

## **Inverse bifurcation analysis for reverse engineering of gene systems**

James Lu

Johann Radon Institute for Computational and Applied Mathematics  
Linz, Austria

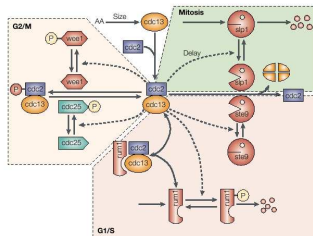
*Joint work with Heinz W. Engl, Peter Schuster*

# Overview

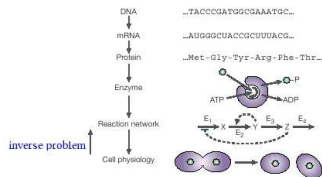
- Motivation
  - gene regulatory systems key to understanding the workings of organisms
- Aim
  - inferring network characteristics from dynamical behavior and engineer desired properties
- Method
  - bifurcation analysis
  - adjoint analysis
  - identification/regularization techniques
  - ...
- Applications
  - regulatory mechanisms for homeostasis
  - engineer/design of genetic switches
- Outlook



# Motivation

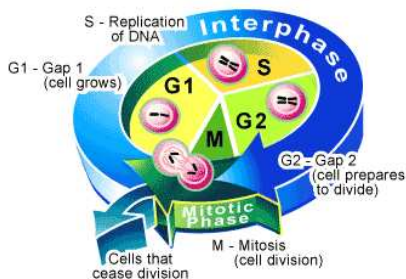


Schematic of cell cycle model (Tyson et al, Mole. Cell. Biol. 2001)



Inverse problem: from cell physiology to infer reaction network

# Motivation



Schematic of phases of cell cycle

- Concentrations of regulatory species typically modelled by **parameterized** ODE systems:

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

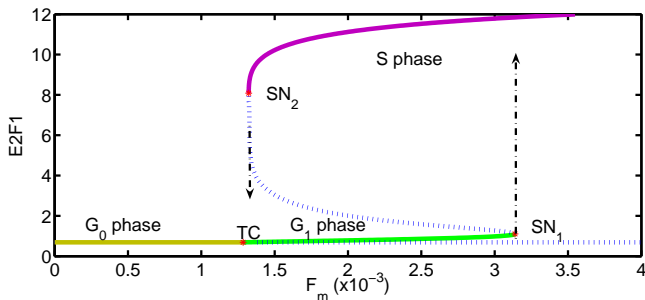
- What is the mathematical correspondence for changes in cycle phase when cell **mass** increases? → **bifurcation** of solution

# Solution continuation and bifurcation analysis

- Equilibrium solution  $x(p)$  of ODE system:

$$0 = f(x(p), p)$$

- One-parameter continuation: the tracing of solution  $x(p)$  as a component of  $p$  is varied.



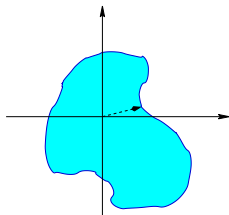
Bifurcation diagram of model for mammalian cell cycle regulation

# Inverse bifurcation analysis

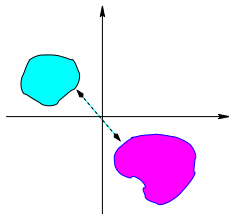
- *Forward* bifurcation analysis:
  - mapping **parameters** → **bifurcation diagrams**
- *Inverse* bifurcation analysis:
  - from certain **observed** or **desired** qualitative behavior, deduce conditions on parameters
  - mapping **geometry of bifurcation diagrams** → **parameters**
- Provides a tool for
  - understanding homeostasis
  - computational design of gene systems

# Geometric properties of bifurcation diagrams

- Parametric distance to bifurcation



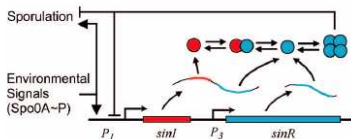
- Parametric distance between qualitative regions



- Biological applications: robustness, evolvability ('tuning' to environment, to perform different function)

