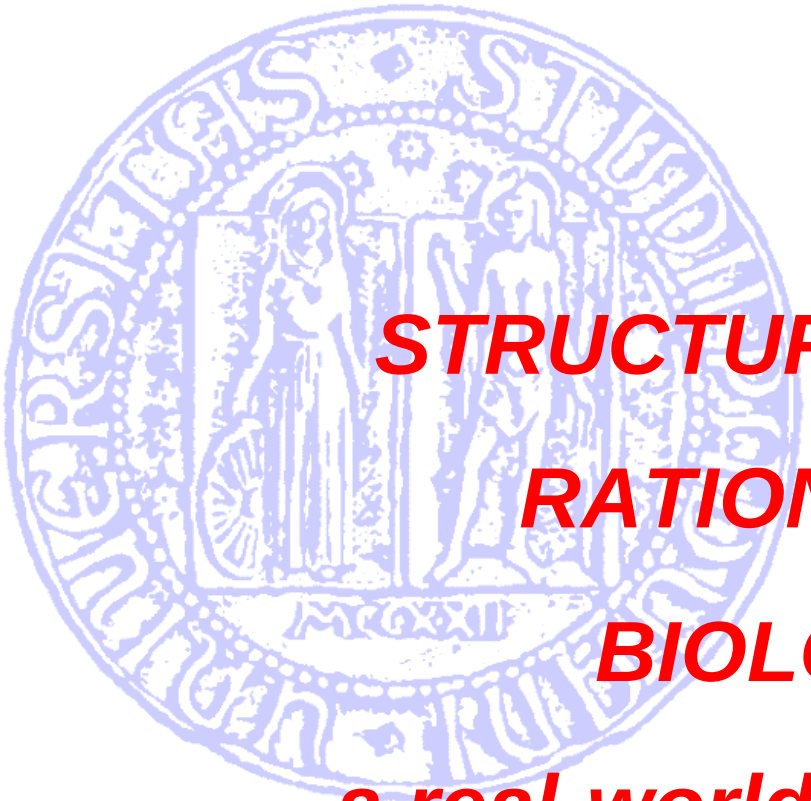
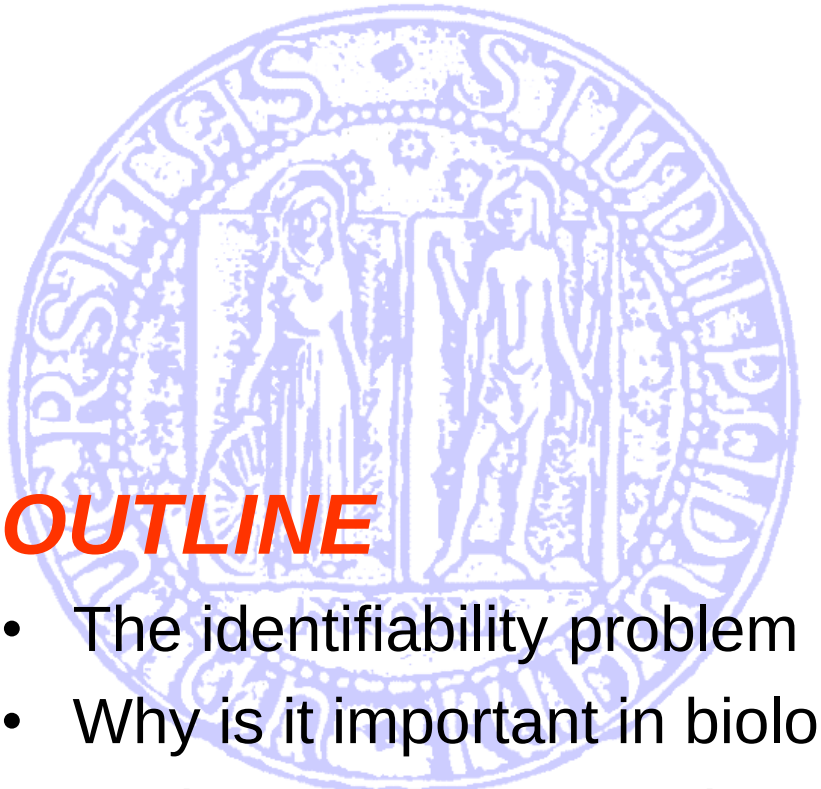


STRUCTURAL IDENTIFIABILITY OF RATIONAL ODE MODELS IN BIOLOGICAL SYSTEMS: a real-world application of differential algebra

M.P. Saccomani
Dept. Information Engineering, Padova, Italy



OUTLINE

- The identifiability problem
- Why is it important in biological and biomedical studies?
- Mathematical formulation
- A differential algebra identifiability method for rational dynamic models

  -Identifiability of models with given initial conditions
- Checking *structural* identifiability in combination with *practical* id.
- Case studies
- Conclusions

THE CONTEXT OF THE PROBLEM

- Complex dynamics, non-linear interaction mechanisms in cellular processes are modeled according to physico-chemical laws.
- Ordinary differential equations (ODE) involving parameters such as reaction rates are commonly used.

- In general ODE parameters can only be measured indirectly. Their recovery can then only be approached indirectly as a parameter estimation problem starting from external input-output measurements.

- *Structural Identifiability* is the first step in model identification, necessary to correctly solve the parameter estimation problem from the experimental data.

MATHEMATICAL FORMULATION

The system-experiment model

$$\begin{cases} \dot{x}(t) = f[x(t), p, u(t)] \\ y(t) = h[x(t), p] \end{cases}$$

Initial conditions:

$$x(t_0) = x_0$$

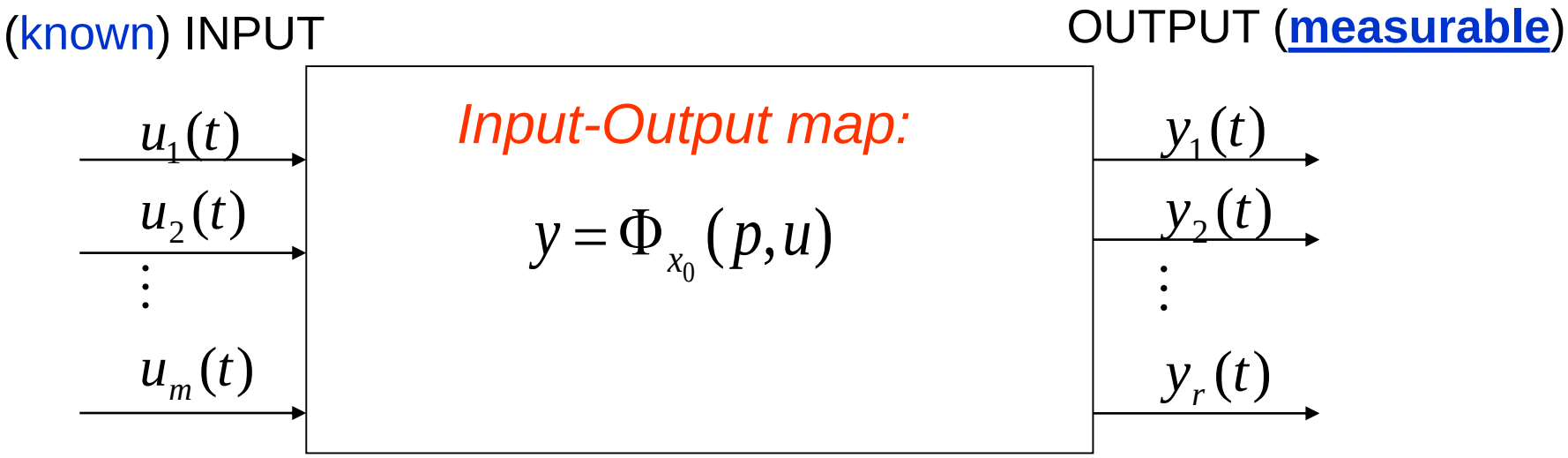
where

- x is the n-dimensional state variable,
- u is the m-dimensional input function
- y is the r-dimensional output function
- p is the constant unknown k-dimensional parameter vector
- f and h polynomial or rational functions (i.e. saturation process, Michaelis-Menten kinetics), for the time being.

NOTE: Our ident. method can be applied to study some **nonpolynomial** equations models, e.g. exponential models.

