \geq \gamma > 0 \ \forall i$$

Papachristodoulou, Prajna CDC02

How to Construct Polynomial Lyapunov Functions

$$\dot{x} = f(x), \ f(0) = 0$$

$$D = \left\{ x \in R^n \,\middle|\, a(x) \triangleq \|x\|_2^2 - \gamma \leq 0 \right\}$$



Assume for now that f is polynomial and look for V polynomial.

Given $\dot{x} = f(x)$ with $f(0) = 0$ and $a(x) \leq 0$, find a polynomial $V(x)$ of order ≥ 2, positive definite polynomials $\varphi(x)$, $\psi(x)$ and $p(x) \geq 0$ such that

$$V(x) - \varphi(x) \geq 0 \quad \Rightarrow V(x) \geq \varphi(x) > 0$$

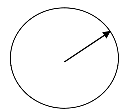
$$-\frac{\partial V(x)}{\partial x} f(x) + p(x)a(x) - \psi(x) \geq 0 \quad \Rightarrow -\dot{V}(x) \geq \psi(x) > 0 \text{ on } D$$

Then $x = 0$ is asymptotically stable and V is a Lyapunov function.

How to Construct Polynomial Lyapunov Functions

$$\dot{x} = f(x), \ f(0) = 0$$

$$D = \left\{ x \in R^n \,\middle|\, a(x) \triangleq \|x\|_2^2 - \gamma \leq 0 \right\}$$



Assume for now that f is polynomial and look for V polynomial.

Given $\dot{x} = f(x)$ with $f(0) = 0$ and $a(x) \leq 0$, find a polynomial $V(x)$ of order ≥ 2, positive definite polynomials $\varphi(x)$, $\psi(x)$ and $p(x)$ SOS such that

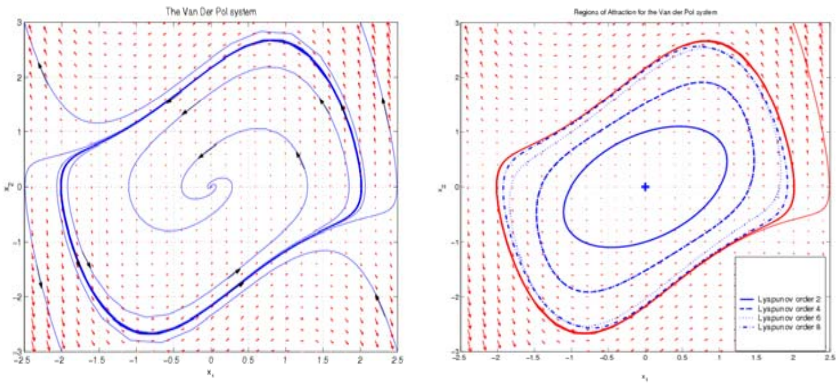
$$V(x) - \varphi(x) \text{ is SOS}$$

$$-\frac{\partial V(x)}{\partial x} f(x) + p(x)a(x) - \psi(x) \text{ is SOS}$$

Then $x = 0$ is asymptotically stable and V is a Lyapunov function.

Van der Pol Oscillator

$$\dot{x}_1 = -x_2$$
$$\dot{x}_2 = x_1 - (1 - x_1^2)x_2$$



Papachristodoulou, 2005

Problems Solved Using SOS

From Systems Theory:

- Robust Stability/Performance analysis for nonlinear systems;
- Analysis of hybrid/switched, time-delay and PDE systems;
- Model invalidation, hybrid verification, controller synthesis.

From Optimization:

- In conjunction with positivstellensatz, provides a nested family of relaxations for NP-hard problems.

Applications:

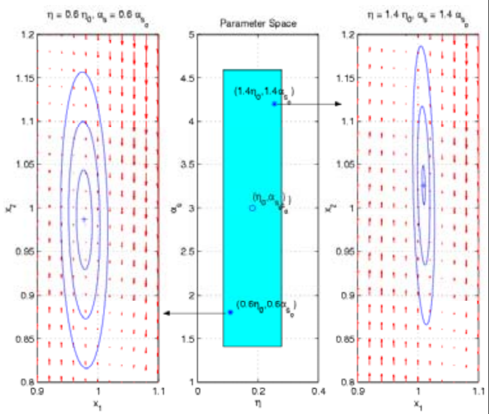
- From Biology to Aerospace, from Fluid Mechanics to the Internet and from Multi-agent Systems to Process Control.

