# On the Identifiability of Linear Compartmental Model Parameters

Tyler Huneke and Tegan Keen

July 17, 2025

**Abstract**

In many applications, including in ecology and pharmacokinetics, linear compartmental models are used to model transfer between "compartments" which may represent populations, drug concentration, etc. Such models are represented by directed graphs in which the edges represent the transfers between compartments. An important feature of such models is the identifiability degree, which summarizes the extent to which it is possible to recover the transfer rates from noiseless experimental data. More precisely, the identifiability degree of a parameter is 1 if the transfer rate can be recovered uniquely, and is greater than 1 if the transfer rate can be recovered only up to a finite set (this size is equal to the degree).

In this paper, we investigate the effects of adding leaks (edges directed out of the model) on the identifiability degree. We show that in a model represented by a strongly connected graph, if exactly one leak is in the same compartment as an output, then that leak parameter is uniquely identifiable. We investigate improvements to this result, looking at the preservation of the identifiability degree of the non-leak parameters and the applicability to non-strongly connected graphs, like in the case of directed path models. Finally, we build on this result to find the identifiability degree of cycle and mammillary models.

## 1 Introduction

In this paper, we will investigate the identifiability of *linear compartmental models*, models that are commonly used and described by a parameterized system of linear ODEs. These models have applications in pharmacokinetics, ecology, and more to understand the interactions between body systems, populations, etc., which we call compartments.

We will begin by giving a background of linear compartmental models, describing the graphs that represent them, the ODEs that represent them, and the input-output equation and coefficient map [1]. Then, we will define identifiability, a foundation for the remainder of the paper, along with identifiability degree. In section 2, we discuss a method to prove the unidentifiability of parameters in linear compartmental models.

In section 3, we prove a main result of parameter identifiability. We show that in a strongly connected model with the output and leak in the same compartment, the leak parameter will be globally identifiable (Proposition 4.5). We also conjecture further generalizations of this result.

In section 5, we prove the identifiability degree of some classes of cycle and mammillary models. In particular, we develop a method using the coefficient map which enables us to prove the identifiability degree of all cycle models with at most one leak, as well as the identifiability degree of the parameters and leaks therein.

In the final section, we define tree models as seen in [3] and generalize a formula conjectured in [3] for the identifiability degree of tree models (Conjecture 6.3).

**Remark 1.1.** Many of our results were enabled by a database we extended [4], originally created by Alexis Edozie, Odalys Garcia-Lopez, and Viridiana Neri.

## 2 Background

Often, biological systems can be represented in a directed graph $G$ with vertices representing organs or body systems and the edges between these vertices representing the rate of transfer between those vertices, denoted as $(V_G, E_G) \in G$. We write an edge $j \to i \in E_G$ as $k_{ij}$ and a leak, a rate of flow leaving the system, as $k_{0i}$.

A linear compartmental model contains input edges (i.e. the input of a drug into the body), output edges (the concentration of the drug), and can contain leak edges (the drug moving from the system or being broken down) directed outwards from a vertex. We can define a linear compartmental model as $\mathcal{M} = (G, In, Out, Leak)$ where $In$, $Out$, and $Leak$ are the sets of compartments containing an input, output, and leak, respectively. The size of the model is given by $n = |V_G|$. Three types of models, catenary, cycle, and mammillary, are given in Figures 1, 2, and 3, respectively. Any of these models can contain leaks from any compartment, labeled as $k_{0i}$.

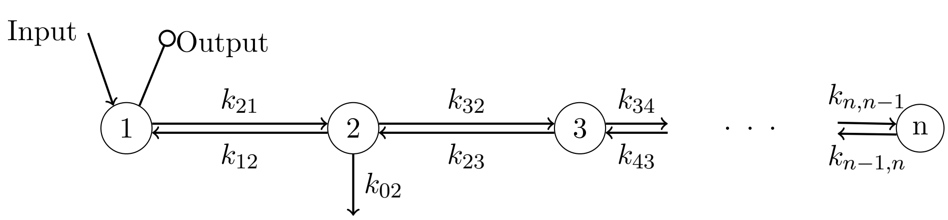



Figure 1: A *catenary* (bidirected path graph) model has a bidirected graph with compartments 1 through $n$.

