

# Inverse Bifurcation Analysis for Reverse Engineering of Cell Cycle and Circadian Rhythm Models

James Lu<sup>1</sup>

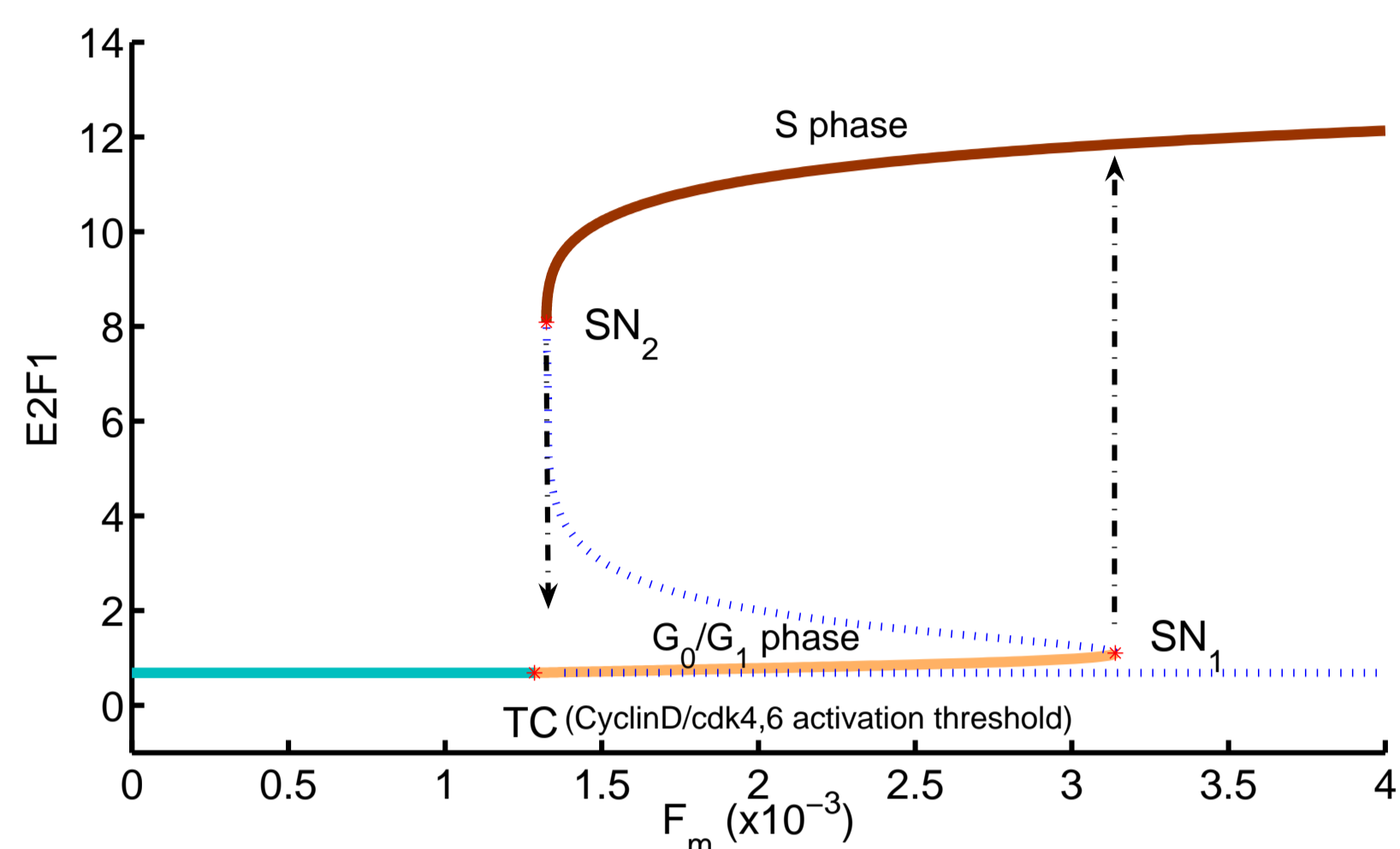
<sup>1</sup> Johann Radon Institute for Computational and Applied Mathematics (RICAM)

Collaborators: Prof. Heinz W. Engl, Prof. Peter Schuster



## Forward Bifurcation

Consider models of cellular processes by systems of first-order ordinary differential equation: denoting  $x$  and  $\alpha$  as the biochemical concentrations and parameters, respectively, the time evolution of  $x$  is given by  $\dot{x} = f(x, \alpha)$ . Bifurcation analysis maps **parameter values**  $\rightarrow$  **bifurcation diagram**. Via the bifurcation diagram, the parameter space is partitioned into "qualitatively different" regions.



