Identifiability of Linear Compartmental Tree Models

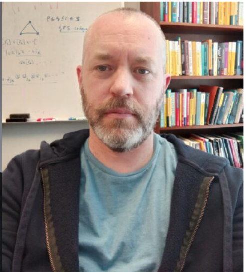
Cash Bortner

California State University, Stanislaus
cbortner@csustan.edu

New Directions in Algebraic Statistics
Institute for Mathematical and Statistical Innovation
July 24, 2025

Collaborators



Dr. Elizabeth Gross
U. of Hawai'i at Mānoa

Dr. Nicolette Meshkat
Santa Clara University

Dr. Anne Shiu
Texas A&M University

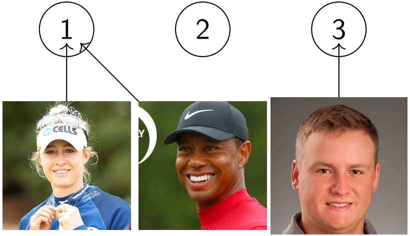
Dr. Seth Sullivant
North Carolina State U.

Motivating Example

"Compartments"

1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

Motivating Example



Nelly Korda   Tiger Woods   Not Nelly or Tiger

1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

Motivating Example

1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

Moving between compartments

Motivating Example

$1 :=$ Good Golfers
$2 :=$ Average Golfers
$3 :=$ Bad Golfers

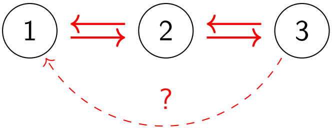
Moving between compartments

Motivating Example

1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

Moving between compartments

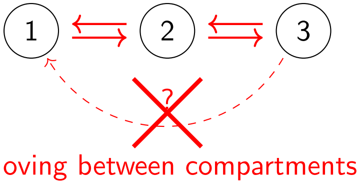
Motivating Example



Moving between compartments

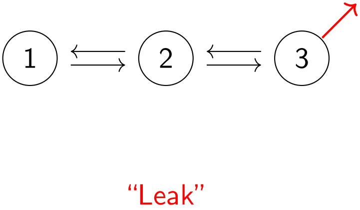
1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

Motivating Example



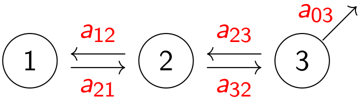
$1 :=$ Good Golfers
$2 :=$ Average Golfers
$3 :=$ Bad Golfers

Motivating Example



1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

Motivating Example

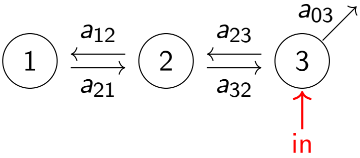


1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

"Flow Rate Parameters"

$a_{(destination)(source)}$

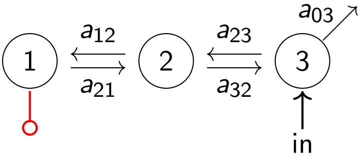
Motivating Example



"Input into the system"

1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

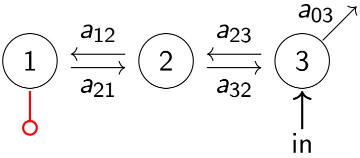
Motivating Example



"Measured Compartment"

1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

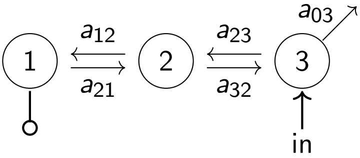
Motivating Example



"**Output** Compartment"

1 := Good Golfers
2 := Average Golfers
3 := Bad Golfers

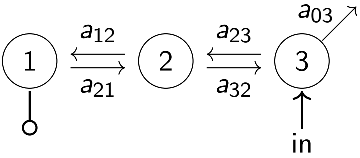
Motivating Example



Linear Compartmental Model

$$\mathcal{M} = (G, In, Out, Leak)$$
$$= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).$$

Motivating Example



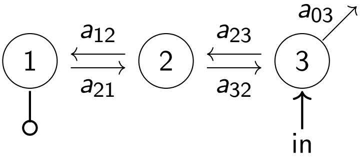
Linear Compartmental Model

$$\mathcal{M} = (G, In, Out, Leak)$$
$$= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).$$

Motivating Question: Identifiability

Given information about the input and output compartment, can we **find** all flow rate parameters?

Motivating Example



Linear Compartmental Model

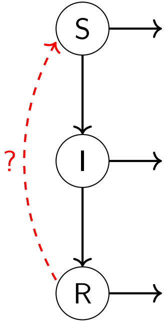
$$\mathcal{M} = (G, In, Out, Leak)$$
$$= (\text{Cat}_3, \{3\}, \{1\}, \{3\}).$$

Motivating Question: Identifiability

Given information about the input and output compartment, can we **identify** all flow rate parameters?

Compartmental Models in the Wild

- SIR Model for spread of a virus in Epidemiology
- SIV Model for vaccine efficiency in Epidemiology
- Modeling Pharmacokinetics for absorption, distribution, metabolism, and excretion in the blood
- Modeling different biological systems

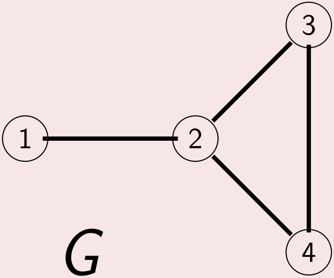


Background - Graph Theory

Definition (by Example)

An *undirected graph* $G = (V, E)$ consists of two sets:

- vertices V, e.g. $\{1, 2, 3, 4\}$
- edges E, e.g. $\underbrace{\{\{1, 2\}, \{2, 3\}, \{2, 4\}, \{4, 3\}\}}_{\text{unordered pairs of vertices}}$

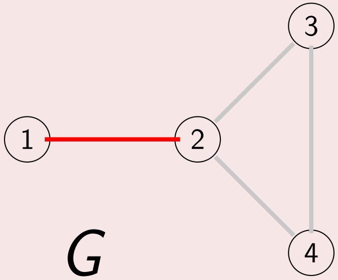


Background - Graph Theory

Definition (by Example)

An *undirected graph* $G = (V, E)$ consists of two sets:

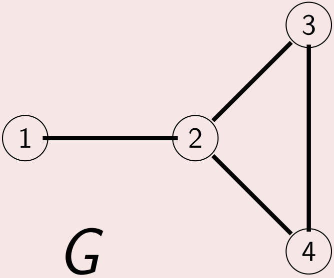
- vertices V, e.g. $\{1, 2, 3, 4\}$
- edges E, e.g. $\underbrace{\{\{\mathbf{1,2}\}, \{2, 3\}, \{2, 4\}, \{4, 3\}\}}_{\text{unordered pairs of vertices}}$



Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

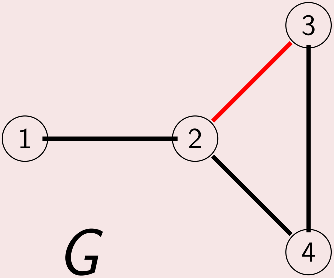


Does this graph contain a cycle?

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

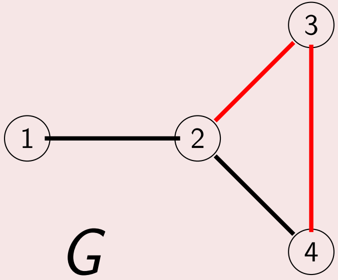


Does this graph contain a cycle?

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

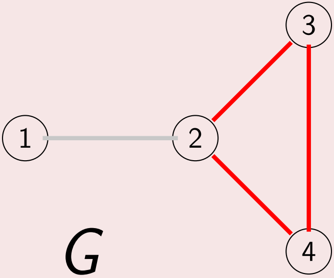


Does this graph contain a cycle?

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

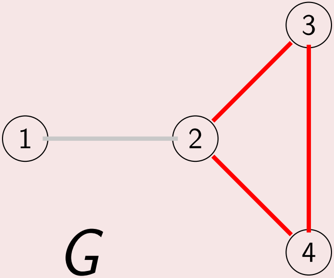


Does this graph contain a cycle?

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

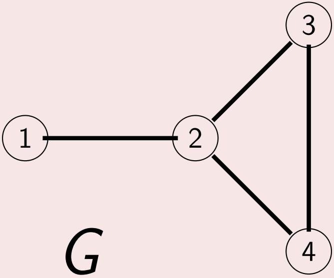


Does this graph contain a cycle? YES!

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

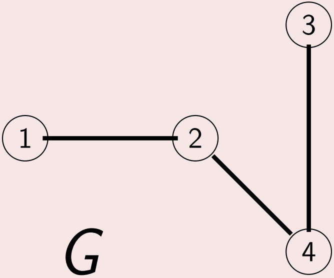


How could we make this graph a forest?

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

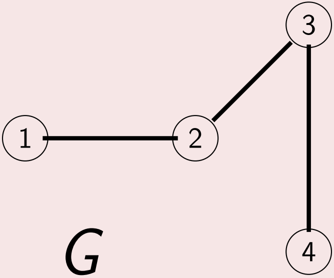


How could we make this graph a forest? Remove {2,3}

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

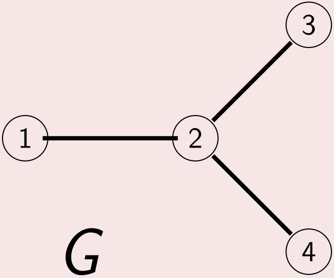


How could we make this graph a forest? Remove {2,4}

Background - Graph Theory

Definition (by Example)

An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.

