Systems biology:

-System identification and parameter identifiability

-Modelling of metabolic systems

People:

FCC: Mats Jirstrand, H. Schmidth, G. Cedersund …

CMR: Martin Adiels (J. Borén)

Mathematical Sciences: Bernt Wennberg, Milena Anguelova,
Martin Berglund, Olle Nerman, Peter Gennemark, Sven Nelander….

Chalmers: Carl Johan Franzen, people at Physics, ComputerSciences,
Jens Nielsen

GU: Anders Blomberg, Stefan Hohmann, Per Sunnerhagen …

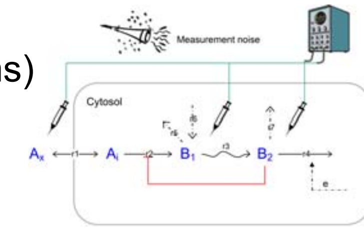
With all the names and organisations mentioned on the title page, I wanted to say that the Systems biology activities within the GMMC is closely connected with many other Systems biology activities at Chalmers and Göteborg university.

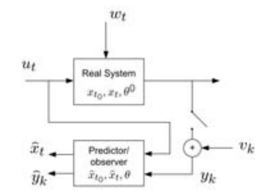
Within the GMMC, Mats Jirstrand and I are working on very similar problems. The first four slides of the presentation are related to Mats's activities.

GMMC Related Systems Biology at FCC – An Overview

Projects

- *A Joint Systems Biology Research Platform for FCC and MV*
  (system identification methods for biological and biochemical systems)
  - Parameter estimation using prediction error minimization
  - Parameter estimation in nonlinear mixed effects models described by stochastic differential equations
  - Modeling the RAS/cAMP/PKA pathway in *S. cerevisiae*
- Model Reduction Based on Time Scale Separation and State Aggregation
- The Systems Biology Toolbox for Matlab





Publications

- M. Jirstrand, ‘Parameter Estimation in Biochemical Reaction Networks – Observer Based Prediction Error Minimization’, 8th International Conference on Systems Biology, Long Beach, USA, 2007.
- M. Sunnåker and M. Jirstrand, ‘A Nonlinear Mixed Effects Modeling Toolbox for Matlab’, 3rd BioSim conference, Potsdam, 2007.
- M. Sunnåker, G. Cedersund, and H. Schmidt, ‘A New Method for Model Reduction Based on Time Scale Separation and Lumping’, 2nd Workshop on Mathematical Aspects of Systems Biology Göteborg, March 21-24, 2007.
- G. Cedersund, P. Strålfors and M. Jirstrand, 'Core-box modeling for biosimulation of drug action', In Eds. Bertau et al. Biosimulation in Drug Development. Wiley-VCH, Weinheim, 2007.
- M. Benson, L.-O. Cardell, S. Hohmann, M. Jirstrand, M. Langston, R. Mobini, and O. Nerman, 'Systems Biology May Radically Change Healthcare - Individually Based Prediction, Prevention, and Treatment', Läkartidningen, 104(42): 3037-3041, 2007.
- H. Schmidt, G. Drews, J. Vera, Wolkenhauer, ‘SBML Export Interface for the Systems Biology Toolbox for MATLAB’, Bioinformatics, 23(10), 1297-1298.
- H. Schmidt, 'SBaddon: High Performance Simulation for the Systems Biology Toolbox for MATLAB', Bioinformatics, Bioinformatics, 23(5), 646-647.

Sample Project:

The RAS/cAMP/PKA Pathway: A Systems Biology Approach

Background

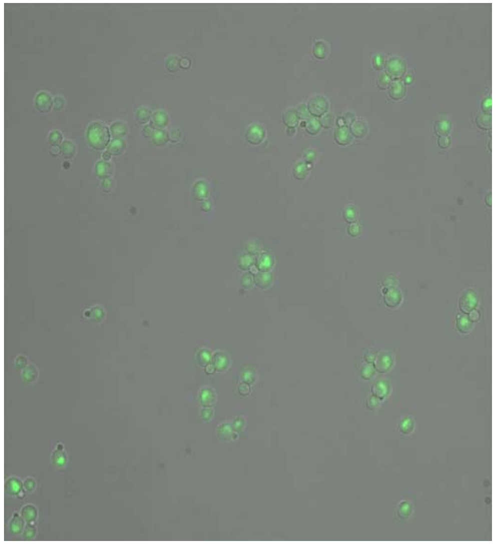
In the budding yeast *S. cerevisiae*, the Ras/cAMP/PKA signal transduction pathway regulates many cellular and physiological processes, such as growth, resting state and sporulation, carbohydrate and nitrogen metabolism, stress tolerance, and cell wall resistance to lyticase digestion, making the pathway central to cell life.

Problem

- The aim of the project is to integrate quantitative microscopy data into systems biology models that provide a deeper understanding of the processes directly involved in, but also linked to, the pathway.

