An Efficient Method for Structural Identifiability Analysis of Large Dynamic Systems

Pharmacokinetics UK 2014 Meeting
November 5-7, 2014, Bath, UK

Fraunhofer

CHALMERS

Research Centre Industrial Mathematics

Johan Karlsson, Milena Anguelova, Mats Jirstrand
Department of Systems and Data Analysis

Outline

Structural Identifiability Analysis

- Forward look: Structural identifiability analysis for $x, \theta \in \mathbb{R}^{100}$ on standard desktop PC
- Structural Identifiability
- A Jacobian Matrix
  - Naive approach (bad computational complexity)
  - Alternative approach (good computational complexity)
- `IdentifiabilityAnalysis` – a *Mathematica* application package
- Timings of a Few Large Models
- Conclusions

Fraunhofer ITWM
CHALMERS
Research Centre
Industrial Mathematics

Structural Identifiability

Can the parameters of a set of differential equations be computed from perfect knowledge of the input and a given set of variables assumed to be measurable?

- Idealized situation but a necessary condition for any estimation procedure to be able to return a sensible result.
- If not identifiable - what parameters needs to be fixed 'to values from elsewhere' or what additional variables needs to measured?
- 'Structural' denotes a property of the symbolic form of the equations regardless of specific numerical values for parameters or initial conditions.

We have implemented and extended a highly efficient method (Sedoglavic 2002) in terms of a *Mathematica* package capable of handling system in the order of a hundred states and equally many parameters on a regular laptop.

Local algebraic identifiability!

IdentifiabilityAnalysis – A *Mathematica* Package

$$
\begin{aligned}
\dot{M} &= \frac{v_s {K_I}^4}{{K_I}^4+{P_N}^4} - \frac{v_m M}{K_m+M}, \\
\dot{P}_0 &= k_s M - \frac{V_1 P_0}{K_1+P_0} + \frac{V_2 P_1}{K_2+P_1}, \\
\dot{P}_1 &= \frac{V_1 P_0}{K_1+P_0} + \frac{V_4 P_2}{K_4+P_2} - P_1\left(\frac{V_2}{K_2+P_1} + \frac{V_3}{K_3+P_1}\right), \\
\dot{P}_2 &= \frac{V_3 P_1}{K_3+P_1} - P_2\left(\frac{V_4}{K_4+P_2} + k_1 + \frac{v_d}{K_d+P_2}\right) + k_2 P_N, \\
\dot{P}_N &= k_1 P_2 - k_2 P_N, \\
y &= P_N.
\end{aligned}
$$

n=5 state variables
d=21 parameters
m=1 inputs
p=1 outputs

IdentifiabilityAnalysis – A *Mathematica* Package

$$\dot M = \frac{v_s K_I{}^4}{K_I{}^4+P_N{}^4} - \frac{v_m M}{K_m+M},$$
$$\dot P_0 = k_s M - \frac{V_1 P_0}{K_1+P_0} + \frac{V_2 \ldots}{K_2+\ldots}$$
$$\dot P_1 = \frac{V_1 P_0}{K_1+P_0} + \frac{V_4 P_2}{K_4+P_2} - P_1 \ldots$$
$$\dot P_2 = \frac{V_3 P_1}{K_3+P_1} - P_2 \left(\frac{V_4}{K_4+P_2} \ldots \right.$$
$$\dot P_N = k_1 P_2 - k_2 P_N,$$
$$y = P_N.$$

deq = {

$$M'[t] == \frac{v_s K_i^4}{K_i^4 + P_N[t]^4} - \frac{v_m M[t]}{K_m + M[t]}, \; M[0] == m_0,$$

$$P_0'[t] == k_s M[t] - \frac{V_1 P_0[t]}{K_1 + P_0[t]} + \frac{V_2 P_1[t]}{K_2 + P_1[t]}, \; P_0[0] == p_0,$$

$$P_1'[t] == \frac{V_1 P_0[t]}{K_1 + P_0[t]} + \frac{V_4 P_2[t]}{K_4 + P_2[t]} - P_1[t]\left(\frac{V_2}{K_2 + P_1[t]} + \frac{V_3}{K_3 + P_1[t]}\right), \; P_1[0] == p_1,$$

$$P_2'[t] == \frac{V_3 P_1[t]}{K_3 + P_1[t]} - P_2[t]\left(\frac{V_4}{K_4 + P_2[t]} + k_1 + \frac{v_d[t]}{K_d + P_2[t]}\right) + k_2 P_N[t], \; P_2[0] == p_2,$$

$$P_N'[t] == k_1 P_2[t] - k_2 P_N[t], \; P_N[0] == p_N\};$$

params = $\{v_s, v_m, k_s, k_1, k_2, K_d, K_i, K_m, V_1, V_2, V_3, V_4, K_1, K_2, K_3, K_4, m_0, p_0, p_1, p_2, p_N\}$;

9 s ⟶ iad = IdentifiabilityAnalysis[{deq, P_N[t]}, {M, P_0, P_1, P_2, P_N}, params, t, v_d]

IdentifiabilityAnalysisData[False, <>]

iad["NonIdentifiableParameters"]

$\{k_s, m_0, v_m, v_s, K_m\}$

iad["DegreesOfFreedom"]

1

”In a nutshell”

The method (Sedoglavic 2002) works by

- Specialization of symbolic parameter values and initial conditions to random integers
- Specialization of input/perturbation signals to truncated random integer coefficient power series
- Computations of truncated power series representations of the solution to the sensitivity differential equations of the system
- Exact rank computations of a observability/identifiability matrix
- All computations are done modulus a large prime, which returns correct analysis with very high probability.

Truncated random integer coefficient power series ←

Integer matrix Modular rank ←

Fraunhofer ITWM
CHALMERS
Research Centre
Industrial Mathematics

System Description

A parametrized class of models in state space form

$$\dot{x}(t) = f(x(t), u(t), \theta), \qquad x(0) = x^0(\theta)$$
$$y(t) = g(x(t), u(t), \theta)$$

where $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m, \theta \in \mathbb{R}^d, y(t) \in \mathbb{R}^p$ and f and g are rational functions of x, u, and θ.

Output Derivatives

Higher-order derivatives of the output w.r.t. time $y^{(\nu)}$ can be obtained by repeated use of the chain rule and replacing $\dot{x}$ using the system dynamics (a.k.a. extended *Lie*-derivative along f)

$$y = g$$

$$\dot{y} = \frac{\partial g}{\partial x} f + \frac{\partial g}{\partial u} \dot{u} = \mathcal{L}_f g$$

$$\ddot{y} = \mathcal{L}_f(\mathcal{L}_f g) = \mathcal{L}_f^2 g$$

$$\vdots$$

$$y^{(\nu)} = \mathcal{L}_f^\nu g \qquad \nu = n + d - 1$$

This is true for any point in time and in particular *t=0*.

y(t), u(t) known $\Rightarrow$
LHS known
RHS fcn of *x*(0) and θ
($u^{(i)}(0)$ is a known quantity)

where $\mathcal{L}_f = \sum_{i=1}^n f_i \frac{\partial}{\partial x_i} + \sum_{i=0}^\infty u^{(i+1)} \frac{\partial}{\partial u^{(i)}}$

$$\mathcal{Y} = \mathcal{Y}(x, \theta)$$

It can be shown that $y^{(\nu)}, \nu \geq n + d$ are agebraically dependent, i.e., could be written as solutions to algebraic equations with coefficients made up of expressions in lower order derivatives.

Output Derivatives

$$\mathcal{Y} = \mathcal{Y}(x, \theta)$$

i.e., y and all its derivatives can be expressed in terms of the state and parameters (and the input and its derivatives).

What about the reverse relation? ← Structural identifiability (and observability)

A System of Nonlinear Equations

The system of equations

$$\mathcal{Y} = \mathcal{Y}(x, \theta)$$

can be uniquely solved (locally) for x and θ iff the Jacobian

$$J(x,\theta) = \frac{\partial \mathcal{Y}(x,\theta)}{\partial(x,\theta)} = \begin{pmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_n} & \cdots & \frac{\partial y}{\partial \theta_1} & \cdots & \frac{\partial y}{\partial \theta_d} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \frac{\partial y^{(\nu)}}{\partial x_1} & \cdots & \frac{\partial y^{(\nu)}}{\partial x_n} & \cdots & \frac{\partial y^{(\nu)}}{\partial \theta_1} & \cdots & \frac{\partial y^{(\nu)}}{\partial \theta_d} \end{pmatrix}$$

$\nu = n + d - 1$

is non-singular, i.e., $\det(J) \neq 0$ (the inverse function theorem).