STRUCTURAL IDENTIFIABILITY PROBLEM

HYPOTHESIS: noise-free data.



PROBLEM:

Given u and y, how many parameter values *p* satisfy the I-O map?

- one (***Global identifiability***)
- > one
  - p has a finite number of solutions (***Local Id***.)
  - p has an infinite number of solutions (***Non Id***.)

DEFINITIONS

$$y = h[x(t, u, p, x_0), p] := \phi_{x_0}(p, u)$$ *Input-Output Map*

DEFINITION 1

– The system is *structurally globally (or uniquely) identifiable from input-output data* if, for at least a generic set of points $p^* \in P$, there exists at least one input function u such that

$$\phi_{x_0}(p, u) = \phi_{x_0}(p^*, u) \qquad (1)$$

has only one solution $p = p^*$ for all x_0 in a generic subset of R^n.

DEFINITION 2

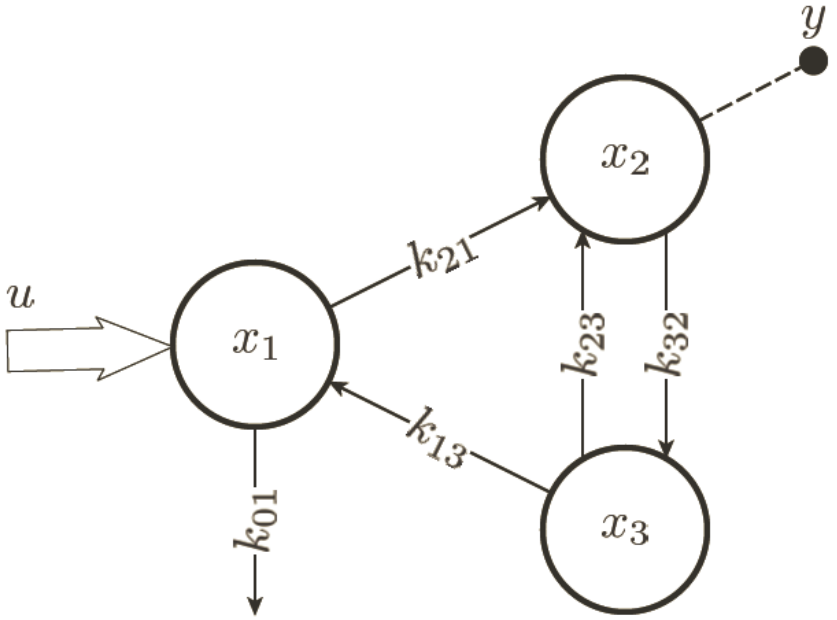
The system is *structurally locally identifiable from input-output data at* $p^* \in P$, if there exists at least one input function u *and* an open neighborhood U_{p^*} of p^*, such that eq. (1) has a unique solution $p^* \in U_{p^*}$ for all x_0 in a generic subset of $\in R^n$.

For a system which is not even locally identifiable, equation (1) has generically an infinite number of solutions for at least one input u. This is called *non-identifiability*.

Why check structural id. of biological models?

A simple locally identifiable model:

$$\begin{cases} \dot{x}_1(t) = -(k_{01} + k_{21})x_1(t) + k_{13}x_3(t) + u(t) \\ \dot{x}_2(t) = k_{21}x_1(t) - k_{32}x_2(t) + k_{23}x_3(t) \\ \dot{x}_3(t) = k_{32}x_2(t) - (k_{13}x_2 + k_{23})x_3(t) \\ y(t) = x_2(t) \end{cases}$$

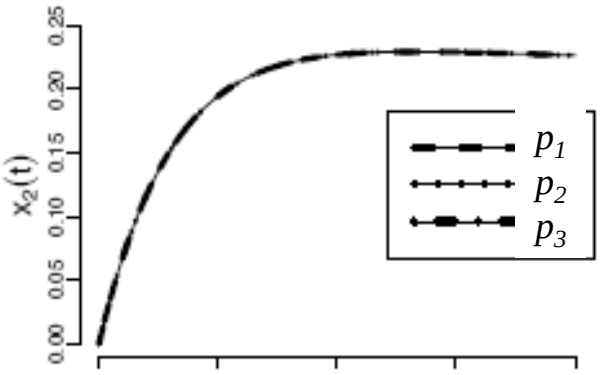


The unknown parameter vector is

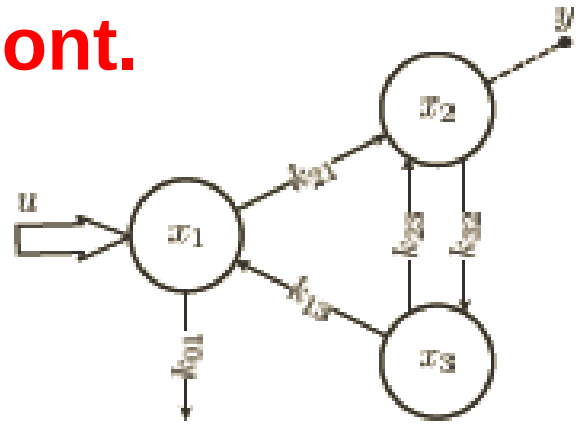
$p = [k_{01}, k_{21}, k_{13}, k_{32}, k_{23}]$.

Differential algebra identifiability results:

the model has 3 parameter solutions equivalently describing the I/O experimental data.

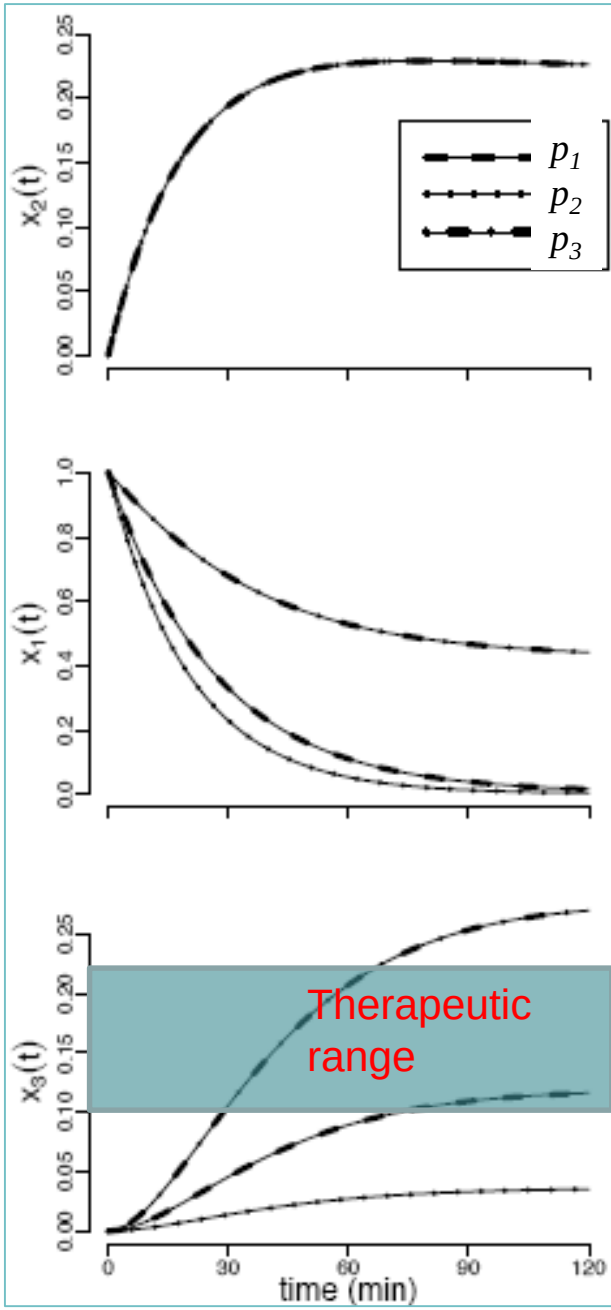


cont.



State trajectories of the two unobservable compartments determined for the three locally identifiable parameterization.

