# Identifiability of linear compartmental models

<u>Nicolette Meshkat</u>

North Carolina State University

Parameter Estimation Workshop – NCSU August 9, 2014 78 slides

## Structural Identifiability Analysis

- Linear Model:
  - x state variable
  - *u* input
  - y output

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$

- *A*, *B*, *C* matrices with unknown parameters
- Finding which unknown parameters of a model can be quantified from given inputoutput data

#### But why linear compartment models?

- Used in many biological applications, e.g. pharmacokinetics
- Very often unidentifiable!
- Nice algebraic structure

– Can actually prove some general results!

## Unidentifiable Models

• Question 1: Can we always "reparametrize" an unidentifiable model into an identifiable one?

#### **Motivation:** Question 1

• <u>Model 1</u>:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix}$$
$$y = x_1$$

• <u>Model 2</u>:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & 0 & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix}$$
$$y = x_1$$

#### **Motivation:** Question 1

<u>Model 1</u>: <u>No</u> ID scaling reparametrization!



• <u>Model 2</u>: ID scaling reparametrization:

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & 1 & 0 \\ a_{12}a_{21} & a_{22} & 1 \\ a_{12}a_{31}a_{23} & 0 & a_{33} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix}$$
$$y = X_1$$

## Unidentifiable Models

• Question 1: Can we always "reparametrize" an unidentifiable into an identifiable one?

 Question 2: If a reparametrization exists, can we instead modify the original model to make it identifiable?

## **Motivation: Question 2**

• <u>Model 2</u>:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & 0 & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix}$$
$$y = x_1$$

- Starting with Model 2, how should we adjust model to obtain identifiability?
  - Decrease # of parameters?
  - Add input/output data?





#### Linear Compartment Models

• System equations:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -(k_{01} + k_{21}) & k_{12} \\ k_{21} & -(k_{02} + k_{12}) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$
$$y = x_1$$

• Can change to form  $\dot{x} = Ax + u$ 

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$

$$y = x_1$$

#### Larger class of models to investigate

- Assumptions:
  - I/O in first compartment
  - Leaks from every compartment

$$\dot{x} = Ax + u$$
$$y = x_1$$

where  $u = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$  and diagonal elements =  $a_{ii}$ 

## Useful tool: Directed Graph

- A directed graph G is a set of:
  - Vertices
  - Edges
- Ex 1: 1 = 2
  - Vertices: {1, 2}
  - Edges:  $\{1 \rightarrow 2, 2 \rightarrow 1\}$

## Useful tool: Directed Graph

- A directed graph G is a set of:
  - Vertices
  - Edges
- Ex 2:  $1 \rightleftharpoons 2 \longrightarrow 3$ 
  - Vertices: {1, 2, 3}
  - Edges:  $\{1 \rightarrow 2, 2 \rightarrow 1, 2 \rightarrow 3\}$
- A graph is *strongly connected* if there exists a path from each vertex to every other vertex

## Useful tool: Directed Graph

- A directed graph G is a set of:
  - Vertices
  - Edges
- Ex 3:

- Edges:  $\{1 \rightarrow 2, 2 \rightarrow 1, 2 \rightarrow 3, 3 \rightarrow 1\}$ 

• A graph is *strongly connected* if there exists a path from each vertex to every other vertex

 $1 \rightleftharpoons 2 \longrightarrow 3$ 

## Convert to graph

- Let G be directed graph with m edges, n vertices
- Associate a matrix A to the graph G:

$$A(G)_{ij} = \begin{cases} a_{ii} & \text{if } i = j \\ a_{ij} & \text{if } j \to i \text{ is an edge} \\ 0 & \text{otherwise} \end{cases}$$

where each  $a_{ij}$  is an independent real parameter

• Look only at strongly connected graphs

#### 2-compartment model as graph

Model: 
$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$
  
 $y = x_1$ 





- Cycle: *a*<sub>12</sub>*a*<sub>21</sub>
  "Self" cycles: *a*<sub>11</sub>, *a*<sub>22</sub>

## **Identifiability Analysis**

• Model: 
$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$
  
 $y = x_1$ 

• Unknown parameters:  $\{a_{11}, a_{12}, a_{21}, a_{22}\}$ 

- <u>Identifiability</u>: Which parameters of model can be quantified from given input-output data?
  - Must first determine *input-output equation*