- An interesting down-stream target of the RAS/cAMP/PKA pathway are the transcription factors Msn2p and Msn4p, which during low cAMP levels enters the nucleus and activates specific gene promoter targets. The localisation of Msn2/4p, which can be time-resolved visually, thus becomes good reporters for the activity in the pathway.

- The project brings together a unique mix of competences (applied physics, cell and molecular biology, mathematical statistics, image analysis and systems biology) in an area that is central to contemporary biotechnology.



Microscopy imaging of a cell culture.

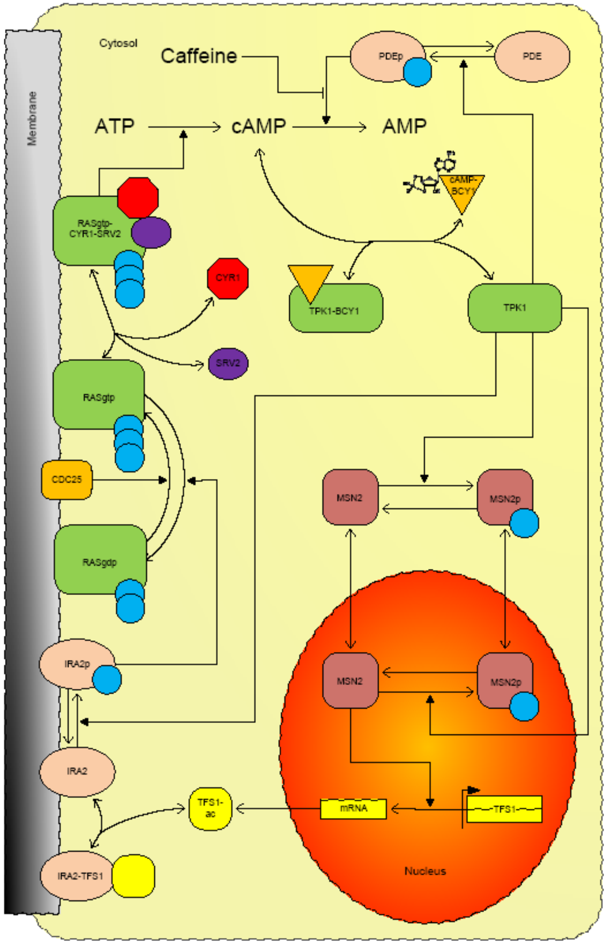
External Partners
Mikael Käll, Chalmers tekniska högskola,
Anders Blomberg, Göteborgs universitet

The RAS/cAMP/PKA Pathway: A Systems Biology Approach

Methods

- Dynamic models will be built from quantitative *in vivo* data in order to encapsulate obtained knowledge about causal relationships between the involved molecular species on the mechanistic level.

- Nonlinear differential equations describing the interactions between components can be formulated and methods from system identification can be used to utilize the available data to estimate unknown parameters and to validate different hypothesis.



PathwayLab model structure for the RAS/cAMP/PKA pathway

Fraunhofer
ITWM
CHALMERS
Research Centre
Industrial Mathematics

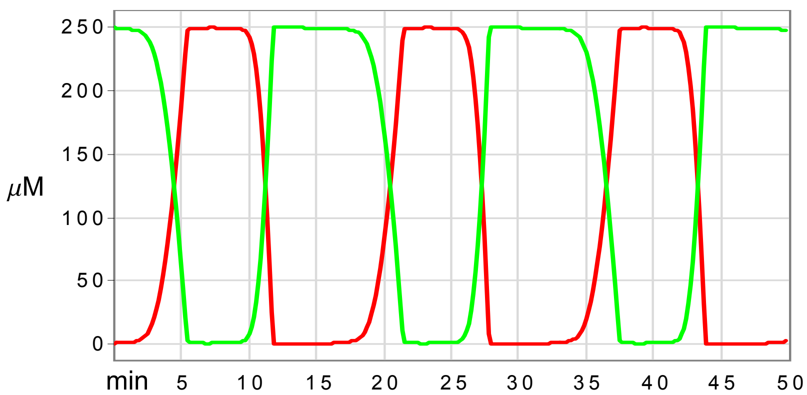
Sample Project:

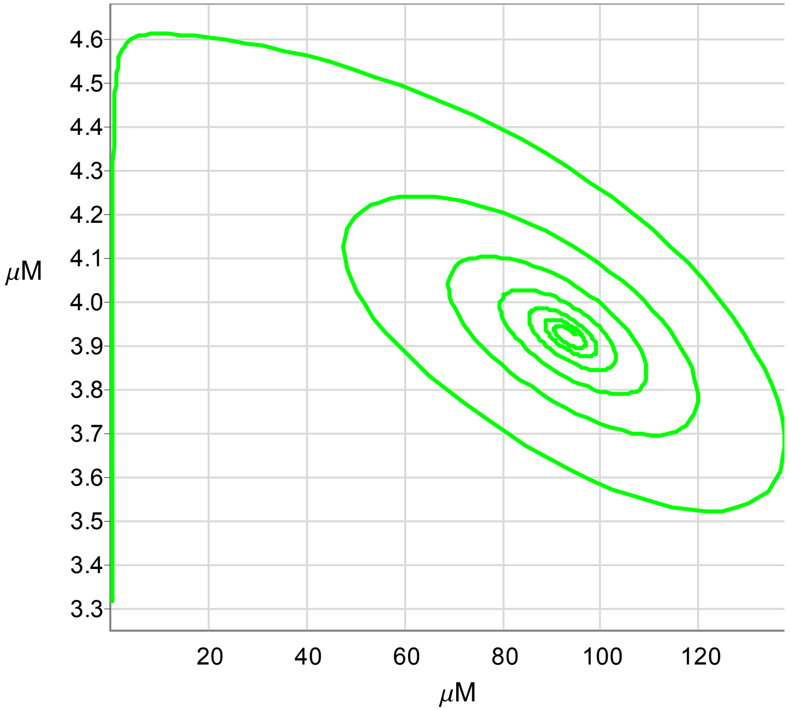
The RAS/cAMP/PKA Pathway: A Systems Biology Approach

Status/Results

- Based on different types of negative feedback regulation, the RAS/cAMP/PKA pathway has a possible role as an oscillator, responsible for the observed periodic nucleocytoplasmic shuttling of the transcription factors Msn2p and Msn4p.

- By inhibiting the phosphodiesterases involved in the degradation of cAMP, caffeine has the potential of modulating the characteristics of such oscillations.





Simulations of a model for the RAS/cAMP/PKA pathway. Sustained oscillations of the two RAS forms GDP and GTP (left) and a phase plane plot of damped oscillations in RAS_{GDP} and cAMP (right).

Mats Jirstrand

ITWM
Fraunhofer
CHALMERS
Research Centre
Industrial Mathematics

The CMR is another SSF-funded research center, of which all principal investigators except me are medical scientists.

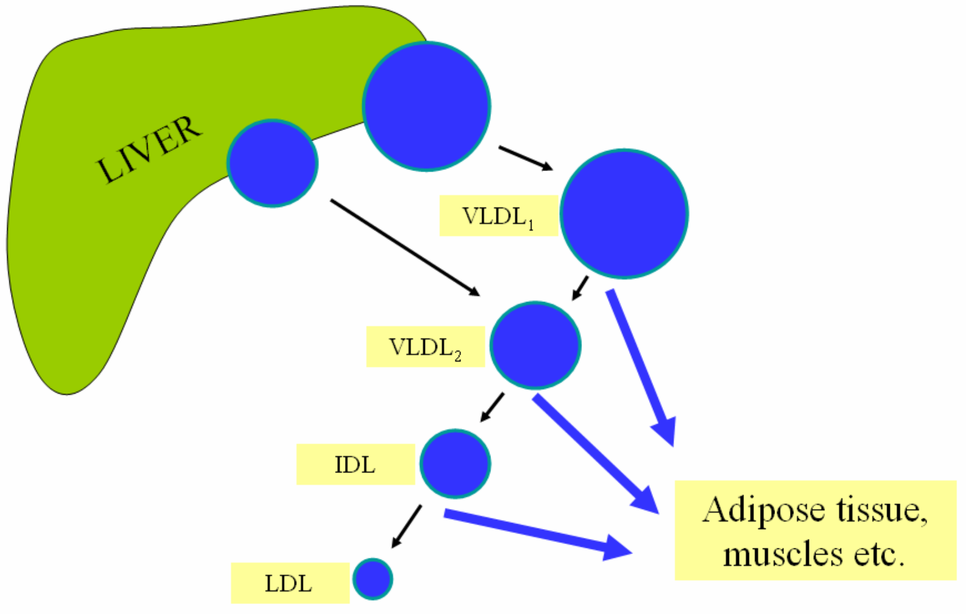
There are at least two connections between the GMMC and the CMR:

-my activities related to compartmental modelling of the apolipoprotein metabolism, which is very briefly described in the following three slides (the medical PI is Jan Borén, but the real main collaborator is Martin Adiels who did his PhD at the department of Mathematics). A new PhD-student, Martin Berglund will continue on that project.