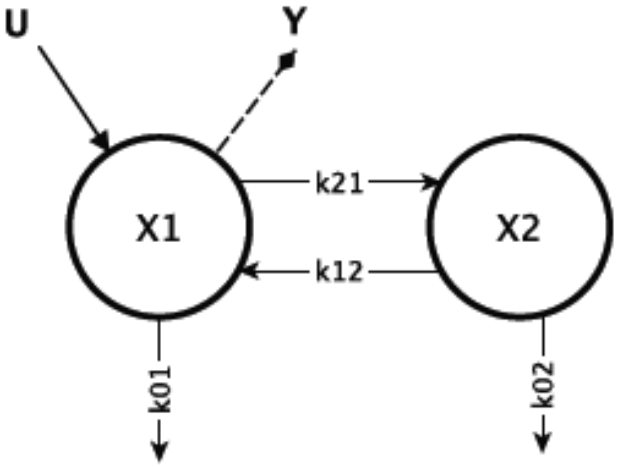
For example, this can lead to an erroneous therapeutic decision.



cont.

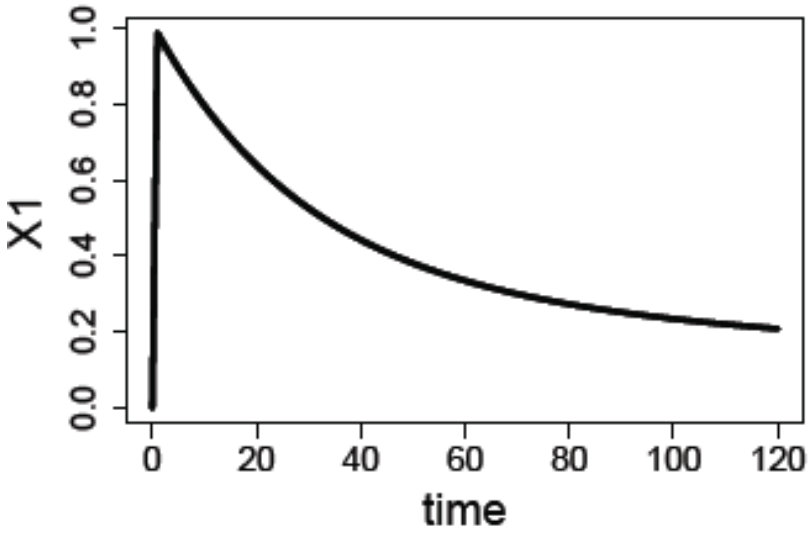
A simple non-identifiable model:

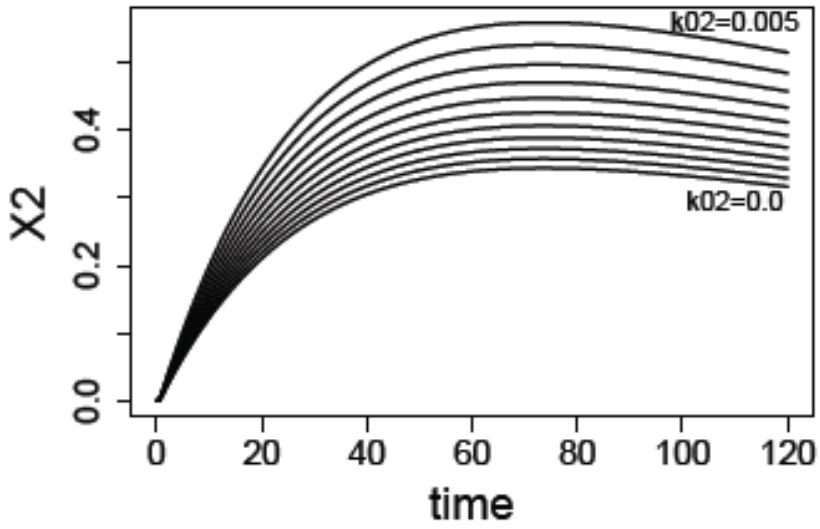
$$\begin{cases} \dot{x}_1(t) = -(k_{01} + k_{21})x_1(t) + k_{12}x_2(t) + u(t) & x_1(0) = x_{10} \\ \dot{x}_2(t) = k_{21}x_1(t) - (k_{02} + k_{12})x_2(t) & x_2(0) = 0 \\ y(t) = x_1(t) \end{cases}$$



The unknown parameter vector is

$$p = [k_{01}, k_{21}, k_{12}, k_{02}].$$





For example, this can lead to an erroneous therapeutic decision.

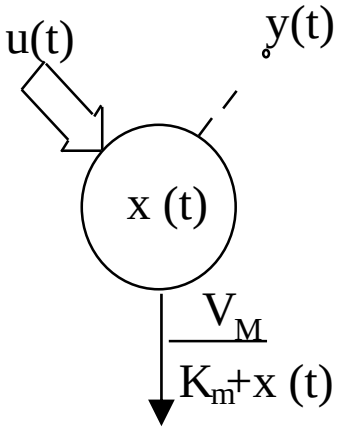
STATE OF THE ART FOR NONLINEAR MODELS

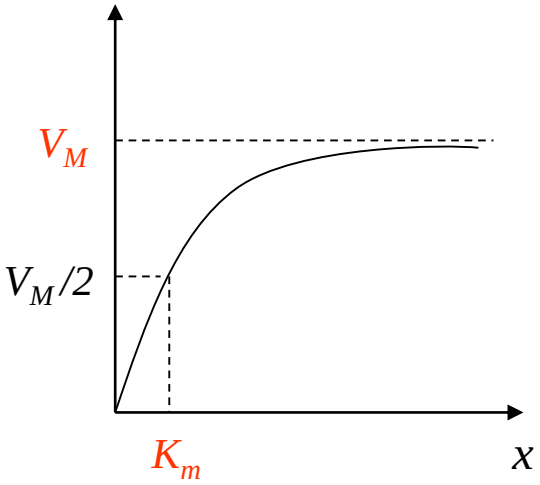
STRUCTURAL IDENTIFIABILITY

- Taylor series expansion of y(p,t) (Pohjampalo 1978)
- Generating Series (Walter et al. 1982)
- Similarity transformation approach (Godfrey et al. 1989).
- **Differential algebra** methods (Ollivier 1990, Ljung et al. 1994, Joly-Blanchard et al. 1998, Margaria et al. 2001, IEEE Trans. Biomed. Eng. 2001).
  - Identifiability of nonlinear models from given **initial conditions** (Automatica 2003).
- Structural identifiability software
  - DAISY (Comp. Meth. and Progr. in Biomedicine 2007)
  - GenSSI (Chis et al. 2011)
  - EAR (Anguelova, 2012)
  - COMBOS (Meshkat et al. 2014)
  - SIAN (Hong et al. 2019 to appear)

EXAMPLE: A SIMPLE NONLINEAR MODEL (1) (Michaelis-Menten kinetics)

$$\begin{cases} \dot{x}(t) = -\dfrac{V_M\, x(t)}{K_m + x(t)} + u(t) \qquad x(0) = d \\ y(t) = x(t)/V \end{cases}$$





where $p = (K_m, V_M, V)$

K_m and V_m are the Michaelis-Menten parameters,

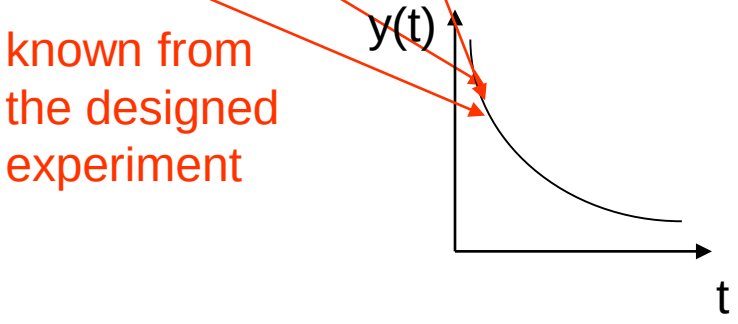
V denotes the distribution volume of the accessible pool.