## **Find Input-Output Equation**

- Rewrite system eqns as  $(\partial I A)x = u$
- Cramer's Rule:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \qquad A_1 = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$$
$$x_1 = \frac{\det(\partial I - A_1)u_1}{\det(\partial I - A)}$$

• I/O eqn:  $det(\partial I - A)y = det(\partial I - A_1)u_1$ 

$$\begin{split} y^{(n)} + c_1 y^{(n-1)} + \cdots + c_n y \\ &= u_1^{(n-1)} + c_{n+1} u_1^{(n-2)} + \cdots + c_{2n-1} u_1 \end{split}$$

## Identifiability

- Can recover coefficients from data
- <u>Identifiability</u>: is it possible to recover the parameters of the original system, from the coefficients of I/O eqn?
  - Two sets of parameter values yield same coefficient values?
  - Is coeff map 1-to-1?

$$y^{(n)} + c_1 y^{(n-1)} + \dots + c_n y$$
  
=  $u_1^{(n-1)} + c_{n+1} u_1^{(n-2)} + \dots + c_{2n-1} u_1$ 

## 2-compartment model

• I/O eqn

$$\ddot{y} - (a_{11} + a_{22})\dot{y} + (a_{11}a_{22} - a_{12}a_{21})y$$
  
=  $\dot{u}_1 - a_{22}u_1$ 

• Coefficient map  $c: \mathbb{R}^4 \to \mathbb{R}^3$ 

$$(a_{11}, a_{12}, a_{21}, a_{22})$$
  
 $\mapsto (-(a_{11} + a_{22}), a_{11}a_{22} - a_{12}a_{21}, -a_{22})$ 

• Identifiability: Is the coefficient map 1-to-1?

## Identifiability from I/O eqns

- Question of injectivity of the coefficient map  $c \colon \mathbb{R}^{m+n} \to \mathbb{R}^{2n-1}$
- If *c* is one-to-one: <u>globally identifiable</u> finite-to-one: <u>locally identifiable</u> infinite-to-one: <u>unidentifiable</u>

#### **One-to-one Example**

• Map  $c: \mathbb{R}^2 \to \mathbb{R}^2$ 

$$(p_1, p_2) \mapsto (p_1 + p_2, p_1 - p_2)$$

• 2 equations:

$$p_1 + p_2 = p_1^* + p_2^*$$
  
$$p_1 - p_2 = p_1^* - p_2^*$$

• One-to-one:

$$p_1 = p_1^* \qquad p_2 = p_2^*$$

#### Finite-to-one Example

• Map  $c: \mathbb{R}^2 \to \mathbb{R}^2$ 

$$(p_1, p_2) \mapsto (p_1 + p_2, p_1 p_2)$$

• 2 equations:

$$p_1 + p_2 = p_1^* + p_2^*$$
$$p_1 p_2 = p_1^* p_2^*$$

• Finite-to-one:

$$p_1 = p_1^*$$
  $p_2 = p_2^*$   
or  
 $p_1 = p_2^*$   $p_2 = p_1^*$ 

#### Our example

• 3 equations:

$$a_{11} + a_{22} = a_{11}^* + a_{22}^*$$
$$a_{11}a_{22} - a_{12}a_{21} = a_{11}^*a_{22}^* - a_{12}^*a_{21}^*$$
$$a_{22} = a_{22}^*$$

• Infinite-to-one!

$$a_{11} = a_{11}^*$$
  $a_{22} = a_{22}^*$   $a_{12} = \frac{a_{12}^* a_{21}^*}{a_{21}}$ 

#### Our example

• 3 equations:

$$a_{11} + a_{22} = a_{11}^* + a_{22}^*$$
$$a_{11}a_{22} - a_{12}a_{21} = a_{11}^*a_{22}^* - a_{12}^*a_{21}^*$$
$$a_{22} = a_{22}^*$$

• Infinite-to-one!

$$a_{11} = a_{11}^*$$
  $a_{22} = a_{22}^*$   $a_{12}a_{21} = a_{12}^*a_{21}^*$ 

## Testing identifiability in practice

- Check dimension of image of coefficient map  $c \colon \mathbb{R}^{m+n} \to \mathbb{R}^{2n-1}$
- If *dim im c = m+n,* then locally identifiable
- If *dim im c < m+n*, then unidentifiable
- Linear Ex:  $c: \mathbb{R}^2 \to \mathbb{R}^2$

