IDENTIFIABILITY OF LINEAR COMPARTMENT MODELS

Anne Shiu
Texas A&M University

ICERM
15 November 2018
Outline

- Introduction: Linear compartment models
- Identifiability (via differential algebra)
- The singular locus

Joint work with
Elizabeth Gross, Heather Harrington, and Nicolette Meshkat

INTRODUCTION
Motivation: biological models

Drug input → Measured drug concentration → Drug exchange → Loss from organ → Loss from blood
Compartment model

\[
\begin{align*}
\dot{x}_1 &= -(k_{01} + k_{21})x_1 + k_{12}x_2 + u_1 \\
\dot{x}_2 &= k_{21}x_1 - (k_{02} + k_{12})x_2 \\
y &= x_1
\end{align*}
\]

Example: Linear 2-Compartment Model

Structural identifiability: Recover parameters $k_{ij}$ from perfect input-output data $u_1(t)$ and $y(t)$? (Bellman & Astrom 1970)
Identifiability via differential algebra\textsuperscript{1}:

Which models are identifiable?

\textsuperscript{1}Ljung and Glad 1994
**Input-output equations**

- **Setup:** a linear compartment model
- \( m \) = number of compartments
- **Input-output equation:** an equation that holds along any solution of the ODEs,
Input-output equations

- **Setup:** a linear compartment model
- \( m \) = number of compartments
- **Input-output equation:** an equation that holds along any solution of the ODEs, involving only input variables \( u_i \) and output variables \( y_i \) (and parameters \( k_{ij} \)), and their derivatives
**Input-output equations**

- Setup: a linear compartment model
- \( m = \) number of compartments
- **Input-output equation**: an equation that holds along any solution of the ODEs, involving only input variables \( u_i \) and output variables \( y_i \) (and parameters \( k_{ij} \)), and their derivatives
- Example, continued:

\[
y_1^{(2)} + (k_{01} + k_{02} + k_{12} + k_{21}) y_1' + (k_{01} k_{12} + k_{01} k_{02} + k_{02} k_{21}) y_1 = (k_{02} + k_{12}) u_1
\]
**Input-output equations**

- **Setup:** a linear compartment model
- \( m \) = number of compartments
- **Input-output equation:** an equation that holds along any solution of the ODEs, involving only input variables \( u_i \) and output variables \( y_i \) (and parameters \( k_{ij} \)), and their derivatives
- Example, continued:

\[
\begin{align*}
y_1^{(2)} + (k_{01} + k_{02} + k_{12} + k_{21}) y_1' + (k_{01}k_{12} + k_{01}k_{02} + k_{02}k_{21}) y_1 &= (k_{02} + k_{12}) u_1
\end{align*}
\]

- **Input-output equations** come from the elimination ideal:

\[
\langle \text{differential eqns., output eqns. } y_i = x_j, \text{ their } m \text{ derivatives} \rangle \cap \mathbb{C}(k_{ij})[u_i^{(k)}, y_i^{(k)}]
\]
Input-output equations, continued

\[ A = \begin{pmatrix} -k_{01} - k_{21} & k_{12} \\ k_{21} & -k_{02} - k_{12} \end{pmatrix} \quad \quad x'(t) = Ax(t) + u(t) \]

**Proposition** (Meshkat, Sullivant, Eisenberg 2015):
For a linear compartment model with input and output in compartment-1 only, the input-output equation is:

\[ \det(\partial I - A)y_1 = \det((\partial I - A)_{11})u_1. \]
Input-output equations, continued

\[ A = \begin{pmatrix} -k_{01} - k_{21} & k_{12} \\ k_{21} & -k_{02} - k_{12} \end{pmatrix} \]

\[ x'(t) = Ax(t) + u(t) \]

\[ \textbf{Proposition} \ (\text{Meshkat, Sullivant, Eisenberg 2015}): \]
For a linear compartment model with input and output in compartment-1 only, the input-output equation is:

\[ \det(\partial I - A)y_1 = \det((\partial I - A)_{11})u_1. \]

\[ \textbf{Proof uses Cramer's Rule and Laplace expansion} \]
Input-output equations, continued