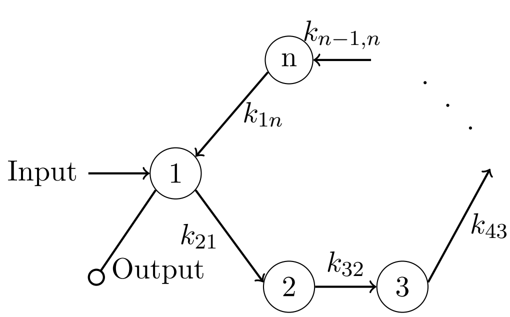



Figure 2: A *cycle* model has a directed graph with compartments 1 through $n$.

Following the notation of [1], a model with $n$ vertices produces the $n \times n$ matrix, $A$, with entries

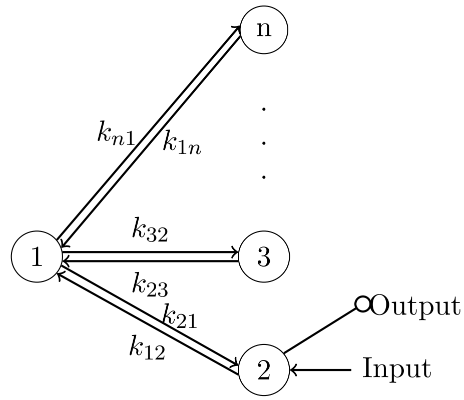



Figure 3: A *mammillary* (star graph) model is a bidirected graph with compartments 1 through $n$ in which the input and output are always in the same compartment with each "branch" containing one compartment.

as follows.

$$A_{i,j} = \begin{cases} -\sum_{m:i\to m\in E_G} k_{mi}, & i=j, i\notin Leak \\ -k_{0i} - \sum_{m:i\to m\in E_G} k_{mi}, & i=j, i\in Leak \\ k_{ij}, & i\neq j, (j,i)\in E_G \\ 0, & i\neq j, (j,i)\notin E_G \end{cases}$$

Using this matrix, the model defines the following ODE system where $u_i(t)$ and $y_i(t)$ are the concentration of input and output compartments, respectively, and $x(t) = (x_1(t), x_2(t), ..., x_n(t))$ is the vector of concentrations of all compartments:

$$\frac{dx}{dt} = Ax(t) + u(t),$$

$$y_i(t) = x_i(t) \text{ for all } i \in Out$$

**Proposition 2.1.** [1, Proposition 2.8] Let $\mathcal{M} = (G, In, Out, Leak)$ be a linear compartmental model with $n$ compartments and at least one input. Define $\partial I$ to be the $n \times n$ matrix in which every diagonal entry is the differential operator $\frac{d}{dt}$ and every off-diagonal entry is 0. Let $A$ be the compartmental matrix. Then, the following equations are input-output equations of $\mathcal{M}$:

$$det(\partial I - A)y_i = \sum_{j\in In} (-1)^{i+j} det((\partial I - A)^{j,i})u_j \tag{1}$$

**Remark 2.2.** In the right hand side of (1), $det(\partial I - A)^{j,i}$ represents the determinant of the matrix $\partial I - A$ with the $j^{th}$ row (the input compartment) and the $i^{th}$ column (the output compartment) removed.