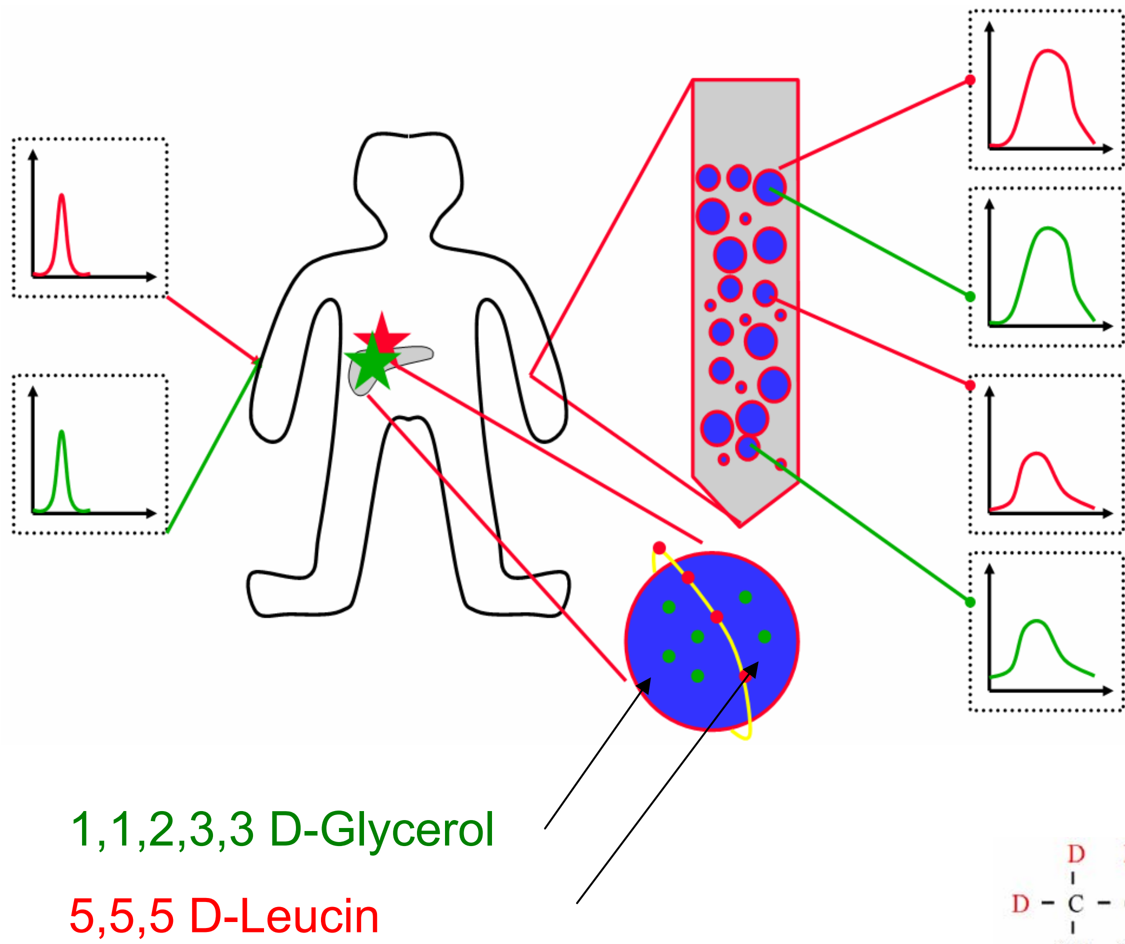
-a collaboration between Lena Carlsson at the CMR and Mats Rudemo on gene expression data.

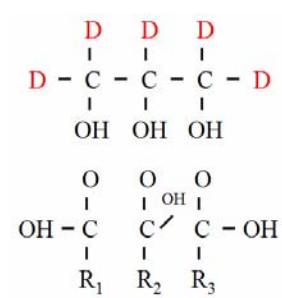
Modelling of the lipo protein metabolism