$$(p_1, p_2) \mapsto (p_1 + p_2, p_1 - p_2)$$

• Jacobian has rank 2:  $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ 

## Testing identifiability in practice

- Check dimension of image of coefficient map  $c \colon \mathbb{R}^{m+n} \to \mathbb{R}^{2n-1}$
- If *dim im c = m+n,* then locally identifiable
- If *dim im c < m+n*, then unidentifiable
- Our Ex:  $c: \mathbb{R}^4 \to \mathbb{R}^3$

$$(a_{11}, a_{12}, a_{21}, a_{22})$$
  
 $\mapsto (-(a_{11} + a_{22}), a_{11}a_{22} - a_{12}a_{21}, -a_{22})$ 

• Jacobian has rank 3:  $\begin{pmatrix} -1 & 0 & 0 & -1 \\ a_{22} & -a_{21} & -a_{12} & a_{11} \\ 0 & 0 & 0 & -1 \end{pmatrix}$ 

## Unidentifiable models

 Cannot determine individual parameters, but can we determine some combination of the parameters?

Ex:  $p_1 + p_2$  or  $p_1 p_2$ 

• A function  $f : \mathbb{R}^{m+n} \to \mathbb{R}$  is called identifiable from c if  $f = \Phi(c_1, \dots, c_{2n-1})$ 

## Identifiable functions

- Coefficients:  $c_1 = -(a_{11} + a_{22})$   $c_2 = a_{11}a_{22} - a_{12}a_{21}$  $c_3 = -a_{22}$
- Identifiable functions (cycles):  $a_{11} = -(c_1 - c_3)$   $a_{22} = -c_3$  $a_{12}a_{21} = (c_1 - c_3)c_3 - c_2$
- Coefficients can be written in terms of identifiable functions:

$$(a_{11}, a_{12}, a_{21}, a_{22})$$
  
 $\mapsto (-(a_{11} + a_{22}), a_{11}a_{22} - a_{12}a_{21}, -a_{22})$ 

## Unidentifiable model

• Model 
$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$
  
 $y = x_1$ 

- Identifiable functions  $a_{11}$ ,  $a_{22}$ ,  $a_{12}a_{21}$ i.e.  $-(k_{01} + k_{21}), -(k_{02} + k_{12}), k_{12}k_{21}$
- <u>Reparametrize</u>: 4 independent parameters → 3 independent parameters?

#### Identifiable reparametrization

#### Let $c: \mathbb{R}^{m_1} \to \mathbb{R}^{m_2}$ be a coefficient map

An identifiable reparametrization of a model is a map  $q: \mathbb{R}^{m_3} \to \mathbb{R}^{m_1}$  such that:

- $c \circ q : \mathbb{R}^{m_3} \to \mathbb{R}^{m_2}$  has the same image as c
- *c q* is identifiable (finite-to-one)

## Scaling reparametrization

• Choice of functions  $f_1, ..., f_n : \mathbb{R}^{m+n} \to \mathbb{R}$  where we replace  $x_1, ..., x_n$  with

$$X_i = f_i(A) x_i$$

- Set  $f_1 = 1$  since  $y = x_1$  is observed
- Since model is  $\dot{x} = Ax + u$ , each parameter  $a_{ij}$ is replaced with  $\frac{a_{ij}f_i(A)}{f_j(A)}$
- Only graphs with at most 2n-2 edges

#### Reparametrize original model

• Use scaling:  $X_1 = x_1$   $X_2 = a_{12}x_2$ 

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & 1 \\ a_{12}a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$
$$y = X_1$$

• Re-write:

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \end{pmatrix} = \begin{pmatrix} q_{11} & 1 \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$
$$y = X_1$$

• Map  $c \circ q$  has same image as c and is 1-to-1  $(q_{11}, q_{21}, q_{22})$  $\mapsto (-(q_{11} + q_{22}), q_{11}q_{22} - q_{21}, -q_{22})$ 

## Motivation: Unidentifiable models

• <u>Model 1</u>: <u>No</u> ID scaling reparametrization!



• <u>Model 2</u>: ID scaling reparametrization:



## Main question:

#### Which graphs with ≤ 2n-2 edges admit an identifiable scaling reparametrization?
# Main result<sup>1</sup>:

Let G be a strongly connected graph. Then TFAE:

The model has an identifiable scaling reparametrization

- ↔ The model has an identifiable scaling
   reparametrization by monomial functions of
   the original parameters
- ⇔ All the monomial cycles in G are identifiable functions
- $\Leftrightarrow$  dim im c = m+1

<sup>1</sup>N. Meshkat and S. Sullivant, Identifiable reparametrizations of linear compartment models, Journal of Symbolic Computation 63 (2014) 46-67.