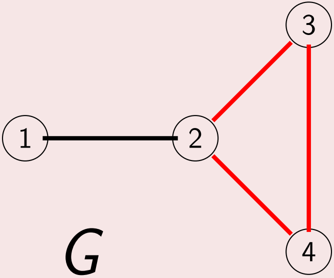


How could we make this graph a forest? Remove {4,3}

Background - Graph Theory

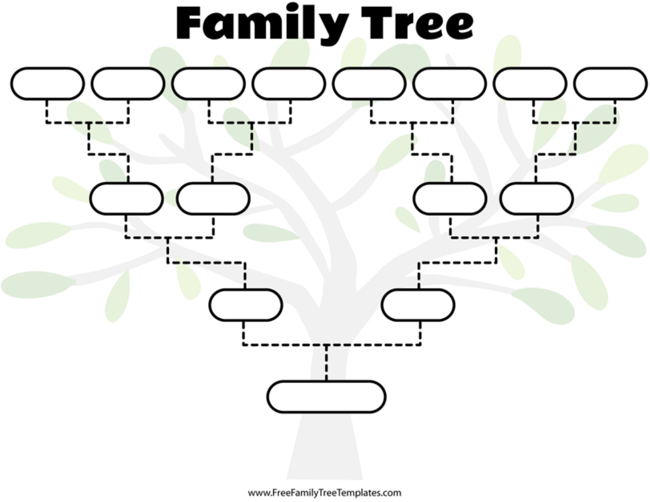
Definition (by Example)

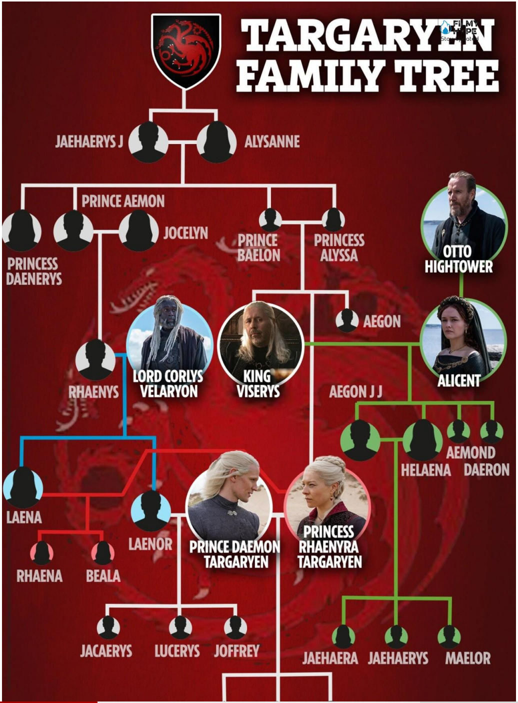
An undirected graph $G = (V, E)$ is a *forest* if it contains no cycles, i.e. a path starting and ending at the same vertex without using an edge twice.



Note: A forest is a collection of trees (connected graph with no cycles)!

Family Trees



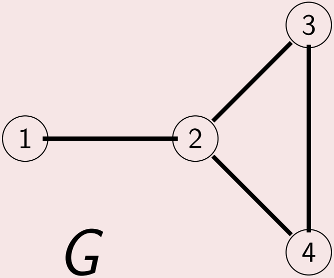

TARGARYEN FAMILY TREE
JAEHAERYS J
ALYSANNE
PRINCE AEMON
JOCELYN
PRINCESS DAENERYS
PRINCE BAELON
PRINCESS ALYSSA
OTTO HIGHTOWER
AEGON
RHAENYS
LORD CORLYS VELARYON
KING VISERYS
ALICENT
AEGON J J
AEMOND
HELAENA
DAERON
LAENA
LAENOR
PRINCE DAEMON TARGARYEN
PRINCESS RHAENYRA TARGARYEN
RHAENA
BEALA
JACAERYS
LUCERYS
JOFFREY
JAEHAERA
JAEHAERYS
MAELOR

Background - Graph Theory

Definition (by Example)

An *undirected graph* $G = (V, E)$ consists of two sets:

- vertices V, e.g. $\{1, 2, 3, 4\}$
- edges E, e.g. $\underbrace{\{\{1,2\},\{2,3\},\{2,4\},\{4,3\}\}}_{\text{unordered pairs of vertices}}$

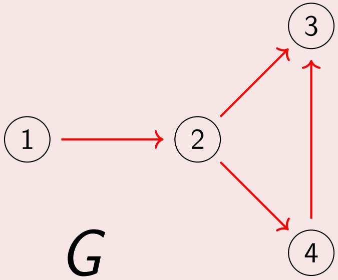


Background - Graph Theory

Definition (by Example)

An *directed* *graph* $G = (V, E)$ consists of two sets:

- vertices V, e.g. $\{1, 2, 3, 4\}$
- edges E, e.g. $\underbrace{\{(1, 2), (2, 3), (2, 4), (4, 3)\}}_{\text{ordered pairs of vertices}}$

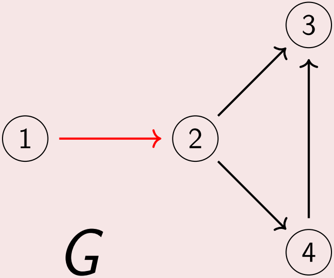


Background - Graph Theory

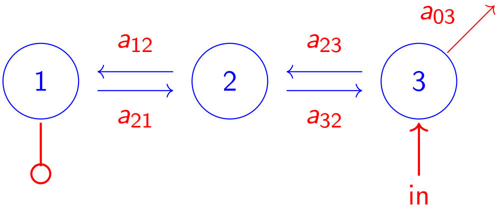
Definition (by Example)

An *directed* *graph* $G = (V, E)$ consists of two sets:

- vertices V, e.g. $\{1, 2, 3, 4\}$
- edges E, e.g. $\underbrace{\{(1, 2), (2, 3), (2, 4), (4, 3)\}}_{\text{ordered pairs of vertices}}$



Motivating Example



$$\begin{aligned}\mathcal{M} &= (G, In, Out, Leak) \\ &= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).\end{aligned}$$

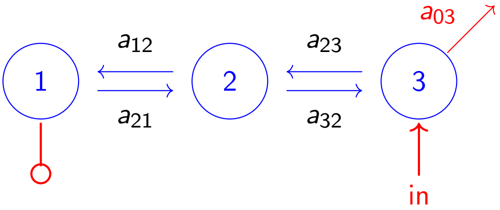
Motivating Example

$$\mathcal{M} = (G, In, Out, Leak)$$
$$= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).$$

With

$$\mathrm{Cat}_3 = (\{1, 2, 3\}, \{(1, 2), (2, 1), (2, 3), (3, 2)\})$$

Motivating Example

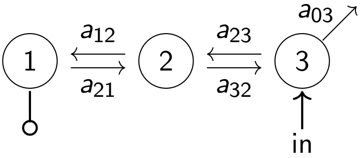


$$\mathcal{M} = (G, In, Out, Leak) = (\text{Cat}_3, \{3\}, \{1\}, \{3\}).$$

With

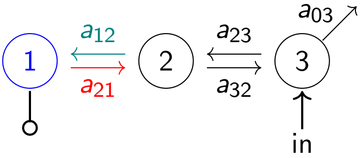
$$\text{Cat}_3 = (\{1,2,3\}, \{(1,2),(2,1),(2,3),(3,2)\})$$

Motivating Example



$$\begin{aligned}\mathcal{M} &= (G, In, Out, Leak)\\ &= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).\end{aligned}$$

Motivating Example

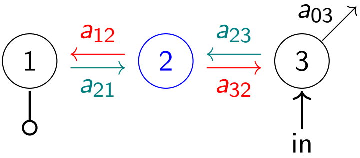


$$\mathcal{M} = (G, In, Out, Leak) = (\text{Cat}_3, \{3\}, \{1\}, \{3\}).$$

ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$x_1'(t) = -a_{21}x_1(t) \qquad +a_{12}x_2(t)$$

Motivating Example

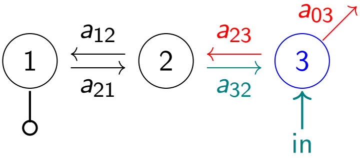


$$\mathcal{M} = (G, In, Out, Leak) = (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).$$

ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$\begin{aligned} x_1'(t) &= -a_{21}x_1(t) + a_{12}x_2(t) \\ x_2'(t) &= a_{21}x_1(t) - (a_{12} + a_{32})x_2(t) + a_{23}x_3(t) \end{aligned}$$

Motivating Example

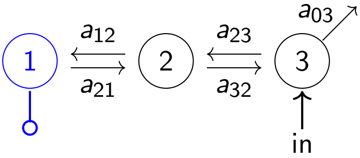


$$\mathcal{M} = (G, In, Out, Leak) = (\text{Cat}_3, \{3\}, \{1\}, \{3\}).$$

ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$\begin{aligned}
x_1'(t) &= -a_{21}x_1(t) + a_{12}x_2(t) \\
x_2'(t) &= a_{21}x_1(t) - (a_{12} + a_{32})x_2(t) + a_{23}x_3(t) \\
x_3'(t) &= a_{32}x_2(t) - (a_{03} + a_{23})x_3(t) + u_3(t)
\end{aligned}$$