Bifurcation for mammalian  $G_1/S$  module [8]

## Inverse Bifurcation

**Inverse Problems** is a field of applied mathematics where one looks for causes of **observed** or **desired** effects [2]. Mathematically, inverse problems may be formulated in terms of operator equations of the form: determine  $q \in Q$  from  $y \in Y$ , such that

$$F(q) = y,$$

where  $F : Q \rightarrow Y$  is a (in general, nonlinear) parameter-to-output map acting between Hilbert spaces  $Q$  and  $Y$ . Such problems are typically ill-posed and hence regularization techniques have to be developed [2]. The aim of inverse bifurcation analysis is to map the geometry of **bifurcation diagrams**  $\rightarrow$  **biochemical parameters** (see [1, 5, 6] for applications in biology; refer to [3, 7] for various applications of bifurcation control in engineering). Such problems may be broadly divided into two classes:

- **design** type, e.g. finding parameter configurations so that the distance to bifurcation is as large as possible
- **identification** type, e.g. finding parameter configurations such that the bifurcation diagram exhibits desired behavior.

To **reverse-engineer** gene regulatory systems, we focus on the latter type.

### Sparsity-promoting regularization

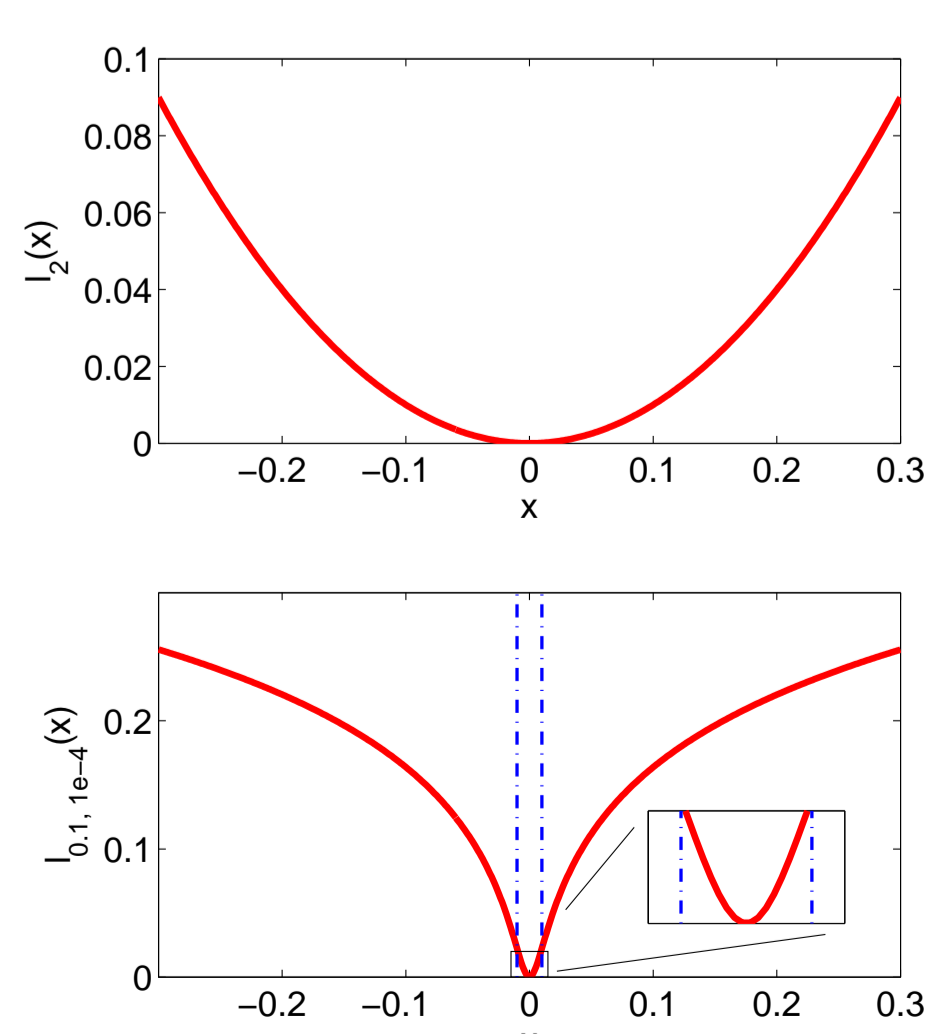
Mathematical techniques, called **regularization methods**, have to be used to cope with the ill-posedness of inverse problems [2]. While stabilizing ill-posed problems, regularization methods typically bias the solution. Depending on the mathematical properties of the problem and the application of interest, different regularization techniques may be appropriate. Does one want to obtain a solution of the minimum Euclidean norm? Or is it more desirable to obtain a solution that is **sparse**, i.e., has as few non-zeros as possible? For biological applications, sparsity is useful when one wishes to identify a **small number of parameters** whose variation can be mapped to a **wide range** of system behavior.