ACC Tutorial session on "SOS in Industry"
CDC Workshop on Biomolecular Circuit Analysis and Design

Biologically Driven Problem: Heat-Shock in *E-Coli*

$$\frac{dS_t}{dt} = \eta - \alpha_0 S_t - \alpha_s \frac{K_s D_t}{1 + K_u U_f + K_s D_t} S_t$$
$$\frac{dD_t}{dt} = K_d \frac{S_t(1 + K_u U_f)}{1 + K_u U_f + K_s D_t} - \alpha_d D_t$$
$$0 = K\left[P_t - U_f\right] - [K + K_{fold}]D_t$$

Robustly stable with respect to η and α_s.



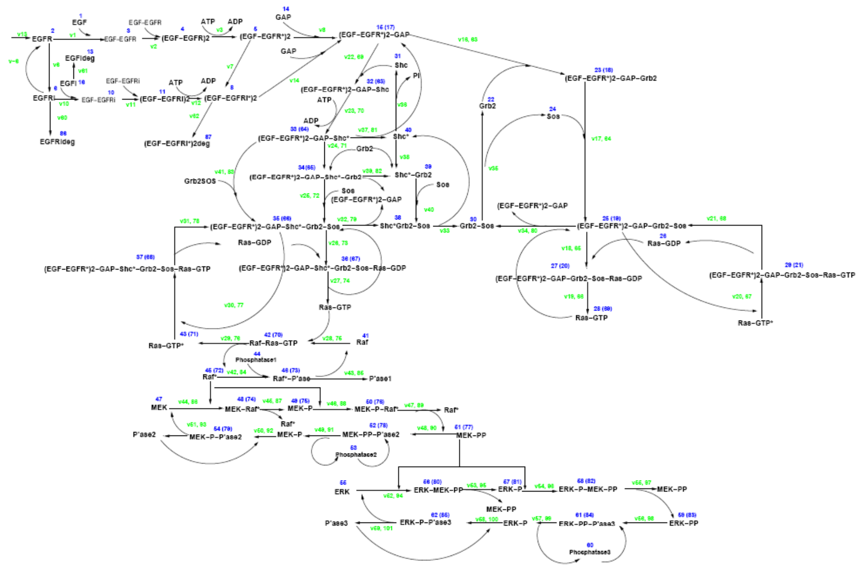
System features:

- Rational vector field;
- Parametric uncertainty;
- Equilibrium moves as parameters change.

Robust stability can NEVER be verified using simulation.

El-Samad et al., Proceedings of the IEEE, 2006

EGF-MAPK



Schoeberl et al, Nat. Biot. 2002

What restricts SOS methods?

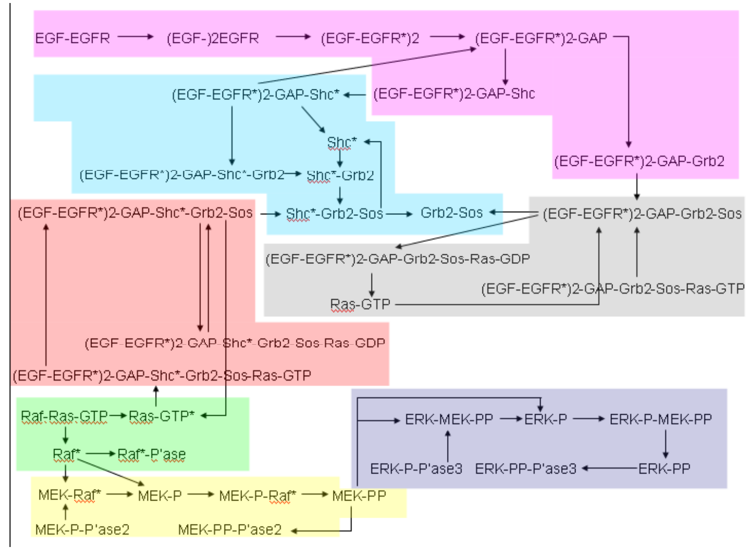
Computation

$$p(x) = \sum_{i=1}^{m} f_i^2(x) = Z(x)^T Q Z(x) = Z(x)^T \left(Q_0 + \sum_{i=1}^{L} \lambda_i Q_i \right) Z(x)$$

Require
$$Q_0 + \sum_{i=1}^{L} \lambda_i Q_i \geq 0$$
$$Z(x)^T Q_i Z(x) = 0, \; i = 1, \ldots, L$$

