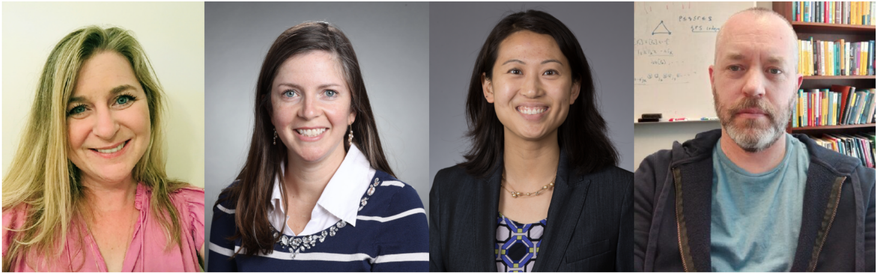
# Identifiability of Linear Compartmental Tree Models

Cash Bortner

CSU-Stanislaus
cbortner@csustan.edu

DART XI
Queen Mary University of London
6/7/2023 or 7/6/2023

# Collaborators



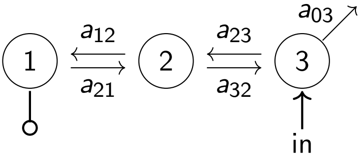
Dr. Elizabeth Gross
U. of Hawai'i at Mānoa

Dr. Nicolette Meshkat
Santa Clara University

Dr. Anne Shiu
Texas A&M University

Dr. Seth Sullivant
North Carolina State U.

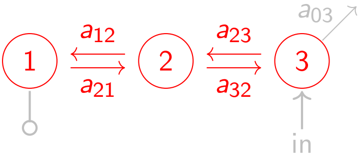
# Motivating Example




**Linear Compartmental Model**

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

# Motivating Example

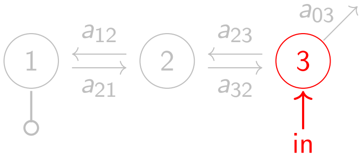



Directed Graph: $G = \mathrm{Cat}_3$

**Linear Compartmental Model**

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

# Motivating Example

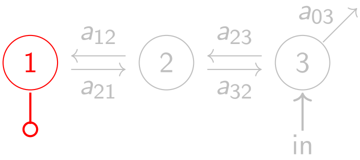



Input Compartment: $In = \{3\}$

**Linear Compartmental Model**

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

# Motivating Example

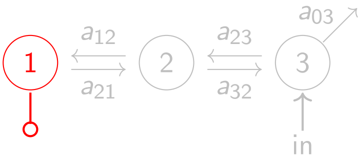



Measured Compartment: $Out = \{1\}$

**Linear Compartmental Model**

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

# Motivating Example

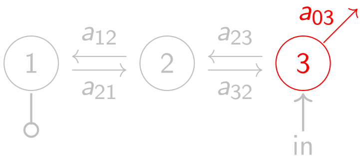



"Output" Compartment: $\mathit{Out} = \{1\}$

**Linear Compartmental Model**

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

# Motivating Example

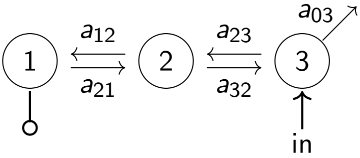



Leak Compartment: $Leak = \{3\}$

**Linear Compartmental Model**

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

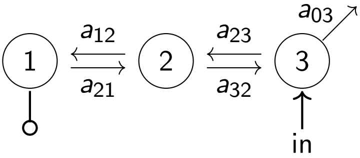
# Motivating Example




**Linear Compartmental Model**

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

# Motivating Example



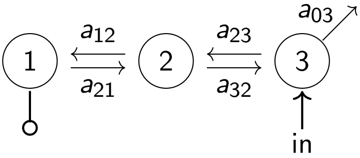

**Linear Compartmental Model**

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

## Motivating Question: Identifiability

Given information about the input and output compartment[s], can we **recover** all flow rate parameters?

# Motivating Example




**Linear Compartmental Model**

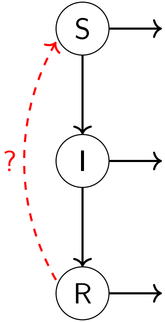
$$\mathcal{M} = (G, In, Out, Leak) \\ = (\text{Cat}_3, \{3\}, \{1\}, \{3\}).$$

## Motivating Question: Identifiability

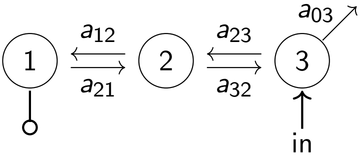
Given information about the input and output compartment[s], can we **identify** all flow rate parameters?