**Example 2.3.** The model in Figure 4 produces the following $A$ matrix:

$$A = \begin{pmatrix} -k_{21} & k_{12} & 0 \\ k_{21} & -k_{02} - k_{12} - k_{32} & k_{23} \\ 0 & k_{32} & -k_{23} \end{pmatrix}$$

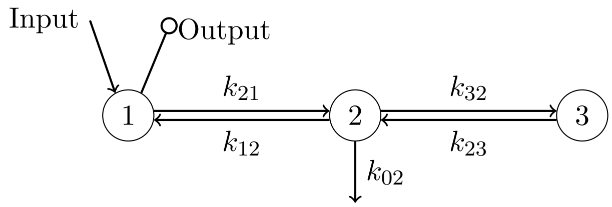



Figure 4: A linear compartmental model with $n = 3$, $In = Out = \{1\}$, and $Leak = \{2\}$.

which we can use to find our input-output equation (1) for $\mathcal{M}$

$$det\begin{pmatrix} \frac{d}{dt}+k_{21} & -k_{12} & 0 \\ -k_{21} & \frac{d}{dt}+k_{02}+k_{12}+k_{32} & -k_{23} \\ 0 & -k_{32} & \frac{d}{dt}+k_{23} \end{pmatrix} = det\begin{pmatrix} \frac{d}{dt}+k_{02}+k_{12}+k_{32} & -k_{23} \\ -k_{32} & \frac{d}{dt}+k_{23} \end{pmatrix}$$

$$y_1^{(3)} + (k_{02}+k_{12}+k_{21}+k_{23}+k_{32})y_1^{(2)} + (k_{02}k_{21}+k_{02}k_{23}+k_{12}k_{23}+k_{21}k_{23}+k_{21}k_{32})y_1' + (k_{02}k_{21}k_{23})y_1$$
$$= u_1^{(2)} + (k_{02}+k_{12}+k_{23}+k_{32})u_1' + (k_{02}k_{23}+k_{12}k_{23})u_1$$

## 2.1 Identifiability

An important aspect of linear compartmental models is understanding which parameters ($k_{ij}$) have solutions. Specifically, we are interested in the number of solutions we can find for each parameter, the identifiability degree. These are defined more carefully below. In order to understand identifiability, it is important to define the following.

**Definition 2.4.** (Coefficient Map) Let $\mathcal{M}$ be a model with $p$ parameters. Let $\Sigma$ be the set of input-output equations for $\mathcal{M}$. The *coefficient map* of $\Sigma$ is the function $\bar{c} : \mathbb{R}^p \to \mathbb{R}^l$ that is the vector of all nonmonic coefficient functions of every differential monomial term in every input-output equation in $\Sigma$, where $l$ is the number of nonmonic coefficients.

**Example 2.5.** (Example 2.3 cont.) Returning to the model in Figure 4 with $In = Out = \{1\}$ and $Leak = \{2\}$, we use the input-output equation computed in Example 2.3 to identify the coefficient map of $\mathcal{M}$, $\bar{c} : \mathbb{R}^5 \to \mathbb{R}^5$

$$\begin{pmatrix} k_{02} \\ k_{12} \\ k_{21} \\ k_{23} \\ k_{32} \end{pmatrix} \mapsto \bar{c} = \begin{pmatrix} k_{02}+k_{12}+k_{21}+k_{23}+k_{32} \\ k_{02}k_{21}+k_{02}k_{23}+k_{12}k_{23}+k_{21}k_{23}+k_{21}k_{32} \\ k_{02}k_{21}k_{23} \\ k_{02}+k_{12}+k_{23}+k_{32} \\ k_{02}k_{23}+k_{12}k_{23} \end{pmatrix}$$

**Definition 2.6.** (Model Identifiability)[1, Definition 2.13] Consider a strongly connected linear compartmental model $\mathcal{M} = (G, In, Out, Leak)$ with at least one input. Assume that $|E_G|+|Leak| \geq 1$. Let $\bar{c} : \mathbb{R}^{|E_G|+|Leak|} \to \mathbb{R}^m$ be the coefficient map arising from the input-output equations (1). Then $\mathcal{M}$ is:

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

**Remark 2.7.** A model $\mathcal{M}$ is *strongly connected* if the directed graph $G$ representing the model is strongly connected (that is, every pair of vertices is mutually reachable).