To obtain sparse solutions via gradient-based methods, we consider **differentiable penalty functions** the form

$$l_{p,\epsilon}(x) = \sum_i (x_i^2 + \epsilon)^{p/2}, \quad 1 \geq p > 0, \epsilon > 0,$$

where  $\epsilon$  is used to remove non-differentiability at  $x_i = 0$ . Note that for  $p < 1$ , the above functional is **non-convex**.

Comparison of  $l_2$  and  $l_{p,\epsilon}$  functionals



## Hierarchical identification strategy

In general there are **multiple distinct solutions** to inverse bifurcation problems; hierarchical strategy identifies a **sequence** of parameters, of possibly **increasing cardinality**.

Alg: HIER-PARAM-IDENT( $\alpha_s^0$ , MaxLev,  $p, \epsilon$ )

- Initialize:  $s \leftarrow \{1, \dots, m\}$ ,  $I_{\text{identified}} \leftarrow \emptyset$
- FOR  $j = 1, \dots, \text{MaxLev}$ 
  - $I_{\text{rem}} \leftarrow s \setminus I_{\text{identified}}$
  - Solve  $\alpha_{I_{\text{rem}}}^j \leftarrow \text{ConMin}(\alpha_{I_{\text{rem}}}^0, p, \epsilon)$
  - $I_j \leftarrow \{i : |(\alpha_{I_{\text{rem}}}^j)_i| > \sqrt{\epsilon}\}$
  - $I_{\text{identified}} \leftarrow I_{\text{identified}} \cup I_j$
- Return  $\{\alpha_{I_1}^1, \alpha_{I_2}^2, \alpha_{I_3}^3, \dots\}$