# Compartmental Models in the Wild

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

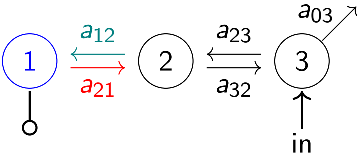


# Motivating Example




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

# Motivating Example

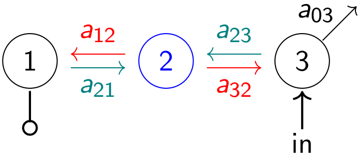



$$\begin{aligned}\mathcal{M} &= (G, In, Out, Leak) \\ &= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).\end{aligned}$$

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

$$\dot{x_1} = -a_{21}x_1(t) \qquad +a_{12}x_2(t)$$

# Motivating Example

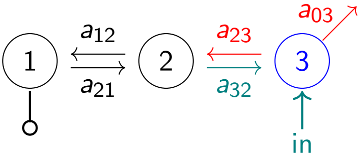



$$\mathcal{M} = (G, In, Out, Leak) = (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).$$

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

$$\dot{x}_1 = -a_{21}x_1(t) \qquad + a_{12}x_2(t)$$
$$\dot{x}_2 = a_{21}x_1(t) - (a_{12} + a_{32})x_2(t) \qquad + a_{23}x_3(t)$$

# Motivating Example

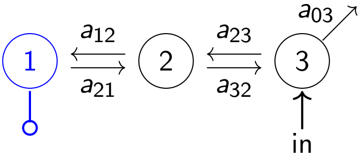



$$\mathcal{M} = (G, In, Out, Leak) = (\text{Cat}_3, \{3\}, \{1\}, \{3\}).$$

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

$$\begin{aligned}
\dot{x}_1 &= -a_{21}x_1(t) + a_{12}x_2(t) \\
\dot{x}_2 &= a_{21}x_1(t) - (a_{12} + a_{32})x_2(t) + a_{23}x_3(t) \\
\dot{x}_3 &= a_{32}x_2(t) - (a_{03} + a_{23})x_3(t) + u_3(t)
\end{aligned}$$

# Motivating Example




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

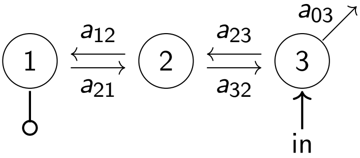
ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$\begin{aligned}
\dot{x}_1 &= -a_{21}x_1(t) + a_{12}x_2(t) \\
\dot{x}_2 &= a_{21}x_1(t) - (a_{12} + a_{32})x_2(t) + a_{23}x_3(t) \\
\dot{x}_3 &= a_{32}x_2(t) - (a_{03} + a_{23})x_3(t) + u_3(t)
\end{aligned}$$

with

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

# Motivating Example




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

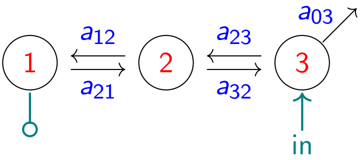
ODEs in terms of concentrations $x_i(t)$, input $u_3(t)$, and output $y_1(t)$:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \underbrace{\begin{pmatrix} -a_{21} & a_{12} & 0 \\ a_{21} & -a_{12} - a_{32} & a_{23} \\ 0 & a_{32} & -a_{03} - a_{23} \end{pmatrix}}_{\text{compartmental matrix } A} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

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

# Motivating Example




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

<u>*Goal*</u>: *Identify the parameters $a_{ji}$ from the measurable variables.*

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \underbrace{\begin{pmatrix} -a_{21} & a_{12} & 0 \\ a_{21} & -a_{12} - a_{32} & a_{23} \\ 0 & a_{32} & -a_{03} - a_{23} \end{pmatrix}}_{\text{compartmental matrix A}} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

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

# Motivating Example

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

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} -a_{21} & a_{12} & 0 \\ a_{21} & -a_{12} - a_{32} & a_{23} \\ 0 & a_{32} & -a_{03} - a_{23} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