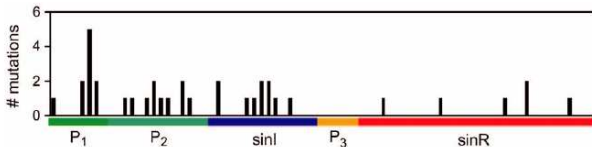
# Example: evolvability of *Bacillus subtilis*

- Sporulation: response of *Bacillus subtilis* to environment stress



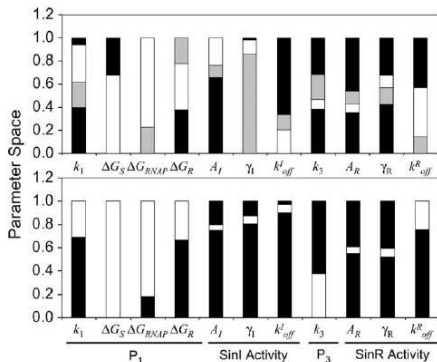
Sporulation switch (Voigt *et al.*, Genetics 2005)

- Distinct species evolved to colonize different hosts (e.g., insects, mammals)
- Diversification of sporulation response (condition and timing) according to environmental niche



Evolutionary divergence of *sin* operon (Voigt *et al.*, Genetics 2005)

# Example: evolvability of *Bacillus subtilis*

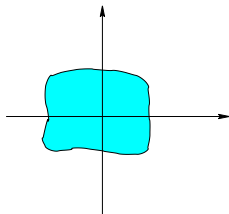


Robustness study by series of one-parameter bifurcations (Voigt *et al.*, Genetics 2005)

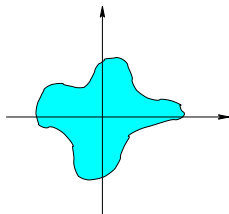
- $P_1$  promoter robust to change. Supports adaptation to tune thresholds of sporulation switch.
- Parameters associated with gene *SinR* fragile. Most conserved across species

# Robustness to simultaneous parameter variation

- One-parameter bifurcation diagrams do not account for simultaneous parameter variation
- Good predictor for robustness when:



- But not when:



# Robustness to simultaneous parameter variation

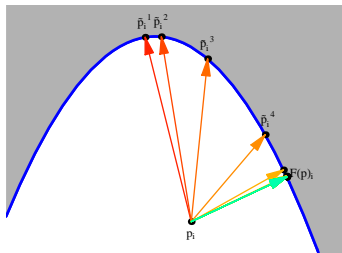
- Need to compute distance to bifurcation under **simultaneous** parameter variation
- Iterative method for computing (locally minimum) distance to bifurcation manifold (I. Dobson 1993, M. Mönnigmann and W. Marquardt 2002):
  - based on performing a series of one-param continuation
  - geometric convergence if conditions on magnitude of principal curvatures are satisfied

# Iteratively computing (locally) closest point

ALGORITHM: LOCMINDIST( $x, (p_i, p_s), v, \varepsilon, )$

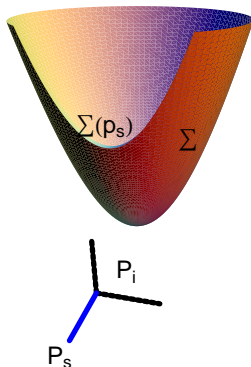
- Set initial parameter:  $p^0 \leftarrow p, x^0 \leftarrow x$
- FOR  $j = 1, \dots, j_{max}$ 
  - 1 From  $p$  and  $x$ , continue along parameter ray  $\{(p_i + rv, p_s): r \in \mathbb{R}_+\}$ , until bifurcation point  $p^b$  detected
  - 2 Compute normal vector at  $p^b$ :  $v \leftarrow N_i(p^b)$
  - 3 Update: parameter  $p^j \leftarrow p^b$ , ODE solution  $x^j \leftarrow x(p^b)$
  - 4 Terminate if  $\|p^j - p^{j-1}\| / \|p^0\| < \varepsilon$