\[
\begin{align*}
\det \left( \frac{\partial I}{\partial t} - A \right) y_1 &= \det \left( \left( \frac{\partial I}{\partial t} - A \right) u_1 \right) \\
\det \left( \frac{\partial I}{\partial t} + k_{01} + k_{12} - k_{12} + k_{21} - k_{23} + k_{32} - k_{23} \right) u_1 &= y_1(3) + (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}) y_1(2) + (k_{01} k_{12} + k_{01} k_{01} k_{23} + k_{01} k_{32} + k_{12} k_{23} + k_{21} k_{23} + k_{21} k_{32}) y_1' + (k_{01} k_{12} k_{23}) u_1(2) + (k_{12} k_{23}) u_1(1) + (k_{12} k_{23}) u_1(1).
\end{align*}
\]
**Input-output equations, continued**

\[
\begin{align*}
\det(\partial I - A)y_1 &= \det ((\partial I - A)_{11})u_1 \\
\det \begin{pmatrix}
  \frac{d}{dt} + k_{01} + k_{21} & -k_{12} & 0 \\
  -k_{21} & \frac{d}{dt} + k_{12} + k_{32} & -k_{23} \\
  0 & -k_{32} & \frac{d}{dt} + k_{23}
\end{pmatrix} y_1 \\
&= \det \begin{pmatrix}
  \frac{d}{dt} + k_{12} + k_{32} & -k_{23} \\
  -k_{32} & \frac{d}{dt} + k_{23}
\end{pmatrix} u_1
\end{align*}
\]
\[ \det(\partial I - A)y_1 = \det ((\partial I - A)_{11}) u_1 \]

\[
\begin{vmatrix}
\frac{d}{dt} + k_{01} + k_{21} & -k_{12} & 0 \\
-k_{21} & \frac{d}{dt} + k_{12} + k_{32} & -k_{23} \\
0 & -k_{32} & \frac{d}{dt} + k_{23}
\end{vmatrix} y_1 \\
= \det \begin{pmatrix}
\frac{d}{dt} + k_{12} + k_{32} & -k_{23} \\
-k_{32} & \frac{d}{dt} + k_{23}
\end{pmatrix} u_1
\]

... expands to the input-output equation:

\[
y_1^{(3)} + (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}) y_1^{(2)} + (k_{01} k_{12} + k_{01} k_{23} + k_{01} k_{32} + k_{12} k_{23} + k_{21} k_{23} + k_{21} k_{32}) y'_1 + (k_{01} k_{12} k_{23}) y_1
\]

\[
= u_1^{(2)} + (k_{12} + k_{23} + k_{32}) u'_1 + (k_{12} k_{23}) u_1 .
\]
\[ y_1^{(3)} + (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}) y_1^{(2)} \\
+ (k_{01} k_{12} + k_{01} k_{23} + k_{01} k_{32} + k_{12} k_{23} + k_{21} k_{23} + k_{21} k_{32}) y_1' + (k_{01} k_{12} k_{23}) y_1 \\
= u_1^{(2)} + (k_{12} + k_{23} + k_{32}) u_1' + (k_{12} k_{23}) u_1. \]
COEFFICIENTS OF INPUT-OUTPUT EQUATIONS

\[
y_1^{(3)} + (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}) y_1^{(2)} \\
+ (k_{01}k_{12} + k_{01}k_{23} + k_{01}k_{32} + k_{12}k_{23} + k_{21}k_{23} + k_{21}k_{32}) y_1' + (k_{01}k_{12}k_{23}) y_1 \\
= u_1^{(2)} + (k_{12} + k_{23} + k_{32}) u_1' + (k_{12}k_{23}) u_1.
\]

- coefficient of \( y_1^{(i)} \) corresponds to forests with \((3 - i)\) edges and \(\leq 1\) outgoing edge per compartment
- coefficient of \( u_1^{(i)} \) corresponds to \((n - i - 1)\)-edge forests:

\[\text{Thm 1: The coefficients correspond to forests in model.}\]
Identifiability

\[ y_1^{(3)} + (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}) y_1^{(2)} \]
\[ + (k_{01}k_{12} + k_{01}k_{23} + k_{01}k_{32} + k_{12}k_{23} + k_{21}k_{23} + k_{21}k_{32}) y_1' + (k_{01}k_{12}k_{23}) y_1 \]
\[ = u_1^{(2)} + (k_{12} + k_{23} + k_{32}) u_1' + (k_{12}k_{23}) u_1. \]