EXAMPLE: A SIMPLE NONLINEAR MODEL (2)

(Michaelis-Menten kinetics)

$$\begin{cases} \dot{x}(t) = -\dfrac{V_M x(t)}{K_m + x(t)} + u(t) \qquad x(0) = d \\ y(t) = x(t)/V \end{cases}$$

Taylor series expansion: $y(t) = y(0) + t\,\dot{y}(0) + \dfrac{t^2}{2}\ddot{y}(0) + \ldots$



$$y(0) = \frac{x_1(0)}{V} = \frac{d}{V} \quad \Rightarrow \quad V \text{ has one solution}$$

Cont.

$$\dot{y}(0)=\frac{\dot{x}_1(0)}{V}=-\frac{1}{V}\frac{V_M x_1(0)}{K_m+x_1(0)}+u(0)=\frac{1}{V}\frac{V_M d}{K_m+d}+u(0)$$

$$\ddot{y}(0)=\frac{\ddot{x}_1(0)}{V}=-\frac{1}{V}\frac{V_M \dot{x}_1(0)[K_m+x_1(0)]-V_M x_1(0)\dot{x}_1(0)}{(K_m+x_1(0))^2}+\dot{u}(0)=$$

$$-\frac{1}{V}\frac{K_m V_M \dot{x}_1(0)}{(K_m+x_1(0))^2}+\dot{u}(0)=-\frac{K_m V_M \dot{y}(0)}{(K_m+d)^2}+\dot{u}(0)$$

$$\dddot{y}(0)=\frac{\dddot{x}_1(0)}{V}=-\frac{1}{V}\frac{V_M K_m \ddot{x}_1(0)[K_m+x_1(0)]^2-2V_M K_m \dot{x}_1^2(0)[K_m+x_1(0)]}{(K_m+x_1(0))^4}+\ddot{u}(0)=$$

$$-\frac{K_m V_M[\ddot{y}(0)(K_m+d)-2\dot{y}(0)^2]}{(K_m+d)^3}+\ddot{u}(0)$$

K_m, V_M, V are globally identifiable.

ATTENTION: only a sufficient condition!

HOW DO WE CHECK IDENTIFIABILITY?

1. Compute the I/O map of the system (in its implicit form).

   This is formed by a set of differential polynomials in the variables

   $u, \dot{u}, \ddot{u}, \ldots, y, \dot{y}, \ddot{y}, \ldots$

HOW? By elimination of *x*.

TOOL (from differential algebra): *Ritt Algorithm (1950)* to calculate the *characteristic set* of the ideal generated by the polynomials defining our dynamic model.

2. Check that the parameterisation of the I/O relation(s) is injective.

HOW?

By calculating the *Gröbner basis* of the algebraic system to be solved.

TOOL *(*from computer algebra*): Buchberger Algorithm.*

DIFFERENTIAL ALGEBRA

Differential polynomial ring: $K[z_1,\ldots,z_n]$

- Ranking among the variables and their derivatives
- Ranking among the polynomials A_i, $a_j \in K[z_1,\ldots,z_n]$

DEFINITIONS

- The *leader* u_j of a polynomial A_j is the highest ranking derivative of the variables appearing in A_j.
- The order of the leader is its maximum order of derivation.
- A_j is *of lower rank* than A_i if $u_j<u_i$ or, whenever $u_j=u_i$ and $deg(u_j)\leq deg(u_i)$.
- A_i is reduced with respect to A_j if does not contain neither the leader of A_j with equal or greater algebraic degree, nor its derivatives.
- If A_i is not reduced w.r.t. to A_j it can be reduced by using the pseudodivision algorithm among polynomials (suggested by Ritt, 1950).
- A set of differential polynomials $A=\{A_1, A_2,\ldots, A_r\}$ that are all reduced with respect to each other, is called an *autoreduced set*.
- A lowest rank autoreduced set that can be formed with polynomials from a given set S of differential polynomials, is called a ***characteristic set*** of S.

DIFFERENTIAL ALGEBRA AND IDENTIFIABILITY

The differential polynomials

$$\begin{cases} \dot{x} - f(x, p, u) \\ y - h(x, p) \end{cases}$$

generate a (*prime*) *differential ideal I* in the ring R(**p**)[**u**,**y**,**x**].

Ranking: $u < \dot{u} < \ddot{u} < \ldots < y < \dot{y} < \ddot{y} < \ldots < x_1 < x_2 < \ldots < x_n < \dot{x}_1 < \dot{x}_2 < \ldots < \dot{x}_n < \ldots$

Characteristic set (Ritt algorithm, 1950)

For simplicity scalar output:

$A_0(p,u,y)$ → ***INPUT-OUTPUT RELATION***

$A_1(p,u,y,x_1)$

$A_2(p,u,y,x_1,x_2)$

$\vdots$

$A_n(p,u,y,x_1,x_2,\ldots,x_n)$

Differential polynomial obtained after elimination of the state variables, hence represents exactly the pairs ***(u,y)*** described by the original system.

Cont.

- The *input-output relation* is:

$$A_0(p,u,y) = \sum_{i=1}^{\upsilon} c_i(p)\, f_i(u,\dot{u},\ldots,y,\dot{y},\ldots) \qquad (1)$$

polynomials in *p* — known monomials in I/O variables

- The *exhaustive summary* (Ollivier) of the model is:

$$c_i(p) \qquad i = 1,\ldots,\nu \qquad (2)$$

To test *global identifiability* check the unique solvability (***injectivity***) of the algebraic nonlinear system:

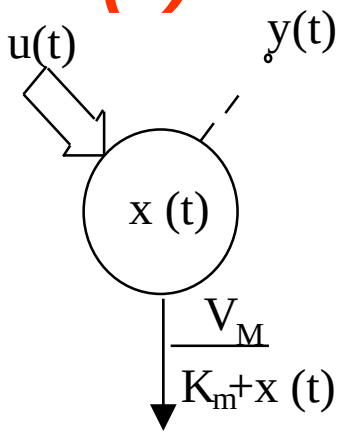
$$c_i(p) = c_i^* \qquad i = 1,\ldots,\nu \qquad (3)$$

Cont.

- We use the *Buchberger algorithm*, a computer algebra algorithm for computing with multivariate polynomials $K[p_1,p_2,\ldots,p_k]$
- Generalizes well-known algorithms:
  - Gaussian elimination
  - Euclidean algorithm
- Calculates the reduced Gröbner basis of system (2)
- A *Gröbner basis* for an ideal *I* is a set of generator for *I* having certain property with respect to an ordering < on the monomials.
- The *reduced Gröbner basis* minimally generates *I*.
- If the term order < is fixed , then every ideal *I* in $K[p_1,p_2,\ldots,p_k]$ has a unique reduced *Gröbner basis*.

EXAMPLE: A SIMPLE NONLINEAR MODEL (3)

$$\begin{cases} \dot{x} + \dfrac{V_M}{K_m + x} x - u \\ y - x/V \end{cases}$$



Standard ranking: $u < \dot{u} < \ddot{u} < y < \dot{y} < \ddot{y} < x < \dot{x} < \ldots$

The *characteristic set* is:

$$\begin{cases} \dot{y}\, y + \underbrace{\dfrac{K_m}{V}}_{c_1} \dot{y} \underbrace{- \dfrac{1}{V}}_{c_2} u\, y \underbrace{- \dfrac{K_m}{V^2}}_{c_3} u + \underbrace{\dfrac{V_M}{V}}_{c_4} y \\ -x + yV \end{cases}$$

(normalized) input-output relation

Checking the ***injectivity*** of the exhaustive summary leads to a system of nonlinear algebraic equations in the unknown K_m, V_M and V (Buchberger algorithm to calculate the reduced Gröbner basis).

$$\begin{cases} V_M = c_4 / c_2 \\ K_m = c_1 / c_2 \\ V = 1/c_2 \end{cases} \Longrightarrow$$

One solution (*global id.*)

IDENTIFIABILITY WITH INITIAL STATE CONDITIONS (1/2)

HYPOTHESIS: Algebraic Observability

DEFINITION: A state component is *algebraically observable* if its derivative does not appear in the last n equations of the characteristic set.

$$A_0(p,u,y)$$
$$A_1(p,u,y,x_1)$$
$$A_2(p,u,y,x_1,x_2)$$
$$\vdots$$
$$A_n(p,u,y,x_1,x_2,\ldots,x_n)$$

These last n polynomials are evaluated at time t = 0.

The identifiability test with given initial conditions is based on the exhaustive summary *c(p)* together with polynomials:

$$A_1(p,u(0),y(0),x_1(0))$$
$$A_2(p,u(0),y(0),x_1(0),x_2(0))$$
$$\vdots$$
$$A_n(p,u(0),y(0),x_1(0),x_2(0),\ldots,x_n(0))$$

(*)

"Augmented" exhaustive summary

IDENTIFIABILITY WITH INITIAL STATE CONDITIONS (1/2)

1. Known initial conditions

Polynomials (*) are in the unknown *p* with coefficients which are monomials in the known data $x(0), u(0), \dot{u}(0), ..., y(0), \dot{y}(0), ...$

2. Some information on initial conditions

Polynomials (*) are in the unknowns *p* and x_0.