Now that identifiability in general is understand, we can define the identifiability degree of both individual parameters and whole models.

**Definition 2.8.** (Identifiability Degree of a Parameter) The *identifiability degree* of a parameter $k_{ij}$ is $m$ if exactly $m$ parameter values map to a generic input-output data vector.

**Definition 2.9.** (Identifiability Degree of a Model)[6] The *identifiability degree* of a model $\mathcal{M}$ is $m$ if exactly $m$ sets of parameter values map to a generic input-output data vector.

**Remark 2.10.** A model is said to be *globally identifiable* if $m$ in Definition 2.9 is equal to 1.

To determine if a model is identifiable, you can solve the coefficient map for each parameter or determine the rank of the Jacobian matrix, $Jac(\bar{c})$. The model is generically locally identifiable if and only if $Jac(\bar{c})$ has full rank (or $rk(Jac(\bar{c})) = P$, where $P$ is the number of parameters) [1, Proposition 2.14].

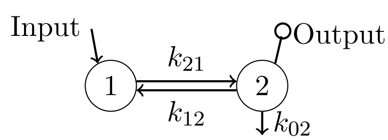



Figure 5: A catenary model with $n = 3$, $In = \{1\}$, and $Out = Leak = \{2\}$.

**Example 2.11.** Consider the model in Figure 5. For this model we can obtain the following coefficient map:

$$(k_{02}, k_{12}, k_{21}) \mapsto \bar{c} = (k_{02} + k_{21} + k_{12}, k_{21}k_{02}, k_{21})$$

Since we can determine $k_{21}$ uniquely from the coefficient map ($\bar{c}$) it is globally identifiable.
$k_{02}$ is in the coefficient $k_{21}k_{02}$, and we know $k_{21}$, so $k_{02}$ can also be determined uniquely.
$k_{12}$ is globally identifiable for similar reasoning using the coefficient $k_{02} + k_{21} + k_{12}$.
Since all of the parameters in the model are globally identifiable, we call the model globally identifiable.

**Remark 2.12.** In Example 2.11, if one of the parameters was instead locally identifiable, the model as a whole would be locally identifiable. Similarly, if one of the parameters was instead unidentifiable, the model as a whole would be unidentifiable.

## 3 Parameter Unidentifiability

While model identifiability is a generally well-understood problem, it is difficult to determine the identifiability degree (or unidentifiability) of individual parameters within the model. The unidentifiability of parameters, especially, is an elusive problem. Despite this, we develop the method exhibited in Example 3.1 to prove the unidentifiability of parameters in some unidentifiable models.

When unidentifiability of a model is known, by definition, we are guaranteed that one of the parameters in that model is itself unidentifiable. Using this fact, we often can prove a relationship between parameters of some models which would guarantee that, if one parameter is identifiable, then every parameter must be identifiable. We illustrate this method using the following example:

**Example 3.1.** Consider the model pictured in Figure 6.

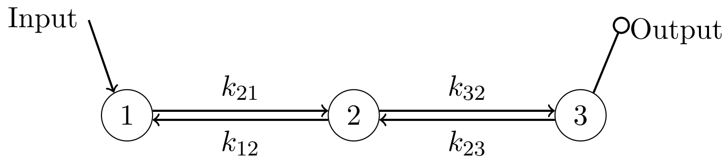



Figure 6: 3-compartment catenary model with input in Compartment 1, output in Compartment 3, and no leaks

It is routine to determine that this model has coefficient map $\mathbf{c} : \mathbb{R}^4 \to \mathbb{R}^3$ given by:

$$\begin{pmatrix} k_{12} \\ k_{21} \\ k_{23} \\ k_{32} \end{pmatrix} \mapsto \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} k_{12} + k_{21} + k_{23} + k_{32} \\ k_{12}k_{23} + k_{21}k_{23} + k_{21}k_{32} \\ k_{21}k_{32} \end{pmatrix}$$

From here, it is possible to show that being able to identify any of the 4 parameters will guarantee the (local) identifiability of all parameters in the model. We assume that all parameters are positive, so the result will hold generically.