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

# Motivating Example

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

$$\begin{pmatrix} \partial_t & 0 & 0 \\ 0 & \partial_t & 0 \\ 0 & 0 & \partial_t \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} -a_{21} & a_{12} & 0 \\ a_{21} & -a_{12} - a_{32} & a_{23} \\ 0 & a_{32} & -a_{03} - a_{23} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

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

# Motivating Example

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

$$\begin{pmatrix} \partial_t & 0 & 0 \\ 0 & \partial_t & 0 \\ 0 & 0 & \partial_t \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} - \begin{pmatrix} -a_{21} & a_{12} & 0 \\ a_{21} & -a_{12} - a_{32} & a_{23} \\ 0 & a_{32} & -a_{03} - a_{23} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

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

# Motivating Example

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

$$\left(\begin{pmatrix}\partial_t & 0 & 0\\ 0 & \partial_t & 0\\ 0 & 0 & \partial_t\end{pmatrix} - \begin{pmatrix}-a_{21} & a_{12} & 0\\ a_{21} & -a_{12}-a_{32} & a_{23}\\ 0 & a_{32} & -a_{03}-a_{23}\end{pmatrix}\right)\begin{pmatrix}x_1(t)\\ x_2(t)\\ x_3(t)\end{pmatrix} = \begin{pmatrix}0\\ 0\\ u_3(t)\end{pmatrix}$$

with

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

# Motivating Example

*<u>Goal</u>: Identify the parameters $a_{ji}$ from the measurable variables.*

$$\begin{pmatrix} \partial_t + a_{21} & -a_{12} & 0 \\ -a_{21} & \partial_t + a_{12} + a_{32} & -a_{23} \\ 0 & -a_{32} & \partial_t + a_{03} + a_{23} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

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

# Motivating Example

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

$$\begin{pmatrix} \partial_t + a_{21} & -a_{12} & 0 \\ -a_{21} & \partial_t + a_{12} + a_{32} & -a_{23} \\ 0 & -a_{32} & \partial_t + a_{03} + a_{23} \end{pmatrix} \begin{pmatrix} y_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

with

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

# Motivating Example

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

$$\begin{pmatrix} \partial_t + a_{21} & -a_{12} & 0 \\ -a_{21} & \partial_t + a_{12} + a_{32} & -a_{23} \\ 0 & -a_{32} & \partial_t + a_{03} + a_{23} \end{pmatrix} \begin{pmatrix} y_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

via Cramer's Rule:

$$\det \begin{pmatrix} \partial_t + a_{21} & -a_{12} & 0 \\ -a_{21} & \partial_t + a_{12} + a_{32} & -a_{23} \\ 0 & -a_{32} & \partial_t + a_{03} + a_{23} \end{pmatrix} y_1(t)$$

$$= \det \begin{pmatrix} 0 & -a_{12} & 0 \\ 0 & \partial_t + a_{12} + a_{32} & -a_{23} \\ u_3(t) & -a_{32} & \partial_t + a_{03} + a_{23} \end{pmatrix}$$

# Motivating Example

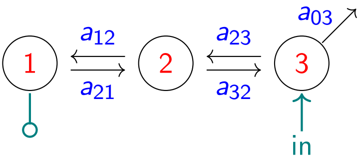
*Goal: Identify the parameters $a_{ji}$ from the measurable variables.*

$$\begin{pmatrix} \partial_t + a_{21} & -a_{12} & 0 \\ -a_{21} & \partial_t + a_{12} + a_{32} & -a_{23} \\ 0 & -a_{32} & \partial_t + a_{03} + a_{23} \end{pmatrix} \begin{pmatrix} y_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ u_3(t) \end{pmatrix}$$

via Cramer's Rule:

$$\det \begin{pmatrix} \partial_t + a_{21} & -a_{12} & 0 \\ -a_{21} & \partial_t + a_{12} + a_{32} & -a_{23} \\ 0 & -a_{32} & \partial_t + a_{03} + a_{23} \end{pmatrix} y_1(t) = \det \begin{pmatrix} -a_{12} & 0 \\ \partial_t + a_{12} + a_{32} & -a_{23} \end{pmatrix} u_3(t)$$

# Motivating Example



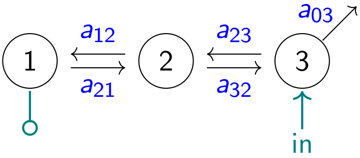

$$\begin{aligned}\mathcal{M} &= (G, In, Out, Leak)\\ &= (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\}).\end{aligned}$$

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

$$\det(\partial_t I - A) y_1 = \underbrace{\det(\partial_t I - A)^{(3,1)}}_{\text{remove row 3 and col 1}} u_3$$

by Cramer's Rule and substitution.