In both cases:

1. the corresponding equations of the augmented exhaustive summary is solved by the Buchberger algorithm
2. the new reduced Gröbner basis provides the parameter identifiability results from input-state-output data.

A HIV model (Wodarz et al., 2002)

$$\dot{x} = \lambda - d\,x - \beta\,x\,v,$$

$$y\dot{y} = \beta\,x\,v - a\,yy,$$

$$\dot{v} = k\,yy - uu\,v,$$

$$\dot{w} = c\,x\,yy\,w - c\,q\,yy\,w - b\,w,$$

$$\dot{z} = c\,q\,yy\,w - h\,z,$$

$$y_1 = w,$$

$$y_2 = z$$

$\boldsymbol{x}=[x, yy, v, w, z]$ state variables;

y_1 and y_2 measured outputs;

$\boldsymbol{p}=[\beta, \lambda, a, b, c, d, h, k, q, uu]$ unknown parameters;

$\boldsymbol{x}(0)=[\underbrace{x^*, yy^*, v^*}_{\text{unknown}}, \underbrace{w^*, z^*}_{\text{known}}]$ initial conditions.

A chemical reaction model (Conradi et al., 2018)

$$\dot{x}_1 = -k_1 x_1 x_2 + k_2 x_4 + k_4 x_6,$$

$$\dot{x}_2 = k_1 x_1 x_2 + k_2 x_4 + k_3 x_4,$$

$$\dot{x}_3 = k_3 x_4 + k_5 x_6 - k_6 x_3 x_5,$$

$$\dot{x}_4 = k_1 x_1 x_2 - k_2 x_4 - k_3 x_4,$$

$$x_5 = k_4 x_6 + k_5 x_6 - k_6 x_3 x_5,$$

$$\dot{x}_6 = -k_4 x_6 - k_5 x_6 - k_6 x_3 x_5,$$

$$y_1 = x_2,$$

$$y_2 = x_3$$

where $\boldsymbol{x}=[x_1, x_2, x_3, x_4, x_5, x_6]$ state variables;

y_1 and y_2 measured outputs;

$\boldsymbol{p}=[k_1, k_2, k_3, k_4, k_5, k_6]$ unknown parameters;

$\boldsymbol{x}(0)=[x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*]$ initial conditions

(x_2^* and x_3^* known from the output functions).

PROBLEM

The input-state-output identifiability method may not work when the system is started at “special” initial conditions.

OSS: This happens whenever method is used to check identifiability

IMPORTANT STRUCTURAL PROPERTIES

MINIMALITY

OBSERVABILITY: if the model is non algebraically observable, the known initial conditions cannot be used in the identifiability test.

ACCESSIBILITY: Accessibility may fail for many biosystems starting from known but “non generic” initial conditions

DEFINITION: A system is *accessible from* x_0 if, for suitable input $u(t)$, the state $x(t)$ can reach an open set of full dimension of the state space.

ACCESSIBILITY RANK CONDITION

For analytic, in particular polynomial, models

$$\begin{cases} \dot{x} = f[x(t), p] + \sum_{i=1}^{m} g_i[x(t), p]u_i(t) \\ y(t) = h[x(t), u(t), p] \end{cases} \qquad x(t) \in R^n$$

a necessary and sufficient condition for *accessibility from* x_0 is

$$\dim \Delta_C(x_0) = n$$

where Δ_C is the distribution associated to the Control Lie Algebra, i.e. the smallest Lie algebra C containing the vector fields $f, g_1, \ldots, g_m$ and invariant under Lie bracketing with $f, g_1, \ldots, g_m$.

IDENTIFIABILITY AND ACCESSIBILITY (1/2)

THEOREM 1. If the system is accessible from all initial states x_0 from which it may have been started, then the characteristic set does not change and provides the correct identifiability test.

THEOREM 2. If the system is accessible except from a “thin set" T of measure zero and $x_0 \in T$.

Let $\{\phi_k(x) = 0;\ k=1,\ldots,d\}$ be the set of algebraic equations defining the invariant manifold T containing x_0.

Then the characteristic set can change.

The correct answer is given by the identifiability test applied to the original model equations plus equations $\phi_k(x) = 0$.

EXAMPLE

$$\begin{cases} \dot{x}_1 = -p_0 u - p_2 x_1 - p_3 x_2 & x_1(0) = x_{10} \\ \dot{x}_2 = p_3 x_1 x_2 - p_1 x_1 & x_2(0) = x_{20} \\ y = x_1 \end{cases}$$

where p_0, p_1, p_2, p_3, are the unknown parameters.

The system is generically accessible, i.e. $\dim \Delta_C(x_0) = 2 \ \textit{for all} \ x_0 \notin T$

where T is the invariant manifold defined by $\phi(x) = \{p_3 x_2 - p_1 = 0\}$

Thus, if $x_0 \notin T$, the characteristic set of the ideal generated by the polynomials defining the system shows that the model is glob. identifiable.

If $x_0 \in T$ the system solution evolves in the invariant set T.

From Theorem 2, equation $p_3 x_2 - p_1 = 0$ has to be added to the polynomials defining the original dynamic system, in order to correctly check identifiability. In this case the model is nonidentifiable.

IDENTIFIABILITY AND ACCESSIBILITY (1/2)

In case of systems non-accessible from everywhere, the system evolves in a lower dimension manifold defined by a co-distribution orthogonal to Δ_C.

THEOREM 3. If the system is non-accessible from all initial states x_0 from which it may have been started, then the characteristic set does not change and provides the correct identifiability test.

A Model of erythroproietin (Epo) receptor

V. Becker et al. *Science* 2010.

The model describes the nonlinear processes of ligand-receptor (Epo-EpoR) interaction and trafficking kinetics. In particular the biochemical processes underlying the EpoR endocytosis, that is the process of engulfing substances outside the cell with a membrane and transporting them into cytoplasm.

Six species are incorporated in the model, x_i $i = 1,\ldots,6$ being the relative concentrations, and all interactions are modeled by mass-action kinetics.

$$
\begin{cases}
\dot{x}_1(t) = b_{\max} k_1 - k_1 x_1 - k_2 x_1 x_2 + k_3 x_3 + k_4 x_4 \\
\dot{x}_2(t) = -k_2 x_1 x_2 + k_3 x_3 + k_4 x_4 \qquad \Rightarrow \quad \dot{x}_2 = -\dot{x}_3 \Rightarrow \quad x_2 = -x_3 + p_2 \\
\dot{x}_3(t) = k_2 x_1 x_2 - k_3 x_3 - k_5 x_3 \\
\dot{x}_4(t) = k_5 x_3 - (k_4 + k_6 + k_7) x_4 \\
\dot{x}_5(t) = k_6 x_4 \\
\dot{x}_6(t) = k_7 x_4 \qquad \Rightarrow \quad \dot{x}_6 = \frac{k_7}{k_6}\dot{x}_5 \Rightarrow \quad x_6 = \frac{k_7}{k_6} x_5 \\
y_1(t) = x_2 + x_6 \\
y_2(t) = x_3 \\
y_3(t) = x_4 + x_5
\end{cases}
$$

with initial condition $x(0)=[x_{10},x_{20},0,0,0,0]$.

DAISY input file

```
WRITE "CORE MODEL, Becker SCIENCE 2010 Suppl. mat. pg.17"$

% B_ IS THE VARIABLE VECTOR
B_:={y1,y2,y3,x3,x5,x4,x1}$
FOR EACH EL_ IN B_ DO DEPEND EL_,T$

%B1_ IS THE UNKNOWN PARAMETER VECTOR
B1_:={k1,k2,k3,k4,k5,k6,k7,bmax,p2}$

%NUMBER OF STATE(S)
NX_:=4$
%NUMBER OF INPUT(S)
NU_:=0$
%NUMBER OF OUTPUT(S)
NY_:=3$

%MODEL EQUATION(S)
C_:={df(x1,t)=bmax*k1-k1*x1-k2*x1*(-x3+p2)+k3*x3+k4*x4,
 %   df(x2,t)=-k2*x1*(-x3+p2)+k3*x3+k4*x4,
    df(x3,t)=k2*x1*(-x3+p2)-k3*x3-k5*x3,
    df(x4,t)=k5*x3-(k4+k6+k7)*x4,
    df(x5,t)=k6*x4,
%    df(x6,t)=k7*x4,
    y1=(-x3+p2)+(k7/k6)*x5,
    y2=x3,
    y3=x4+x5}$

FLAG_:=1$
DAISY()$

%VALUES OF INITIAL CONDITIONS ARE GIVEN
ICK_:={X3=0,X4=0,X5=0}$
ICUNK_:={X1=X10}$
CONDINIZ()$
END$
```

NON-POLYNOMIAL MODELS?