First, suppose $k_{21}$ or $k_{32}$ is identifiable. Then, $c_3 = k_{21}k_{32}$ implies that both $k_{21}$ and $k_{32}$ are identifiable. From here, it is straightforward to check that

$$c_2 - c_3 = (k_{12} + k_{21})k_{23} = c_1 k_{23} - k_{23}^2 - k_{23}k_{32} \tag{2}$$

So, we obtain the quadratic equation

$$k_{23}^2 + (k_{32} - c_1)k_{23} + (c_2 - c_3) = 0$$

and conclude that $k_{23}$ is generically locally identifiable if $k_{32}$ is identifiable. With this, 3 of our total 4 parameters are identifiable, and so the last parameter is also identifiable.

It remains to show that if $k_{12}$ or $k_{23}$ are identifiable, then every parameter must be identifiable. We begin with $k_{23}$. Indeed, it follows from equation 2 that if $k_{23}$ is identifiable, then so is $k_{32}$, and hence by the argument above that every parameter is identifiable.

To develop the $k_{12}$ case, consider the following equations:

$$c_1 k_{21} = (k_{12} + k_{21})k_{21} + k_{21}k_{23} + k_{21}k_{32}$$

$$c_2 = k_{12}k_{23} + k_{21}k_{23} + k_{21}k_{32}$$

Then, we obtain the equality:

$$c_1 k_{21} - (k_{12} + k_{21})k_{21} = c_2 - k_{12}k_{23}$$

Notice that we can solve for $k_{23}$ in terms of $k_{12}$ and $k_{21}$, so

$$k_{23} = \frac{c_2 + (k_{12} + k_{21})k_{21} - c_1 k_{21}}{k_{12}}$$

Finally, substituting this value for $k_{23}$ in the equation $(k_{12} + k_{21})k_{23} + c_3 = c_2$ yields a polynomial equation for $k_{21}$ in terms of $k_{12}$, so we conclude that the identifiability of $k_{12}$ implies that of $k_{21}$, and hence of each parameter as discussed earlier.

Recall, finally, that this model is unidentifiable, and has at least one unidentifiable parameter as a result. But, if any one parameter is identifiable, then all four of the model's parameters are identifiable. So, to avoid contradiction, we conclude that each of the model's parameters are unidentifiable.

**Remark 3.2.** We can use the method shown in Example 3.1 to prove the unidentifiability of a subset of parameters in a model, also. While the proof will be omitted, the model pictured in Figure 7 has coefficient map:

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} k_{03} + k_{12} + k_{21} + k_{23} + k_{32} \\ k_{03}(k_{12} + k_{21} + k_{32}) + k_{12}k_{23} + k_{21}k_{23} + k_{21}k_{32} \\ k_{03}k_{21}k_{32} \\ k_{21}k_{32} \end{pmatrix}$$

Here, $\frac{c_3}{c_4} = k_{03}$, so $k_{03}$ is globally identifiable. Despite this, readers may verify using methods similar to those in Example 3.1 that each of the other parameters in the model is unidentifiable.

## 4 Identifiability of Leak Parameters

**Definition 4.1.** (Graphs associated to linear compartmental models) [1, Section 2.2] We define the following auxiliary graphs arising from a linear compartmental model $\mathcal{M} = (G, In, Out, Leak)$.