CMR Sahlgrenska Center for Cardiovascular and Metabolic Research



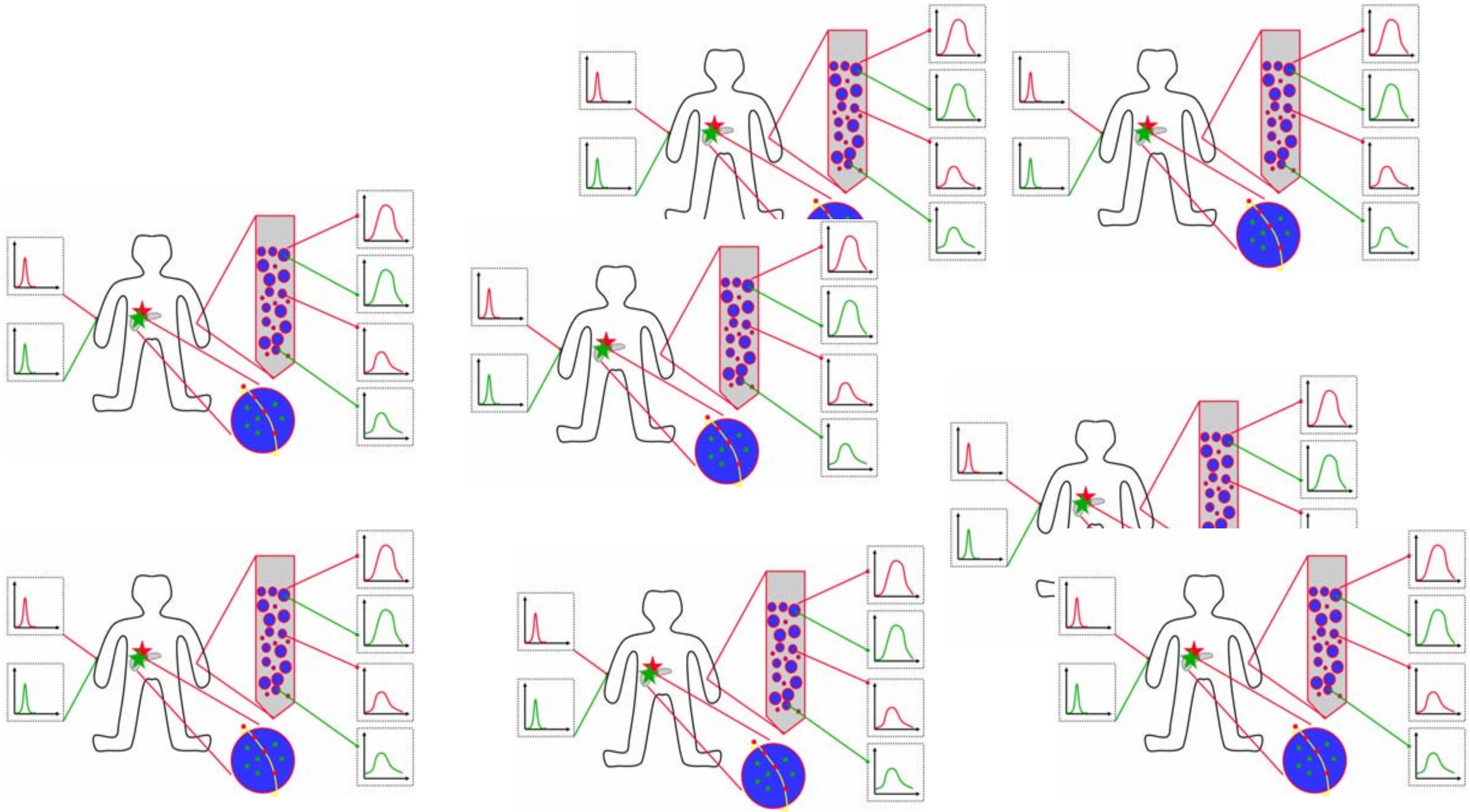
Compartmental modelling and tracer/tracee experiments





Next step: modelling of individual—population

(Martin Berglund)



The main part of my talk was supposed to be on the problem of identifiability and observability of nonlinear delay systems.

Very briefly the problems is the following:

Given a state space model of the following kind,

$$\begin{aligned}\dot{x}(t) &= f(x(t), x(t-\tau), u, u(t-\tau)) \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \\ y(t) &= h(x(t), x(t-\tau))\end{aligned}$$

-is it possible to find a so-called input-output relation, i.e. an equation that involves only the measured function, y(t), and its derivatives, possibly evaluated at present time, t, and at earlier times?

- can all input-output relations be reduced to involving only one timepoint (the present t), or is it necessary to include also delayed time points? (in the latter case, it is possible to identify the delay parameters from measured data, otherwise it is not.

The origin of our interest in this problem was a paper by Timmer et al, on a signalling pathway.

The main ideas are then presented in a couple of simple examples, where all calculations can be carried out explicitly. However, one of the objectives with the work is to find algorithms that can be implemented e.g. for computer algebra programs.

On page 19, an input output relation is shown, and on page 20 a simulation is shown that illustrates how such a relation in principle could be used to estimate the delay parameter(s) using the measured data.

Then the mathematical framework is described. This part is probably not so easy to follow from the slides and the comments that I add now.

The main paper has been accepted for publication in Automatica, and I can send a preprint or a copy of Milena’s thesis to anyone who is interested.