Motivating Example



$$\mathcal{M} = (G, In, Out, Leak) = (\text{Cat}_3, \{3\}, \{1\}, \{3\}).$$

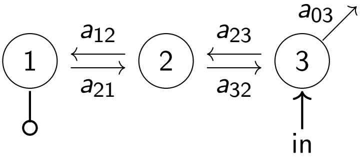
ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$\begin{aligned}
x_1'(t) &= -a_{21}x_1(t) + a_{12}x_2(t) \\
x_2'(t) &= a_{21}x_1(t) - (a_{12} + a_{32})x_2(t) + a_{23}x_3(t) \\
x_3'(t) &= a_{32}x_2(t) - (a_{03} + a_{23})x_3(t) + u_3(t)
\end{aligned}$$

with

$$y_1(t) = x_1(t).$$

Motivating Example



$$\begin{aligned}\mathcal{M} &= (G, In, Out, Leak)\\ &= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).\end{aligned}$$

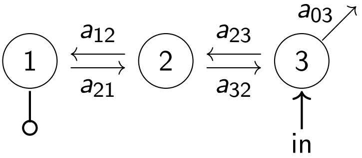
ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{pmatrix} = \underbrace{\begin{pmatrix} -a_{21} & a_{12} & 0 \\ a_{21} & -a_{12} - a_{32} & a_{23} \\ 0 & a_{32} & -a_{03} - a_{23} \end{pmatrix}}_{\text{compartmental matrix } A} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

$$y_1(t) = x_1(t).$$

Motivating Example



$$\begin{aligned}\mathcal{M} &= (G, In, Out, Leak)\\ &= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).\end{aligned}$$

ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$\begin{aligned}x_1'(t) &= -a_{21}x_1(t) + a_{12}x_2(t)\\ x_2'(t) &= a_{21}x_1(t) - (a_{12}+a_{32})x_2(t) + a_{23}x_3(t)\\ x_3'(t) &= a_{32}x_2(t) - (a_{03}+a_{23})x_3(t) + u_3(t)\end{aligned}$$

with

$$y_1(t) = x_1(t).$$

Goal: Identify the parameters a_{ji} from the measurable variables.

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Differential Substitution/Elimination:

$$
\begin{aligned}
\underline{x_1'(t)} &= -a_{21}\underline{x_1(t)} + a_{12}x_2(t) \\
x_2'(t) &= a_{21}\underline{x_1(t)} - (a_{12} + a_{32})x_2(t) + a_{23}x_3(t) \\
x_3'(t) &= a_{32}x_2(t) - (a_{03} + a_{23})x_3(t) + u_3(t)
\end{aligned}
$$

with

$$\underbrace{\underline{y_1(t) = x_1(t)}}_{y_1'(t) = x_1'(t)}$$

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Differential Substitution/Elimination:

$$\begin{aligned}
y_1'(t) &= -a_{21}y_1(t) && +a_{12}\underline{x_2(t)} \\
x_2'(t) &= \quad a_{21}y_1(t) && -(a_{12}+a_{32})x_2(t) && +a_{23}x_3(t) \\
x_3'(t) &= && \quad a_{32}x_2(t) && -(a_{03}+a_{23})x_3(t) && +u_3(t)
\end{aligned}$$

with

$$\underbrace{y_1(t) = x_1(t)}_{y_1'(t)=x_1'(t)}$$

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Differential Substitution/Elimination:

$$
\begin{aligned}
x_2(t) &= \frac{1}{a_{12}} y_1'(t) + \frac{a_{21}}{a_{12}} y_1(t) \\
\underline{x_2'(t)} &= a_{21} y_1(t) - (a_{12} + a_{32}) \underline{x_2(t)} + a_{23} x_3(t) \\
x_3'(t) &= a_{32} \underline{x_2(t)} - (a_{03} + a_{23}) x_3(t) + u_3(t)
\end{aligned}
$$

with

$$
\underbrace{y_1(t) = x_1(t)}_{y_1'(t) = x_1'(t)}
$$

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Differential Substitution/Elimination:

$$x_2(t) = \frac{1}{a_{12}}y_1'(t) + \frac{a_{21}}{a_{12}}y_1(t)$$

$$\frac{1}{a_{12}}y_1''(t) + \frac{a_{21}}{a_{12}}y_1'(t) = a_{21}y_1(t) - (a_{12}+a_{32})\left(\frac{1}{a_{12}}y_1'(t) + \frac{a_{21}}{a_{12}}y_1(t)\right) + a_{23}\underline{x_3(t)}$$

$$x_3'(t) = a_{32}\left(\frac{1}{a_{12}}y_1'(t) + \frac{a_{21}}{a_{12}}y_1(t)\right) - (a_{03}+a_{23})x_3(t) + u_3(t)$$

with

$$\underbrace{y_1(t) = x_1(t)}_{y_1'(t)=x_1'(t)}$$

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Differential Substitution/Elimination:

$$
\begin{aligned}
x_2(t) &= \frac{1}{a_{12}} y_1'(t) + \frac{a_{21}}{a_{12}} y_1(t) \\
x_3(t) &= \frac{1}{a_{12}a_{23}} y_1''(t) + \left(\frac{a_{21} + a_{12} + a_{21}}{a_{12}a_{23}} \right) y_1'(t) + \left(\frac{a_{12}a_{21} + a_{21}a_{32}}{a_{12}a_{23}} - \frac{a_{21}}{a_{23}} \right) y_1(t) \\
\underline{x_3'(t)} &= \frac{a_{32}}{a_{12}} y_1'(t) + \frac{a_{21}a_{32}}{a_{12}} y_1(t) - (a_{03} + a_{23}) \underline{x_3(t)} + u_3(t)
\end{aligned}
$$

with

$$
\underbrace{y_1(t) = x_1(t)}_{y_1'(t) = x_1'(t)}
$$

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Differential Substitution/Elimination:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})y_1'' + (a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})y_1' + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$

an ODE in only the measurable variables and the parameters:

Input/Output Equation

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Differential Substitution/Elimination:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})y_1'' + (a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})y_1' + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$

an ODE in only the measurable variables and the parameters:

Input/Output Equation

Theorem (Meshkat, Sullivant, Eisenburg, 2015)

Can also be done via Cramer's rule:

$$\det(\partial I - A)y_1 = \overbrace{\det(\partial I - A)^{3,1}}^{\text{remove row 3, col. 1}} u_3$$

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Evaluate at many time instances: $t_1, t_2, t_3, \ldots, t_m$

$$\begin{aligned}
y_1^{(3)}(t_1) + c_2 y_1''(t_1) + c_1 y_1'(t_1) + c_0 y_1(t_1) &= d_0 u_3(t_1) \\
y_1^{(3)}(t_2) + c_2 y_1''(t_2) + c_1 y_1'(t_2) + c_0 y_1(t_2) &= d_0 u_3(t_2) \\
y_1^{(3)}(t_3) + c_2 y_1''(t_3) + c_1 y_1'(t_3) + c_0 y_1(t_3) &= d_0 u_3(t_3) \\
&\vdots \\
y_1^{(3)}(t_m) + c_2 y_1''(t_m) + c_1 y_1'(t_m) + c_0 y_1(t_m) &= d_0 u_3(t_m)
\end{aligned}$$

and recover each coefficient c_2, c_1, c_0, d_0 uniquely.

Linear Compartmental Model Motivating Example

Goal: Identify the parameters a_{ji} from the measurable variables.

Consider the injectivity (invertibility) of the coefficient map:

Example (Continued)

For $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$, the input/output equation is:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})y_1'' + (a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})y_1' + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$

The *coefficient map* corresponding to $\mathcal{M}$ is the map from the space of parameters to the space of coefficients of the input-output equation:

$$\phi \colon \mathbb{R}^5 \to \mathbb{R}^4$$

$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

Identifiability Analysis: Structural vs. Practical

Overview

We want to recover (identify) parameters of ODE models from measured variables.

Identifiability Analysis: Structural vs. Practical

Overview

We want to recover (identify) parameters of ODE models from measured variables.

Structural vs. Practical

Identifiability Analysis: Structural vs. Practical

Overview

We want to recover (identify) parameters of ODE models from measured variables.

Structural vs. Practical

- *structural identifiability analysis*, is done *a priori*, assumes perfect conditions, and does not provide numerical parameter estimates

Identifiability Analysis: Structural vs. Practical

Overview

We want to recover (identify) parameters of ODE models from measured variables.

Structural vs. Practical

- *structural identifiability analysis*, is done *a priori*, assumes perfect conditions, and does not provide numerical parameter estimates
- *practical identifiability analysis*, is done *a posteriori*, assumes some error exists, and generally computes parameter estimates from measured data

Identifiability Analysis: Structural vs. Practical

Overview

We want to recover (identify) parameters of ODE models from measured variables.

Structural vs. Practical

- *structural identifiability analysis*, is done *a priori*, assumes perfect conditions, and does not provide numerical parameter estimates
- *practical identifiability analysis*, is done *a posteriori*, assumes some error exists, and generally computes parameter estimates from measured data

Remark

Structural identifiability is a **necessary** condition for practical identifiability.

Identifiability Analysis: Structural vs. Practical

Overview

We want to recover (identify) parameters of ODE models from measured variables.

