Inverse bifurcation analysis for reverse engineering of gene systems

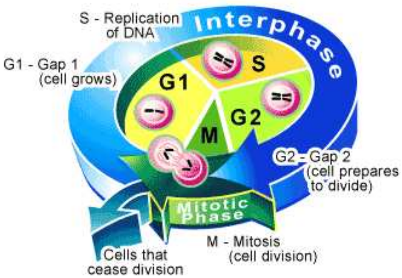
James Lu

Johann **R**adon **I**nstitute for **C**omputational and **A**pplied **M**athematics (RICAM) Linz, Austria

Joint work with Heinz W. Engl (Inverse Problems Group, RICAM)

Peter Schuster (Theoretical Biochemistry Group, University of Vienna)

Motivation



Schematic of phases of cell cycle

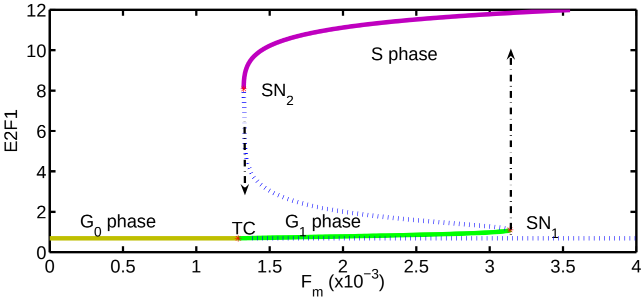
- Concentrations of regulatory species x typically modelled by parameterized ODE systems:

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

- Mathematical correspondence for changes in cycle phase: bifurcation of solution

Bifurcation analysis in molecular biology

- Bifurcation diagram for genetic "switch" regulating mammalian cell cycle



- Biologically important geometric characterization of bifurcation diagram:
  - x-location of SN_2: whether switch of state is permanent
  - length of "nose" for saddle-node: whether transition of state is delayed

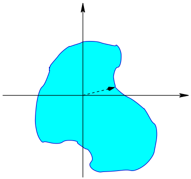
Inverse bifurcation analysis

- In forward problems one looks at predictions of given model
- In inverse problems one looks for causes of observed or desired effects
  - typically ill-posed: the solution may not *exist*, may not be *unique*, or does not *depend continuously* on data
  - $\Rightarrow$ mathematical theory of regularization

- Forward bifurcation analysis:
  - map parameters $\longrightarrow$ bifurcation diagrams
  - e.g.: *is the genetic switch irreversible for given parameter values?*
- Inverse bifurcation analysis:
  - map geometry of bifurcation diagrams $\longrightarrow$ parameters
  - e.g.: *what are the mechanisms for the nose of saddle-node to increase under given signals?*

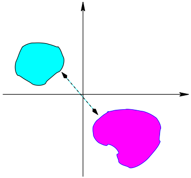
Geometric properties of bifurcation diagrams

- Geometric properties: location, size, and spatial relationship
- Parametric distance to bifurcation



quantify: robustness; ability of 'tuning' to environment

- Parametric distance between qualitative regions

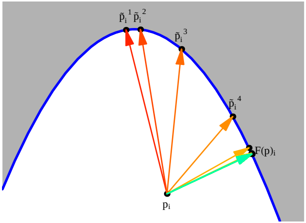


quantify: evolvability to perform different functions

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

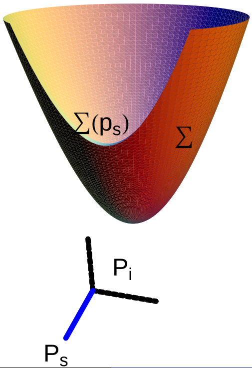


ALGORITHM: LOCMINDIST $\left(x,(p_i,p_s),v,\varepsilon,\right)$

- Set initial parameter: $p^0 \leftarrow p$, $x^0 \leftarrow x$
- FOR $j = 1,\cdots,j_{max}$
  1. From p and x, continue along parameter ray $\{(p_i + rv, p_s)\colon\, r \in \mathbb{R}_+\}$ 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 $\|p^j - p^{j-1}\|/\|p^0\| < \varepsilon$