EXAMPLE

$$\begin{cases} \dot{x}_1 = a\exp(-x_2) + u \\ \dot{x}_2 = -bx_1 \end{cases}$$

where $p = [a, b]$.

By introducing a new state $x_3 = a\ exp(-x_2)$ and by differentiating x_3 one obtains:

$$\dot{x}_3 = -\dot{x}_2 x_3$$

This additional differential equation will turn into the third order system which is trivially globally nonaccessible since the evolution of this augmented system is constrained to take place in some invariant submanifold:

$$\begin{cases} \dot{x}_1 = x_3 + u \\ \dot{x}_2 = -bx_1 \\ \dot{x}_3 = bx_1 x_3 \end{cases}$$

polynomial

globally nonaccessible

EXAMPLE OF TIME-VARYING MODELS

- In many biological and physiological applications, often in the differential equations describing the phenomena, time-varying coefficients appear with a known functional form but depending on some unknown parameters. For example:

$$\dot{x}_1 = a\exp(-bt)x_1 + u$$

By introducing a new state

$$x_2 = \exp(-bt)$$

and differentiating

$$\begin{cases} \dot{x}_1 = ax_1x_2 + u \\ \dot{x}_2 = -bx_2 \end{cases}$$

polynomial
globally nonaccessible
time-invariant

SOME OBSERVATIONS

- Diff. Alg. *Id.* provide the exact number of model parameter solutions.
- In case of nonidentifiability, it provides the analytical form of the functional dependence (*invariants)* of the nonidentifiable parameters.
- Structural identifiability analysis avoids to waste resources in performing useless experiments.
- There is no way that all parameters of a non identifiable model can be numerically estimated with good precision in a real situation.

- Obviously, although necessary, structural id. is not sufficient to guarantee an accurate identification of the model from the real input/output data.
- The analytical nature of the structural id. algorithms imposes restrictions on the size and complexity of the systems.

IDEA

Joint use of *structural* and *practical identifiability*

Practical id. analysis, based on optimization algorithms can take advantage of information provided by structural id.

I. Local identifiability case

- To calculate all the multiple solutions of the model.

II. Nonidentifiability case

- To exploit the analytical relations among nonidentifiable parameters described by the Gröbner bases to use these invariants as constraints in the optimization algorithm.

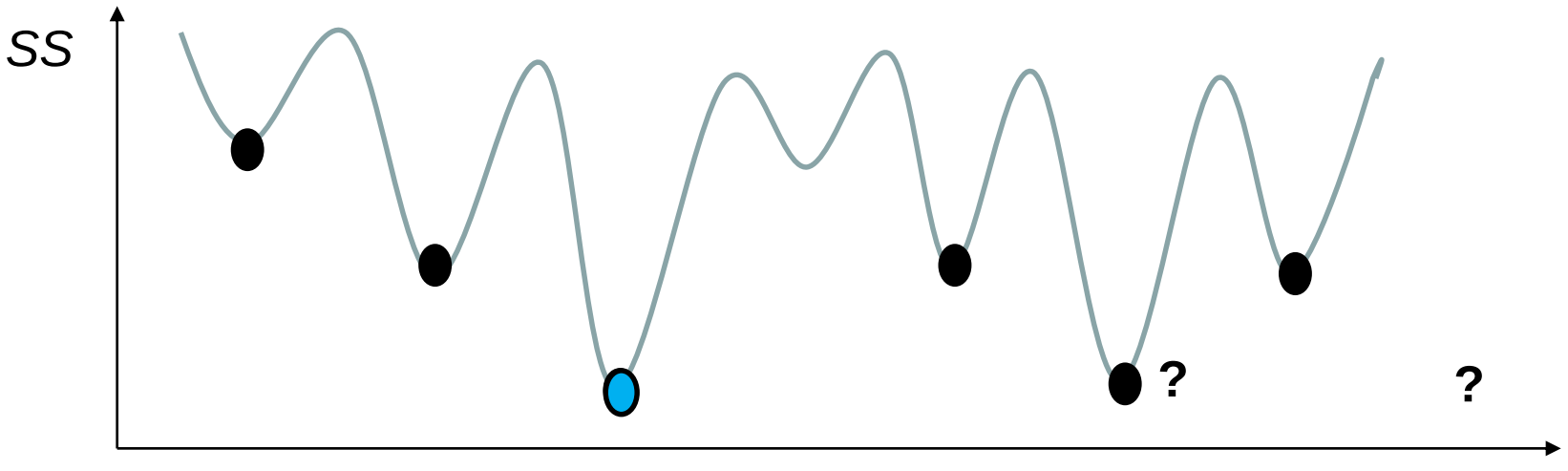
PRACTICAL IDENTIFIABILITY

$$P = \{\hat{p} \mid \nabla_p SS(\hat{p}) = 0 \ \wedge \ \nabla^2_{p^2} SS(\hat{p}) > 0\}$$ global and local minima

where *SS* is a cost function defined in an optimization algorithm.

- For local identifiable models (multiple global minima), multi-start searches require many optimization steps starting from different initial conditions.

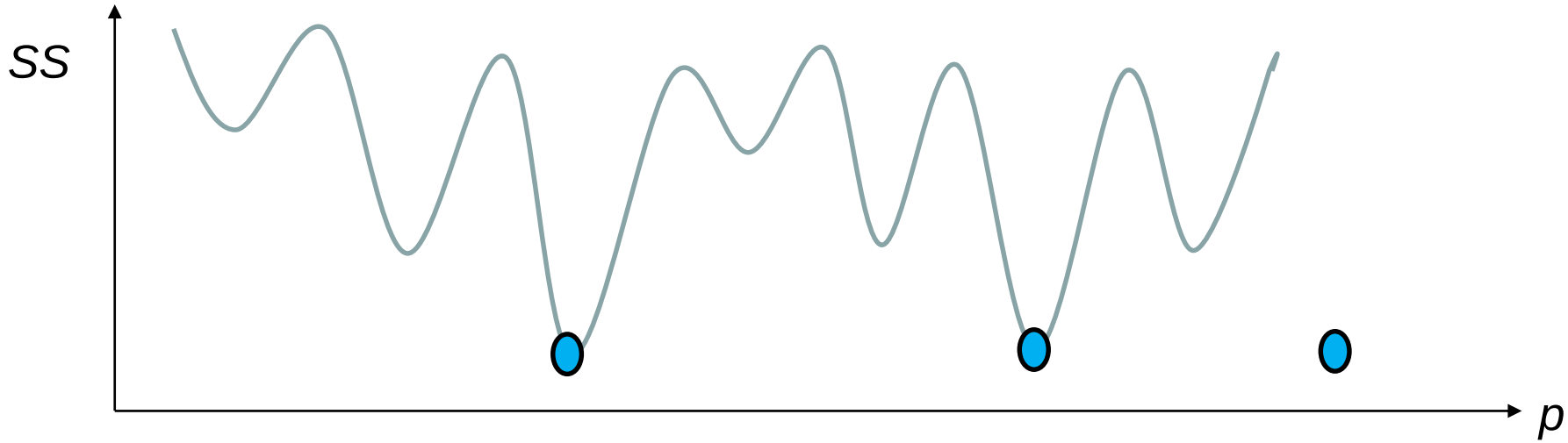
They can avoid local minima but cannot guarantee to find ALL global minima.