## Mammalian $G_1/S$ module

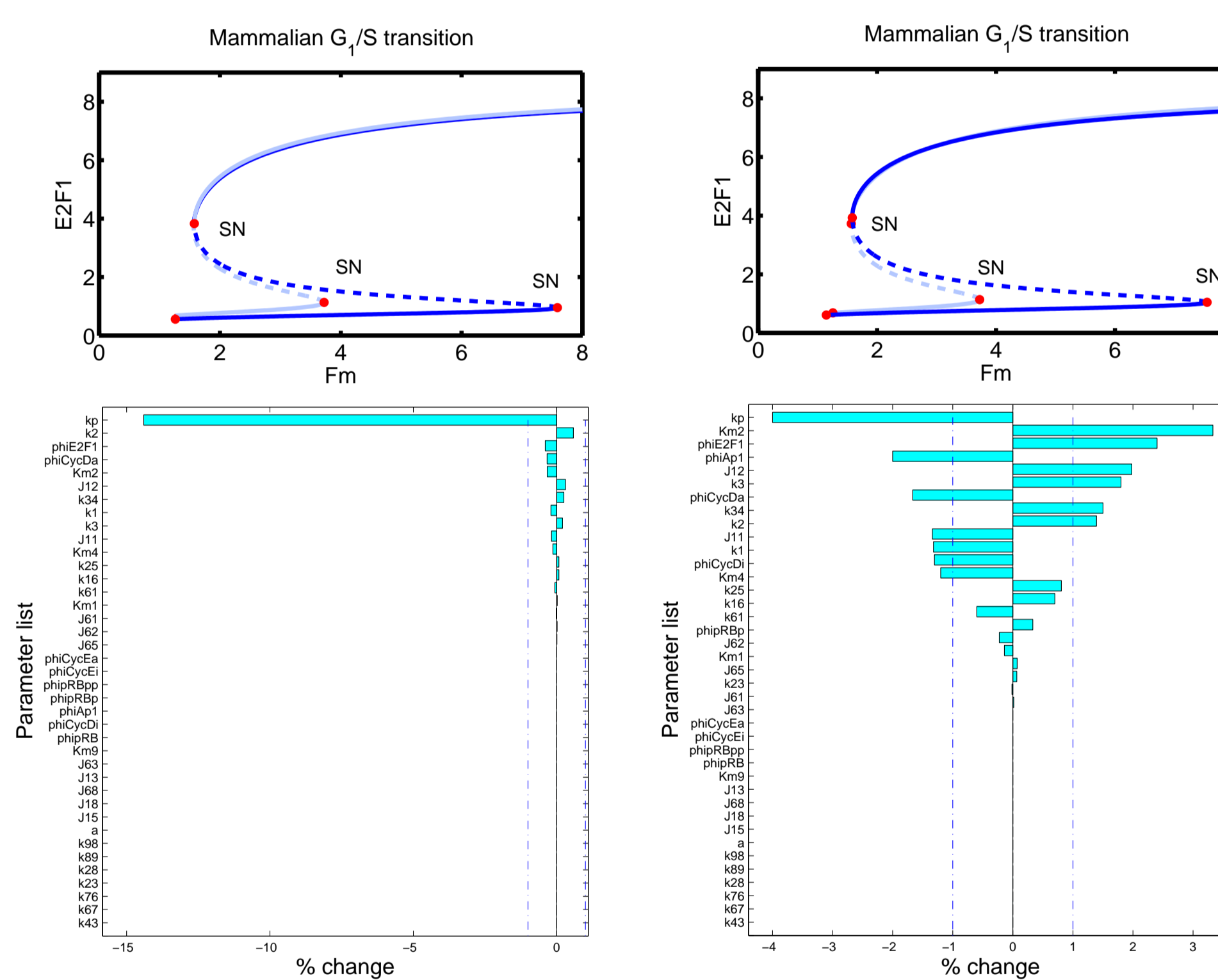
The model of the mammalian  $G_1/S$  transition by Swat *et al.* [8] consists of 9 chemical species, 25 reactions and 40 parameters, representing the transcription factor families AP-1, E2F, pRB, and cyclin/cyclin-dependent kinase complexes cyclin D/Cdk4,6 and cyclin E/Cdk2. To carry out inverse analysis, we consider constrained optimization problems of the following form:

$$\text{ConMin}(\alpha_s, p, \epsilon) : \min_{\alpha_s} l_{p,\epsilon} \left( \frac{\alpha_s - \alpha_s^*}{\alpha_s^*} \right)$$

subject to  $\text{SN}_1(\alpha_s) = \text{SN}_1^*$   
 $\text{SN}_2(\alpha_s) = \text{SN}_2^*$ ,

where the following **3 modes of geometric variation** are mapped to the parameter space:

- Elongating saddle-node nose:  $\text{SN}_1^* \leftarrow \text{SN}_1 + d$
- Moving saddle-nodes right:  $\text{SN}_1^* \leftarrow \text{SN}_1 + d$ ,  $\text{SN}_2^* \leftarrow \text{SN}_2 + d$
- Decreasing bistability:  $\text{SN}_2^* \leftarrow \text{SN}_2 + d$



Sparsity vs.  $l_2$  regularization

The **hierarchical, sparsity-promoting** regularization strategy identifies the following parameters:

Case \ Param.	Level $j = 1$	Level $j = 2$	Level $j = 3$
Elongating $\text{SN}_1$ nose	$k_p \downarrow 14.3\%$	$k_{34} \uparrow 31.7\%$ $K_{m2} \uparrow 6.4\%$	$\phi_{\text{AP-1}} \downarrow 20.9\%$ $\phi_{\text{E2F1}} \uparrow 7.3\%$
Moving $\text{SN}_{1,2}$ right	$K_{m4} \uparrow 269.3\%$	$J_{11} \uparrow 191.7\%$ $k_p \uparrow 17.3\%$	$k_2 \downarrow 39.9\%$ $\phi_{\text{E2F1}} \downarrow 11.7\%$ $K_{m2} \downarrow 10.3\%$
Decreasing bistability	$J_{11} \uparrow 128.5\%$ $k_p \uparrow 33.8\%$	$k_1 \uparrow 169.1\%$ $K_{m2} \downarrow 21.7\%$ $J_{12} \downarrow 20.1\%$	$k_2 \downarrow 43.7\%$ $\phi_{\text{E2F1}} \downarrow 28.3\%$

These correspond to the following species interactions:

$$\frac{d}{dt}[\text{pRB}] = k_1 \frac{[\text{E2F1}]}{K_{m1} + [\text{E2F1}]} J_{11} + [\text{pRB}] J_{61} + [\text{pRB}_p] - k_{16}[\text{pRB}][\text{CycD}_a] + k_{61}[\text{pRB}_p] - \phi_{\text{pRB}}[\text{pRB}],$$

$$\frac{d}{dt}[\text{E2F1}] = k_p + k_2 \frac{a^2 + [\text{E2F1}]^2}{K_{m2}^2 + [\text{E2F1}]^2} J_{12} + [\text{pRB}] J_{62} + [\text{pRB}_p] - \phi_{\text{E2F1}}[\text{E2F1}]$$

$$\frac{d}{dt}[\text{CycD}_i] = -k_{34}[\text{CycD}_i] \frac{[\text{CycD}_a]}{K_{m4} + [\text{CycD}_a]} + \dots$$

$$\frac{d}{dt}[\text{CycD}_a] = k_{34}[\text{CycD}_i] \frac{[\text{CycD}_a]}{K_{m4} + [\text{CycD}_a]} + \dots$$