Mathematical description

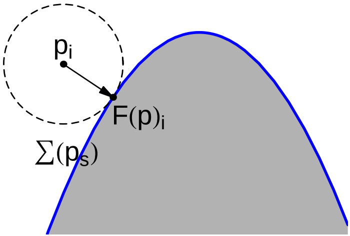
- Consider splitting of parameter space P: $p = (p_i, p_s) \in P_i \times P_s$,
  1. P_i: space of *control inputs*
  2. P_s: space of *system parameters*
- Σ: bifurcation manifold in total parameter space P
- Define $\Sigma(p_s) \equiv \Sigma \cap \{p_s\}$ the intersection of Σ with the p_s-plane:



Forward operator and functional

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

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



- Functional: l_2-distance $J(p) = \|F(p) - p\|_{l_2}$
- Adjoint analysis: allows for efficient implementation by gradient-based solution methods

Sensitivity analysis by adjoint method

- Denote linearization of forward operator and functional as $F'(p)$ and $\langle \cdot, l \rangle$
- Define *adjoint operator* $F'^*(p)$ (as in general Hilbert space setting) by:

$$\langle F'(p)\delta p, \delta \tilde{p} \rangle = \langle \delta p, F'^*(p)\delta \tilde{p} \rangle, \;\; \forall \delta \tilde{p}, \delta \tilde{p} \in \mathbb{R}^m.$$

- By defining the *adjoint solution* $\psi \equiv F'^*(p)l$, obtain equivalence relation:

$$\langle \delta p, \psi \rangle = \langle F'(p)\delta p, l \rangle$$

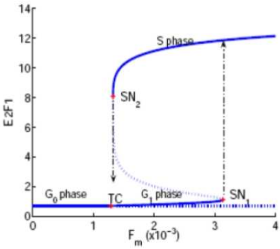
- For codim-1 bifurcations of equilibrium, components of adjoint solutions given by:

$$\begin{aligned} \text{Saddle node} &: \;\; \mathbf{w}^T \mathbf{f_p} \\ \text{Hopf} &: \;\; \mathrm{Re}\{\mathbf{w}^T \mathbf{f_{xp}} \mathbf{v} - (\mathbf{w}^T \mathbf{f_{xx}} \mathbf{v}) \mathbf{f_x}^{-1} \mathbf{f_p}\} \end{aligned}$$

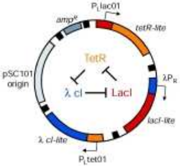
Software implementation for inverse bifurcation

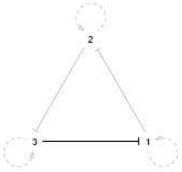
- Packages involved
  - MathSBML (B. E. Shapiro): reads biological models in **S**ystems **B**iology **M**ark-up **L**anguage (**SBML**) format into Mathematica
  - Mathematica Symbolic Toolbox for Matlab (B. E. Barrowes): allows communication between Matlab and Mathematica
  - CL_MATCONT (A. Dhooge, W. Govaerts, Y. Kuznetsov, B. Sautois, etc.): performs bifurcation analysis
  - Matlab Optimization Toolbox (MathWorks): solves inverse bifurcation using Sequential Quadratic Programming (SQP)


```
<model name = "oscillating_MAPK">
  <MAPK.xmllistOfCompartments>
    <compartment name = "uVol" volume = "1"/>
  </listOfCompartments>
  <listOfSpecies>
    <specie name = "MKKK" boundaryCondition = "false" initialAmount = "90"
    <specie name = "MKKK_P" boundaryCondition = "false" initialAmount = "10"
  </listOfSpecies>
  <listOfParameters>
      <parameter name = "V1" value = "2.5"/>
      <parameter name = "Ki" value = "9"/>
      <parameter name = "K1" value = "10"/>
   </listOfParameters>
   <listOfReactions>
    <reaction name = "J0" reversible = "false">
      <listOfReactants>
        <specieReference specie = "MKKK" stoichiometry = "1"/>
      </listOfReactants>
      <listOfProducts>
        <specieReference specie = "MKKK_P" stoichiometry = "1"/>
      </listOfProducts>
      <kineticLaw formula = "V1*MKKK/((1+(MAPK_PP/Ki)^n)*(K1+MKKK))">
        <listOfParameters>
          <parameter name = "n" value = "1"/>
        </listOfParameters>
      </kineticLaw>
```