We will also use the notation:

$$J(x,\theta) = \left(\frac{\partial \left(\mathcal{L}_f^i g \right)_{0 \le i \le \nu}}{\partial(x,\theta)} \right)$$

A System of Nonlinear Equations

Local structural identifiability is a generic property, i.e., should be true for almost any (initial condition and) parameter value.

Check if the Jacobian has full rank for a random specialization of the (initial conditions and) parameter values and input!

$$J(x,\theta) = \frac{\partial \mathcal{Y}(x,\theta)}{\partial (x,\theta)} = \begin{pmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_n} & \cdots & \frac{\partial y}{\partial \theta_1} & \cdots & \frac{\partial y}{\partial \theta_d} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \frac{\partial y^{(\nu)}}{\partial x_1} & \cdots & \frac{\partial y^{(\nu)}}{\partial x_n} & \cdots & \frac{\partial y^{(\nu)}}{\partial \theta_1} & \cdots & \frac{\partial y^{(\nu)}}{\partial \theta_d} \end{pmatrix}$$

$\nu = n + d - 1$

Columns removed without affecting the rank indicates unidentifiable parameters.

Example – The Naive Approach

Consider the following example (Vajda et al, 1989)

$$
\begin{aligned}
\dot{x}_1 &= \theta_1 x_1^2 + \theta_2 x_1 x_2 + u, \qquad & x_1(0) &= x_1^0 \\
\dot{x}_2 &= \theta_3 x_1^2 + \theta_4 x_1 x_2, \qquad & x_2(0) &= x_2^0 \\
y &= x_1
\end{aligned}
$$

1) Compute time derivative of the output (Lie-derivatives)
2) Compute partial derivatives w.r.t. state and parameters
3) Specialize on random integer values (we want to be exact!)
4) Compute rank!

$$\{ \boxed{x_1(0)},\ \boxed{\theta_1 x_1(0)^2 + \theta_2 x_2(0) x_1(0) + u(0)},$$

$$\boxed{2\theta_1^2 x_1(0)^3 + \theta_2^2 x_2(0)^2 x_1(0) + \theta_1 (2u(0) + 3\theta_2 x_1(0) x_2(0)) x_1(0) + \theta_2 (\theta_3 x_1(0)^3 + (\theta_4 x_1(0)^2 + u(0)) x_2(0)) + u'(0)},$$

$$6\theta_1^3 x_1(0)^4 + 4\theta_1^2 (2u(0) + 3\theta_2 x_1(0) x_2(0)) x_1(0)^2 + \theta_2^3 x_2(0)^3 x_1(0) +$$
$$\theta_2^2 x_2(0) (5\theta_3 x_1(0)^3 + (4\theta_4 x_1(0)^2 + u(0)) x_2(0)) + \theta_2 (\theta_3 (\theta_4 x_1(0)^4 + 4u(0) x_1(0)^2) + x_2(0) (\theta_4^2 x_1(0)^3 + 3\theta_4 u(0) x_1(0) + u'(0))) + ,$$
$$\theta_1 (7\theta_2^2 x_1(0)^2 x_2(0)^2 + \theta_2 (6\theta_3 x_1(0)^4 + (5\theta_4 x_1(0)^2 + 8u(0)) x_2(0) x_1(0)) + 2(u(0)^2 + x_1(0) u'(0))) + u''(0)$$

$$24\theta_1^4 x_1(0)^5 + 20\theta_1^3 (2u(0) + 3\theta_2 x_1(0) x_2(0)) x_1(0)^3 + \theta_2^4 x_2(0)^4 x_1(0) +$$
$$\theta_1^2 (50\theta_2^2 x_1(0)^2 x_2(0)^2 + 3\theta_2 (12\theta_3 x_1(0)^4 + (9\theta_4 x_1(0)^2 + 20u(0)) x_2(0) x_1(0)) + 2(8u(0)^2 + 5x_1(0) u'(0))) x_1(0) +$$
$$\theta_2^3 x_2(0)^2 (18\theta_3 x_1(0)^3 + (11\theta_4 x_1(0)^2 + u(0)) x_2(0)) +$$
$$\theta_2^2 (5\theta_3^2 x_1(0)^5 + \theta_3 (17\theta_4 x_1(0)^2 + 25u(0)) x_2(0) x_1(0)^2 + x_2(0)^2 (11\theta_4^2 x_1(0)^3 + 13\theta_4 u(0) x_1(0) + u'(0))) +$$
$$\theta_1 (15\theta_2^3 x_1(0)^2 x_2(0)^3 + \theta_2^2 x_1(0) (53\theta_3 x_1(0)^3 + (37\theta_4 x_1(0)^2 + 22u(0)) x_2(0)) x_2(0) + 6u(0) u'(0) +$$
$$\theta_2 (\theta_3 (9\theta_4 x_1(0)^5 + 40u(0) x_1(0)^3) + 2x_2(0) (4\theta_4^2 x_1(0)^4 + 13\theta_4 u(0) x_1(0)^2 + 5u'(0) x_1(0) + 4u(0)^2 \ldots$$ [illegible]
$$\theta_2 (\theta_3 x_1(0) (\theta_4^2 x_1(0)^4 + 7\theta_4 u(0) x_1(0)^2 + 5u'(0) x_1(0) + 8u(0)^2) + x_2(0) (\theta_4^3 x_1(0)^4 + 6\theta_4^2 u(0) \ldots$$ [illegible] $$\ldots u''(0))) +$$
$$u^{(3)}(0)$$

$$120\theta_1^5 x_1(0)^6 + 120\theta_1^4 (2u(0) + 3\theta_2 x_1(0) x_2(0)) x_1(0)^4 +$$
$$2\theta_1^3 (68u(0)^2 + 195\theta_2^2 x_1(0)^2 x_2(0)^2 + 12\theta_2 (10\theta_3 x_1(0)^4 \ldots$$ [illegible] $$\ldots) + 30 x_1(0) u'(0)) x_1(0)^2 +$$
$$\theta_2^5 x_2(0)^5 x_1(0) + \theta_2^4 x_2(0)^3 (58\theta_3 x_1(0)^3 + (26 \ldots$$ [illegible]
$$\theta_2^3 x_2(0) (61\theta_3^2 x_1(0)^5 + \theta_3 (137\theta_4 x_1 \ldots$$ [illegible] $$\ldots (66\theta_4^2 x_1(0)^3 + 38\theta_4 u(0) x_1(0) + u'(0))) +$$
$$\theta_2^2 ((17\theta_4 x_1(0)^6 + 50u(0) \ldots$$ [illegible] $$\ldots + 140\theta_4 u(0) x_1(0)^2 + 37u'(0) x_1(0) + 58u(0)^2) \theta_3 +$$
$$x_2(0)^2 (26\theta_4^3 x_1 \ldots$$ [illegible] $$\ldots u(0)^2 + 19 x_1(0) u'(0)) + u''(0))) +$$
$$\theta_1^2 (180\theta_2^3 x_1(0)^3 x_2(0) \ldots$$ [illegible] $$\ldots x_1(0)^3 + (319\theta_4 x_1(0)^2 + 292u(0)) x_2(0)) x_2(0) +$$
$$\theta_2 x_1(0) (72\theta_3 (\theta_4 x \ldots$$ [illegible] $$\ldots (0) x_1(0)^3) + x_2(0) (59\theta_4^2 x_1(0)^4 + 220\theta_4 u(0) x_1(0)^2 + 90u'(0) x_1(0) + 136u(0)^2)) +$$
$$4(4u(0)^3 + 13 x_1(0) u'(0) u(0) + 3 x_1(0)^2 u''(0))) +$$
$$\theta_1 (31\theta_2^4 x_1(0)^2 x_2(0)^4 + \theta_2^3 x_1(0) (311\theta_3 x_1(0)^3 + 2(89\theta_4 x_1(0)^2 + 26u(0)) x_2(0)) x_2(0)^2 + 6u'(0)^2 +$$
$$\theta_2^2 (78\theta_3^2 x_1(0)^6 + 6\theta_3 (40\theta_4 x_1(0)^2 + 71u(0)) x_2(0) x_1(0)^3 + x_2(0)^2 (139\theta_4^2 x_1(0)^4 + 220\theta_4 u(0) x_1(0)^2 + 32u'(0) x_1(0) + 22u(0)^2)) +$$
$$8u(0) u''(0) + \theta_2 (\theta_3 (13\theta_4^2 x_1(0)^4 + 92\theta_4 u(0) x_1(0)^2 + 60u'(0) x_1(0) + 136u(0)^2) x_1(0)^2 +$$
$$2x_2(0) (6\theta_4^3 x_1(0)^5 + 35\theta_4^2 u(0) x_1(0)^3 + 10\theta_4 (3u(0)^2 + 2x_1(0) u'(0)) x_1(0) + 6u''(0) x_1(0) + 13u(0) u'(0))) + 2x_1(0) u^{(3)}(0)) +$$
$$\theta_2 (\theta_3 (\theta_4^3 x_1(0)^6 + 11\theta_4^2 u(0) x_1(0)^4 + \theta_4 (24u(0)^2 + 11 x_1(0) u'(0)) x_1(0)^2 + 6u''(0) x_1(0)^2 + 26u(0) u'(0) x_1(0) + 8u(0)^3) +$$
$$x_2(0) (\theta_4^4 x_1(0)^5 + 10\theta_4^3 u(0) x_1(0)^3 + 5\theta_4^2 (3u(0)^2 + 2x_1(0) u'(0)) x_1(0) + 5\theta_4 (2u(0) u'(0) + x_1(0) u''(0)) + u^{(3)}(0))) + u^{(4)}(0) \}$$