Structural vs. Practical

- *structural identifiability analysis*, is done *a priori*, assumes perfect conditions, and does not provide numerical parameter estimates
- *practical identifiability analysis*, is done *a posteriori*, assumes some error exists, and generally computes parameter estimates from measured data

Remark

Structural identifiability is a **necessary** condition for practical identifiability.

Structural Identifiability Analysis: A two part problem

Structural Identifiability via the input-output equation

We consider structural identifiability as a two-step problem:

1. Find an **input/output equation** of the ODE system in terms of measurable variables in a way that's <u>intuitive</u> from the graph
2. Determine the **injectivity** (invertibility) of the coefficient map (i.e. can we recover the parameters)

Structural Identifiability Analysis: A two part problem

Structural Identifiability via the input-output equation

We consider structural identifiability as a two-step problem:

1. Find an **input/output equation** of the ODE system in terms of measurable variables in a way that's intuitive from the graph
2. Determine the **injectivity** (invertibility) of the coefficient map (i.e. can we recover the parameters)

Goal

We want to classify identifiability by the underlying **graph structure**.

Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

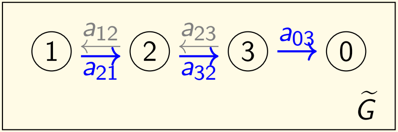
Definitions

A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$



Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Definitions

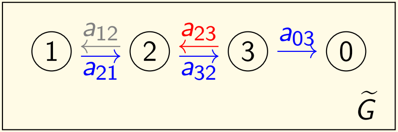
A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$

Can we swap $3 \to 2$ with any edge?



Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Definitions

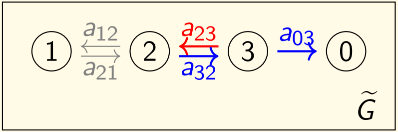
A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$

Can we swap $3 \to 2$ with any edge? $1 \to 2$



Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Definitions

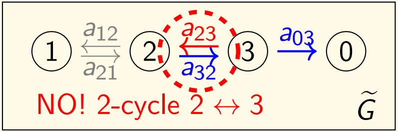
A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$

Can we swap $3 \to 2$ with any edge? $1 \to 2$



Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Definitions

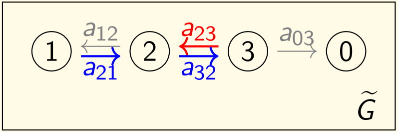
A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$

Can we swap $3 \to 2$ with any edge? $3 \to 0$



Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Definitions

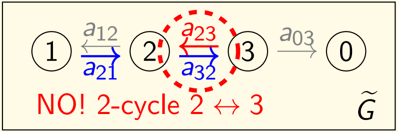
A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$

Can we swap $3 \to 2$ with any edge? $3 \to 0$



Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Definitions

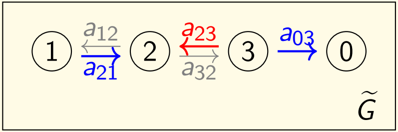
A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$

Can we swap $3 \to 2$ with any edge? $3 \to 2$



Novel Input-Output Equation Characterization

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The coefficients of the input-output equation of an LCM $(G, In, Out, Leak)$ can be generated by *incoming forests* on graphs related to G.

Definitions

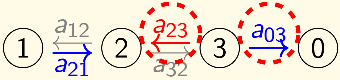
A directed graph H is called an *incoming forest* if

- no vertex has **more than one outgoing edge**, and
- its underlying undirected graph is a forest (**has no cycles**)

Example

The set of incoming forests with 3 edges on $\widetilde{G}$: $\mathcal{F}_3(\widetilde{G}) = \{\{1 \to 2, 2 \to 3, 3 \to 0\}\}$

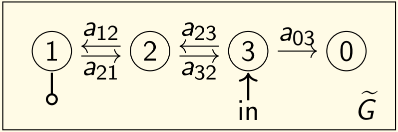
Can we swap $3 \to 2$ with any edge? $3 \to 2$



NO! double outgoing from 3 $\widetilde{G}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

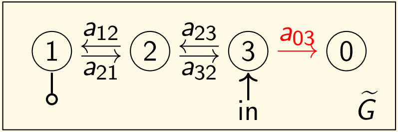
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(2)}$: Incoming forests with 1 edge

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

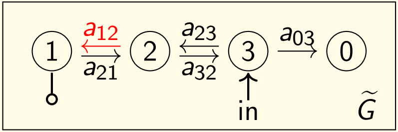
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(2)}$: Incoming forests with 1 edge

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

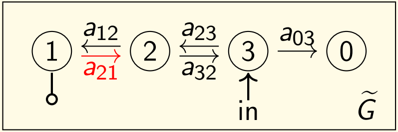
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(2)}$: Incoming forests with 1 edge

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

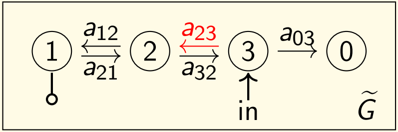
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(2)}$: Incoming forests with 1 edge

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

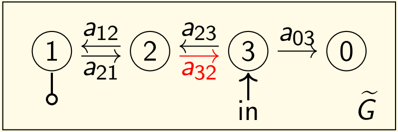
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(2)}$: Incoming forests with 1 edge

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

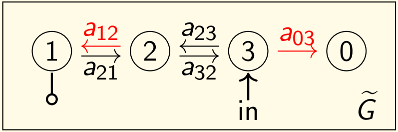
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(2)}$: Incoming forests with 1 edge

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

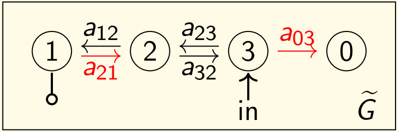
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(1)}$: Incoming forests with 2 edges

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

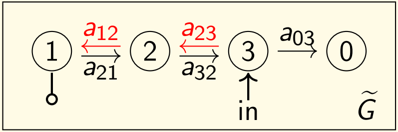
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(1)}$: Incoming forests with 2 edges

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

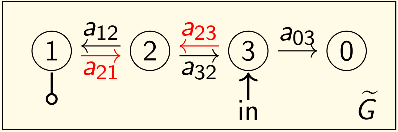
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(1)}$: Incoming forests with 2 edges

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

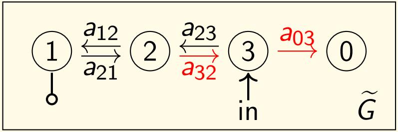
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(1)}$: Incoming forests with 2 edges

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

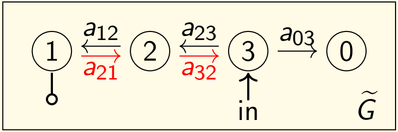
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(1)}$: Incoming forests with 2 edges

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

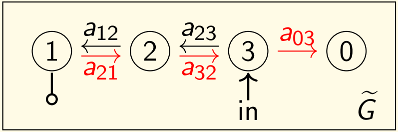
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(1)}$: Incoming forests with 2 edges

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

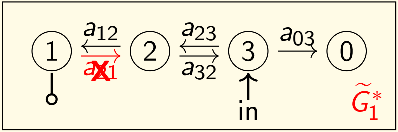
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

LHS coefficients: $y_1^{(0)}$: Incoming forests with 3 edges

Derivative	Coefficient
$y_1^{(3)}$	1
$y_1^{(2)}$	$a_{03} + a_{12} + a_{21} + a_{23} + a_{32}$
$y_1^{(1)}$	$a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32}$
$y_1^{(0)}$	$a_{03}a_{21}a_{32}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



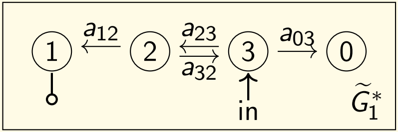
$\widetilde{G}_1^*$: Remove all edges **leaving** the output $\{1\}$.

RHS coeff:

Derivative	Coefficient
$u_3^{(0)}$	

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

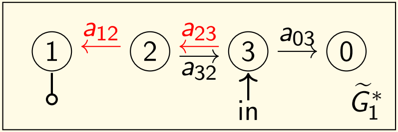
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

RHS coeff: Incoming forests with 2 edges AND a path from $\underbrace{in}_{3}$ to $\underbrace{out}_{1}$

Derivative	Coefficient
$u_3^{(0)}$	

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

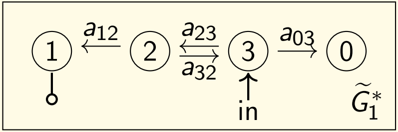
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

RHS coeff: Incoming forests with 2 edges AND a path from $\underbrace{in}_{3}$ to $\underbrace{out}_{1}$

Derivative	Coefficient
$u_3^{(0)}$	$a_{12}a_{23}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

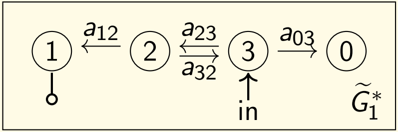
- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

RHS coefficients: Incoming forests with **1 edge** AND a path from 3 to 1?

Derivative	Coefficient
$u_3^{(0)}$	$a_{12}a_{23}$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, we have



Incoming Forests:

- no vertex has more than one outgoing edge, and
- its underlying undirected graph is a forest (no cycles)

RHS coefficients: Incoming forests with **3 edges** AND a path from 3 to 1?