- $\tilde{G}$: the *leak-augmented graph* [6] is obtained from $G$ by adding a new node, labeled by 0 and referred to as the *leak node*, and for every $j \in Leak$, adding an edge $j \to 0$ with label $k_{0j}$
- $\tilde{G}_i^*$: obtained from $\tilde{G}$ by removing all outgoing edges from some node $i$

**Example 4.2.** The model in Figure 8 has the corresponding graph $G$, as shown in Figure 9. Using $G$, it is simple to obtain $\tilde{G}$ by adding the node 0 and the edge $4 \to 0$. Next, to obtain $\tilde{G}_i^*$ we remove the outgoing edges from node 4 (in this case, we chose $i$ to be 4).

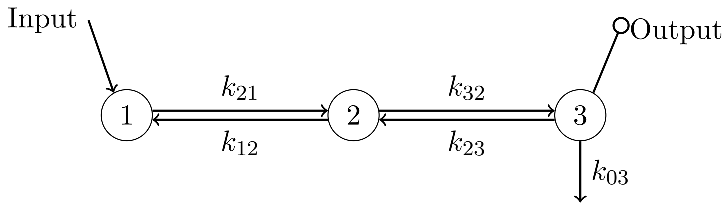



Figure 7: 3-compartment catenary model with input in Compartment 1, output and leak in Compartment 3

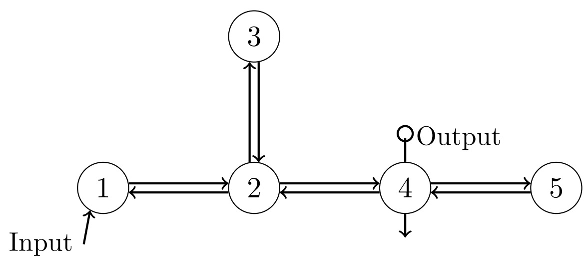



Figure 8: Linear compartmental model with $In = \{1\}$ and $Out = Leak = \{4\}$. This model is used to obtain the graphs in Figure 9.

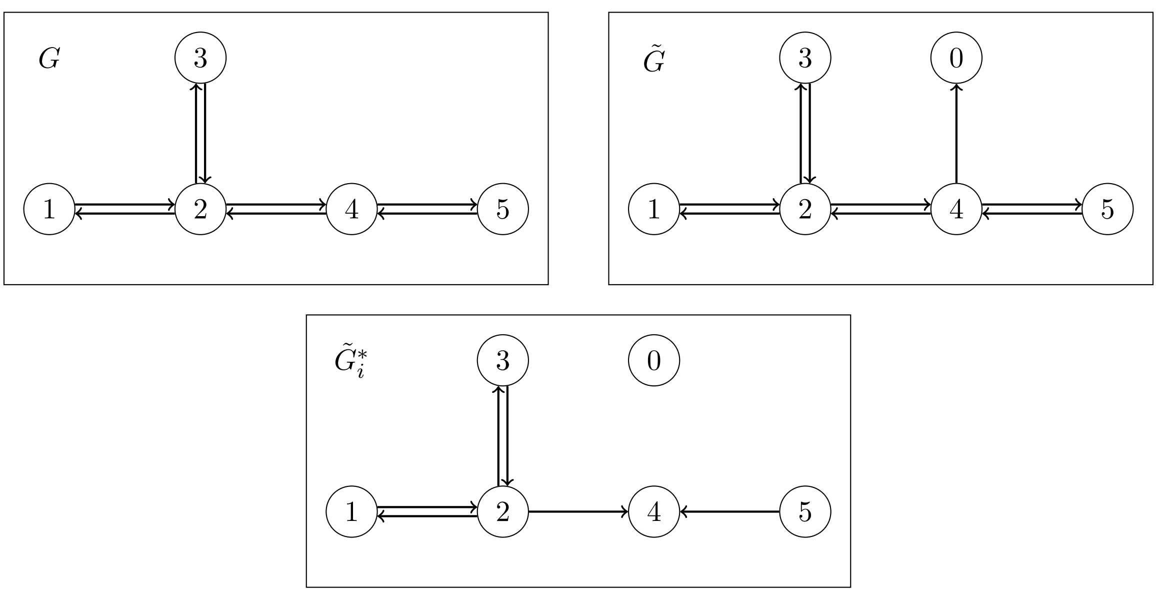



Figure 9: Graphs obtained from the model in Figure 8.