#### Non-Example: Model 1



*dim im c* = 4, so no ID scaling reparametrization!

#### Example: Model 2



Identifiable cycles:  $a_{11}$ ,  $a_{22}$ ,  $a_{33}$ ,  $a_{12}a_{21}$ ,  $a_{12}a_{31}a_{23}$ 

# Algorithm to find identifiable reparametrization

- 1) Form a spanning tree T
- 2) Form the directed incidence matrix *E*(*T*):

$$E(T)_{i(j,k)} = \begin{cases} 1 & \text{if } i = j \\ -1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$$

- 3) Let *E* be E(T) with first row removed
- 4) Columns of  $E^{-1}$  are exponent vectors of monomials  $f_i(A)$  in scaling  $X_i = f_i(A)x_i$

### Identifiable reparametrization



$$X_1 = x_1$$
  $X_2 = a_{12}x_2$   $X_3 = a_{12}a_{23}x_3$ 

Identifiable scaling reparametrization

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & 1 & 0 \\ a_{12}a_{21} & a_{22} & 1 \\ a_{12}a_{31}a_{23} & 0 & a_{33} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix}$$
$$y = X_1$$

# Main result

- A model with
  - I/O in first compartment
  - n leaks
  - Strongly connected graph G

has an identifiable scaling reparametrization

⇔ all the monomial cycles are identifiable

 $\Leftrightarrow$  dim im c = m+1

# Which graphs have this property?

• Not complete characterization:



# Unidentifiable Models

• Question 1: Can we always "reparametrize" an unidentifiable into an identifiable one?

 Question 2: If a reparametrization exists, can we instead modify the original model to make it identifiable?

# Model 2



- Input/Output in compartment 1
- Leaks from every compartment
- *dim im c* = m+1 = 5
- Identifiable cycles  $a_{11}$ ,  $a_{22}$ ,  $a_{33}$ ,  $a_{12}a_{21}$ ,  $a_{12}a_{31}a_{23}$

# **Obtaining Identifiability**



- Starting with Model 2, how should we adjust model to obtain identifiability?
- Two options: Remove leaks or add input/output

# **Removing leaks**



• Remove 2 leaks

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & -a_{12} & a_{23} \\ a_{31} & 0 & -a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix}$$
$$y = x_1$$

• *dim im c* = 5

## Theorem on Removing leaks<sup>2</sup>

- Starting with a model with:
  - I/O in first compartment
  - n leaks
  - Strongly connected graph G
  - *dim im c* = m+1
- Remove n-1 leaks  $\Rightarrow$  Local identifiability
- Ex:



<sup>2</sup> N. Meshkat, S. Sullivant, and M. Eisenberg, Identifiability results for several classes of linear compartment models, In preparation.

## Theorem on Removing leaks<sup>2</sup>

- Starting with a model with:
  - I/O in first compartment
  - n leaks
  - Strongly connected graph G
  - *dim im c* = m+1
- Remove n-1 leaks  $\Longrightarrow$  Local identifiability
- Ex:



<sup>2</sup>N. Meshkat, S. Sullivant, and M. Eisenberg, Identifiability results for several classes of linear compartment models, In preparation.

## Example: Manganese Model<sup>3</sup>



<sup>3</sup> P. K. Douglas, M. S. Cohen, and J. J. DiStefano III, Chronic exposure to Mn inhalation may have lasting effects: A physiologically-based toxicokinetic model in rats, Toxicology and Environmental Chemistry 92 (2) (2010) 279-299.

#### Adding output to leak compartment



 Remove 1 leak and add 1 output to leak compartment

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & 0 & -a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix}$$
$$y_1 = x_1$$
$$y_1 = x_1$$
$$dim \ im \ c = 6 \qquad y_2 = x_2$$

#### Thm: Removing leaks and adding inputs/outputs

- Starting with a model with:
  - I/O in first compartment
  - n leaks
  - Strongly connected graph G
  - dim im c = m+1
- Remove a subset of leaks so that every leak compartment has either input or output ⇒ Local identifiability
- Ex:



#### Thm: Removing leaks and adding inputs/outputs

- Starting with a model with:
  - I/O in first compartment
  - n leaks
  - Strongly connected graph G
  - dim im c = m+1
- Remove a subset of leaks so that every leak compartment has either input or output ⇒ Local identifiability
- Ex:



# Sufficient, not necessary

 Harder to find *general* conditions if I/O not in leak compartment



# Quiz!

• Which of the following models are identifiable?