END



# Mathematical notation

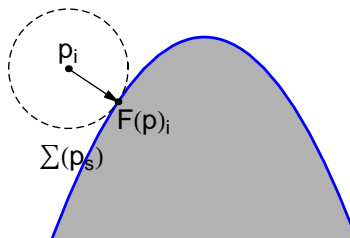
- Consider the splitting:  $p = (p_i, p_s) \in P_i \times P_s$  where
  - 1  $P_i$ : space of *control inputs*
  - 2  $P_s$ : space of *system parameters*
- $\Sigma$ : bifurcation manifold
- Define  $\Sigma(p_s) \equiv \Sigma \cap \{p_s\}$  the intersection of  $\Sigma$  with the  $p_s$ -plane:



# Forward operator

- Consider the orthogonal projection operator  $F : P \rightarrow P$ ,

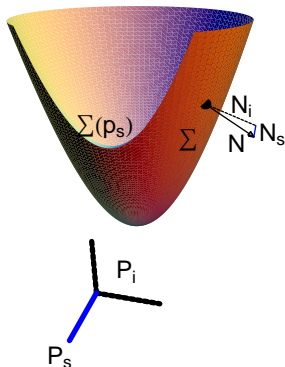
$$\begin{aligned} F(p) &\equiv (F(p)_i, F(p)_s) \\ &= (\mathbb{P}_{\Sigma(p_s)} p_i, p_s). \end{aligned}$$



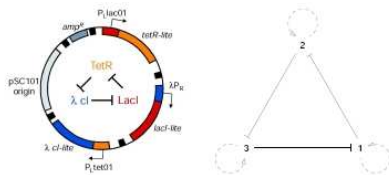
- Measure robustness using  $l_2$ -distance functional  
 $J(p) = \|F(p) - p\|_{l_2}$

# Adjoint analysis

- Consider linearized  $l_2$ -distance functional,  $\langle \cdot, l \rangle = J'(p)(\cdot)$
- Denote adjoint solution  $\psi \equiv F'^*(p)l$
- Adjoint solution may be geometrically interpreted as:  
$$\psi = \pm \frac{1}{\|N_i\|} (N_i, N_s).$$



# Applications: symmetric repressilator (Müller *et. al.*)



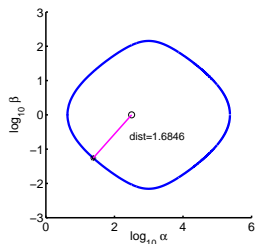
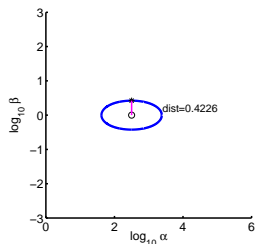
- ODE system for protein and RNA concentrations  $x_i$  and  $y_i$

$$\begin{aligned} \dot{x}_i &= \beta(y_i - x_i) \\ \dot{y}_i &= \alpha \left( \frac{1 - \delta}{1 + x_{i-1}^h \text{ mod } 3} + \delta \right) - y_i, \quad i = 0, \dots, 2. \end{aligned}$$

where  $\alpha$ : efficiency of gene transcription,  $\beta$ : ratio of degradation rates,  $\delta$ : leakiness of the gene transcription,  $h$ : Hill-coefficient.

- Inverse problem: determine  $\delta, h$  to maximize region of oscillations in the  $(\alpha, \beta)$ -plane

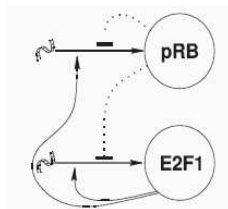
# Applications: symmetric repressilator (Müller *et. al.*)



$$(\delta, h) : (10^{-3}, 1.5) \rightarrow (10^{-4}, 2)$$

# Applications: cell cycle transition

- Model: 9 state, 40 parameter ODE system for mammalian cell phase transition (Swat *et. al.*, 2004)