Derivative	Coefficient
$u_3^{(0)}$	$a_{12}a_{23}$

Number of Coefficients

Corollary ($, Gross, Meshkat, Shiu, Sullivant, [2023])

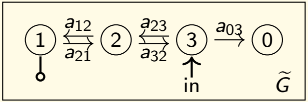
Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where G is strongly connected and $|V_G| = n$. Then the **number of coefficients** in the input/output equation:

$$\#\text{ on LHS} = \begin{cases} n & \text{if } \#\text{Leaks} \geq 1 \\ n-1 & \text{if } \#\text{Leaks} = 0 \end{cases}, \quad \#\text{ on RHS} = \begin{cases} n-1 & \text{if } in = out \\ n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) & \text{if } in \neq out. \end{cases}$$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, the input/output equation is:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})y_1'' + (a_{03}a_{12} + a_{03}a_{21}$$
$$a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})y_1' + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$



$|Leak| = 1$

Number of Coefficients

Corollary ($, Gross, Meshkat, Shiu, Sullivant, [2023])

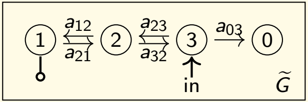
Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where G is strongly connected and $|V_G| = n$. Then the **number of coefficients** in the input/output equation:

$$\# \text{ on LHS} = \begin{cases} n & \text{if } \#\text{Leaks} \geq 1 \\ n-1 & \text{if } \#\text{Leaks} = 0 \end{cases}, \quad \# \text{ on RHS} = \begin{cases} n-1 & \text{if } in = out \\ n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) & \text{if } in \neq out. \end{cases}$$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, the input/output equation is:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})y_1'' + (a_{03}a_{12} + a_{03}a_{21}$$
$$a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})y_1' + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$



$|Leak| = 1 \Rightarrow \# \text{ on LHS} = 3$

Number of Coefficients

Corollary ($, Gross, Meshkat, Shiu, Sullivant, [2023])

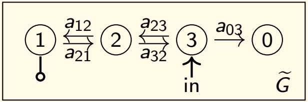
Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where G is strongly connected and $|V_G| = n$. Then the **number of coefficients** in the input/output equation:

$$\#\text{ on LHS} = \begin{cases} n & \text{if } \#\text{Leaks} \geq 1 \\ n-1 & \text{if } \#\text{Leaks} = 0 \end{cases}, \quad \#\text{ on RHS} = \begin{cases} n-1 & \text{if } in = out \\ n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) & \text{if } in \neq out. \end{cases}$$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, the input/output equation is:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})y_1'' + (a_{03}a_{12} + a_{03}a_{21} \\ a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})y_1' + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$



$|Leak| = 1 \Rightarrow \#\text{ on LHS} = 3$

$\mathrm{dist}(3, 1) = 2$

Number of Coefficients

Corollary ($, Gross, Meshkat, Shiu, Sullivant, [2023])

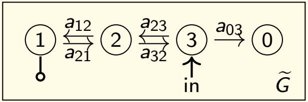
Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where G is strongly connected and $|V_G| = n$. Then the **number of coefficients** in the input/output equation:

$$\#\text{ on LHS} = \begin{cases} n & \text{if } \#\text{Leaks} \geq 1 \\ n-1 & \text{if } \#\text{Leaks} = 0 \end{cases}, \quad \#\text{ on RHS} = \begin{cases} n-1 & \text{if } in = out \\ n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) & \text{if } in \neq out. \end{cases}$$

Example

For $\mathcal{M} = (G, \{3\}, \{1\}, \{3\})$, the input/output equation is:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})y_1'' + (a_{03}a_{12} + a_{03}a_{21}$$
$$a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})y_1' + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$



$|Leak| = 1 \Rightarrow \#$ on LHS $= 3$

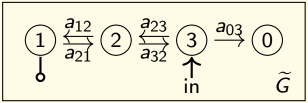
$\mathrm{dist}(3,1) = 2 \Rightarrow \#$ on RHS $= 3-2 = 1$

Example

Example (Continued)

For $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$, the input/output equation is:

$$y_1^{(3)} + (a_{03} + a_{12} + a_{21} + a_{23} + a_{32})\ddot{y}_1 + (a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})\dot{y}_1 + (a_{03}a_{21}a_{32})y_1 = (a_{12}a_{23})u_3.$$



on LHS = 3

on RHS = 1

The *coefficient map* corresponding to $\mathcal{M}$ is the map from the space of parameters to the space of coefficients of the input-output equation:

$$\phi\colon \mathbb{R}^5 \to \mathbb{R}^4$$

$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

Identifiability

Definition

A model is

- *globally identifiable* if its coefficient map is one-to-one
- *locally identifiable* if its coefficient map is finite-to-one
- *unidentifiable* if its coefficient map is infinite-to-one

Identifiability

Definition

A model is
- *globally identifiable* if its coefficient map is one-to-one
- *locally identifiable* if its coefficient map is finite-to-one
- *unidentifiable* if its coefficient map is infinite-to-one

Example (Continued)

For $\mathcal{M} = (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$, the coefficient map corresponding to $\mathcal{M}$:

$$\phi\colon \mathbb{R}^5 \to \mathbb{R}^4$$

$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

Is $\mathcal{M}$ globally/locally identifiable?

Identifiability

Definition

A model is

- *globally identifiable* if its coefficient map is one-to-one
- *locally identifiable* if its coefficient map is finite-to-one
- *unidentifiable* if its coefficient map is infinite-to-one

Example (Continued)

For $\mathcal{M} = (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$, the coefficient map corresponding to $\mathcal{M}$:

$$\phi\colon \mathbb{R}^5 \to \mathbb{R}^4$$

$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

Is $\mathcal{M}$ globally/locally identifiable? $\longleftrightarrow$ Is ϕ one/finite-to-one?

Identifiability

Definition

A model is

- *globally identifiable* if its coefficient map is one-to-one
- *locally identifiable* if its coefficient map is finite-to-one
- *unidentifiable* if its coefficient map is infinite-to-one

Example (Continued)

For $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$, the coefficient map corresponding to $\mathcal{M}$:

$$\phi\colon \mathbb{R}^5 \to \mathbb{R}^4$$

$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

Is $\mathcal{M}$ globally/locally identifiable? $\longleftrightarrow$ Is ϕ one/finite-to-one? NO!!!

Identifiability

Definition

A model is

- *globally identifiable* if its coefficient map is one-to-one
- *locally identifiable* if its coefficient map is finite-to-one
- *unidentifiable* if its coefficient map is infinite-to-one

Example (Continued)

For $\mathcal{M} = (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$, the coefficient map corresponding to $\mathcal{M}$:

$$\phi \colon \mathbb{R}^5 \to \mathbb{R}^4$$

$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

$\mathcal{M}$ is **unidentifiable** since there are more parameters than coefficients!

Identifiability

Definition

A model is

- *globally identifiable* if its coefficient map is one-to-one
- *locally identifiable* if its coefficient map is finite-to-one
- *unidentifiable* if its coefficient map is infinite-to-one

Example (Continued)

For $\mathcal{M} = (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$, the coefficient map corresponding to $\mathcal{M}$:

$$\phi \colon \mathbb{R}^5 \to \mathbb{R}^4$$

$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

$\mathcal{M}$ is **unidentifiable** since there are more parameters than coefficients!

Sufficient Condition for Unidentifiability

Proposition (Sufficient condition for unidentifiability)

A model $\mathcal{M}$ is *unidentifiable* if $\#$ parameters $>$ $\#$ coefficients.

Sufficient Condition for Unidentifiability

Proposition (Sufficient condition for unidentifiability)

A model $\mathcal{M}$ is *unidentifiable* if $\#$ parameters $>$ $\#$ coefficients.

Corollary ($, Gross, Meshkat, Shiu, Sullivant, [2023])

Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where G is *strongly connected* and

$$L = \begin{cases} 0 & \text{if } \#Leak = 0 \\ 1 & \text{if } \#Leak \geq 1 \end{cases} \quad \text{and} \quad d = \begin{cases} 1 & \text{if } \mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0 \\ \mathrm{dist}(\mathrm{in}, \mathrm{out}) & \text{if } \mathrm{dist}(\mathrm{in}, \mathrm{out}) \neq 0. \end{cases}$$

Then $\mathcal{M}$ is **unidentifiable** if $\underbrace{|Leak| + |E_G|}_{\#\text{ parameters}} > \underbrace{2|V_G| - L - d}_{\#\text{ coefficients}}$.

Sufficient Condition for Unidentifiability

Proposition (Sufficient condition for unidentifiability)

A model $\mathcal{M}$ is *unidentifiable* if $\#$ parameters $>$ $\#$ coefficients.

Corollary ($, Gross, Meshkat, Shiu, Sullivant, [2023])

Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where G is *strongly connected* and

$$L = \begin{cases} 0 & \text{if } \# Leak = 0 \\ 1 & \text{if } \# Leak \geq 1 \end{cases} \quad \text{and} \quad d = \begin{cases} 1 & \text{if } \mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0 \\ \mathrm{dist}(\mathrm{in}, \mathrm{out}) & \text{if } \mathrm{dist}(\mathrm{in}, \mathrm{out}) \neq 0. \end{cases}$$

Then $\mathcal{M}$ is **unidentifiable** if $\underbrace{|Leak| + |E_G|}_{\#\text{ parameters}} > \underbrace{2|V_G| - L - d}_{\#\text{ coefficients}}$.