• Answer: B and C

# Identifiability Problem for Nonlinear Models

- What is our model is nonlinear?
- Same process:

– Find I/O equations

Test injectivity of coefficient map

$$F_j(u, \dot{u}, \ddot{u}, \dots, y, \dot{y}, \ddot{y}, \dots, p) = 0$$

# Identifiability Problem for Nonlinear Models

- What is our model is nonlinear?
- Same process:
  - Find I/O equations

- Test injectivity of coefficient map

$$\sum_i c_{ji}(p)\psi_{ji}(u,\dot{u},\ddot{u},\ldots,y,\dot{y},\ddot{y},\ldots)=0$$



• I/O eqn

$$\begin{aligned} -K_M^2 \dot{u} &- 2K_M \dot{u}y - \dot{u}y^2 + K_M^2 \ddot{y} + 2K_M \ddot{y}y + \ddot{y}y^2 \\ &+ (k_{02}K_M^2 + k_{12}K_M^2 + k_{21}K_M^2 + K_M V_M) \dot{y} \\ &+ (2k_{02}K_M + 2k_{12}K_M + 2k_{21}K_M) \dot{y}y \\ &+ (k_{02} + k_{12} + k_{21}) \dot{y}y^2 - (k_{02}K_M^2 + k_{12}K_M^2) u \\ &- (2k_{02}K_M + 2k_{12}K_M) uy - (k_{02} + k_{12}) uy^2 \\ &+ (k_{02}k_{21}K_M^2 + k_{02}K_M V_M + k_{12}K_M V_M) y \\ &+ (2k_{02}k_{21}K_M + k_{02}V_M + k_{12}V_M) y^2 + k_{02}k_{21}y^3 \\ &= 0 \end{aligned}$$

• I/O eqn

$$-K_{M}^{2}\dot{u} - 2K_{M}\dot{u}y - \dot{u}y^{2} + K_{M}^{2}\ddot{y} + 2K_{M}\ddot{y}y + \ddot{y}y^{2} + (k_{02}K_{M}^{2} + k_{12}K_{M}^{2} + k_{21}K_{M}^{2} + K_{M}V_{M})\dot{y} + (2k_{02}K_{M} + 2k_{12}K_{M} + 2k_{21}K_{M})\dot{y}y + (k_{02} + k_{12} + k_{21})\dot{y}y^{2} - (k_{02}K_{M}^{2} + k_{12}K_{M}^{2})u - (2k_{02}K_{M} + 2k_{12}K_{M})uy - (k_{02} + k_{12})uy^{2} + (k_{02}k_{21}K_{M}^{2} + k_{02}K_{M}V_{M} + k_{12}K_{M}V_{M})y + (2k_{02}k_{21}K_{M} + k_{02}V_{M} + k_{12}V_{M})y^{2} + k_{02}k_{21}y^{3} = 0$$

• I/O eqn

$$\begin{aligned} -K_M^2 \dot{u} &- 2K_M \dot{u}y - \dot{u}y^2 + K_M^2 \ddot{y} + 2K_M \ddot{y}y + \ddot{y}y^2 \\ &+ (k_{02}K_M^2 + k_{12}K_M^2 + k_{21}K_M^2 + K_M V_M) \dot{y} \\ &+ (2k_{02}K_M + 2k_{12}K_M + 2k_{21}K_M) \dot{y}y \\ &+ (k_{02} + k_{12} + k_{21}) \dot{y}y^2 - (k_{02}K_M^2 + k_{12}K_M^2) u \\ &- (2k_{02}K_M + 2k_{12}K_M) uy - (k_{02} + k_{12}) uy^2 \\ &+ (k_{02}k_{21}K_M^2 + k_{02}K_M V_M + k_{12}K_M V_M) y \\ &+ (2k_{02}k_{21}K_M + k_{02}V_M + k_{12}V_M) y^2 + k_{02}k_{21}y^3 \\ &= 0 \end{aligned}$$