es)
ters
e exact!)

Computational complexity!

The Sensitivity Equations

Observation: the elements of the Jacobian matrix equals the coefficients of the formal Taylor series expansion around t=0 of the output sensitivity derivatives (wrt initial conditions and parameters)

$$\begin{pmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_n} & \cdots & \frac{\partial y}{\partial \theta_1} & \cdots & \frac{\partial y}{\partial \theta_d} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \frac{\partial y^{(\nu)}}{\partial x_1} & \cdots & \frac{\partial y^{(\nu)}}{\partial x_n} & \cdots & \frac{\partial y^{(\nu)}}{\partial \theta_1} & \cdots & \frac{\partial y^{(\nu)}}{\partial \theta_d} \end{pmatrix}$$

$$y(t) = y + \dot{y}\,t + \ddot{y}\,\frac{t^2}{2!} + \dddot{y}\,\frac{t^3}{3!} + \mathcal{O}(t^4)$$

$$\frac{d}{dx_i^0} y(t) = \frac{\partial y}{\partial x_i^0} + \frac{\partial \dot{y}}{\partial x_i^0} t + \frac{\partial \ddot{y}}{\partial x_i^0}\frac{t^2}{2!} + \frac{\partial y^{(3)}}{\partial x_i^0}\frac{t^3}{3!} + \mathcal{O}(t^4)$$

$$\frac{d}{d\theta_i} y(t) = \frac{\partial y}{\partial \theta_i} + \frac{\partial \dot{y}}{\partial \theta_i} t + \frac{\partial \ddot{y}}{\partial \theta_i}\frac{t^2}{2!} + \frac{\partial y^{(3)}}{\partial \theta_i}\frac{t^3}{3!} + \mathcal{O}(t^4)$$

Idea: compute the power series of y(t) and its sensitivity derivatives directly from a random specialization of initial conditions, parameter values, and input function.

This can be done very computationaly efficiently utilizing random integers, truncated integer coefficient power series, and modular arithmetics.

The Sensitivity Equations

$$\Sigma^* : \begin{cases} \Sigma : \ \dot{x} = f(x,u,\theta), & x(0) = x^0 \\ \dfrac{d}{dt}\dfrac{\partial x}{\partial x_i^0} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial x_i^0}, & \dfrac{\partial x}{\partial x_i^0}(0) = 1_n \\ \dfrac{d}{dt}\dfrac{\partial x}{\partial \theta_i} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \theta_i} + \dfrac{\partial f}{\partial \theta_i}, & \dfrac{\partial x}{\partial \theta_i}(0) = 0_d \end{cases}$$

Random integer specialization:

$$x^0 \to \tilde{x}^0 \qquad \theta \to \tilde{\theta} \qquad u(t) \to \tilde{u}(t)$$

$$x = x(t), \qquad \frac{\partial x}{\partial x_i^0} = \frac{\partial x}{\partial x_i^0}(t), \qquad \frac{\partial x}{\partial \theta_i} = \frac{\partial x}{\partial \theta_i}(t)$$

Truncated power series

The system Σ^* can be solved iteratively generating truncated power series solutions of desired order. Insertion into the output sensitivity expressions gives truncated power series

$$\frac{d}{dx_i^0}y(t) = \frac{\partial g}{\partial x}(x(t),u(t),\theta)\,\frac{\partial x}{\partial x_i^0}(t)$$

$$\frac{d}{d\theta_i}y(t) = \frac{\partial g}{\partial x}(x(t),u(t),\theta)\,\frac{\partial x}{\partial \theta_i}(t) + \frac{\partial g}{\partial \theta_i}(x(t),u(t),\theta)$$

Truncated power series

Mats Jirstrand

Fraunhofer ITWM
CHALMERS
Research Centre
Industrial Mathematics

$$\frac{d}{dx_i^0}y(t) = \frac{\partial y}{\partial x_i^0} + \frac{\partial \dot{y}}{\partial x_i^0}t + \frac{\partial \ddot{y}}{\partial x_i^0}\frac{t^2}{2!} + \frac{\partial y^{(3)}}{\partial x_i^0}\frac{t^3}{3!} + \mathcal{O}(t^4)$$

$$\frac{d}{d\theta_i}y(t) = \frac{\partial y}{\partial \theta_i} + \frac{\partial \dot{y}}{\partial \theta_i}t + \frac{\partial \ddot{y}}{\partial \theta_i}\frac{t^2}{2!} + \frac{\partial y^{(3)}}{\partial \theta_i}\frac{t^3}{3!} + \mathcal{O}(t^4)$$

Numerical Approach

Symbolic

Σ — Symbolic computations (repeated Lie-derivatives) → $\left(\frac{\partial \left(\mathcal{L}_f^i g \right)_{0 \leq i \leq \nu}}{\partial (x, \theta)} \right)$

Σ ↓ Σ^*

$$x \to \tilde{x}, \quad \theta \to \tilde{\theta}, \quad u(t) \to \tilde{u}(t)$$

Σ^* — Numerical computations (truncated power series) → $\left(\frac{\partial \left(\mathcal{L}_f^i g \right)_{0 \leq i \leq \nu}}{\partial (x, \theta)} \right)$

$$x \to \tilde{x}, \quad \theta \to \tilde{\theta}, \quad u(t) \to \tilde{u}(t)$$

Numeric

+ modular arithmetics

IdentifiabilityAnalysis – A Small Example

Load the package

```
In[1]:= Needs["IdentifiabilityAnalysis`"]
```

A model (Vajda et al., 1989) with numerical initial conditions (generic case)

```
In[2]:= deq = {
           x1'[t] == θ1 x1[t]^2 + θ2 x1[t]x2[t] + u[t], x1[0]==1,
           x2'[t] == θ3 x1[t]^2 + θ4 x1[t]x2[t], x2[0]==2
          };
```

An identifiability analysis of the model with output $x_1[t]$:

```
In[3]:= IdentifiabilityAnalysis[
                {deq,x1[t]},{x1,x2},Table[θi,{i,4}],t,u];

Out[3]= IdentifiabilityAnalysis[True,<>]
```


Fraunhofer ITWM
CHALMERS
Research Centre
Industrial Mathematics

IdentifiabilityAnalysis – A Small Example

Load the package

```
In[1]:= Needs["IdentifiabilityAnalysis`"]
```

A model (Vajda et al., 1989) with numerical initial conditions (special case)

```
In[2]:= deq = {
           x1'[t] == θ1 x1[t]^2 + θ2 x1[t]x2[t] + u[t], x1[0]==1,
           x2'[t] == θ3 x1[t]^2 + θ4 x1[t]x2[t], x2[0]==0
          };
```