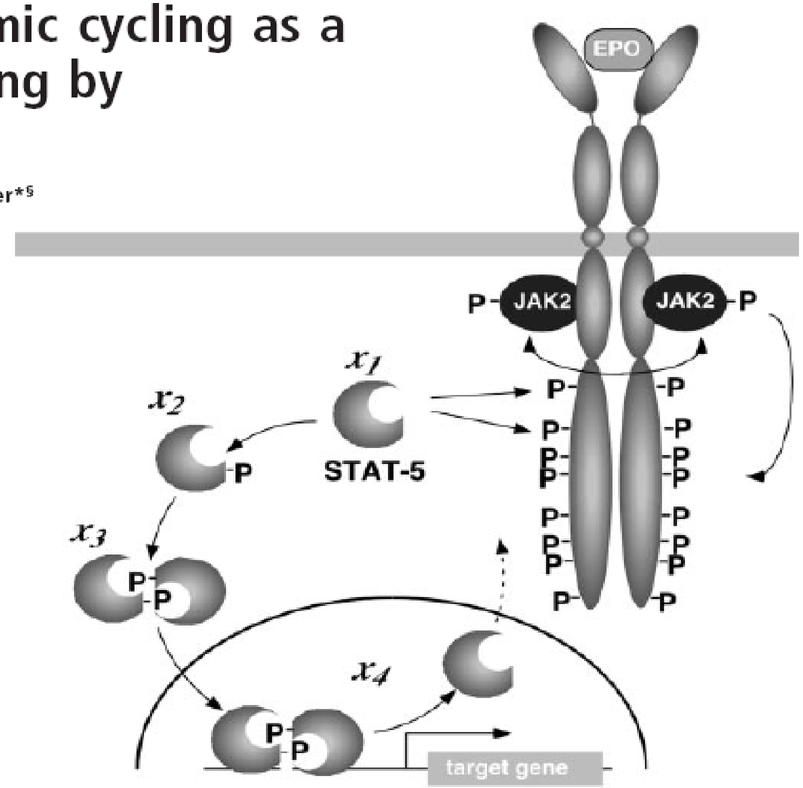
PARAMETER IDENTIFICATION IN DELAY DIFFERENTIAL EQUATIONS

Milena Anguelova (this work is part of her doctoral thesis)

Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling

I. Swameye*†, T. G. Müller†‡, J. Timmer†‡, O. Sandra*, and U. Klingmüller*§

$$
\begin{aligned}
\dot{x}_1 &= -k_1 x_1 \mathrm{EpoR}_A \\
\dot{x}_2 &= -k_2 x_2 + k_1 x_1 \mathrm{EpoR}_A \\
\dot{x}_3 &= -k_3 x_3 + \frac{1}{2} k_2 x_2^2 \\
\dot{x}_4 &= k_3 x_3
\end{aligned}
$$



PNAS 2002

A model of JAK-STAT signalling pathway with recycling

$$
\begin{aligned}
\dot{x}_1 &= -k_1 x_1 \mathrm{EpoR}_A + 2k_4 x_3(t-\tau) \\
\dot{x}_2 &= k_1 x_1 \mathrm{EpoR}_A - k_2 x_2^2 \\
\dot{x}_3 &= -k_3 x_3 + k_2 x_2^2 \\
\dot{x}_4 &= k_3 x_3 - k_4 x_3(t-\tau)
\end{aligned}
$$

$$
\begin{aligned}
y_1 &= k_5(x_2 + 2x_3) \\
y_2 &= k_6(x_1 + x_2 + 2x_3)
\end{aligned}
$$

Timmer J, Müller T G, Swameye I, Sandra O, Klingmüller U

Int. J. Bifurcation and Chaos, 14, 2069-2079

The property of identifiability

- Guarantees that the system parameters can be uniquely determined from measured data

- Well characterised for e.g. ODE-models and some delay systems (in particular for linear systems; Laplace transform methods)

- Identifiability analysed for linear delay systems

STATE SPACE MODELS (ODE)

Identifiability:

to determine k from measured y

Observability:

to determine x from measured y

$$
\begin{aligned}
\dot{x} &= f(x, k, u) \qquad & x \in \mathbf{R}^n \\
x(0) &= x_0 \\
y &= h(x, k) \qquad & y \in \mathbf{R}^p
\end{aligned}
$$

STATE SPACE MODELS (Delay Differential Equations)

$$
\begin{aligned}
\dot{x}(t) &= f(x(t), x(t-\tau), u, u(t-\tau)) \\
x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\
u(t) &= u_0(t), \quad t \in [-T, 0] \\
y(t) &= h(x(t), x(t-\tau))
\end{aligned}
$$

Identifiability: to determine k from measured y

we also want to identify τ

Observability: to determine x from measured y

NB: φ, u_0 not known

(initial data must be specified on [-T,0])

A SIMPLE EXAMPLE

$$
\begin{aligned}
\dot{x}_1 &= k_1 x_2(t-\tau) + u \\
\dot{x}_2 &= k_2 x_2(t-\tau) \\
y &= x_1 \\
x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\
u(t) &= u_0(t), \quad t \in [-T, 0]
\end{aligned}
$$

Differentiate output (and try to find x and the parameters)

$$
\begin{aligned}
\dot{y}(t) &= k_1 x_2(t-\tau) + u(t) \\
\ddot{y}(t) &= k_1 k_2 x_2(t-2\tau) + \dot{u}(t) \\
y^{(3)}(t) &= k_1 k_2^2 x_2(t-3\tau) + \ddot{u}(t)
\end{aligned}
$$

If delay parameter known:

-methods exist to determine observability and identifiability

(Zhang, Xia, Moog, IEEE T. Aut. Cont. 2006)

$$k_2 = \frac{\ddot{y}(t) - \dot{u}(t)}{\dot{y}(t-\tau) - u(t-\tau)}$$

(k_1 and x_1 cannot be observed)

… but what about the delay parameter itself?