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





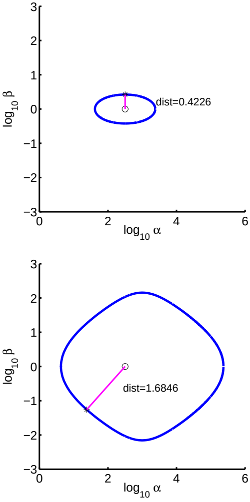
- ODE system for 3 protein and RNA concentrations x_i and y_i

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

  where α: efficiency of gene transcription, β: ratio of degradation rates, δ: leakiness of the gene transcription, h: Hill-coefficient.

- Inverse problem: determine δ, h to maximize region of oscillations in the (α, β)-plane

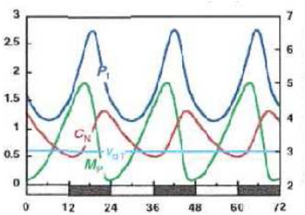
Applications: symmetric repressilator (Müller *et. al.*)



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

Application: circadian rhythm

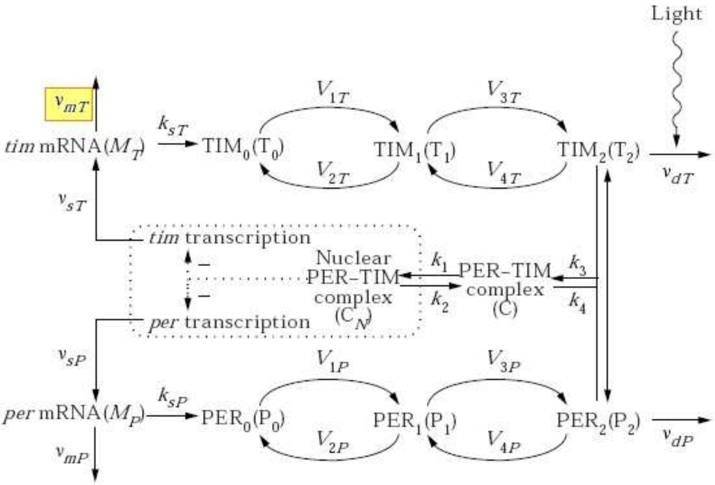
- Sustained, autonomous biochemical oscillations with ≈ 24 hour period



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

Test case: Leloup/Goldbeter model (1999)

- Model: 10 variables, 38 parameters



Limit cycle continuation in CL_MATCONT

- Numerical continuation of limit cycles: reformulation as boundary value problem
- Time transformation $t \to t/T$. 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: $(\phi(t), \sigma) \in C^1(\mathbb{S}) \times \mathbb{R}$,