\( \triangleright \) (Generic, local) identifiability: can the parameters \( k_{ij} \) be recovered from coefficients of input-output equations?
Identifiability

\[
y_1^{(3)} + (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}) y_1^{(2)} \\
+ (k_{01}k_{12} + k_{01}k_{23} + k_{01}k_{32} + k_{12}k_{23} + k_{21}k_{23} + k_{21}k_{32}) y_1' + (k_{01}k_{12}k_{23}) y_1 \\
= u_1^{(2)} + (k_{12} + k_{23} + k_{32}) u_1' + (k_{12}k_{23}) u_1 .
\]

- (Generic, local) identifiability: can the parameters \( k_{ij} \) be recovered from coefficients of input-output equations?

\[
\mathbb{R}^5 \rightarrow \mathbb{R}^5 \\
(k_{01}, k_{12}, k_{21}, k_{23}, k_{32}) \mapsto (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}, \ldots )
\]

- Solve directly, or use ...

- Proposition (Meshkat, Sullivant, Eisenberg 2015): Identifiable \( \iff \) Jacobian matrix of coefficient map has (full) rank = number of parameters
Identifiability

\[ y_1^{(3)} + (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}) y_1^{(2)} \]
\[ + (k_{01} k_{12} + k_{01} k_{23} + k_{01} k_{32} + k_{12} k_{23} + k_{21} k_{23} + k_{21} k_{32}) y_1' + (k_{01} k_{12} k_{23}) y_1 \]
\[ = u_1^{(2)} + (k_{12} + k_{23} + k_{32}) u_1' + (k_{12} k_{23}) u_1 . \]

➤ (Generic, local) identifiability: can the parameters \( k_{ij} \) be recovered from coefficients of input-output equations?

\[ \mathbb{R}^5 \rightarrow \mathbb{R}^5 \]
\[ (k_{01}, k_{12}, k_{21}, k_{23}, k_{32}) \mapsto (k_{01} + k_{12} + k_{21} + k_{23} + k_{32}, \ldots) \]

➤ Solve directly, or use ...

➤ Proposition (Meshkat, Sullivant, Eisenberg 2015):
Identifiable \( \Leftrightarrow \) Jacobian matrix of coefficient map has (full) rank = number of parameters \textit{generically}.
The singular locus
Definition

Focus on the non-identifiable parameters: the **singular locus** is where the Jacobian matrix of coefficient map is rank-deficient.

Example, continued:

The equation of the singular locus is:

\[ \det \text{Jac} = k_{12}^2 k_{21} k_{23} = 0. \]
Identifiable submodels

- Motivation: drug targets
- Thm 2: Let $\mathcal{M}$ be an identifiable linear compartment model, with singular-locus equation $f$. Let $\tilde{\mathcal{M}}$ be obtained from $\mathcal{M}$ by deleting edges $\mathcal{I}$.

If $f \notin \langle k_{ji} \mid (i, j) \in \mathcal{I} \rangle$, then $\tilde{\mathcal{M}}$ is identifiable.
Identifiable submodels

- **Motivation:** drug targets
- **Thm 2:** Let $\mathcal{M}$ be an identifiable linear compartment model, with singular-locus equation $f$. Let $\tilde{\mathcal{M}}$ be obtained from $\mathcal{M}$ by deleting edges $\mathcal{I}$. If $f \notin \langle k_{ji} \mid (i, j) \in \mathcal{I} \rangle$, then $\tilde{\mathcal{M}}$ is identifiable.
- **Example:**

\[
f = k_{12}k_{14}k_{21}k_{32}(k_{12}k_{14} - k_{14}^2 - \ldots)(k_{12}k_{23} + k_{12}k_{43} + k_{32}k_{43}) .
\]
**Motivation:** drug targets

**Thm 2:** Let $\mathcal{M}$ be an identifiable linear compartment model, with singular-locus equation $f$. Let $\tilde{\mathcal{M}}$ be obtained from $\mathcal{M}$ by deleting edges $\mathcal{I}$. If $f \notin \langle k_{ji} \mid (i, j) \in \mathcal{I} \rangle$, then $\tilde{\mathcal{M}}$ is identifiable.