An identifiability analysis of the model with output $x_1[t]$:

```
In[3]:= iad = IdentifiabilityAnalysis[
                  {deq,x1[t]},{x1,x2},Table[θi,{i,4}],t,u];

Out[3]= IdentifiabilityAnalysis[False,<>]

In[4]:= iad["NonIdentifiableParameters"]

Out[4]= {θ2,θ3}

In[5]:= iad["DegreesOfFreedom"]

Out[5]= 1
```

IdentifiabilityAnalysis – A Small Example

Load the package

```
In[1]:= Needs["IdentifiabilityAnalysis`"]
```

A model (Vajda et al., 1989) with parametrized initial conditions

```
In[2]:= deq = {
           x1'[t] == θ1 x1[t]^2 + θ2 x1[t]x2[t] + u[t], x1[0]==θ5,
           x2'[t] == θ3 x1[t]^2 + θ4 x1[t]x2[t], x2[0]==θ6
          };
```

An identifiability analysis of the model with output `x1[t]`:

```
In[3]:= iad = IdentifiabilityAnalysis[
                    {deq,x1[t]},{x1,x2},Table[θi,{i,6}],t,u];

Out[3]= IdentifiabilityAnalysis[False,<>]

In[4]:= iad["NonIdentifiableParameters"]

Out[4]= {θ2,θ3,θ6}

In[5]:= iad["DegreesOfFreedom"]

Out[5]= 1
```

IdentifiabilityAnalysis – Syntax

`IdentifiabilityAnalysis`[{*eqns*, *expr*}, {x_1,x_2, ...}, {θ_1, θ_2, ...}, *t*, *u*]

performs an identifiability analysis of a system defined by the system of ordinary differential equations *eqns* with output given by *expr* in the variables x_i, parameters θ_i, independent variable *t*, and input *u*.

- **`IdentifiabilityAnalysis`** returns an **`IdentifiabilityAnalysisData`** object
- The system *eqns* must be written as a system of first order differential equations including initial conditions.
- The system needs to be rational.
- Initial conditions may be numbers or expressions containing parameters.
- For autonomous systems the last argument denoting the input can be dropped.
- For multi-output systems *expr* is replaced by a list of expressions {$expr_1$, $expr_2$, ...}.
- For multi-input systems *u* is replaced by a list of input symbols {u_1, u_2, ...}.

IdentifiabilityAnalysis – Syntax

`IdentifiabilityAnalysisData`[...]

represents identifiability analysis data generated by **`IdentifiabilityAnalysis`**.

- An **`IdentifiabilityAnalysisData`** object, *iad*, can be used to retrieve additional analysis results and reports through *iad*[”*property*”].
- A list of available properties is given by *iad*["Properties"].
- Available properties are:

`"IdentifiableQ"`	True if the system is identifiable and False otherwise
`"NonIdentifiableParameters"`	list of non-identifiable parameters
`"DegreesOfFreedom"`	number of non-identifiable parameters, which should be assumed to be known to obtain an identifiable system

IdentifiabilityAnalysis – A *Mathematica* Package

Load the package

```
In[1]:= Needs["IdentifiabilityAnalysis`"]
```

A model of circadian oscillations in Drosophila period protein (Goldbeter, 1995)

In[2]:= deq = {

$$M'[t] == v_s\, K_I{}^4/(K_I{}^4+P_N[t]^4)-v_m M[t]/(K_M+M[t]),$$

$$P_0'[t] == k_s\, M[t]-V_1 P_0[t]/(K_1+P_0[t])+V_2 P_1[t]/(K_2+P_1[t]),$$

$$P_1'[t] == V_1 P_0[t]/(K_1+P_0[t])+V_4 P_2[t]/(K_4+P_2[t]) -P_1[t](V_2/(K_2+P_1[t])+ V_3/(K_3+P_1[t])),$$

$$P_2'[t] == V_3 P_1[t]/(K_3+P_1[t]) -P_2[t](V_4/(K_4+P_2[t])+k_1+v_d/(K_d+P_2[t]))+k_2 P_N[t],$$

$$P_N'[t] == k_1 P_2[t]-k_2 P_N[t],$$

$$M[0]== m_0,\ P_0[0]== p_0,\ P_1[0]== p_1,\ P_2[0]== p_2,\ P_N[0]== p_N$$

};

params = $\{v_s, K_I, v_m, K_M, k_s, V_1, K_1, V_2, K_2, V_3, K_3, V_4, K_4, k_1, K_d, k_2, m_0, p_0, p_1, p_2, p_N\}$;

IdentifiabilityAnalysis – A *Mathematica* Package

An identifiability analysis of the model with input v_d and output $P_N[t]$:

```
In[3]:= iad = IdentifiabilityAnalysis[
              {deq,P_N[t]},{M,P_0,P_1,P_2,P_N},params,t,v_d]

Out[3]= IdentifiabilityAnalysis[False,<>]
```

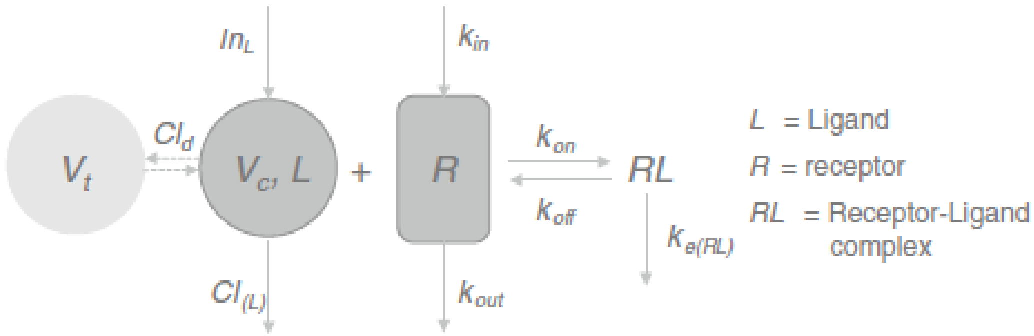
The nonidentifiable parameters of the model:

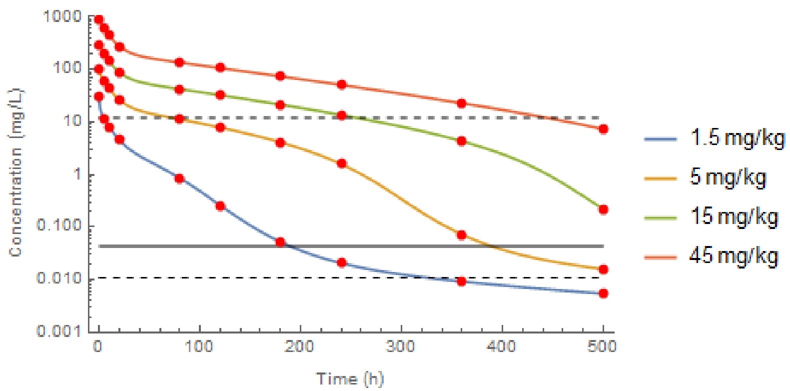
```
In[4]:= iad["NonIdentifiableParameters"]

Out[4]= {K_M,k_s,m_0,v_m,v_s}
```

The number of nonidentifiable parameters, which should be assumed to be known to obtain an identifiable system:

```
In[5]:= iad["DegreesOfFreedom"]

Out[5]= 1
```

IdentifiabilityAnalysis – A *Mathematica* Package





A target mediated drug disposition model (Peletier&Gabrielsson, 2012)

```
In[6]:= deq = {
      L'[t]==(-ClL*L[t] - Cld*(L[t] - T[t]))/Vc -
                                 kon*L[t]*R[t] + koff*RL[t],
      T'[t] == Cld*(L[t] - T[t])/Vt,
      R'[t] == kout*(R0 - R[t]) - kon*L[t]*R[t] + koff*RL[t],
      RL'[t] == kon*L[t]*R[t] - (koff + keRL)*RL[t]};
      L[0] == D/Vc, T[0] == 0, R[0] == R0, RL[0] == 0 };

   params = {Vt, Cld, ClL, kon, koff, kout, keRL, R0};
```