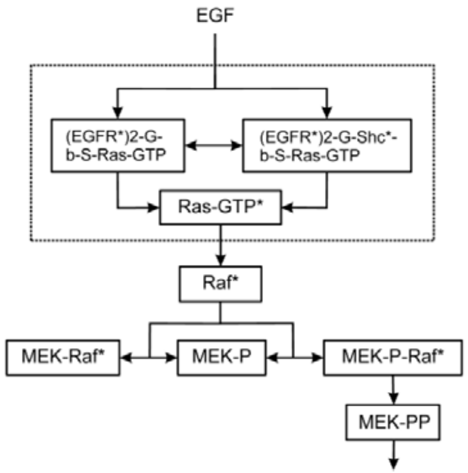
- Reduce the size of $Z(x)$?
- Reduce the number of decision variables?
- Structure the decomposition?

Model Decomposition



Anderson, Chang, Papachristodoulou, to appear

Model Reduction



Anderson, Chang, Papachristodoulou, to appear

Model (Reaction Network) Decomposition

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

For the EGF-MAPK model we consider has $n > 100$.
- This system is difficult to analyze and understand directly.

Instead, we can decompose the dynamical system into interacting subsystems:

$$\dot{x}_1 = f_1(x_1) + g_1(x_1, u_1), \quad y_1 = x_1$$
$$\dot{x}_2 = f_2(x_2) + g_2(x_2, u_2), \quad y_2 = x_2$$
$$u_1 = y_2, \; u_2 = y_1$$

Then we can analyze/understand each subsystem in turn and then attempt to draw global conclusions.

Model (Reaction Network) Decomposition

For static networks (graphs) many decomposition strategies exist:

Given a graph with a weighted Adjacency matrix *A*, assign integer values $z_i = \pm 1$ to each node in the graph so as to minimize:

$$f(z) = \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}(z_i - z_j)^2$$

Using the weighted Laplacian, *L*, solve:

$$\min \quad \frac{1}{2} z^T L z$$

$$\text{s.t.} \quad z_i^2 = 1$$

Relaxation solution is the Fiedler eigenvector, corresponding to the smallest non-zero eigenvalue of *L*.

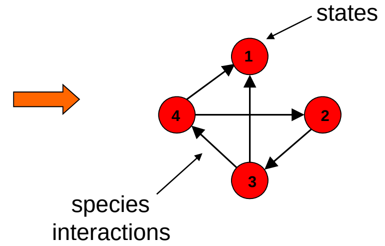
Model (Reaction Network) Decomposition

Can the same ideas be used for partitioning dynamical systems?

Want a decomposition that takes into account structure AND dynamics.

$$\dot{x}_1 = -x_1 + 2x_3^2 x_4 \qquad \dot{x}_3 = -2x_2^4$$

$$\dot{x}_2 = -4x_2^2 - x_4^3 \qquad \dot{x}_4 = x_4^2 - 5x_1^2$$



Let the edge weights represent energy flow between species

Apply spectral partitioning to obtain a decomposition

Partitioning Algorithm

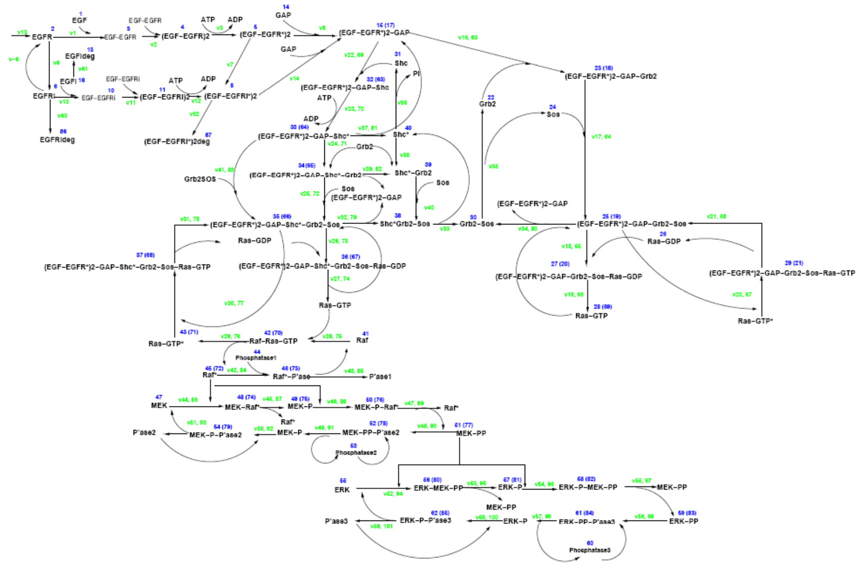
1) Represent $\dot{x} = f(x)$ as a weighted graph $G(V,E)$, where each state is a vertex and an edge connects v_i to v_j if $\dot{x}_i$ is a function of x_j
2) Compute initial condition $\hat{x}$ that maximizes energy input to the system by solving for P:

$$F^T P + PF = -CC^T$$

where F is the linearization of $\dot{x} = f(x)$ and C is the incidence matrix of G.
3) Construct the energy matrix W, where W_{ij} defines the energy flow from v_i to v_j; $W_{ij} \triangleq \|y\|_2^2 = \hat{x}^T X \hat{x}$

where X solves $F^T X + XF = -\delta^T \delta$, where δ is the appropriate column of C.
4) Apply spectral partitioning algorithm on W.

EGF-MAPK



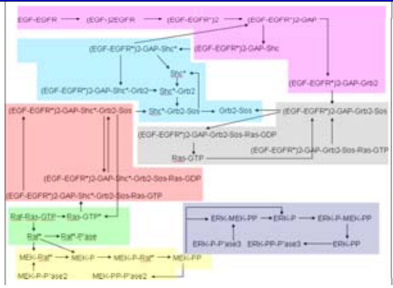
Schoeberl et al, Nat. Biot. 2002

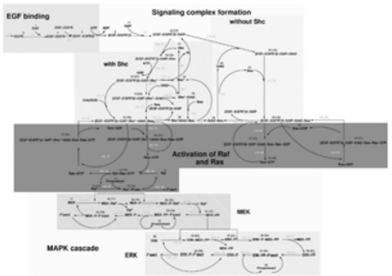
EGF-MAPK

Decomposition of a large network by trying to minimize the energy flow between species in it.

Decomposition of a large network by trying to minimize the retroactivity between the components.

The decompositions look very similar and we are currently investigating the relation between the two approaches





Conzelman et al, IEE Sys Bio, 2004

Dynamical Systems Analysis

$$\dot{x}_1 = f_1(x_1) + g_1(x_1, u_1)$$
$$y_1 = x_1$$

$$\dot{x}_2 = f_2(x_2) + g_2(u_2, x_2)$$
$$y_2 = x_2$$

- Construct a Lyapunov function for each subsystem:

$$V_1(x_1) > 0, \ \frac{\partial V_1}{\partial x_1} f_1(x_1) < 0$$

$$V_2(x_2) > 0, \ \frac{\partial V_2}{\partial x_2} f_2(x_2) < 0$$

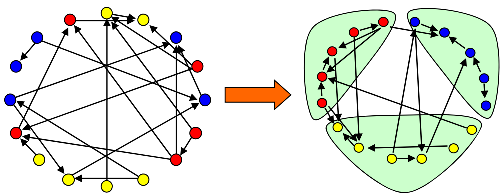
- Test whether the following is a Lyapunov function for the full system

$$V = V_1(x_1) + \alpha V_2(x_2)$$

Numerical Results

Lotka-Voltera Dynamics: 16 states, standard assumptions for scalable analysis do not hold

$$\dot{x}_i = x_i \left(b_i - x_i - \sum_{j=1}^{n} A_{ij} x_j \right)$$



Direct SOS analysis not possible: LMI too larger for SeDuMi to parse

Decompose into 3 subsystems:

$V(x)$ construction	Size of LMI	CPU time
Subsystem for $\mathcal{I}_1$	527 × 146	0.25s
Subsystem for $\mathcal{I}2$	527 × 146	0.25s
Subsystem for $\mathcal{I}3$	902 × 237	0.45s
Composite system	23362 × 4828	1415.23s
Full system	too large to parse	n/a

Can we do better? Yes.