# Motivating Example: Input/Output Equation




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

Via a substitution and application of Cramer's Rule:

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

an ODE in only the measurable variables and the parameters:

Input/Output Equation

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

# 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
2. Determine the injectivity of the coefficient map defined by the input/output equation

# 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
2. Determine the injectivity of the coefficient map defined by the input/output equation

## Goal

*We want to classify structural 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
2. Determine the injectivity of the coefficient map defined by the input/output equation

## Goal

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

# Novel Input-Output Equation Characterization

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

The coefficients of the input-output equation of a 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 [1])

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

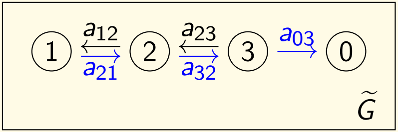
## Definitions

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

## 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 [1])

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

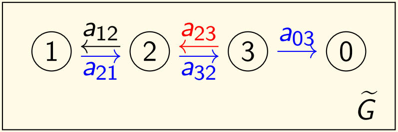
## Definitions

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

## 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 [1])

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

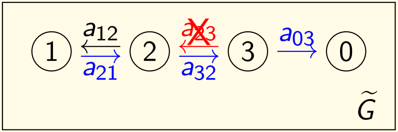
## Definitions

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

## 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 [1])

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

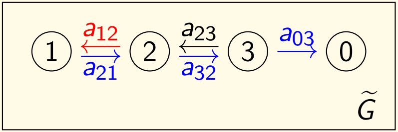
## Definitions

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

## 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 [1])

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

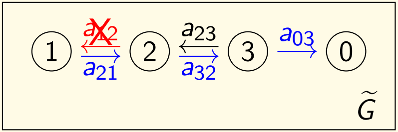
## Definitions

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

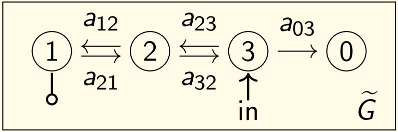
## Example

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



## Example

For $\mathcal{M} = (\mathrm{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

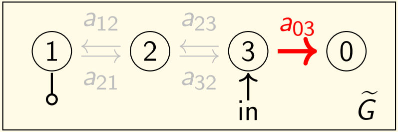
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

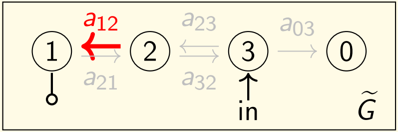
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

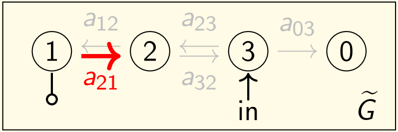
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

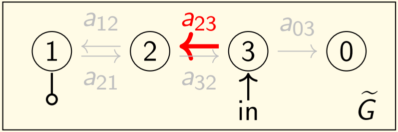
$$c_k \;=\; \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

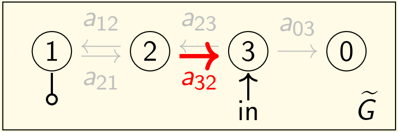
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

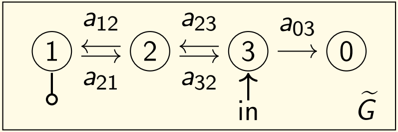
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

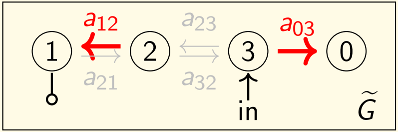
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 2 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

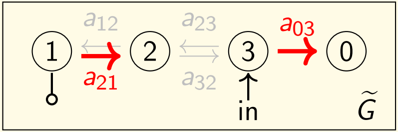
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 2 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

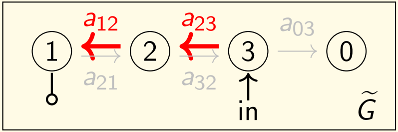
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 2 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