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

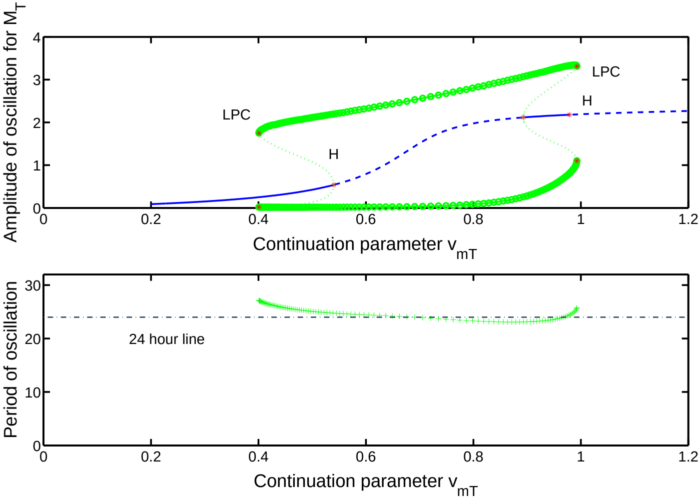
- Sensitivity of distance to LPC is (up to scalar factor):

$$\langle \phi, Tf_p(x,p) \rangle_{L_2(\mathbb{S})}$$

Numerical analysis

- Adjoint analysis can be carried out in either discrete or continuous settings
  - convergence of discrete adjoint to continuous counterpart
- Collocation methods for optimal control:
  - **J. A. Pietz, M. Heinkenschloss** (2003, Rice Univ. thesis): convergence of discrete adjoint approximation in optimal control of ODEs using Legendre pseudospectral collocation
  - **S. Kameswaran, L. T. Biegler** (2004): convergence of optimal control using collocation at Radau points
- CL_MATCONT: Lagrange interpolation at evenly-spaced points, collocation at Legendre points
- If integral constraint $\langle \cdot, f \rangle_{L_2(\mathbb{S})} = 0$ computed using Gauss-Legendre quadrature, discrete adjoint coefficients divided by G-L weights provide approximation to continuous adjoint approximation at Legendre points
- Futher work: prove convergence rates, numerical tests

Bifurcation diagram

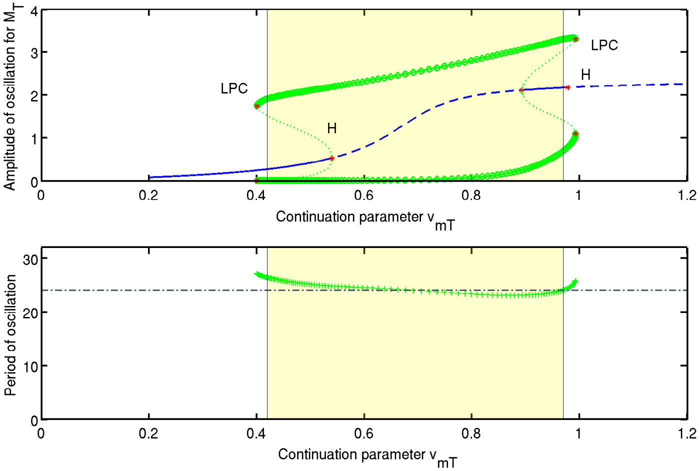


Stable limit cycles within range of parameter v_{mT}

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 ≈ 2-fold change in rate parameter v_{mT}