An input-output relation

(i.e. an expression involving only y and u,

obtained by state elimination)

$$\xi_t(\tau) \equiv (\ddot{y}(t) - \dot{u}(t))(\ddot{y}(t-\tau) - \dot{u}(t-\tau))$$
$$-(y^{(3)}(t) - \ddot{u}(t))(\dot{y}(t-\tau) - u(t-\tau)) = 0$$

-The only unknown quantity: the delay parameter

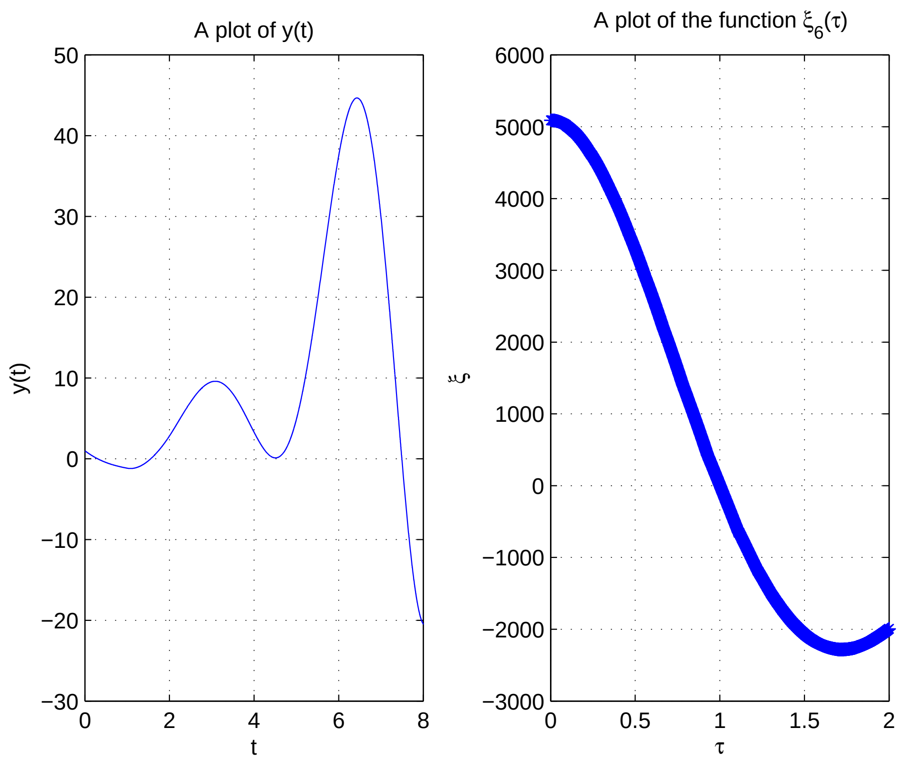
-This is identifiable if and only if such an expression can be obtained

-A mathematical question: when is this possible

In the following slide, the graph to the right shows the function $\xi_\tau(t)$ as a function of the delay parameter τ at a fixed time point t. The delay parameter can be found by finding the τ that makes the expression equal to zero.

This is in principle … in practice it may of course be very difficult to estimate higher derivatives of the measured function y(t) from data. There are new methods obtained by Fliess and coworkers that might be useful, and we would like to study this.

But of greater interest is to implement a computer algebra method for obtainin the input-output relation for larger state-space models, and to determine identifiability and observability in an automatic way. There are very efficient methods for determining observability and identifiability for ordinary differential equations (obtained by Sedoglavic), but is far from trivial to do the same for delay systems, and this is another topic that we would like to pursue.



y(t) and $\xi_t(\tau)$ at t=4

- in principle possible to determine delay this way,
- in practice difficult

ANOTHER SIMPLE EXAMPLE

$$\begin{aligned} \dot{x}_1 &= x_2^2(t-\tau) + u \\ \dot{x}_2 &= x_2(t) \\ y &= x_1 \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \end{aligned}$$

Differentiate output (and try to find x)

$$\begin{aligned} \dot{y}(t) &= x_2^2(t-\tau) \\ \ddot{y}(t) &= 2x_2^2(t-\tau) \end{aligned}$$

Input output relation does not contain the delay parameter:

$$\ddot{y}(t) = 2\dot{y}(t)$$

IDENTIFIABILITY OF THE DELAY PARAMETER

-If there is an input-output relation that contains the delay parameter explicity, then it is identifiable, otherwise not.