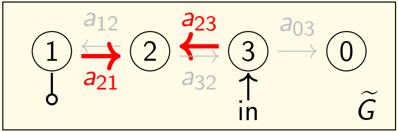
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 2 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

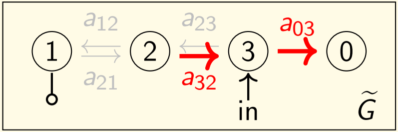
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 2 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

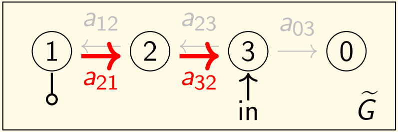
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 2 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

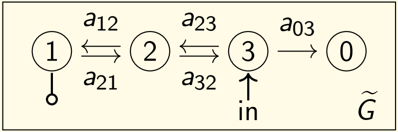
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 2 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

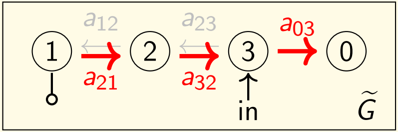
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 3 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of LHS of the i-o equation is:

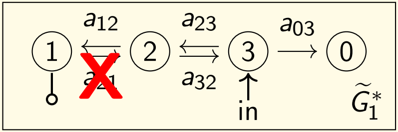
$$c_k = \sum_{F \in \mathcal{F}_{3-k}(\widetilde{G})} \pi_F$$

LHS coefficients: Incoming forests with 3 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 $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of RHS of the i-o equation is:

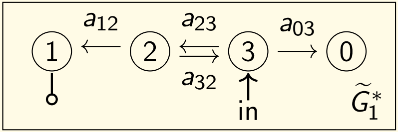
$$d_k \;=\; \sum_{F \in \mathcal{F}^{3,1}_{3-k-1}(\widetilde{G}_1^*)} \pi_F$$

RHS coefficients:

| Derivative | Coefficient |
|---|---|
| $u_3^{(0)}$ | |

## Example

For $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of RHS of the i-o equation is:

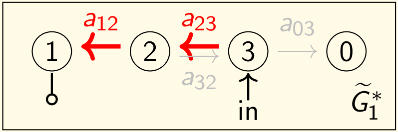
$$d_k = \sum_{F \in \mathcal{F}^{3,1}_{3-k-1}(\widetilde{G}_1^*)} \pi_F$$

RHS coefficients: Incoming forests with 2 edges AND a path from 3 to 1

| Derivative | Coefficient |
|---|---|
| $u_3^{(0)}$ | |

## Example

For $\mathcal{M} = (\text{Cat}_3, \{3\}, \{1\}, \{3\})$:




The $k^{\text{th}}$ coefficient of RHS of the i-o equation is:

$$d_k = \sum_{F \in \mathcal{F}_{3-k-1}^{3,1}(\widetilde{G}_1^*)} \pi_F$$

RHS coefficients: Incoming forests with 2 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 [1])

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

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

# Number of Coefficients

## Corollary ($, Gross, Meshkat, Shiu, Sullivant [1])

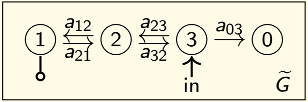
Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where $G$ is strongly connected and $|V_G| = n$. Then the number of non-trivial coefficients in the input/output equation is:

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

## Example

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

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




$\#$ on LHS $= 3$ (since $|Leak| = 1$)

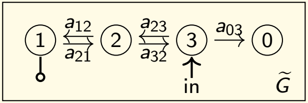
$\#$ on RHS $= 3 - \underbrace{\mathrm{dist}(3,1)}_{2} = 1$

# Example

## Example

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:

$$\phi_{\mathcal{M}} : \mathbb{R}^5 \to \mathbb{R}^4$$

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

# Identifiability

**Definition***

A model $(G, In, Out, Leak)$ with coefficient map $\phi$ is

- *locally identifiable* (identifiable) if, outside a set of measure zero, every point in $\mathbb{R}^{|E_G|+|Leak|}$ has an open neighborhood $U$ for which the restriction $\phi|_U : U \to \mathbb{R}^m$ is **one-to-one**; and
- *unidentifiable* if $c$ is generically **infinite-to-one**.