Remark [not a necessary condition]

There are models that with $\#$parameters $\leq$ $\#$coefficients that are still unidentifiable!

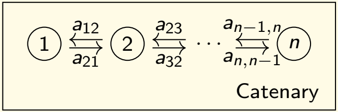
Tree Models

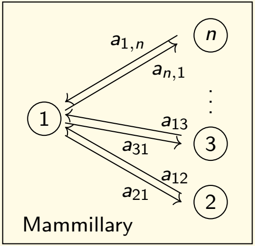
Definition

A (bidirectional) *tree model* $\mathcal{M} = (G, In, Out, Leak)$ has properties

- the edge $i \to j \in E_G$ if and only if the edge $j \to i \in E_G$
- underlying undirected graph of G a [double] tree*

Examples





Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\text{dist}(\text{in}, \text{out}) \geq 2 \text{ or } |\text{Leak}| \geq 2.$$

Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2 \text{ or } |\mathrm{Leak}| \geq 2.$$

Proof idea: Let $n = |V_G|$.

- # parameters: $\underbrace{|E_G|} + |Leak| = \underbrace{2n - 2} + |Leak|$
  since a *tree* always has $n - 1$ edges,' a *double tree* as $2(n - 1)$ edges!

Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2 \text{ or } |\mathrm{Leak}| \geq 2.$$

Proof idea: Let $n = |V_G|$.

- $\#$ parameters: $|E_G| + |Leak| = 2n - 2 + |Leak|$
- $\#$ coefficients:

	$\|Leak\| \geq 2$	$\|Leak\| = 1$	$\|Leak\| = 0$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) - 1$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 1$	$2n - 1$	$2n - 1$	$2n - 2$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0$	$2n - 1$	$2n - 1$	$2n - 2$

Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2 \text{ or } |\mathrm{Leak}| \geq 2.$$

Proof idea: Let $n = |V_G|$.

- # parameters: $|E_G| + |Leak| = 2n - 2 + |Leak| \geq 2n - 2 + 2 = 2n$
- # coefficients:

	$\lvert Leak\rvert \geq 2$	$\lvert Leak\rvert = 1$	$\lvert Leak\rvert = 0$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2$			
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 1$			
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0$	$2n - 1$		

Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2 \text{ or } |\mathrm{Leak}| \geq 2.$$

Proof idea: Let $n = |V_G|$.

- # parameters: $|E_G| + |Leak| = 2n - 2 + |Leak|$
- # coefficients:

	$\lvert Leak\rvert \geq 2$	$\lvert Leak\rvert = 1$	$\lvert Leak\rvert = 0$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) - 1$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 1$	$2n - 1$	$2n - 1$	$2n - 2$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0$	$2n - 1$	$2n - 1$	$2n - 2$

- five red cases have # parameters > # coefficients

Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2 \text{ or } |\mathrm{Leak}| \geq 2.$$

Proof idea: Let $n = |V_G|$.

- # parameters: $|E_G| + |Leak| = 2n - 2 + |Leak|$
- # coefficients:

	$\lvert Leak\rvert \geq 2$	$\lvert Leak\rvert = 1$	$\lvert Leak\rvert = 0$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) - 1$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 1$	$2n - 1$	$2n - 1$	$2n - 2$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0$	$2n - 1$	$2n - 1$	$2n - 2$

- five red cases have # parameters > # coefficients $\implies$ unidentifiability

Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2 \text{ or } |\mathrm{Leak}| \geq 2.$$

Proof idea: Let $n = |V_G|$.

- # parameters: $|E_G| + |Leak| = 2n - 2 + |Leak|$
- # coefficients:

	$\|Leak\| \geq 2$	$\|Leak\| = 1$	$\|Leak\| = 0$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) - 1$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 1$	$2n - 1$	$2n - 1$	$2n - 2$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0$	$2n - 1$	$2n - 1$	$2n - 2$

- five red cases have # parameters > # coefficients $\Longrightarrow$ unidentifiability
- four blue cases have # parameters = # coefficients

Unidentifiability of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is unidentifiable if

$$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2 \text{ or } |\mathrm{Leak}| \geq 2.$$

Proof idea: Let $n = |V_G|$.

- # parameters: $|E_G| + |Leak| = 2n - 2 + |Leak|$
- # coefficients:

	$\|Leak\| \geq 2$	$\|Leak\| = 1$	$\|Leak\| = 0$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) \geq 2$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out})$	$2n - \mathrm{dist}(\mathrm{in}, \mathrm{out}) - 1$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 1$	$2n - 1$	$2n - 1$	$2n - 2$
$\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 0$	$2n - 1$	$2n - 1$	$2n - 2$

- five red cases have # parameters > # coefficients $\Longrightarrow$ unidentifiability
- four blue cases have # parameters = # coefficients,

but that does not guarantee identifiability.

Can we find a counterexample?

Is there an unidentifiable tree model with $\mathrm{dist}(\mathrm{in}, \mathrm{out}) < 2$ and $|Leak| < 2$?

*Brought to you by Cursor

Building [Locally] Identifiable Tree Models

Plan for showing that # parameters = # coefficients implies identifiability:

1. start with some base model that we know is identifiable

Building [Locally] Identifiable Tree Models

Plan for showing that # parameters = # coefficients implies identifiability:

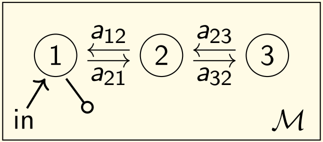
1. start with some base model that we know is identifiable

Proposition ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The tree model $\mathcal{M} = (G, \{i\}, \{i\}, \emptyset)$ is *locally* identifiable.

Example

$\mathcal{M} = (\text{Cat}_3, \{1\}, \{1\}, \emptyset)$ is locally identifiable:



Building [Locally] Identifiable Tree Models

Plan for showing that # parameters = # coefficients implies identifiability:

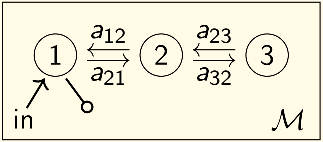
1. start with some base model that we know is identifiable

Proposition ($, Gross, Meshkat, Shiu, Sullivant, [2023])

The tree model $\mathcal{M} = (G, \{i\}, \{i\}, \emptyset)$ is *locally* identifiable.

Example

$\mathcal{M} = (\text{Cat}_3, \{1\}, \{1\}, \emptyset)$ is locally identifiable:

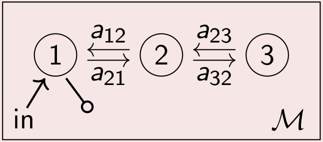


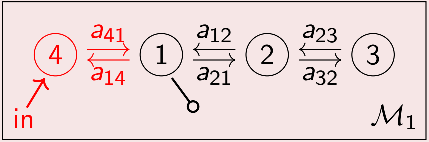
2. from base model, *build* all tree models where $|Leak| \leq 1$ and $\text{dist}(\text{in}, \text{out}) \leq 1$ and **retain local identifiability at each step**

Legal Moves: 1) Moving Input or Output

Prop: Moving the Input or Output

$\mathcal{M} = (G, \{1\}, \{1\}, \emptyset)$ identifiable $\implies$ $\mathcal{M}_1 = (H, \{4\}, \{1\}, \emptyset)$ identifiable:

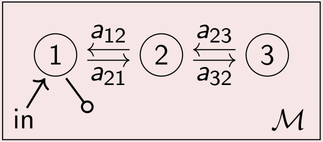


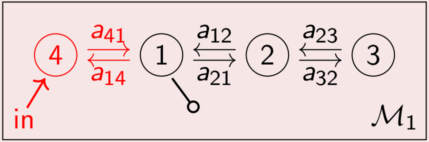


Legal Moves: 1) Moving Input or Output

Prop: Moving the Input or Output

$\mathcal{M} = (G, \{1\}, \{1\}, \emptyset)$ identifiable $\implies$ $\mathcal{M}_1 = (H, \{4\}, \{1\}, \emptyset)$ identifiable:





Proof idea:

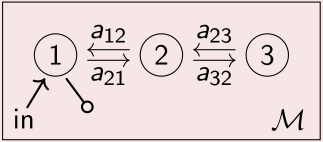
Theorem (Inverse Function Theorem)

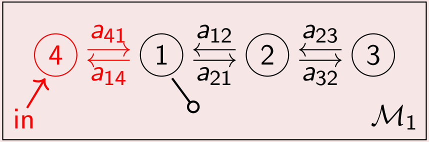
If a function $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ has Jacobian with (generically) nonzero determinant, then that function is locally injective.