IdentifiabilityAnalysis – A *Mathematica* Package

An identifiability analysis of the model with $\mathcal{D}$ as a parameter and output `L[t]`:

```
In[7]:= iad = IdentifiabilityAnalysis[
              {deq/.L0->D/Vc,L[t]},{L,T,R,RL},Append[params, 𝒟],t]

Out[7]= IdentifiabilityAnalysis[False,<>]
```

The nonidentifiable parameters of the model:

```
In[8]:= iad["NonIdentifiableParameters"]

Out[8]= {Cld, ClL, Vc, Vt, 𝒟}
```

The number of nonidentifiable parameters, which should be assumed to be known to obtain an identifiable system:

```
In[9]:= iad["DegreesOfFreedom"]

Out[9]= 1
```

IdentifiabilityAnalysis – A *Mathematica* Package

We can resolve the problem by fixing one of the involved parameters, e.g., the dose $\mathcal{D}$:

```
In[10]:= iad = IdentifiabilityAnalysis[
              {deq/.L0->𝒟/Vc,L[t]}/.𝒟->11,{L,T,R,RL},params,t]

Out[10]= IdentifiabilityAnalysis[True,<>]
```

or the Vc value:

```
In[11]:= iad = IdentifiabilityAnalysis[
              {deq/.L0->𝒟/Vc,L[t]}/.Vc->1/20,{L,T,R,RL},params,t]

Out[11]= IdentifiabilityAnalysis[True,<>]
```

Timings of a Few Selected Models

Model	State	Params	Output	Timing
NF-kB (Lipniacki *et al.*, 2004)	16	28	5	2 min
NF-kB (Lipniacki *et al.*, 2004)	16	28	16	4 s
JAK-STAT (Yamada *et al.*, 2003)	31	36	2	20 min
JAK-STAT (Yamada *et al.*, 2003)	31	36	31	16 s
The Ras-pathway (Wolf *et al.*, 2007)	58	113	1	2 hours
The Ras-pathway (Wolf *et al.*, 2007)	58	113	58	36 min
A MAPK cascade (Schoeberl *et al.*, 2002)	102	99	1	2 days
A MAPK cascade (Schoeberl *et al.*, 2002)	102	99	102	39 min

Conclusions

- Local algebraic identifiability analysis can be perfomed *fast*
- Testing generic rank of a Jacobian matrix, J
- Compute a numeric specialization of J
  - Sensitivity equations, specialization, power series solutions
- Rank computations
- All computations are done modulus a large prime
- Structural identifiability analysis for large systems on standard desktop PC
- IdentifiabilityAnalysis – a *Mathematica* application package

Practical Identifiability Analysis

Mats Jirstrand

Outline

Practical Identifiability Analysis

- Forward look: likelihood based confidence intervals
- Practical Identifiability
- The Profile Likelihood
- Examples
  - Target-Mediated Drug Disposition
- Conclusions

Practical Identifiability

Can the parameters of a set of differential equations be estimated (including measures of finite uncertainty) given knowledge of the input and a given set of variables assumed to be measurable (with noise)?

- Non-practically identifiable parameters may be the result of a *poorly perturbed* system, i.e., a change in design of the input could resolve this situation. *Too few data points* and the *level of measurement noise* are other factors affecting the parameter identifiability.
- ‘Practical' denotes a property of the combination of the model equations and information in the data at hand.

We have implemented a method (Venzon&Moolgavkar 1988) based on likelihood based confidence intervals in an easy to use *Mathematica* package.

The Profile Likelihood

- Assume that we have additive measurement noise

$$dx_t/dt = f(x_t, u_t, \theta), \quad x_{t_0} = x_0(\theta)$$
$$y_k = h(x_k, \theta) + v_k, \ k = 1, \ldots, N$$

where $v_k \sim N(0,\sigma^2)$ independent Gaussian distributed random variables.

- A predictor is given by the expected value of the observation

$$d\widehat{x}_t/dt = f(\widehat{x}_t, u_t, \theta), \quad x_{t_0} = x_0(\theta)$$
$$\widehat{y}_k(\theta) = h(\widehat{x}_k, \theta), \ k = 1, \ldots, N$$

- The maximum likelihood method: maximize the probability of the observed event/data w.r.t. θ.

$$y_k \sim N(\widehat{y}_k, \sigma^2)$$

The Profile Likelihood

The joint probability distribution

y_k are all independent

$$p(y_1, \ldots, y_N|\theta) = p(y_1|\theta) \cdot \ldots \cdot p(y_N|\theta) =$$

$$\prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_k - \hat{y}_k(\theta))^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\sum_{k=1}^{N} \frac{(y_k - \hat{y}_k(\theta))^2}{2\sigma^2}}$$

Taking the negative logarithm: $\sum_{k=1}^{N} \frac{(y_k - \hat{y}_k(\theta))^2}{2\sigma^2} + const.$

$$-2LL(\theta; \mathcal{Y}_N) = \sum_{i=1}^{N} \frac{(y_k - \hat{y}_k(\theta))^2}{\sigma^2} + const. = RSS(\theta) + const.$$

Fraunhofer ITWM
CHALMERS
Research Centre
Industrial Mathematics

Minimizing the sum-of-squares of the *simulation* errors!

The Profile Likelihood

- The maximum likelihood estimate

$$\widehat{\theta}_{ML} = \arg\max_{\theta} LL(\theta; \mathcal{Y}_N) = \arg\min_{\theta} RSS(\theta)$$

- The profile likelihood

$$PL(\theta_i) = \max_{\theta_{j \neq i}} LL(\theta; \mathcal{Y}_N)$$

- A projection of the multivariate log-likelihood onto p-θ_i planes (a univariate entity) – easy to visualize!

Fraunhofer ITWM
CHALMERS
Research Centre
Industrial Mathematics

Profile Likelihood

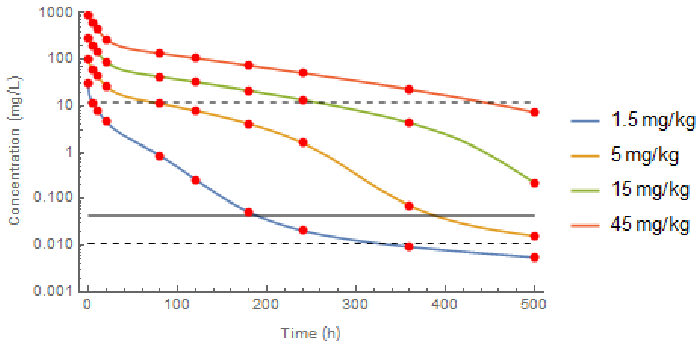
- Profile likelihood based confidence intervals

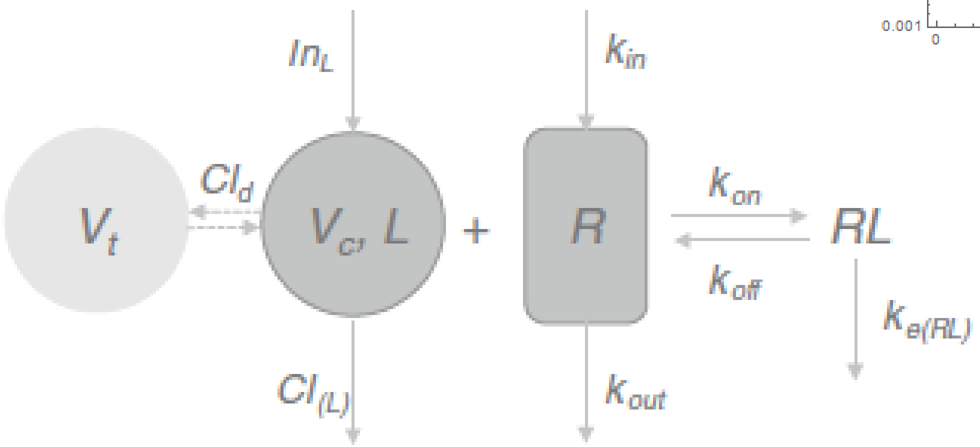
$$CI_\alpha(\theta_i;\mathcal{Y}_N) = \{\theta_i | -2PL(\theta_i) \leq -LL(\mathcal{Y}_N)^* + icdf(\chi_1^2,\alpha)\}$$

- The coverage

$$Prob(\theta_i \in CI_\alpha(\theta_i;\mathcal{Y}_N)) = \alpha$$

A Target-Mediated Drug Disposition Model (TMDD)