**Definition 4.3.** (Spanning incoming forest) [1, Section 2.2] For a graph, a *spanning incoming forest* is a spanning subgraph for which the underlying undirected graph is a forest (has no cycles) and each node has at most one outgoing edge. "Spanning" refers to the fact that every vertex of the graph is included in the forest, which can include isolated vertices.

For a graph $H$:

- $\mathcal{F}_j(H)$ is the set of all spanning incoming forests of $H$ with exactly $j$ edges.
- $\mathcal{F}_j^{k,\ell}(H)$ is the set of all spanning incoming forests of $H$ with exactly $j$ edges, such that some connected component (of the underlying undirected graph) contains both of the vertices $k$ and $\ell$.

**Theorem 4.4.** *[1, Theorem 3.1] Consider a linear compartmental model $\mathcal{M} = (G, In, Out, Leak)$ with at least one input. Let $n$ denote the number of compartments. Write the input-output equation (1) (for some $i \in Out$) as follows:*

$$y_i^{(n)} + c_{n-1}y_i^{(n-1)} + \cdots + c_1 y_i' + c_0 y_i = \sum_{j \in In} (-1)^{i+j}(d_{j,n-1}u_j^{n-1} + \cdots + d_{j1}u_j' + d_{j0}u_j) \tag{3}$$

*Then the coefficients of the input-output equation (3) are as follows:*

$$c_k = \sum_{F \in \mathcal{F}_{n-k}(\tilde{G})} \pi_F \text{ for } k = 0, 1, ..., n-1$$

$$d_{j,k} = \sum_{F \in \mathcal{F}_{n-k-1}^{ji}(\tilde{G}_i^*)} \pi_F \text{ for } j \in In \text{ and } k = 0, 1, ..., n-1$$

**Proposition 4.5.** Assume $n \geq 3$. For an n-compartment strongly connected model with $In = \{j\}$ and $Out = Leak = \{i\}$, the leak parameter, $k_{0i}$, is globally identifiable.

*Proof.* We want to show that $c_0 = k_{0i}d_{j0}$, where $c_0$ is the last coefficient on the left hand side of the input output equation (3) and $d_{j0}$ is the last coefficient on the right hand side of (3).

By definition, $\mathcal{F}_n(\tilde{G})$ is the set of all spanning incoming forests of $\tilde{G}$ with $n$ edges. From Theorem 4.4, we have

$$c_0 = \sum_{F \in \mathcal{F}_n(\tilde{G})} \pi_F$$

We know that $c_0$ is the sum of products of the edges in $F \in \mathcal{F}_n(\tilde{G})$. In the cases we are considering, for $c_0$, $|\mathcal{F}_n(\tilde{G})| = 1$ because, with there being a leak out of $i$ in $\tilde{G}$, there is only one possible forest we are able to obtain that is a *spanning incoming* forest. Since this forest is a spanning incoming forest of $\tilde{G}$, $i$ can have only one outgoing edge which is, in this case, the leak edge $k_{0i}$.

By definition, $\mathcal{F}_{n-1}^{ji}(\tilde{G}_i^*)$ is the set of all spanning incoming forests of $\tilde{G}_i^*$ with $n-1$ edges. From Theorem 4.4, we have

$$d_{j0} = \sum_{F \in \mathcal{F}_{n-1}^{ji}(\tilde{G}_i^*)} \pi_F$$

We know that $d_{j0}$ is the sum of products of the edges of $F \in \mathcal{F}_{n-1}^{ji}(\tilde{G}_i^*)$. In the cases we are considering, for $d_{j0}$, $|\mathcal{F}_{n-1}^{ji}(\tilde{G}_i^*)| = 1$ (for similar reasoning as for $c_0$), and since this is a spanning incoming forest of $\tilde{G}_i^*$, $i$ will have no outgoing edges (in the construction of $\tilde{G}_i^*$ the outgoing edges are removed from the vertice $i$). It is straightforward to check, then, that the only difference between $\mathcal{F}^{ji}_{n-1}(\tilde{G}^*_i)$ and $\mathcal{F}_n(\tilde{G})$ is the leak edge in $\mathcal{F}_n(\tilde{G})$.