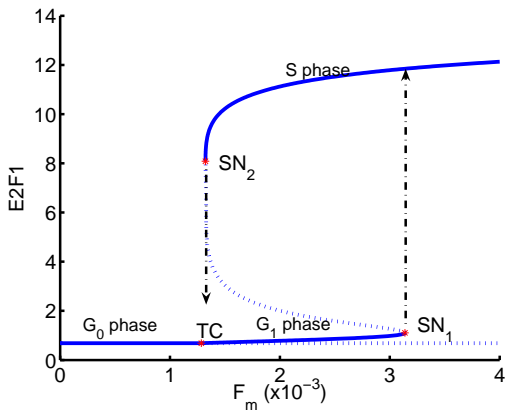


- Core double inhibitor-activator module:

$$\begin{aligned}\frac{d}{dt}[\text{pRB}] &= k_1 \frac{[\text{E2F1}]}{K_{m1} + [\text{E2F1}]} \frac{J_{11}}{J_{11} + [\text{pRB}]} - \phi_{\text{pRB}}[\text{pRB}], \\ \frac{d}{dt}[\text{E2F1}] &= k_p + k_2 \frac{a^2 + [\text{E2F1}]^2}{K_{m2}^2 + [\text{E2F1}]^2} \frac{J_{12}}{J_{12} + [\text{pRB}]} - \phi_{\text{E2F1}}[\text{E2F1}]\end{aligned}$$

# Applications: cell cycle transition

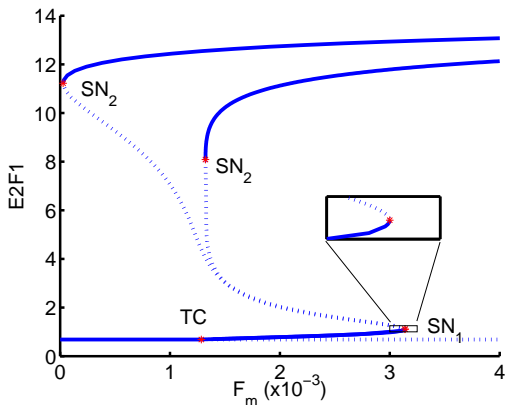
- Bistability wrt strength of mitogenic stimulation,  $F_m$



- Inverse problem: is it possible to design an irreversible switch while fixing  $G_1/S$  transition point?

# Applications: cell cycle transition

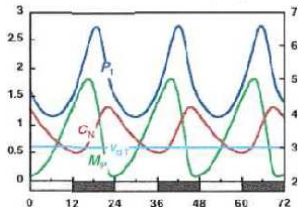
- Design parameters: feedback strength  $k_{25}$ ,  $\phi_{E2F1}$  and  $\phi_{AP-1}$



Bifurcation diagrams for initial and designed systems

# Application: circadian rhythm

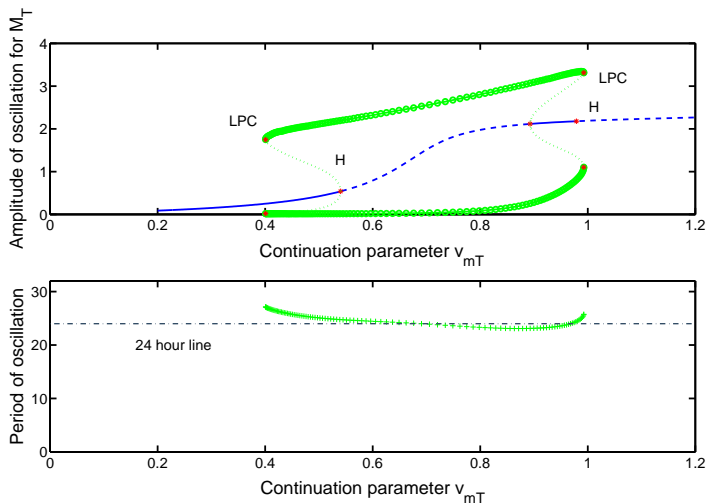
- Sustained, autonomous biochemical oscillations with  $\pm 24$  hour period



- Observed property: **period** of circadian rhythm **insensitive to temperature variations**
- Homeostasis problem: under environmental perturbation,
  - 1 **oscillatory behavior** is conserved
  - 2 **period of oscillation** remains constant