L = Ligand

R = receptor

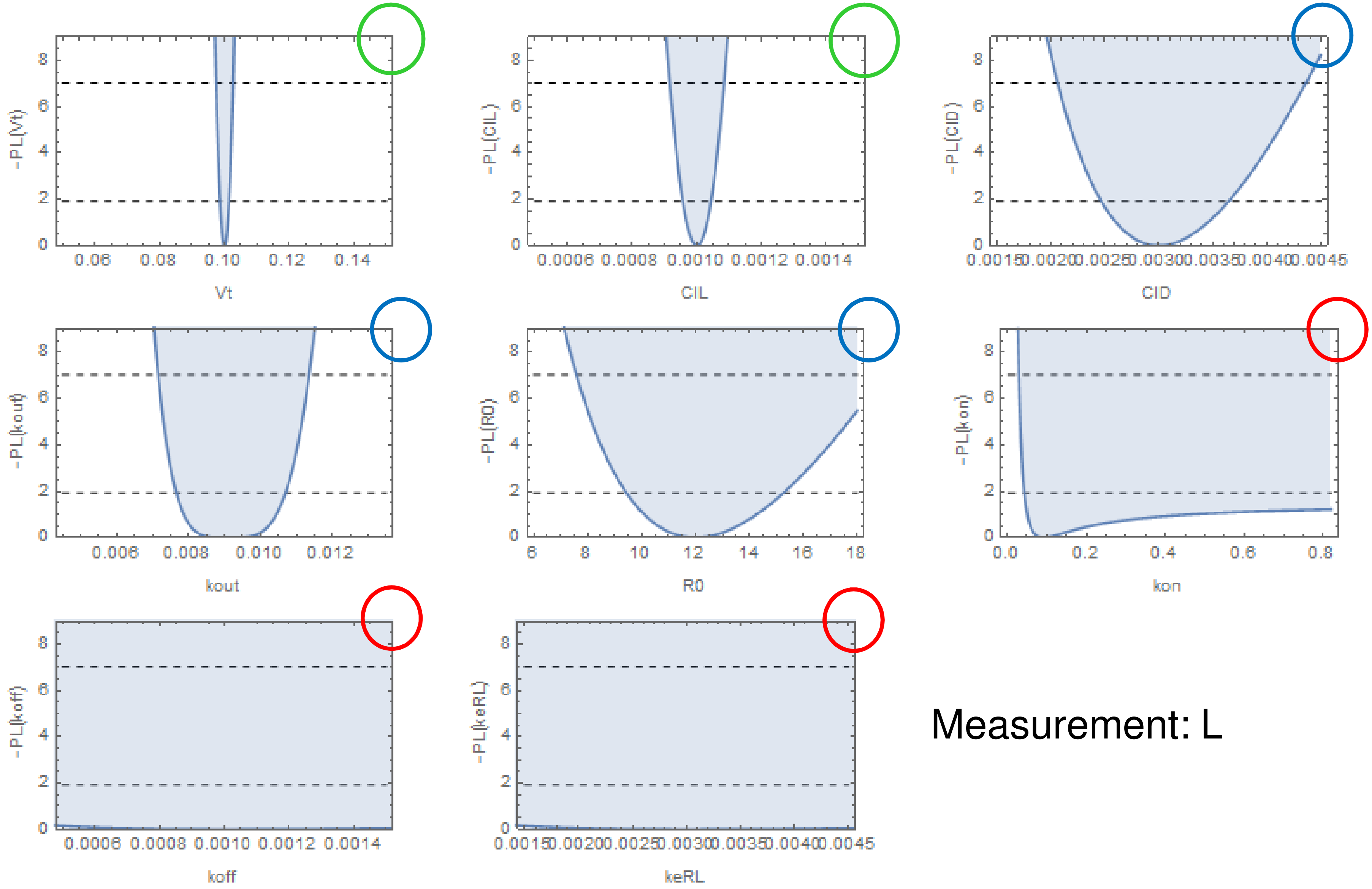
RL = Receptor-Ligand complex

Mats Jirstrand

Fraunhofer ITWM

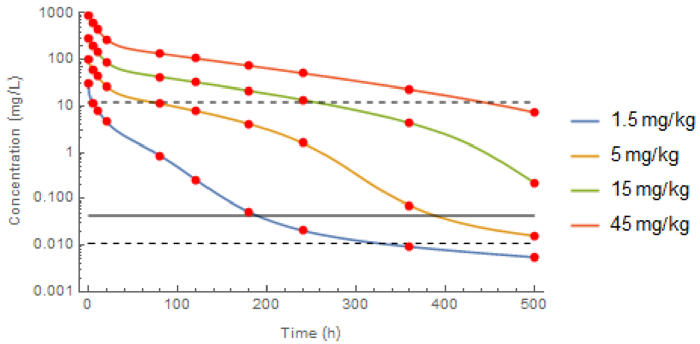
CHALMERS
Research Centre
Industrial Mathematics

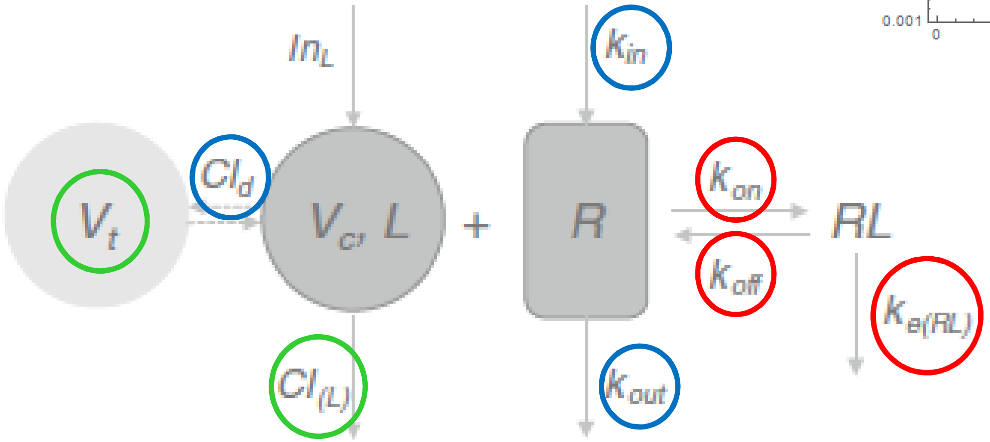
Profile Likelihood – TMDD



Measurement: L

A Target-Mediated Drug Disposition Model (TMDD)