# Identifiability

**Definition***

A model $(G, In, Out, Leak)$ with coefficient map $\phi$ is

- *locally identifiable* (identifiable) if, outside a set of measure zero, every point in $\mathbb{R}^{|E_G|+|Leak|}$ has an open neighborhood $U$ for which the restriction $\phi|_U : U \to \mathbb{R}^m$ is **one-to-one**; and
- *unidentifiable* if $c$ is generically **infinite-to-one**.

**Proposition (Sufficient condition for unidentifiability)**

A model $\mathcal{M} = (G, In, Out, Leak)$ is *unidentifiable* if

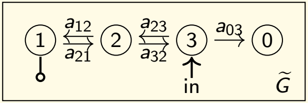
$$\underbrace{\#\text{ parameters}}_{|E_G|+|Leak|} > \#\text{ coefficients.}$$

# Example

## Example

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:

$$\phi_{\mathcal{M}} : \mathbb{R}^5 \to \mathbb{R}^4$$

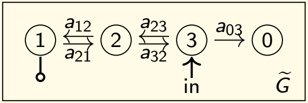
$$\begin{pmatrix} a_{03} \\ a_{12} \\ a_{21} \\ a_{23} \\ a_{32} \end{pmatrix} \mapsto \begin{pmatrix} a_{03} + a_{12} + a_{21} + a_{23} + a_{32} \\ a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32} \\ a_{03}a_{21}a_{32} \\ a_{12}a_{23} \end{pmatrix}$$

# Example

## Example

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:

$\phi_{\mathcal{M}} : \mathbb{R}^5 \to \mathbb{R}^4$ $\quad$ $\mathcal{M}$ is UNIDENTIFIABLE

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

# Unidentifiability

**Corollary ($, Gross, Meshkat, Shiu, Sullivant [1])**

Consider $\mathcal{M} = (G, \{in\}, \{out\}, Leak)$ where $G$ is strongly connected and $|V_G| = n$. Define $L$ and $d$ as follows:

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

Then $\mathcal{M}$ is **unidentifiable** if

$$\underbrace{|Leak| + |E_G|}_{\#\text{ parameters}} > \underbrace{2n - L - d}_{\#\text{ coefficients}}.$$

# The Jacobian

## Proposition

$\mathcal{M} = (G, \{i\}, \{j\}, Leak)$ is locally identifiable if and only if the rank of the Jacobian matrix of its coefficient map is equal to # parameters.

## Example

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

$$y_1^{(3)} + \underbrace{(a_{03} + a_{12} + a_{21} + a_{23} + a_{32})}_{c_2}\ddot{y}_1 + \underbrace{(a_{03}a_{12} + a_{03}a_{21} + a_{12}a_{23} + a_{21}a_{23} + a_{03}a_{32} + a_{21}a_{32})}_{c_1}\dot{y}_1$$

$$+ \underbrace{(a_{03}a_{21}a_{32})}_{c_0} y_1 = \underbrace{(a_{12}a_{23})}_{d_0} u_3$$

$$J(\phi_{\mathcal{M}}) = \begin{array}{c} \\ c_2 \\ c_1 \\ c_0 \\ d_0 \end{array} \begin{array}{c} \begin{array}{ccccc} a_{03} & a_{12} & a_{21} & a_{23} & a_{32} \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ a_{12} + a_{21} + a_{32} & a_{03} + a_{23} & a_{03} + a_{23} + a_{32} & a_{12} + a_{21} & a_{03} + a_{21} \\ a_{21}a_{32} & 0 & a_{03}a_{32} & 0 & a_{03}a_{21} \\ 0 & a_{23} & 0 & a_{12} & 0 \end{pmatrix} \end{array}$$

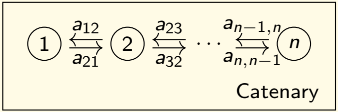
# Tree Models

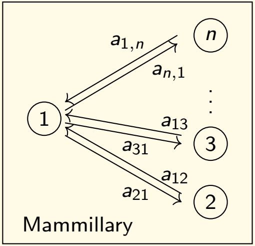
## Definition

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 [1])**

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: $|E_G| + |Leak| = 2n - 2 + |Leak|$

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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 [1])**

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:

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

Top-Left:

- # parameters $\geq 2n$ (since $|Leak| \geq 2$)
- # coefficients $= 2n - \underbrace{\mathrm{dist}(\mathrm{in}, \mathrm{out})}_{\geq 2}$

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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:

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

Top-Left: UNIDENTIFIABLE

- # parameters $\geq 2n$ (since $|Leak| \geq 2$)
- # coefficients $= 2n - \underbrace{\mathrm{dist}(\mathrm{in},\mathrm{out})}_{\geq 2}$

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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:

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

Bottom-Right:

- $\#$ parameters $= 2n - 2$ (since $|Leak| = 0$)
- $\#$ coefficients $= 2n - 2$

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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

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

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

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

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

Bottom-Right: IDENTIFIABLE???