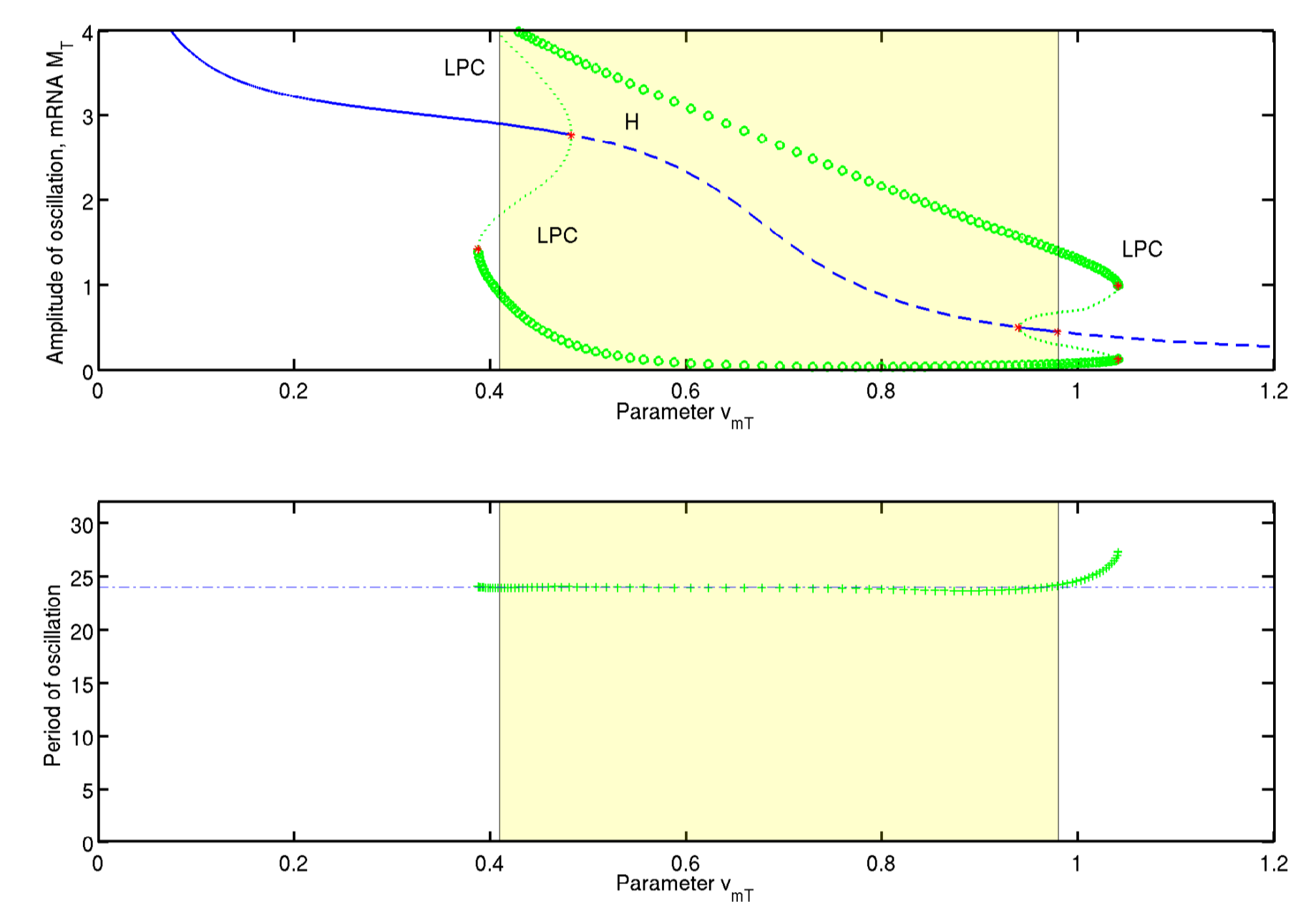
## Circadian Rhythm

In the study of circadian rhythms an important biological question is how the **constancy of period** arises, with respect to variations in biochemical parameters and environmental factors (e.g., **temperature compensation**). The forward analysis entails the continuation and bifurcation of **limit cycles**, which may be formulated as boundary value problems on the unit circle. **Adjoint** method for sensitivity analysis has been carried out to enable inverse analysis for the location of **limit point for cycles** (LPC) and **total variation** (TV) of period:

$$\min_{\alpha} \text{TV}(\text{period} - 24\text{hr}) + \mu \|\alpha - \alpha^*\|_{l_{p,\epsilon}}$$

s.t.  $\text{LPC}_{\text{left}} \leq \alpha_{\text{left}}, \text{LPC}_{\text{right}} \geq \alpha_{\text{right}}$

Together with sparsity regularization, this may be used to deduce **influential parameters** and infer **biological design principles** underlying gene regulatory systems. The following shows the result of inverse analysis for a Leloup and Goldbeter model for circadian rhythm in *Drosophila* [4].



Analysis for constancy of oscillation period

## Acknowledgements

We gratefully acknowledge Lukas Endler, Philipp Kügler, Rainer Machné, Stefan Müller, Maciej Swat for helpful discussions. This research is supported by the WWTF, project MA 05.

## References

- [1] E. CONRAD *Bifurcation analysis and qualitative optimization of models in molecular cell biology with applications to the circadian clock*, PhD dissertation, Virginia Polytechnic Institute and State University, USA, 2006.
- [2] H.W. ENGL, M. HANKE, A. NEUBAUER *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.
- [3] I. DOBSON, *Computing a closest bifurcation instability in multidimensional parameter space*, J. Nonlinear Sci., 3 (1993), pp. 307–327.
- [4] J.-C. LELOUP, A. GOLDBETER, *Chaos and birhythmicity in a model for circadian oscillations of the PER and TIM proteins in Drosophila*, J. Theoret. Biol., 198 (1999), pp. 445–459.
- [5] J. LU, H. W. ENGL, R. MACHNÉ, AND P. SCHUSTER, *Inverse bifurcation analysis of a model for the mammalian  $G_1/S$  regulatory module*, submitted.
- [6] J. LU, H. W. ENGL, AND P. SCHUSTER, *Inverse bifurcation analysis: application to simple gene systems*, Algor. Mole. Biol., 1/11 (2006).
- [7] M. MÖNNIGMANN, AND W. MARQUARDT, *Normal vectors on manifolds of critical points for parametric robustness of equilibrium solutions of ODE systems*, J. Nonlinear Sci., 12 (2002), pp. 85–112.
- [8] M. SWAT, A. KEL, H. HERZEL, *Bifurcation analysis of the regulatory modules of the mammalian  $G_1/S$  transition*, Bioinformatics, 20 (2004), pp. 1506–1511.