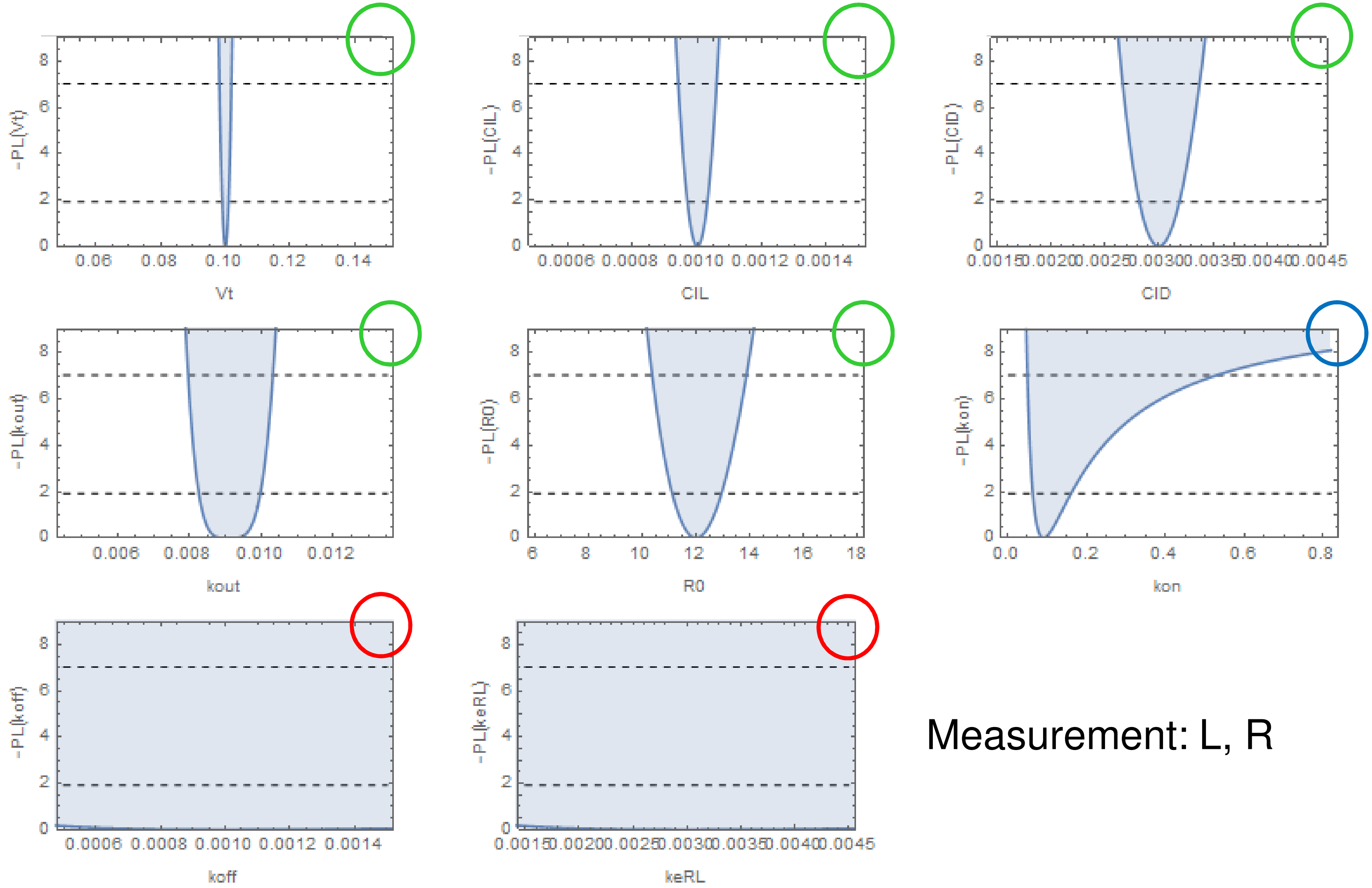
L = Ligand

R = receptor

RL = Receptor-Ligand complex

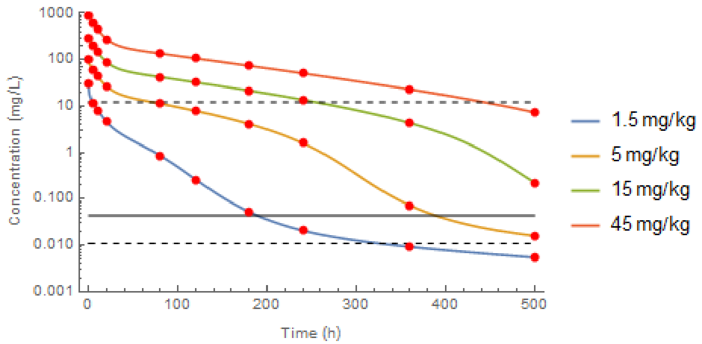
Measurement: L

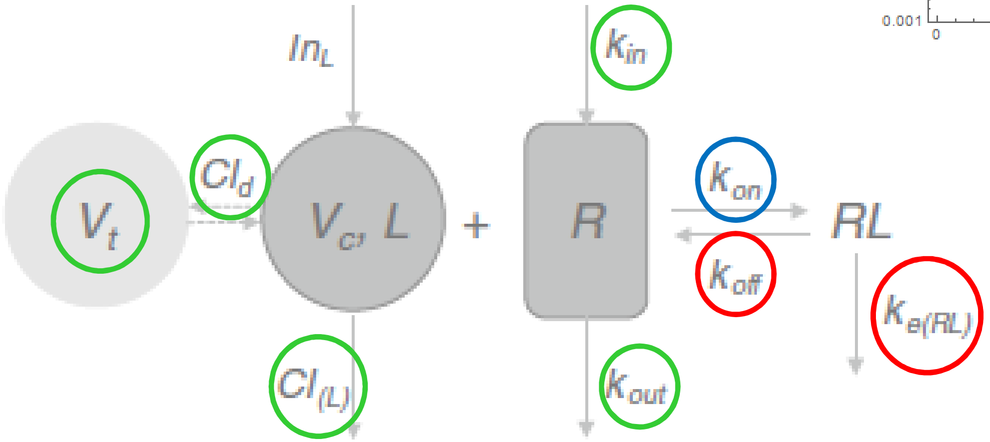
Profile Likelihood – TMDD



Measurement: L, R

A Target-Mediated Drug Disposition Model (TMDD)