IDEA

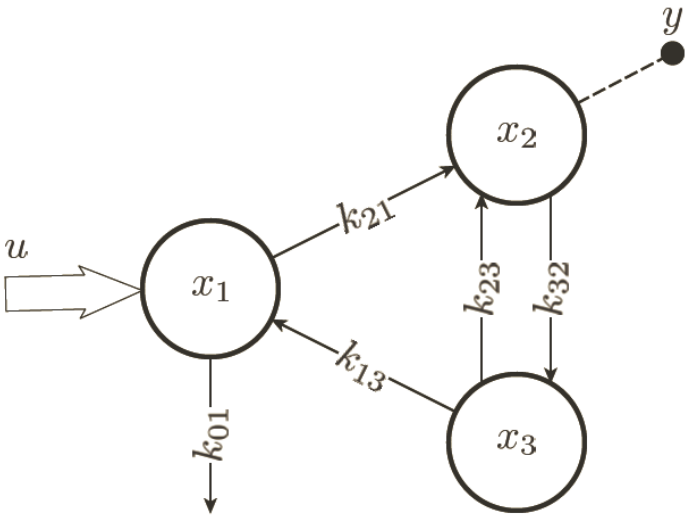
Joint use of *structural* and *practical identifiability*

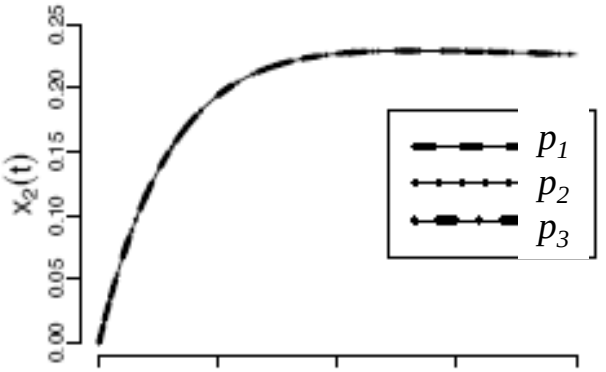
1. structural id. calculates the exact number of parameter solutions
2. find a numerical solution p^* with the practical identifiability
3. use p^* to analytically calculate all the equivalent parameter solutions p_i , with the differential algebra method.

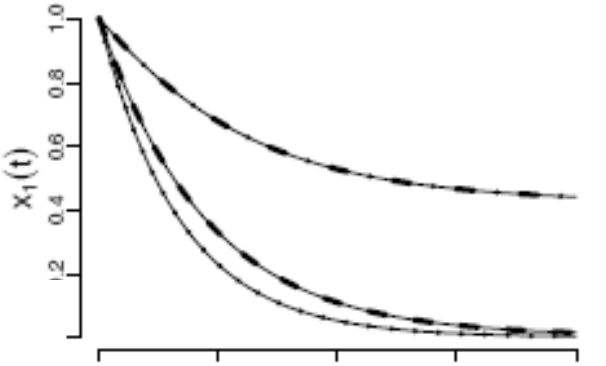


When possible, checking structural identifiability first, decrease the number of searches generally required by practical id.

cont. A simple locally identifiable model







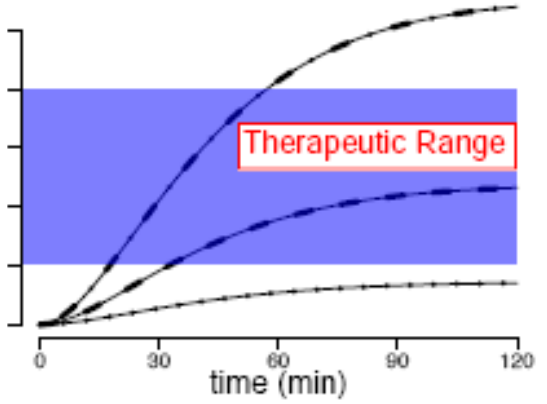
Suppose:

p = [0.02374, 0.00181, 0.01331, 0.03089, 0.01729]*

Table 1. Admissible solutions for the first randomized parameter vector

	p_1	p_2	p_3
k_{01}	0.02374	0.03581	0.0008737
k_{13}	0.00181	0.003971	0.02117
k_{21}[a]	0.01331	0.01331	0.01331
k_{23}	0.03089	0.02873	0.01153
k_{32}	0.01729	0.005225	0.04016

[a] globally identifiable parameter.



cont. A simple locally identifiable model

Suppose now:

p = [0.02324, 0.001834, 0.1202, 0.03072, 0.01632]*

Table 2. Non admissible solutions for the second randomized parameter vector

	p_1	p_2	p_3
k_{01}	0.02324	−0.000422	−0.07089
k_{13}	0.001834	−0.04121	$-8.88 \cdot 10^{-5}$
k_{21} [a]	0.1202	0.1202	0.1202
k_{23}	0.03072	0.07377	0.03264
k_{32}	0.01632	0.03998	0.1104

[a] globally identifiable parameter.

ATT.: The additional two solutions do not belong to the admissible parameter space.

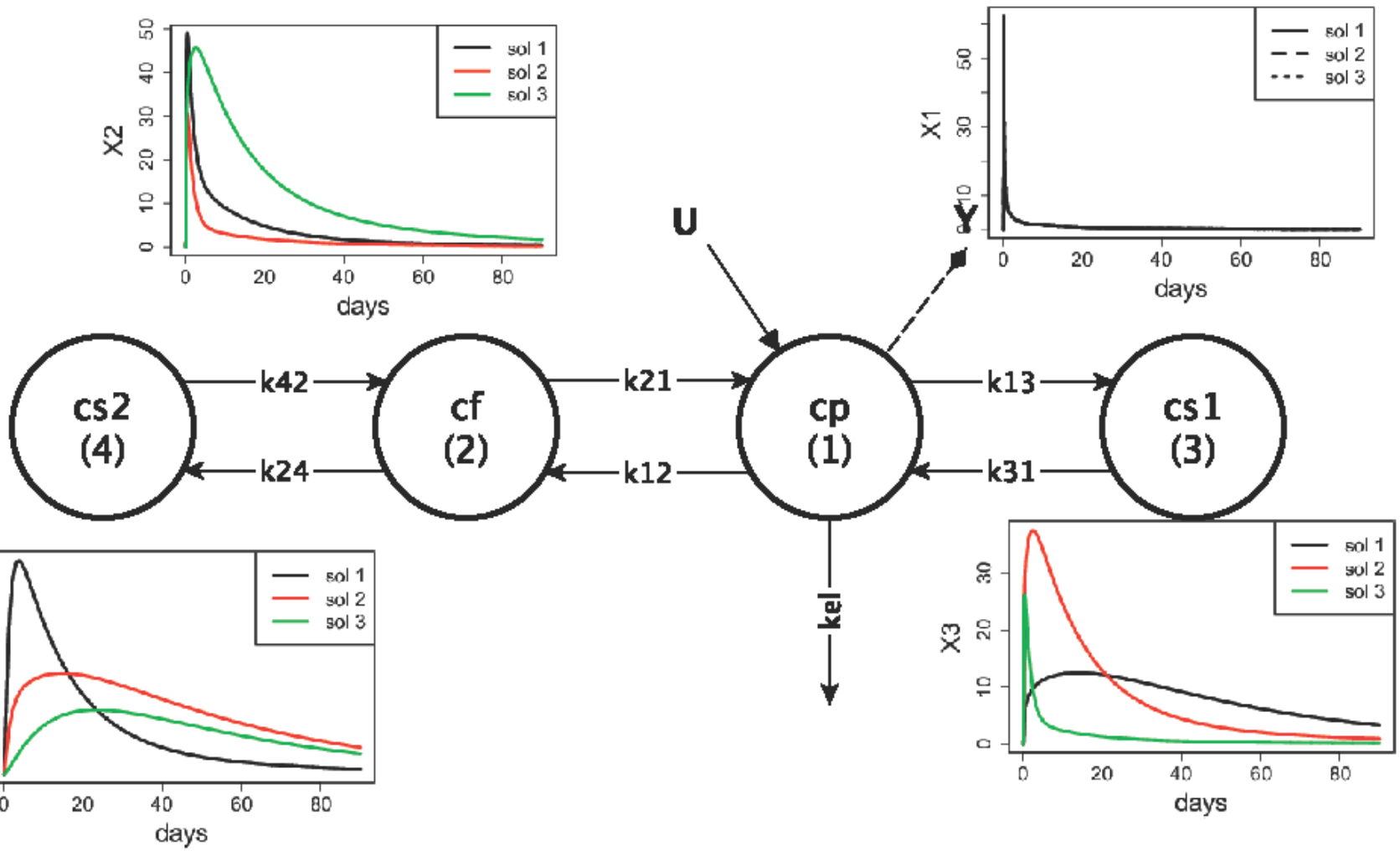
This is a favourable situation in which additional solutions can be rejected, showing, in practice, model global identiability.