• I/O eqn

$$\begin{aligned} -K_M^2 \dot{u} &- 2K_M \dot{u}y - \dot{u}y^2 + K_M^2 \ddot{y} + 2K_M \ddot{y}y + \ddot{y}y^2 \\ &+ (k_{02}K_M^2 + k_{12}K_M^2 + k_{21}K_M^2 + K_M V_M) \dot{y} \\ &+ (2k_{02}K_M + 2k_{12}K_M + 2k_{21}K_M) \dot{y}y \\ &+ (k_{02} + k_{12} + k_{21}) \dot{y}y^2 - (k_{02}K_M^2 + k_{12}K_M^2) u \\ &- (2k_{02}K_M + 2k_{12}K_M) uy - (k_{02} + k_{12}) uy^2 \\ &+ (k_{02}k_{21}K_M^2 + k_{02}K_M V_M + k_{12}K_M V_M) y \\ &+ (2k_{02}k_{21}K_M + k_{02}V_M + k_{12}V_M) y^2 + k_{02}k_{21}y^3 \\ &= 0 \end{aligned}$$

• I/O eqn

$$\begin{aligned} -K_M^2 \dot{u} &- 2K_M \dot{u}y - \dot{u}y^2 + K_M^2 \ddot{y} + 2K_M \ddot{y}y + \ddot{y}y^2 \\ &+ (k_{02}K_M^2 + k_{12}K_M^2 + k_{21}K_M^2 + K_M V_M) \dot{y} \\ &+ (2k_{02}K_M + 2k_{12}K_M + 2k_{21}K_M) \dot{y}y \\ &+ (k_{02} + k_{12} + k_{21}) \dot{y}y^2 - (k_{02}K_M^2 + k_{12}K_M^2) u \\ &- (2k_{02}K_M + 2k_{12}K_M) uy - (k_{02} + k_{12}) uy^2 \\ &+ (k_{02}k_{21}K_M^2 + k_{02}K_M V_M + k_{12}K_M V_M) y \\ &+ (2k_{02}k_{21}K_M + k_{02}V_M + k_{12}V_M) y^2 + k_{02}k_{21}y^3 \\ &= 0 \end{aligned}$$

• I/O eqn

$$\begin{aligned} -K_M^2 \dot{u} &- 2K_M \dot{u}y - \dot{u}y^2 + K_M^2 \ddot{y} + 2K_M \ddot{y}y + \ddot{y}y^2 \\ &+ (k_{02}K_M^2 + k_{12}K_M^2 + k_{21}K_M^2 + K_M V_M) \dot{y} \\ &+ (2k_{02}K_M + 2k_{12}K_M + 2k_{21}K_M) \dot{y}y \\ &+ (k_{02} + k_{12} + k_{21}) \dot{y}y^2 - (k_{02}K_M^2 + k_{12}K_M^2) u \\ &- (2k_{02}K_M + 2k_{12}K_M) uy - (k_{02} + k_{12}) uy^2 \\ &+ (k_{02}k_{21}K_M^2 + k_{02}K_M V_M + k_{12}K_M V_M) y \\ &+ (2k_{02}k_{21}K_M + k_{02}V_M + k_{12}V_M) y^2 + k_{02}k_{21}y^3 \\ &= 0 \end{aligned}$$

• I/O eqn

$$\begin{aligned} -K_M^2 \dot{u} &- 2K_M \dot{u}y - \dot{u}y^2 + K_M^2 \ddot{y} + 2K_M \ddot{y}y + \ddot{y}y^2 \\ &+ (k_{02}K_M^2 + k_{12}K_M^2 + k_{21}K_M^2 + K_M V_M) \dot{y} \\ &+ (2k_{02}K_M + 2k_{12}K_M + 2k_{21}K_M) \dot{y}y \\ &+ (k_{02} + k_{12} + k_{21}) \dot{y}y^2 - (k_{02}K_M^2 + k_{12}K_M^2) u \\ &- (2k_{02}K_M + 2k_{12}K_M) uy - (k_{02} + k_{12}) uy^2 \\ &+ (k_{02}k_{21}K_M^2 + k_{02}K_M V_M + k_{12}K_M V_M) y \\ &+ (2k_{02}k_{21}K_M + k_{02}V_M + k_{12}V_M) y^2 + k_{02}k_{21}y^3 \\ &= 0 \end{aligned}$$

# **Differential algebra**

- How to find I/O equations for nonlinear models?
- Differential elimination
  - Differentiation + Gröbner Basis
  - Differential Gröbner Basis
    - Rosenfeld-Gröbner in Maple
  - Ritt's pseudo-division

## **Example on Lotka-Volterra**

• Equations:

$$\dot{x}_1 = ax_1 - bx_1x_2$$
$$\dot{x}_2 = -cx_2 + dx_1x_2$$
$$y = x_1$$

• Commands in Maple:

with(DifferentialAlgebra)

$$sys := [-diff(x1(t), t) + a \cdot x1(t) - b \cdot x1(t) \cdot x2(t), -diff(x2(t), t) + d \\ \cdot x1(t) \cdot x2(t) - c \cdot x2(t), y(t) - x1(t)]$$