Legal Moves: 1) Moving Input or Output

Prop: Moving the Input or Output

$\mathcal{M} = (G, \{1\}, \{1\}, \emptyset)$ identifiable $\implies$ $\mathcal{M}_1 = (H, \{4\}, \{1\}, \emptyset)$ identifiable:





Proof idea:

- write the coefficients of $\mathcal{M}_1$ in terms of $\mathcal{M}$ and the new parameters
- manipulate the Jacobian of $\mathcal{M}_1$ to "find" the Jacobian of $\mathcal{M}$, which by assumption has full rank:

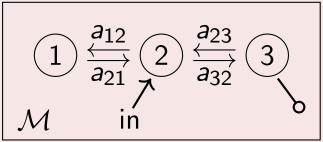
$$J(\mathcal{M}_1) = \begin{pmatrix} J(\mathcal{M}) & 0 \\ * & C \end{pmatrix}$$

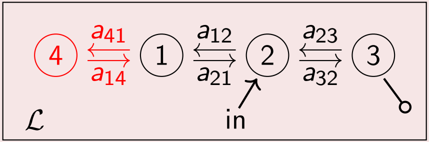
- show that C has full rank using properties of the graph

Legal Moves: 2) Adding Leaves

Prop: Adding Leaves

$\mathcal{M} = (G, \{2\}, \{3\}, \emptyset)$ identifiable $\implies$ $\mathcal{L} = (H, \{2\}, \{3\}, \emptyset)$ identifiable:

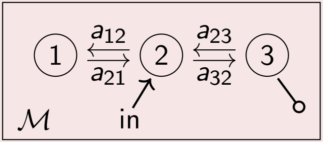


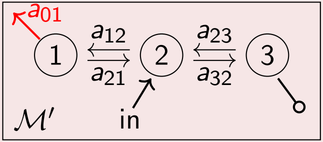


Legal Moves: 3) Adding a Leak

Prop: Adding a Leak

$\mathcal{M} = (G, \{2\}, \{3\}, \emptyset)$ identifiable $\Longrightarrow \mathcal{M}' = (G, \{2\}, \{3\}, \{1\})$ identifiable:



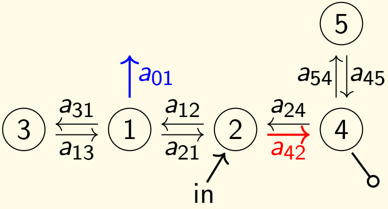


Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

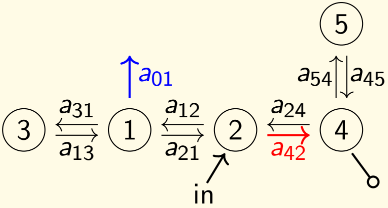
Identifiable?

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

Identifiable?

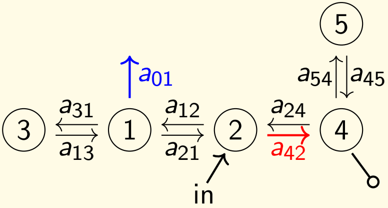
- $\mathrm{dist}(\underbrace{2}_{\text{In}}, \underbrace{4}_{\text{Out}}) =$

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

Identifiable?

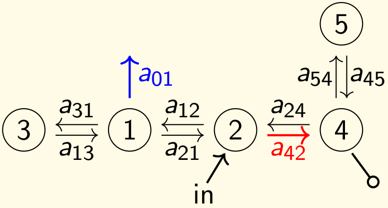
- $\mathrm{dist}(\underbrace{2}_{\text{In}}, \underbrace{4}_{\text{Out}}) = 1$

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

Identifiable?

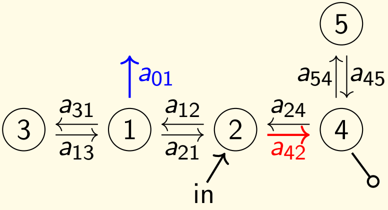
- $\mathrm{dist}(\underbrace{2}_{\text{In}}, \underbrace{4}_{\text{Out}}) = 1$
- $|\underbrace{Leak}_{\{1\}}| =$

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

Identifiable?

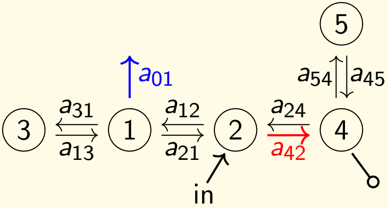
- $\mathrm{dist}(\underbrace{2}_{\mathrm{In}}, \underbrace{4}_{\mathrm{Out}}) = 1$
- $|\underbrace{Leak}_{\{1\}}| = 1$

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\text{dist}(\text{in}, \text{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

Identifiable?

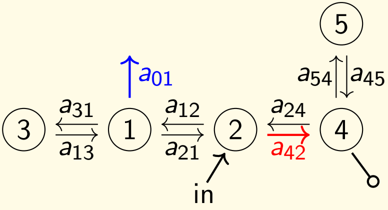
- $\text{dist}(\underbrace{2}_{\text{In}}, \underbrace{4}_{\text{Out}}) = 1 \leq 1$
- $|\underbrace{Leak}_{\{1\}}| = 1$

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

Identifiable?

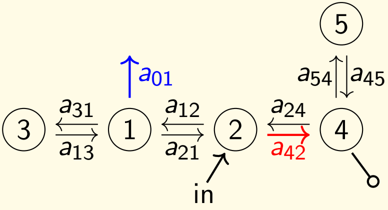
- $\mathrm{dist}(\underbrace{2}_{\mathrm{In}}, \underbrace{4}_{\mathrm{Out}}) = 1 \leq 1$
- $|\underbrace{Leak}_{\{1\}}| = 1 \leq 1$

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

Identifiable?

- $\mathrm{dist}(\underbrace{2}_{\text{In}}, \underbrace{4}_{\text{Out}}) = 1 \leq 1$
- $|\underbrace{Leak}_{\{1\}}| = 1 \leq 1$

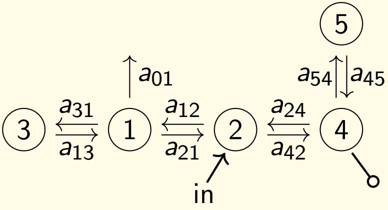
YES!

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G, \{2\}, \{4\}, \{1\})$

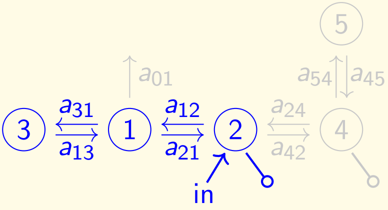
a. Identifiable base model:
$$(\underbrace{G_0}_{\text{graph}}, \underbrace{\{i\}}_{\text{In}}, \underbrace{\{i\}}_{\text{Out}}, \underbrace{\emptyset}_{\text{Leak}})$$
b. Move the input/output
c. Add "leaves"
d. Add a single leak

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(\mathrm{Cat}_3, \{2\}, \{2\}, \emptyset)$

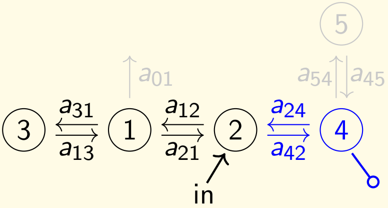
a. Identifiable base model: $(\underbrace{G_0}_{\text{graph}}, \underbrace{\{i\}}_{\text{In}}, \underbrace{\{i\}}_{\text{Out}}, \underbrace{\emptyset}_{\text{Leak}})$
b. Move the input/output
c. Add "leaves"
d. Add a single leak

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G_4, \{2\}, \{4\}, \emptyset)$

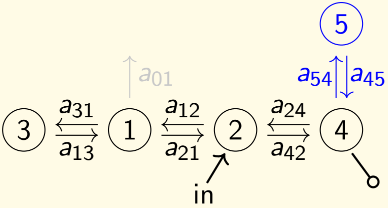
a. Identifiable base model: $(\underbrace{G_0}_{\text{graph}}, \underbrace{\{i\}}_{\text{In}}, \underbrace{\{i\}}_{\text{Out}}, \underbrace{\emptyset}_{\text{Leak}})$
b. Move the input/output
c. Add "leaves"
d. Add a single leak

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G_5, \{2\}, \{4\}, \emptyset)$

a. Identifiable base model: $(\underbrace{G_0}_{\text{graph}}, \underbrace{\{i\}}_{\text{In}}, \underbrace{\{i\}}_{\text{Out}}, \underbrace{\emptyset}_{\text{Leak}})$

b. Move the input/output

c. Add "leaves"

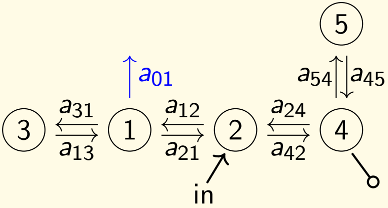
d. Add a single leak

Classification of Tree Models

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example



$(G_5, \{2\}, \{4\}, \{1\})$

a. Identifiable base model: $(\underbrace{G_0}_{\text{graph}}, \underbrace{\{i\}}_{\text{In}}, \underbrace{\{i\}}_{\text{Out}}, \underbrace{\emptyset}_{\text{Leak}})$

b. Move the input/output

c. Add "leaves"

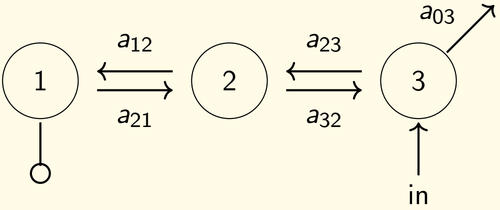
d. Add a single leak

Example

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example $(\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$

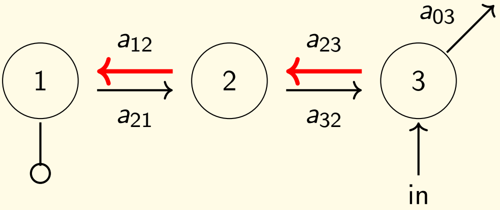


Example

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example $(\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$