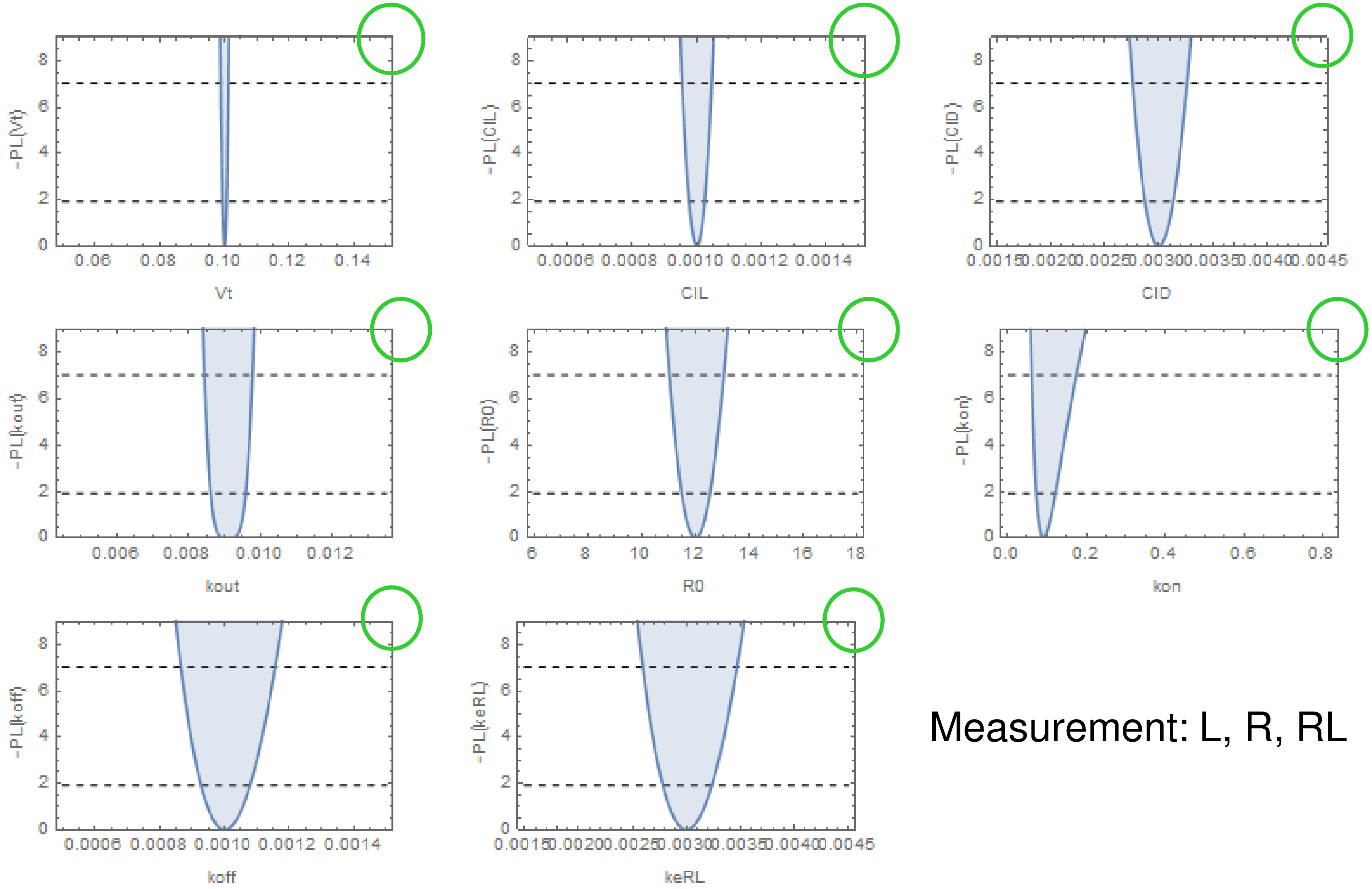
L = Ligand

R = receptor

RL = Receptor-Ligand complex

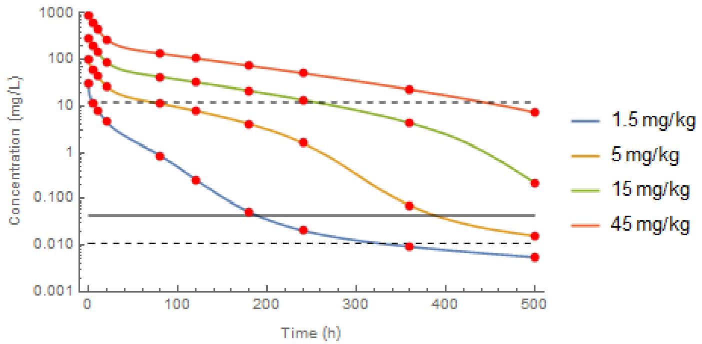
Measurement: L, R

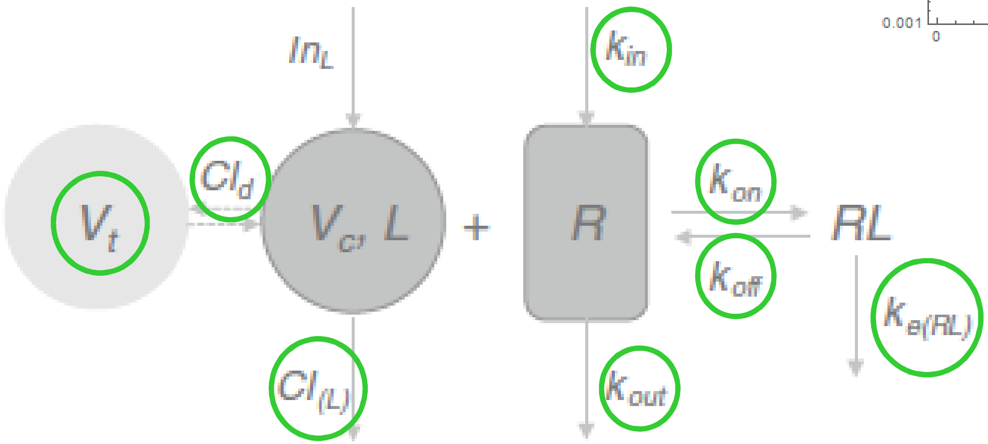
Profile Likelihood – TMDD



Measurement: L, R, RL

A Target-Mediated Drug Disposition Model (TMDD)





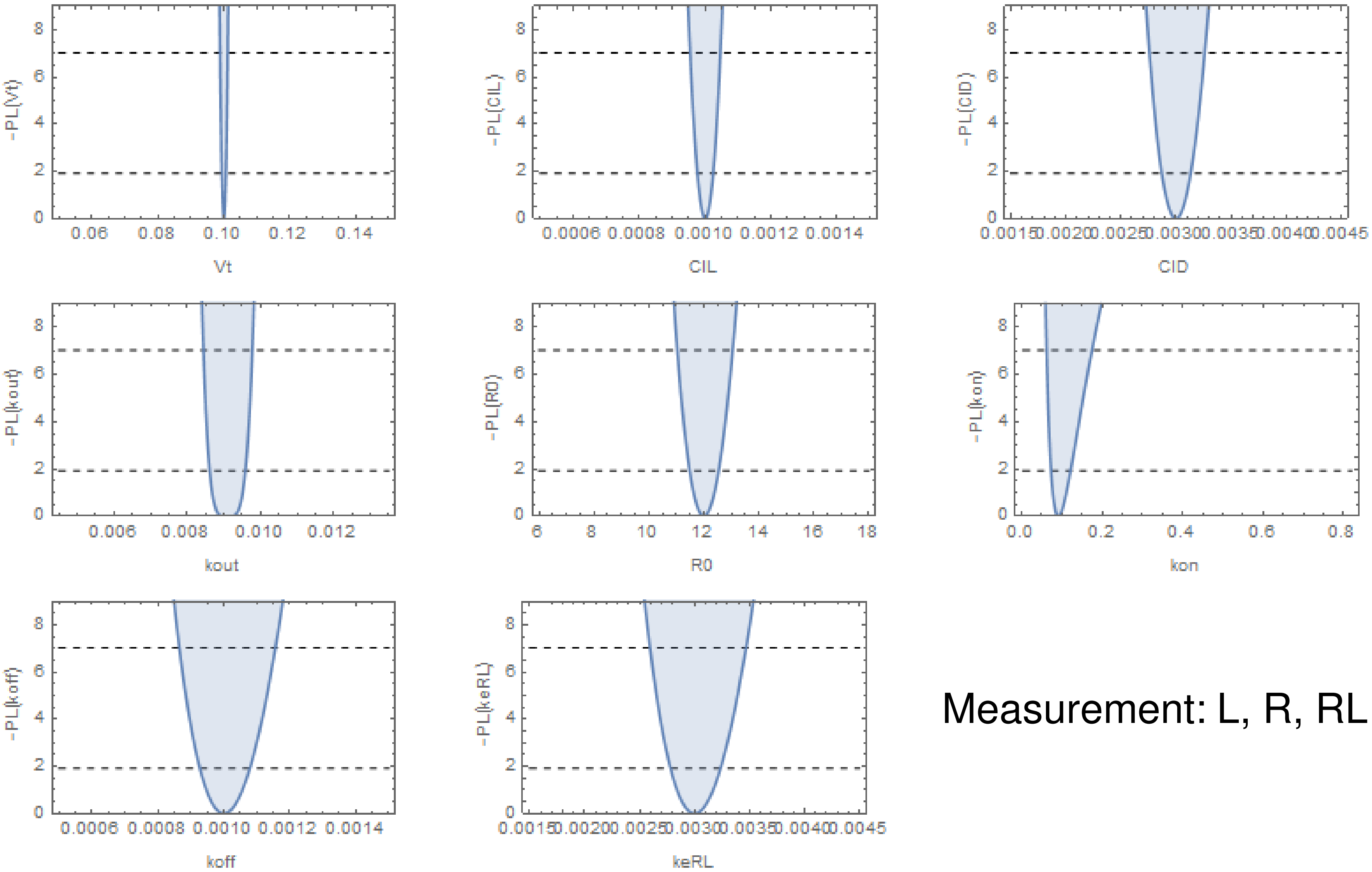
L = Ligand

R = receptor

RL = Receptor-Ligand complex

Measurement: L, R, RL

Profile Likelihood – TMDD



Measurement: L, R, RL

Conclusions

- Quantitative Systems Pharmacology
  - Non-standard model structures (unexplored parametrizations)
  - What is feasible/non-feasible to estimate?
  - Much time in front of tools-for-regression can be saved
- Structural Identifiability Analysis
  - Fast, Yes/No, "NonIdentifiableParameters", "DegreesOfFreedom"
- Practical Identifiability Analysis
  - Local (around point – initial guess)
  - Visual tool

Acknowledgements

Alexandre Sedoglavic, LIFL, University Lille

Johan Gabrielsson, Swedish University of Agricultural Sciences

unicellsys

Eukaryotic unicellular organism biology – systems biology of the control of cell growth and proliferation

CancerSys

Mathematical modeling of β-catenin and RAS Signaling in liver and its impact on proliferation, tissue organization and formation of hepatocellular carcinomas.

WEB:
www.fcc.chalmers.se/software/other-software/identifiabilityanalysis

Comparison of approaches for parameter identifiability analysis of biological systems. A. Raue, J. Karlsson, M.P. Saccomani, M. Jirstrand, J. Timmer. Bioinformatics, 2014 May;30(10):1440-8.

An Efficient Method for Structural Identifiability Analysis of Large Dynamic Systems System Identification. Karlsson, M. Anguelova, and M. Jirstrand. 16th IFAC Symposium on System Identification (SYSID 2012),16:1, 2012, Brussels, Belgium, July 11-13, 2012.

Fraunhofer ITWM

CHALMERS
Research Centre
Industrial Mathematics