II. Case of local identifiability

A pharmacokinetics model of Zalypsis®

a cytotoxic agent having a significant killing action in several tumour sites.

Craig et al.J. Theor. Biol. (2015).



A four-dimensional HIV\AIDS model

(Perelson et al., Math. Biosci. 1993)

$$\begin{cases} \dot{T} = s - \mu_T T + rT\left(1 - \dfrac{T + T^* + T^{**}}{T_{\max}}\right) - k_1 VT \\ \dot{T}^* = k_1 VT - \mu_T T^* - k_2 T^* \\ \dot{T}^{**} = k_2 T^* - \mu_b T^{**} \\ \dot{V} = N\mu_b T^{**} - k_1 VT - \mu_V V \\ y_1 = T \\ y_2 = V \end{cases}$$

T concentration of uninfected cells;

T^* concentration of latently infected cells;

T^{**} concentration of actively infected cells;

V concentration of free infectious virus particles;

y_1 and y_2 measured outputs;

$\boldsymbol{p}=[s, \mu_T, r, T_{max}, k_1, k_2, \mu_b, N, \mu_V]$ unknown parameters.

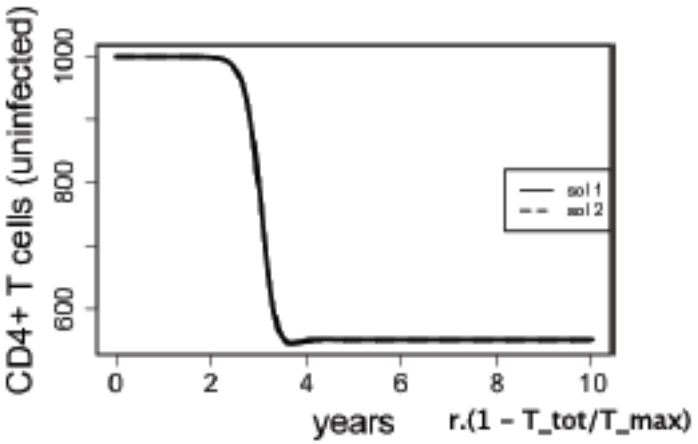
cont.

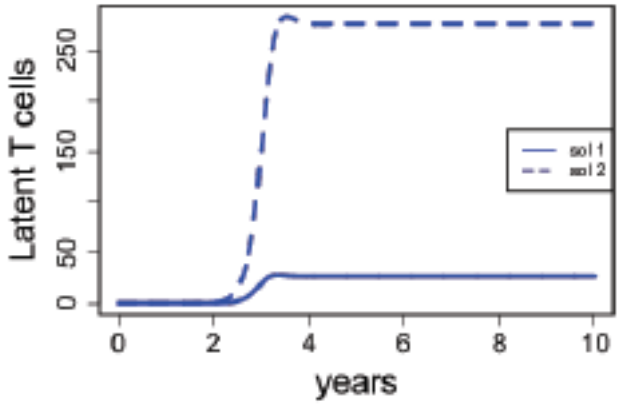
A four-dimensional HIV\AIDS model

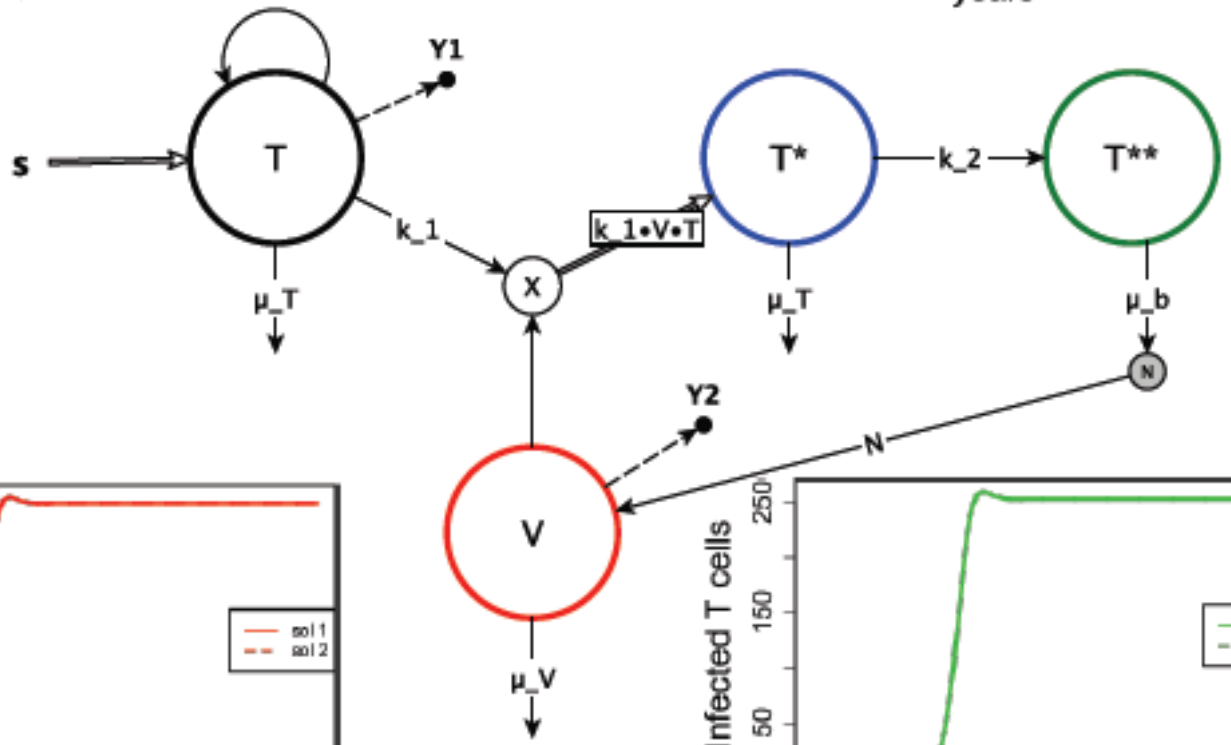
Table 3. The two solutions of the HIV model

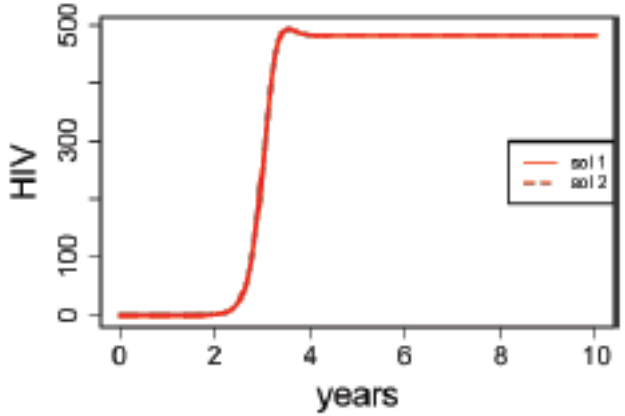
Parameter	Units	p_1	p_2
s	$(\text{day}^{-1}\text{mm}^{-3})$	10	10
r	(day^{-1})	0.03	0.03
T_{max}	(mm^{-3})	1500	1500
μ_T	(day^{-1})	0.02	0.02
μ_b	(day^{-1})	0.24	0.023
μ_v	(day^{-1})	2.4	2.4
k_1	$(\text{mm}^3\ \text{day}^{-1})$	$2.4{\cdot}10^{-5}$	$2.4{\cdot}10^{-5}$
k_2	(day^{-1})	0.003	0.22
N		1400	199.21

Cont. A four-dimensional HIV\AIDS model









THE DAISY SOFTWARE

- Our differential algebra method has been implemented in the software package DAISY, coded in the symbolic language Reduce.

- High-level programming languages, mathematics and computer algebra will *not* be prerequisites for using the software.

- A new beta version is now available with a user-friendly interface on the temporary website

https://daisy-reduce.shinyapps.io/daisy/

the final one will be: http://daisy.dei.unipd.it/