- # parameters $= 2n - 2$ (since $|Leak| = 0$)
- # coefficients $= 2n - 2$

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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 $\Longrightarrow$ unidentifiability

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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
- four blue cases have # parameters = # coefficients

# Unidentifiability of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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

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

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

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

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

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

*but that does* **not** *guarantee identifiability.*

# Building Identifiable Tree Models

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

- start with some base model that we know is identifiable (Prop*)
- from base model, *build* all tree models where $|Leak| \leq 1$ and $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and retain identifiability at each step

**Proposition* ($, Gross, Meshkat, Shiu, Sullivant [1])**

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

# Building Identifiable Tree Models

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

- start with some base model that we know is identifiable (Prop*)
- from base model, *build* all tree models where $|Leak| \leq 1$ and $\mathrm{dist}(\mathrm{in}, \mathrm{out}) \leq 1$ and retain identifiability at each step

**Proposition* ($, Gross, Meshkat, Shiu, Sullivant [1])**

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

**Proposition (Gross, Harrington, Meshkat, Shiu [2])**

Let $\mathcal{M} = (G, In, Out, \emptyset)$ be strongly connected and identifiable. Then, the model $\mathcal{M}' = (G, In, Out, \{k\})$ is also identifiable.

# Moving the Input/Output

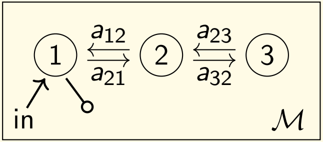
**Proposition ($, Gross, Meshkat, Shiu, Sullivant [1])**

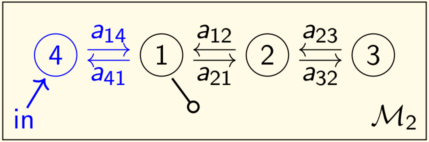
Let $\mathcal{M} = (G, \{i\}, \{i\}, \emptyset)$ be an identifiable tree model. Let $H$ be the graph $G$ with the added node $n$ and edges $i \to n$ and $n \to i$. Then following models are also identifiable:

- $\mathcal{M}_1 = (H, \{i\}, \{n\}, \emptyset)$
- $\mathcal{M}_2 = (H, \{n\}, \{i\}, \emptyset)$.

**Example**

Here, $\mathcal{M} = (\text{Cat}_3, \{1\}, \{1\}, \emptyset)$ and $\mathcal{M}_2 = (\text{Cat}_4^*, \{4\}, \{1\}, \emptyset)$:



# Proof of Moving the Input/Output

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

Let $\mathcal{M} = (G, \{i\}, \{i\}, \emptyset)$ be an identifiable tree model. Let $H$ be the graph $G$ with the added node $n$ and edges $i \to n$ and $n \to i$. Then following models are also identifiable:

- $\mathcal{M}_1 = (H, \{i\}, \{n\}, \emptyset)$
- $\mathcal{M}_2 = (H, \{n\}, \{i\}, \emptyset)$.

*Proof idea:*

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

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

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

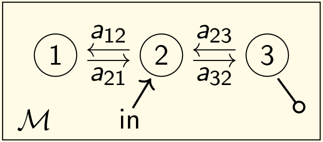
# Adding a Leaf

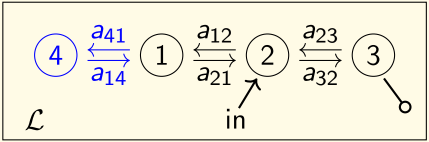
**Proposition ($, Gross, Meshkat, Shiu, Sullivant [1])**

Let $\mathcal{M} = (G, \{i\}, \{j\}, \emptyset)$ be an identifiable tree model. Define $\mathcal{L} = (H, \{i\}, \{j\}, \emptyset)$ where $H$ is the graph $G$ with the added node $n$ and edges $k \to n$ and $n \to k$ for some $k \in V_G$. Then, $\mathcal{L}$ is identifiable.