-For very simple systems, this can be calculated explicitly

-For realistic systems, this is usually not possible

-We need an effective algorithm for determining

1) if an input-output relation exist

2) whether or not it can be expressed without the delay parameter

Here: a linear algebraic method based on techniques developed by

Moog, Castro-Lineares, Velasco-Villa (IEEE Trans Aut Cont, 2000)

Márques-Martínes, Moog, Velasco-Villa (Kybernitica, 2000)

Xia, Márques-Villa, Zagalak, Moog (Automatica, 2002)

THE MATHEMATICAL FRAME WORK

$$\begin{aligned} \dot{x}(t) &= f(x(t), x(t-\tau), u, u(t-\tau)) \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \\ y(t) &= h(x(t), x(t-\tau)) \end{aligned}$$

$$h = (h_1, h_2, \ldots., h_p)$$

$$\delta(\xi(t)) = \xi(t-\tau)$$ the time delay operator

$\mathcal{K}$ the field of meromorphic functions of a finite numer of variables from

$$\{x(t-k\tau), u(t-k\tau), \ldots, u^{(l)}(t-k\tau), \quad k, l \in \mathbb{Z}^+\}$$

$\mathcal{E}$ the vector space $\mathcal{E} = \mathrm{span}_{\mathcal{K}}\{d\xi : \xi \in \mathcal{K}\}$

A non commutative ring:

$\mathcal{K}(\delta]$ polynomials of the form

$$a(\delta] = a_0(t) + a_1(t)\delta + \cdots + a_{r_a}(t)\delta^{r_a} \qquad a_j \in \mathcal{K}$$

Multiplication:

$$a(\delta]b(\delta] = \sum_{k=0}^{r_a+r_b} \sum_{i+j=k}^{i\leq r_a, j\leq r_b} a_i(t)b_j(t-i\tau)\delta^k$$

Properties of $\mathcal{K}(\delta]$

- Noetherian, and a left Ore ring: $\mathcal{K}(\delta]a(\delta] \cap \mathcal{K}(\delta]b(\delta] \neq 0$

- It is possible to carry out row elimination in a matrix with elements from $\mathcal{K}(\delta]$

- The rank of submodules of $\mathcal{M} = \mathrm{span}_{\mathcal{K}(\delta]}\{d\xi : \xi \in \mathcal{K}\}$ is well defined

In the following slide, we introduce some unusual notation. The function F is vector valued, where the components are the mesured data and derivatives of the measured data. These functions may implitly depend on the state variables, and their derivatives evaluated at t and at delayed times. A matrix is then defined, whose entries are polynomials in the delay operator δ.

The methods of the paper a based on computing the rank of this matrix. This is delicate even to define for matrices with entries in a non-commutative ring, but we are saved by the fact that we deal with an Ore-ring. The expression

$$\mathcal{K}(\delta]a(\delta] \cap \mathcal{K}(\delta]b(\delta] \neq 0$$

implies that for any two elements a(δ] and b(δ] one can find other elements

m(δ] and n(δ] such that m(δ] a(δ] + n(δ] b(δ] =0, and this is exactly what is needed to carry out row elimination in the matrix.

Of course, in general it is very difficult to actually carry out such calculations without using some program for symbolic calculation.

Existence of an input-output relation

A definition: let $f \in \mathcal{K}^r$

Then: $\dfrac{\partial f}{\partial x}$ is the matrix with entries

$$\left(\frac{\partial f}{\partial x}\right)_{j,i} = \sum_k \frac{\partial f_j}{\partial x_i(t-k\tau)}\delta^k \in \mathcal{K}(\delta] .$$

Denote by s_1 the smallest integer such that

$$\mathrm{rank}_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1-1)})}{\partial x} = \mathrm{rank}_{\mathcal{K}(\delta]} \frac{\partial(h_1, ..., h_1^{(s_1)})}{\partial x}$$

and continue with the other h_j to form

$$S = (h_1, ..., h_1^{(s_1-1)}, \ldots, h_p, \ldots, h_p^{(s_p-1)})$$

S can be completed with functions g_k so that

$$\mathrm{rank}_{\mathcal{K}(\delta]} \frac{\partial(S, g_1, \ldots, g_{n-K})}{\partial x} = n$$

Using

-Poincaré's lemma

-The Ore property, and definitions of the closure of the submodule

$$\mathrm{span}_{\mathcal{K}(\delta]} \left\{ \frac{\partial(h_1, \ldots, h_1^{(s_1-1)}, \ldots, h_i, \ldots, h_i^{(s_i-1)})}{\partial x} \right\}$$

one can prove that:

There exists a neigborhood of initial data and input functions so that an input output relation exists.

…. But does it contain the delay parameter?

Theorem:

Consider a system formulated in state space form, and the vector S, as before. Then there exists an input-output equation

$$\phi(\delta, y, \ldots, y^{(l)}, u, \ldots, u^{(k)}) = 0$$

that involves δ in an essential way if and only if one of the following is true:

1) A delayed input variable $u_r^{(k)}$ occurs in some of the functions in

$$\{S, h_1^{(s_1)}, \ldots, h_p^{(s_p)}\}$$

2)

$$\mathrm{rank}_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} \neq \mathrm{rank}_{\mathcal{K}} \frac{\partial(S, h_1^{(s_1)}, \ldots, h_p^{(s_p)})}{\partial x}$$