# Bifurcation diagram



Stable **limit cycles** within range of parameter  $v_{mT}$

# Numerical continuation of limit cycles

- Numerical continuation of limit cycles: reformulation as boundary value problem, discretization via orthogonal collocation
- Time transformation  $t \rightarrow t/T$ ; find on unit circle  $\mathbb{S}$ ,  $(x(t), T) \in C^1(\mathbb{S}) \times \mathbb{R}$  satisfying

$$\begin{aligned} \dot{x} - Tf(x, p) &= 0 \\ \langle x, \tilde{f} \rangle_{L_2(\mathbb{S})} &= 0 \end{aligned}$$

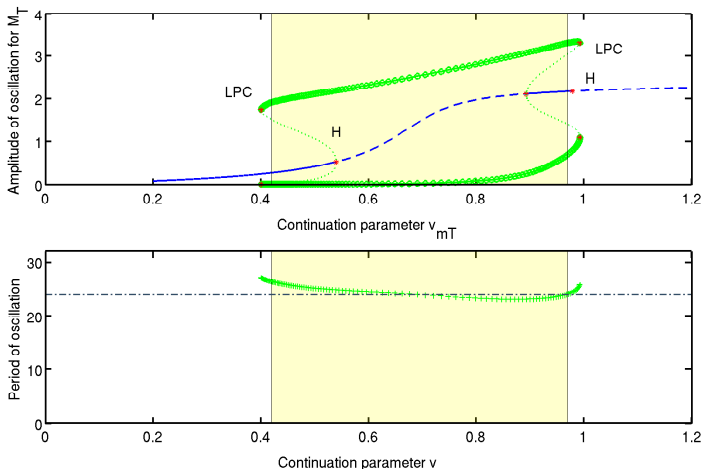
- At limit point for cycles (**LPC**), existence of null space for the adjoint operator:  $(v(t), S) \in C^1(\mathbb{S}) \times \mathbb{R}$ ,

$$\begin{aligned} \dot{v} + Tf_x(x, p)^t v + Sf(x, p) &= 0 \\ \langle v, f(x, p) \rangle_{L_2(\mathbb{S})} &= 0 \end{aligned}$$

# Inferring robustness of oscillation period

- Intrinsic period of circadian rhythm robust to temperature variations
- Inverse problem: what is the underlying homeostatic mechanism?
- First test problem: **identify parameters** resulting in constancy of period under  $\approx 2$ -fold change in rate parameter  $v_{mT}$

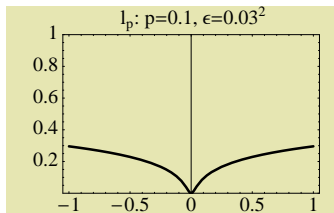
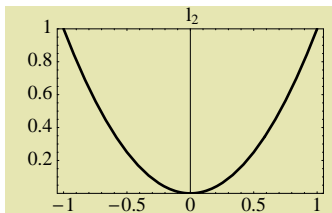
# Test case: constancy of period



Goal: eliminate variation of period within a **window of values** for  $v_{mT}$

# Objective function for identifying period constancy

- Right functional for measuring deviation of period: **total variation**?
- Large number (38) of parameters in the model
- Goal: identify **sparse** subset of parameters **impacting** the robustness of period
- Use  $l_p$ -functional,  $p \leq 1$ , for identifying sparse parameter set