Anderson, Papachristodoulou, ACC 2010

Model Reduction – "Complexity Reduction"

- Given a system

$$\dot{x} = f(x), \ \ y = g(x), \ \ x \in R^n, \ \ x(0) = x_0, \ \ y \in R^m$$

we want to construct a system

$$\dot{\hat{x}} = \hat{f}(\hat{x}), \ \ \hat{y} = \hat{g}(\hat{x}), \ \ \hat{x} \in R^{\hat{n}}, \ \ \hat{x}(0) = \hat{x}_0, \ \ \hat{y} \in R^m$$

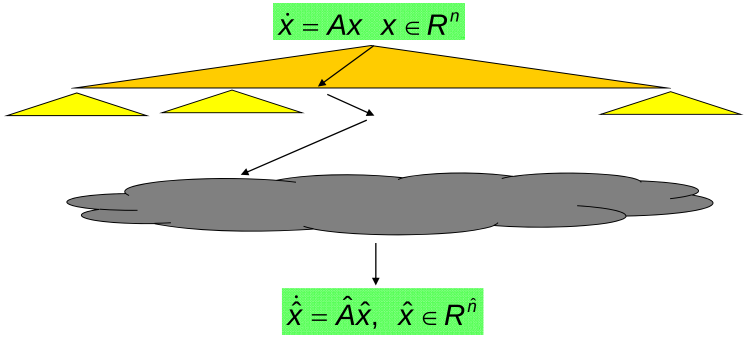
where $\hat{n} << n$ which makes $\|y - \hat{y}\|$ as small as possible.

- In this talk we will aim to "collapse states" - which amounts to reducing the number of differential equations.

- Very similar to Singular Perturbation.

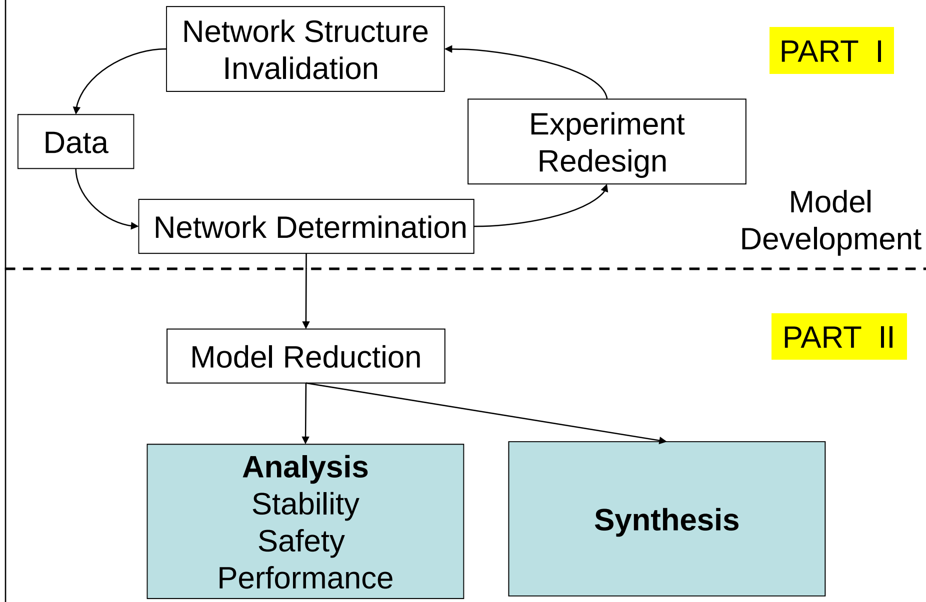
Anderson, Chang, Papachristodoulou, to appear

Our Approach – Linear Case



- Finding the “ground state” is exponential.
- But can still get
  - A bound on the error from the original system
  - A suggested ordered list of states to be removed

Motivation and Outline



Network in Synthetic Biology- the Engineering of Biology

- Establish links between electrical/control/ computer engineers, mathematicians, physicists, biologists and ELSI members.
- A major challenge that Synthetic Biology faces is that new designs have to function in uncertain environments.
- Launched on 6th July 2009.
- First 3-day workshop was on 14-16 September 2009.
- Second 3-day workshop was on 12-14 July 2010.
- Interested? Visit

www.rosbnet.org

RoSBnet: A Network in Synthetic Biology co-funded by BBSRC, EPSRC, AHRC and ESRC

BBSRC bioscience for the future — EPSRC Engineering and Physical Sciences Research Council — Arts & Humanities Research Council — E·S·R·C Economic & Social Research Council

Acknowledgements

Current DPhil (PhD) Students:

- **Yo-Cheng (Mark) Chang, Bence Melykuti, James Anderson, Edward Hancock**

(Previous) PDRAs:

- **Elias August, Mark Roberts, Abdullah Hamadeh, Christoph Maier**

Oxford Collaborators:

- **Judith Armitage, George Wadhams, Biochemistry**
- **Philip Maini, Mathematical Institute**

bbsrc — EPSRC Engineering and Physical Sciences Research Council — E·S·R·C Economic & Social Research Council — Arts & Humanities Research Council