**Example:**

\[
f = k_{12}k_{14}k_{21}^2k_{32}(k_{12}k_{14} - k_{14}^2 - \ldots)(k_{12}k_{23} + k_{12}k_{43} + k_{32}k_{43}) .
\]

**Converse is false:** deleting $k_{12}$ and $k_{23}$ is identifiable!
**Thm 3:**

- The singular-locus equation for the Cycle model is
  \[ k_{32}k_{43} \ldots k_{n,n-1}k_{1,n} \prod_{2 \leq i < j \leq n} (k_{i+1,i} - k_{j+1,j}) \].

- The singular-locus equation for the Mammillary model is
  \[ k_{12}k_{13} \ldots k_{1,n} \prod_{2 \leq i < j \leq n} (k_{1i} - k_{1j})^2 \].
CATENARY (PATH) MODELS

Conjecture:
For catenary models, the exponents in the singular-locus equation generalize the pattern above.
Conjecture: For catenary models, the exponents in the singular-locus equation generalize the pattern above.
Tree conjecture

Conj.: (Hoch, Sweeney, Tung) For tree models, the exponents in the singular-locus equation generalize the pattern above.

\[
(2+1)+1 = 4
\]

\[
2+1 = 3
\]
**Tree Conjecture**

Conj.: (Hoch, Sweeney, Tung) For tree models, the exponents in the singular-locus equation generalize the pattern above.
Identifiable submodels (again)

- **Thm 4**: Let \( \tilde{\mathcal{M}} \) be obtained by:
  - adding a leak to a strongly connected model \( \mathcal{M} \) with no leaks, or
  - deleting the leak from a strongly connected model \( \mathcal{M} \) with input, output, and leak in one compartment.

Then, if \( \mathcal{M} \) is identifiable, then so is \( \tilde{\mathcal{M}} \).

\[ ^2 \text{Can delete edges without making the singular-locus equation } = 0. \]
**Identifiable submodels (again)**

- **Thm 4:** Let $\tilde{M}$ be obtained by:
  - adding a *leak* to a strongly connected model $M$ with *no* leaks, or
  - deleting the *leak* from a strongly connected model $M$ with input, output, and leak in *one* compartment.

Then, if $M$ is identifiable, then so is $\tilde{M}$.

<table>
<thead>
<tr>
<th><strong>Operation</strong></th>
<th><strong>Preserves identifiability?</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Add input</td>
<td>Yes</td>
</tr>
<tr>
<td>Add output</td>
<td>Yes</td>
</tr>
<tr>
<td>Add <em>leak</em></td>
<td>Not always (and see above)</td>
</tr>
<tr>
<td>Add edge</td>
<td>Not always</td>
</tr>
<tr>
<td>Delete input</td>
<td>Not always</td>
</tr>
<tr>
<td>Delete output</td>
<td>Not always</td>
</tr>
<tr>
<td>Delete <em>leak</em></td>
<td>Open (and see above)</td>
</tr>
<tr>
<td>Delete edge</td>
<td>Not always (recall Thm 2(^2))</td>
</tr>
</tbody>
</table>

\(^2\)Can delete edges *without* making the singular-locus equation $= 0$. 
**Future work**

Nonlinear models

---

*From Processive phosphorylation: mechanism and biological importance, Patwardhan and Miller, *Cell Signal.* 2007.*
The **singular locus** is an interesting mathematical object that can help us answer the question, *which linear compartment models are identifiable?*
Thank you.
Identifiability degree

- The identifiability degree of a model is the number of parameter vectors that match (generic) input-output data.
Identifiability degree

- the **identifiability degree** of a model is the number of parameter vectors that match (generic) input-output data.

- **Proposition** (Cobelli, Lepschy, Romanin Jacur 1979)

<table>
<thead>
<tr>
<th>Model</th>
<th>Identifiability degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catenary (path)</td>
<td>1</td>
</tr>
<tr>
<td>Mammillary (star)</td>
<td>$(n - 1)!$</td>
</tr>
</tbody>
</table>

- **Thm 5**

<table>
<thead>
<tr>
<th>Model</th>
<th>Identifiability degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle</td>
<td>$(n - 1)!$</td>
</tr>
</tbody>
</table>