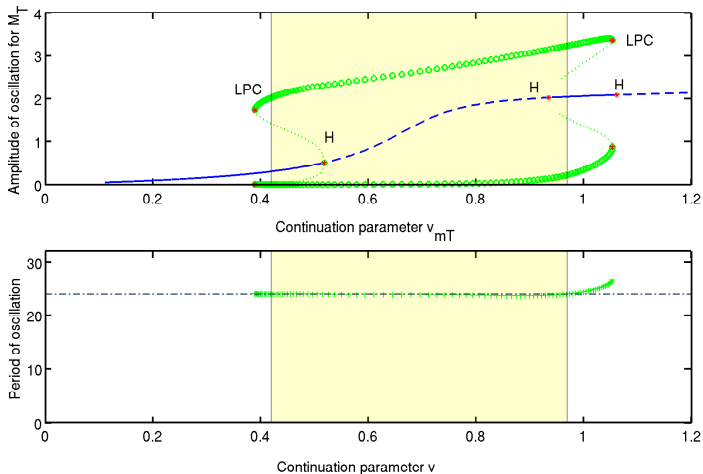


$$l_2(x) = \sum_i x_i^2, \quad \text{and} \quad l_p(x) = \sum_i (x_i^2 + \epsilon)^{p/2}$$

- Minimizing objective function:

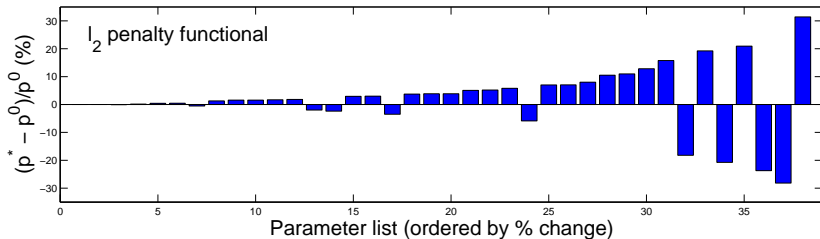
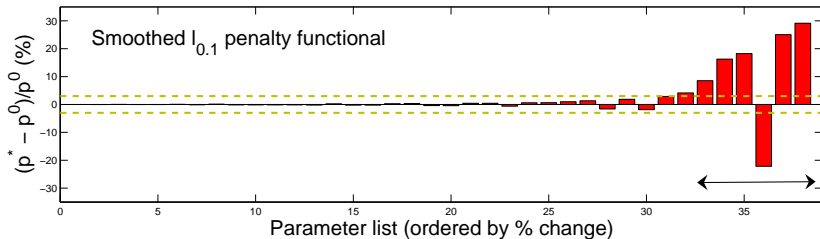
$$\begin{aligned} \min_p J(p) &= TV(\text{period} - 24hr) + \alpha \|p - p^0\|_{l_p} \\ \text{s.t.} \quad &LPC_{\text{left}} \leq p_{\text{left}}, LPC_{\text{right}} \geq p_{\text{right}} \end{aligned}$$

# Test case: result

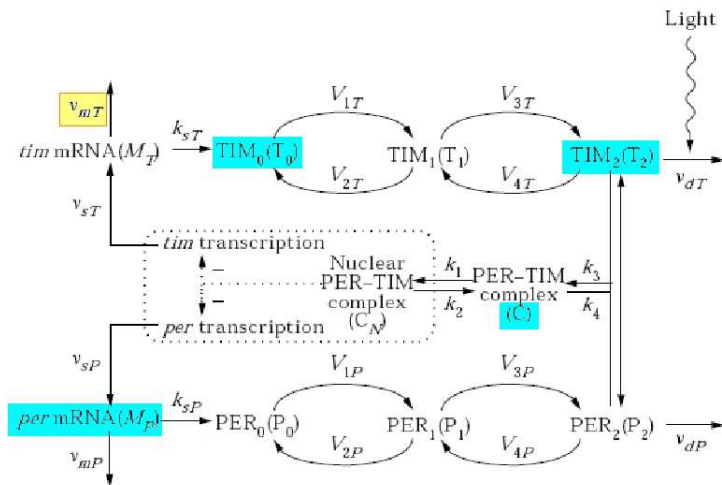


Bifurcation diagram of system for computed system

# Effectiveness of $l_p$ functional for sparsity



# Identification of parameters



Rate equations affected by inferred parameters

# Conclusions

- Inverse bifurcation analysis: methodology for studying how qualitative and quantitative behavior arises
- Distance to bifurcation quantifies
  - robustness to perturbation
  - evolvability of genes
- Sparsity useful for identifying crucial parameters in high-dimensional systems
- Future directions: infer system dynamics arising from choice of network topology