- R := DifferentialRing(blocks = [x1, x2, y], derivations = [t])
- G := RosenfeldGroebner(sys, R)

#### Example on Lotka-Volterra

• Equations:

$$\dot{x}_1 = ax_1 - bx_1x_2$$
$$\dot{x}_2 = -cx_2 + dx_1x_2$$
$$y = x_1$$

• Rosenfeld-Gröbner gives:

$$\begin{bmatrix} x1(t) = y(t), x2(t) = -\frac{\frac{d}{dt}y(t) - y(t)a}{y(t)b}, \frac{d^2}{dt^2}y(t) \\ = \frac{1}{y(t)} \left( \left(\frac{d}{dt}y(t)\right)^2 + \left(\frac{d}{dt}y(t)\right)y(t)^2d - \left(\frac{d}{dt}y(t)\right)y(t)c \\ -y(t)^3ad + y(t)^2ac \right) \end{bmatrix}$$

# Example: Nonlinear HIV Model<sup>4</sup>

• Model equations:

$$\dot{x}_{1} = -bx_{1}x_{4} - dx_{1} + s$$
  

$$\dot{x}_{2} = bq_{1}x_{1}x_{4} - k_{1}x_{2} - m_{1}x_{2}$$
  

$$\dot{x}_{3} = bq_{2}x_{1}x_{4} + k_{1}x_{2} - m_{2}x_{3}$$
  

$$\dot{x}_{4} = -cx_{4} + k_{2}x_{3}$$
  

$$y_{1} = x_{1}$$
  

$$y_{2} = x_{4}$$

• Parameter vector:  $p = \{b, d, s, m_2, c, m_1, k_1, q_1, k_2, q_2\}$ 

<sup>4</sup>X. Xia and C. H. Moog, Identifiability of nonlinear systems with application to HIV/AIDS models, *IEEE Trans Autom Contr* 48 (2003), 330-336.

#### **Input-Output equations**

• Rosenfeld-Gröbner gives two equations:

$$\dot{y}_1 + by_1y_2 + dy_1 - s = 0$$

$$\ddot{y}_2 + (c + k_1 + m_1 + m_2)\ddot{y}_2 - bq_2k_2\dot{y}_2y_1 + (ck_1 + cm_1 + cm_2 + k_1m_2 + m_1m_2)\dot{y}_2 + b^2q_2k_2y_1y_2^2 + bk_2(dq_2 - k_1q_1 - k_1q_2 - m_1q_2)y_1y_2 + (-bq_2k_2s + ck_1m_2 + cm_1m_2)y_2 = 0$$

• Coefficient map  $c: \mathbb{R}^{10} \to \mathbb{R}^9$ 

- Unidentifiable!

# How to find identifiable functions?

- Injectivity test: If  $c(p) = c(p^*)$ , does  $p = p^*$ ?
- Amounts to solving a system of polynomial equations
- Find Gröbner Basis of  $c(p) c(p^*) = 0$ 
  - Gives system of equations  $c(p) = c(p^*)$  in "triangular form"
    - Analogous to Gaussian elimination for systems of linear equations
  - Must give an "ordering" of parameters to do the elimination

#### **Example of Gröbner Basis**

#### • Set up $\boldsymbol{c}(\boldsymbol{p}) - \boldsymbol{c}(\boldsymbol{p}^*)$ , for $\boldsymbol{p} = \{m_1, k_1, m_2, q_1, k_2, q_2, b, d, s, c\}$ $\boldsymbol{p}^* = \{\zeta, \eta, \delta, \theta, \rho, \kappa, \alpha, \lambda, \gamma, \epsilon\}$

 $b - \alpha$   $d - \lambda$   $-s + \gamma$   $-b k2 q2 + \alpha \kappa \rho$   $b^{2} k2 q2 - \alpha^{2} \kappa \rho$   $c k1 m2 + c m1 m2 - b k2 q2 s - \delta \epsilon \zeta - \delta \epsilon \eta + \alpha \gamma \kappa \rho$   $c + k1 + m1 + m2 - \delta - \epsilon - \zeta - \eta$   $-b k1 k2 q1 + b d k2 q2 - b k1 k2 q2 - b k2 m1 q2 + \alpha \eta \theta \rho + \alpha \zeta \kappa \rho + \alpha \eta \kappa \rho - \alpha \kappa \lambda \rho$   $c k1 + c m1 + c m2 + k1 m2 + m1 m2 - \delta \epsilon - \delta \zeta - \epsilon \zeta - \delta \eta - \epsilon \eta$ 