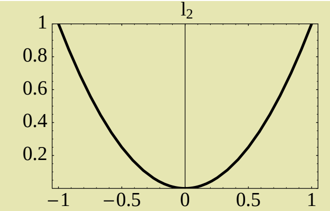
Test case: constancy of period

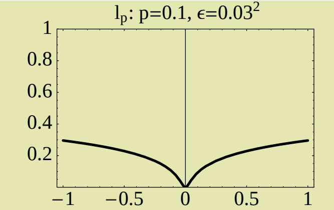


Goal: eliminate variation of period within a window of values for v_{mT}

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



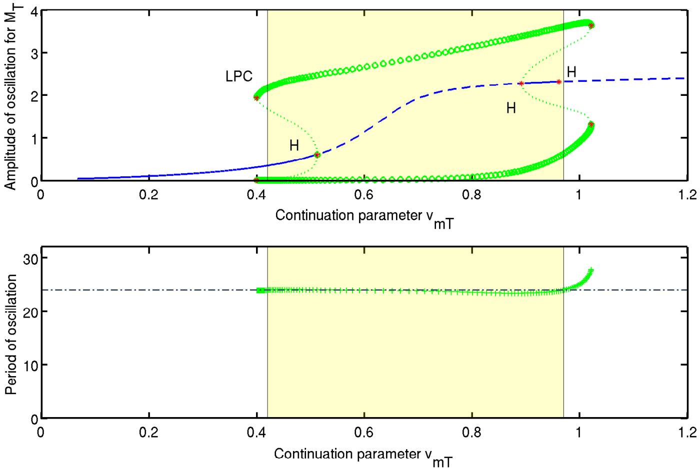


$l_2(x) = \sum_i x_i^2$, and $l_p(x) = \sum_i (x_i^2 + \varepsilon)^{p/2}$

- Minimizing objective function:

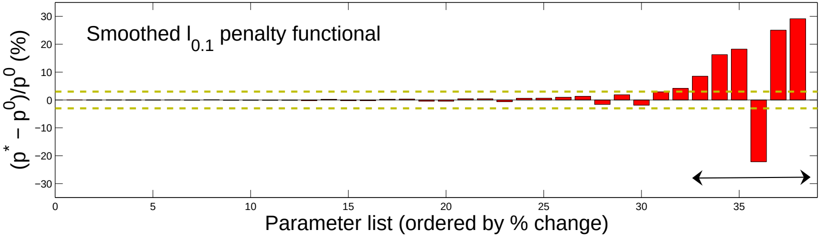
$$\min_p J(p) = TV(\text{period} - 24hr) + \alpha \|p - p^0\|_{l_p}$$
$$s.t. \quad LPC_{\text{left}} \leq p_{\text{left}}, LPC_{\text{right}} \geq p_{\text{right}}$$

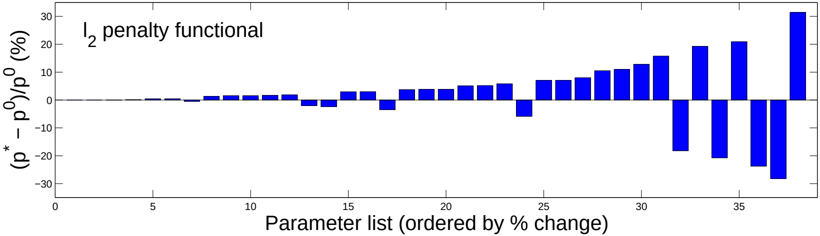
Test case: result



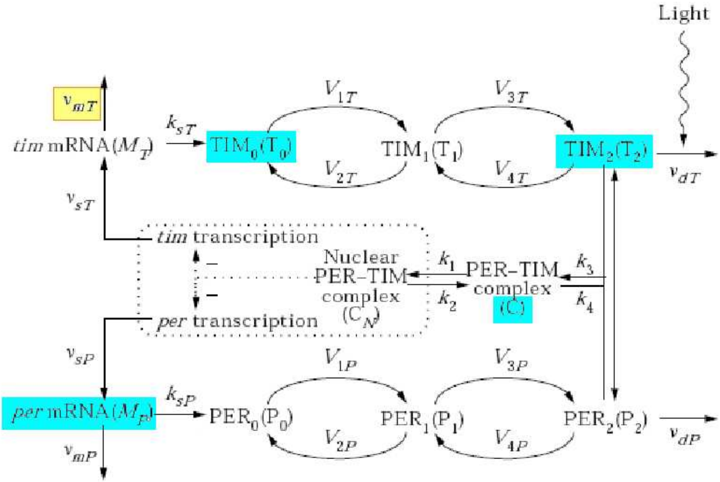
Bifurcation diagram of system for computed system

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
- Sparsity useful for identifying crucial parameters in high-dimensional systems
- Future directions:
  - upscaling: regularization methods, summarizing bifurcation diagrams in high dimensions
  - infer aspects of network topology from system dynamics