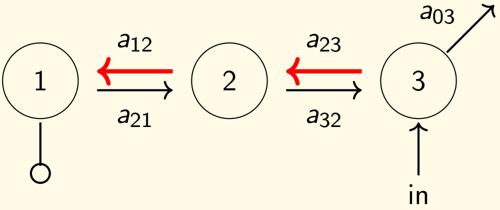
UNIDENTIFIABLE,

Example

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example $(\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$



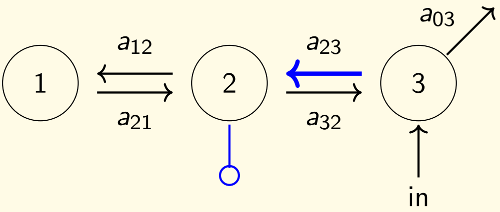
UNIDENTIFIABLE, since $\mathrm{dist}(3, 1) = 2 > 1$

Example

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example $(\mathrm{Cat}_3, \{3\}, \{2\}, \{3\})$

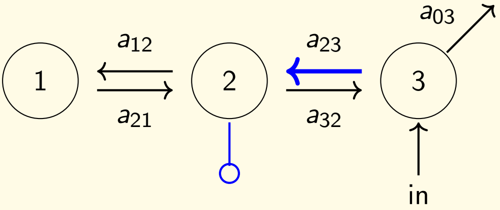


Example

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example $(\mathrm{Cat}_3, \{3\}, \{2\}, \{3\})$



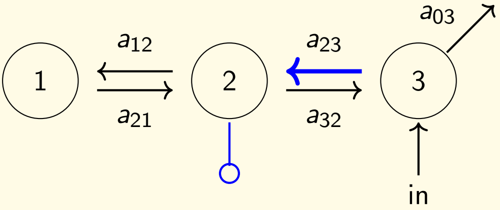
IDENTIFIABLE,

Example

Theorem ($, Gross, Meshkat, Shiu, Sullivant, [2023])

A tree model $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ is identifiable **if and only if** $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and $|Leak| \leq 1$.

Example ($\mathrm{Cat}_3, \{3\}, \{2\}, \{3\}$)



IDENTIFIABLE, since $\mathrm{dist}(3, 2) = 1 \leq 1$ and $|Leak| = 1 \leq 1$.

Conclusion

Theorem

For **ALL** linear compartmental models*, we can generate defining input-output equations from the underlying graph.

Conclusion

Theorem

For **ALL** linear compartmental models*, we can generate defining input-output equations from the underlying graph.

Theorem

For **tree models** with a single input and output, we completely classify *local* structural identifiability by the underlying graph structure.

Conclusion

Theorem

For **ALL** linear compartmental models*, we can generate defining input-output equations from the underlying graph.

Theorem

For **tree models** with a single input and output, we completely classify *local* structural identifiability by the underlying graph structure.

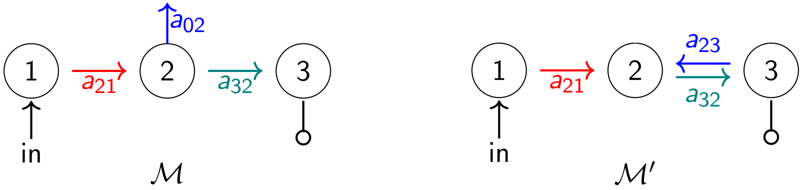
Takeaway

Biologists/modelers can use this information to design models which are structurally identifiable in the hope that they are practically identifiable.

Future Work/Open Problems

- generalize results on tree models to other linear compartmental models (classify identifiability of other families by their graphs)
- consider the problem of determining identifiability when multiple inputs/outputs are present
- expand to *global* identifiability
- find more applications for new characterization of coefficients
  - look for patterns in the singular locus for *dividing edges*
  - expand *indistinguishability* ($-Meshkat, [2024]), i.e. the problem of determining whether two or more linear compartmental models fit a given set of measured data

Motivating Example: Indistinguishability



$$\mathcal{M} : y_3^{(3)} + (a_{21} + a_{02} + a_{32})\ddot{y}_3 + (a_{02}a_{21} + a_{21}a_{32})\dot{y}_3 = (a_{21}a_{32})u_1$$

$$\mathcal{M}' : y_3^{(3)} + (a_{21} + a_{23} + a_{32})\ddot{y}_3 + (a_{23}a_{21} + a_{21}a_{32})\dot{y}_3 = (a_{21}a_{32})u_1$$

Renaming:

$$a_{21} \leftrightarrow a_{21}$$

$$a_{02} \leftrightarrow a_{23}$$

$$a_{32} \leftrightarrow a_{32}$$

Skeletal Path Moves: Walking the Leak

Question

What "*leak moves*" can you perform on a basic skeletal path model resulting in an indistinguishable model?

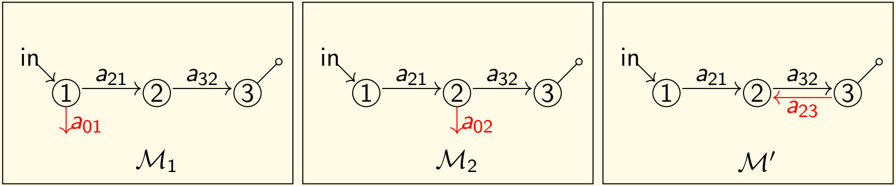
Skeletal Path Moves: Walking the Leak

Theorem ($ & Meshkat [2024]; $, Gilliana, Patel, & Tamras [2025*])

The following skeletal path models are indistinguishable:

- $\mathcal{M}_i = (\overrightarrow{P_n}, \{1\}, \{n\}, \{i\})$ for any $i \in \{1, 2, 3, \ldots, n-1\}$
- $\mathcal{M}' = (\overrightarrow{P_n} \cup \{n \to n-1\}, \{1\}, \{n\}, \emptyset)$.

Example



Skeletal Path Moves: Walking the Leak

Theorem ($ & Meshkat [2024]; $, Gilliana, Patel, & Tamras [2025*])

The following skeletal path models are indistinguishable:

- $\mathcal{M}_i = (\overrightarrow{P_n}, \{1\}, \{n\}, \{i\})$ for any $i \in \{1, 2, 3, \ldots, n-1\}$
- $\mathcal{M}' = (\overrightarrow{P_n} \cup \{n \to n-1\}, \{1\}, \{n\}, \emptyset)$.

Proof idea 1: Left-hand side of the input/output equation of $\mathcal{M}_i$ given by:

$$\det(\partial I - A_i) y_n = \det \begin{pmatrix} \partial + a_{21} & 0 & \cdots & \cdots & 0 & 0 \\ -a_{21} & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \ddots & \partial + a_{0i} + a_{i(i-1)} & \ddots & \vdots & \vdots \\ \vdots & \ddots & -a_{i(i-1)} & \ddots & 0 & 0 \\ 0 & \cdots & \ddots & \ddots & \partial + a_{n(n-1)} & 0 \\ 0 & \cdots & \cdots & 0 & -a_{n(n-1)} & \partial \end{pmatrix} y_n$$

Skeletal Path Moves: Walking the Leak

Theorem ($ & Meshkat [2024]; $, Gilliana, Patel, & Tamras [2025])

The following skeletal path models are indistinguishable:

- $\mathcal{M}_i = (\overrightarrow{P_n}, \{1\}, \{n\}, \{i\})$ for any $i \in \{1, 2, 3, \ldots, n-1\}$
- $\mathcal{M}' = (\overrightarrow{P_n} \cup \{n \to n-1\}, \{1\}, \{n\}, \emptyset)$.

Proof idea 2:

- Under a renaming of the parameters, the incoming forests of each $\mathcal{M}_i$ are exactly the same as the incoming forests of each $\mathcal{M}_j$ (and $\mathcal{M}'$).

Skeletal Path Moves: Walking the Leak

Theorem ($ & Meshkat [2024]; $, Gilliana, Patel, & Tamras [2025])

The following skeletal path models are indistinguishable:

- $\mathcal{M}_i = (\overrightarrow{P_n}, \{1\}, \{n\}, \{i\})$ for any $i \in \{1, 2, 3, \ldots, n-1\}$
- $\mathcal{M}' = (\overrightarrow{P_n} \cup \{n \to n-1\}, \{1\}, \{n\}, \emptyset)$.

Proof idea 2:

- Under a renaming of the parameters, the incoming forests of each $\mathcal{M}_i$ are exactly the same as the incoming forests of each $\mathcal{M}_j$ (and $\mathcal{M}'$).
- Thus, each of the coefficients of the respective input/output equations are indistinguishable.

Acknowledgments and References

Thank you to the American Institute of Mathematics (AIM) for providing a productive work environment. This work was partially supported by the US National Science Foundation (DMS 1615660).

- Cashous Bortner, Elizabeth Gross, Nicolette Meshkat, Anne Shiu, and Seth Sullivant.
  Identifiability of linear compartmental tree models and a general formula for the input-output equations.
  Advances in Applied Mathematics, 146, May 2023.

- Cashous Bortner and Nicolette Meshkat.
  Graph-based sufficient conditions for the indistinguishability of linear compartmental models.
  SIAM Journal on Applied Dynamical Systems, 23(3):2179–2207, 2024.

Thank you!!! - cbortner@csustan.edu

arXiv Link