**Example**

Here, $\mathcal{M} = (\text{Cat}_3, \{2\}, \{3\}, \emptyset)$ and $\mathcal{L} = (\text{Cat}_4^*, \{2\}, \{3\}, \emptyset)$:



# Classification of Tree Models

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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

*Proof outline:*

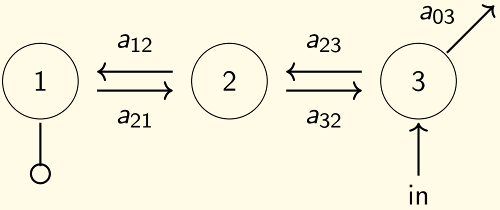
- $\mathcal{M}$ is unidentifiable if either $\mathrm{dist}(\mathrm{in}, \mathrm{out}) > 1$ or $|Leak| > 1$
- $\mathcal{M}$ is identifiable if $in = out$ and $|Leak| = 0$
- $\mathcal{M}$ is identifiable if $\mathrm{dist}(\mathrm{in}, \mathrm{out}) = 1$ and $|Leak| = 0$
- if $\mathcal{M}$ is identifiable with $|Leak| = 0$, then it is identifiable with $|Leak| = 1$

# Example

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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**

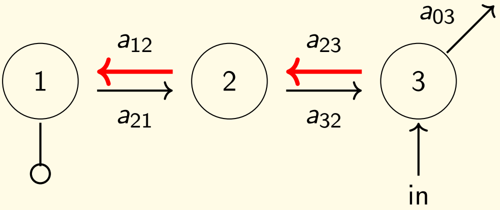


# Example

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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**



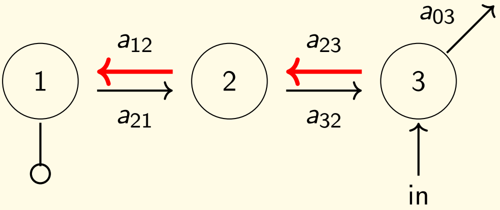

**UNIDENTIFIABLE**,

# Example

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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**



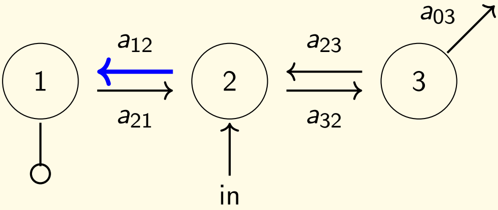

**UNIDENTIFIABLE**, since $\mathrm{dist}(3, 1) = 2 > 1$

# Example

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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**

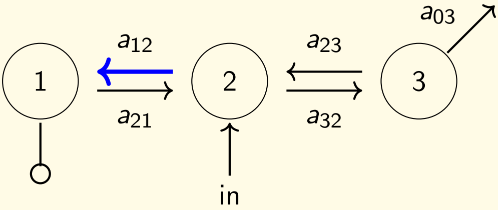


# Example

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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**



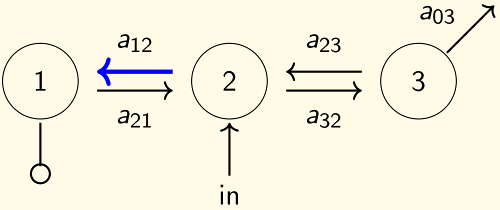

**IDENTIFIABLE**,

# Example

**Theorem ($, Gross, Meshkat, Shiu, Sullivant [1])**

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**




**IDENTIFIABLE**, since $\mathrm{dist}(2, 1) = 1 \leq 1$ 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.

# 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.

**Remark**

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

# Future Work

- generalize results on tree models to other linear compartmental models
- find more applications for new characterization of coefficients
  - consider *distinguishability*, i.e. the problem of determining whether two or more linear compartmental models fit a given set of measured data
  - look for patterns in the singular locus for *dividing edges*
  - consider identifiability versus observability relationship
- consider the problem of determining identifiability when multiple inputs/outputs are present

# 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.

Elizabeth Gross, Heather A. Harrington, Nicolette Meshkat, and Anne Shiu.
Linear compartmental models: input-output equations and operations that preserve identifiability.
*SIAM J. Appl. Math.*, 79(4):1423–1447, 2019.

# Thank you!!!