• Gröbner basis for  $p = \{m_1, k_1, m_2, q_1, k_2, q_2, b, d, s, c\}$ 

 $\begin{pmatrix} -(c-\delta) & (c-\xi-\eta) \\ & -s+\gamma \\ & d-\lambda \\ & b-\alpha \\ & -\alpha & (k2 q2 - \kappa\rho) \\ -c^2 - c m2 - m2^2 + c \delta + m2 \delta + c \epsilon + m2 \epsilon - \delta \epsilon + c \xi + m2 \xi - \delta \xi - \epsilon \xi + c \eta + m2 \eta - \delta \eta - \epsilon \eta \\ & -\alpha & (q2 \eta \theta - k1 q1 \kappa + c q2 \kappa + m2 q2 \kappa - q2 \delta \kappa - q2 \epsilon \kappa) \rho \\ & \alpha & (k1 k2 q1 - \eta \theta \rho - c \kappa \rho - m2 \kappa \rho + \delta \kappa \rho + \epsilon \kappa \rho) \\ & c + k1 + m1 + m2 - \delta - \epsilon - \xi - \eta \\ \end{pmatrix}$
## Algorithm <sup>5</sup>

- Find Gröbner Bases of c(p) c(p\*) for different orderings of the parameter vector p
- Look for elements of the form  $f(p) f(p^*)$  in Gröbner Basis
- Implies *f(p)* is identifiable
- Must find N identifiable functions in order to reparametrize, where N = dim im c

<sup>5</sup> N. Meshkat, M. Eisenberg, and J. J DiStefano III, An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Groebner Bases, *Math. Biosci.* 222 (2009)

#### **Examine Gröbner Bases**



## **Identifiable Functions**

- ID parameters: {*c*,*s*,*d*,*b*, *m*<sub>2</sub>}
  - Globally identifiable (one solution)  $\{s, d, b\}$
  - Locally identifiable (three solutions)  $\{c, m_2\}$
- ID parameter combinations:  $\{q_2k_2, q_1k_1k_2, m_1 + k_1\}$ – Globally identifiable (one solution)  $\{q_2k_2\}$ 
  - Locally identifiable (three solutions)  $\{q_1k_1k_2, m_1 + k_1\}$

#### **Identifiable Reparametrization**

- Identifiable:  $\{c, s, d, b, q_2k_2, m_2, q_1k_1k_2, m_1 + k_1\}$
- Use scaling:  $X_2 = k_1 k_2 x_2$   $X_3 = \frac{x_3}{q_2}$

$$\dot{x}_{1} = -bx_{1}x_{4} - dx_{1} + s$$
  

$$\dot{X}_{2} = b(q_{1}k_{1}k_{2})x_{1}x_{4} - (m_{1} + k_{1})X_{2}$$
  

$$\dot{X}_{3} = bx_{1}x_{4} + \frac{X_{2}}{q_{2}k_{2}} - m_{2}X_{3}$$
  

$$\dot{x}_{4} = -cx_{4} + q_{2}k_{2}X_{3}$$
  

$$y_{1} = x_{1}$$
  

$$y_{2} = x_{4}$$

## Implementation: COMBOS

- Collaboration with Christine Kuo, Joe DiStefano III (UCLA)
- Finds identifiable combinations for unidentifiable models
- http://biocyb1.cs.ucla.edu/combos

# Available software to test identifiability

- Differential Algebra Methods:
  - DAISY
    - Available online at http://www.dei.unipd.it/~pia
  - COMBOS
    - Soon available at http://biocyb1.cs.ucla.edu/combos
- Other methods:
  - GenSSI
    - Available online at http://www.iim.csic.es/~genssi/

## Acknowledgements

- Collaborators:
  - Seth Sullivant (NCSU)
  - Marisa Eisenberg (Univ. of Michigan)
  - Joe DiStefano III (UCLA)
  - Christine Kuo (Harvard)

## Summary

- Nec. and suff. for identifiable scaling reparam
- Suff. conditions for obtaining identifiability
- Algorithm to find identifiable functions in nonlinear models using Gröbner Bases
- COMBOS

#### Thank you for your attention!