Since $\mathcal{F}_n(\tilde{G})$ contains the leak edge and $\mathcal{F}^{ji}_{n-1}(\tilde{G}^*_i)$ does not, $c_0 = k_{0i}d_{j0}$. Hence, the leak parameter $k_{0i}$ is globally identifiable. □

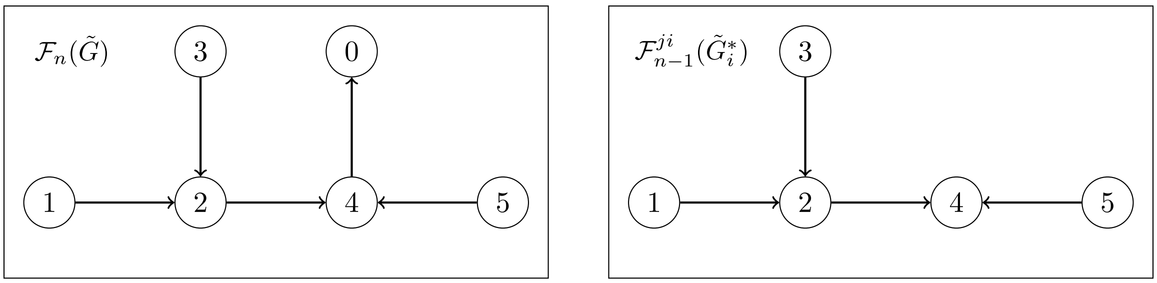



Figure 10: Graphs in $\mathcal{F}_n(\tilde{G})$ and $\mathcal{F}^{ji}_{n-1}(\tilde{G}^*_i)$, with $\tilde{G}$ and $\tilde{G}^*_i$ from Figure 9.

**Example 4.6.** (Example 4.2 cont.) For the model in Figure 8, we can obtain $\mathcal{F}_n(\tilde{G})$ and $\mathcal{F}^{ji}_{n-1}(\tilde{G}^*_i)$ (shown in Figure 10). Using these graphs, we can see

$$c_0 = \sum_{F \in \mathcal{F}_5(\tilde{G})} \pi_F = k_{04}k_{21}k_{23}k_{42}k_{45}$$

and

$$d_{10} = \sum_{F \in \mathcal{F}^{14}_4(\tilde{G}^*_4)} \pi_F = k_{21}k_{23}k_{42}k_{45}$$

As we expected, $c_0 = k_{04}d_{10}$ so $k_{04}$ is globally identifiable.

In order to improve this result, we make a few conjectures.

**Conjecture 4.7.** *Assume $n \geq 3$. For an n-compartment model where every compartment has a directed path to $i$ with $In = \{j\}$ and $Out = Leak = \{i\}$, the leak parameter, $k_{0i}$, is globally identifiable.*

This conjecture changes the condition in Proposition 4.5 of a "strongly connected model" to a looser condition: that every compartment has a directed path to $i$ (the output compartment). We can prove this in the case of certain directed path models, models where there is a directed (not bi-directed) path from input to output.

**Proposition 4.8.** Assume $n \geq 3$. For an n-compartment directed path model with $In = \{j\}$ and $Out = Leak = \{n\}$, the leak parameter, $k_{0n}$, is globally identifiable.

*Proof.* Similar to the proof of Proposition 4.5, we want to show that $c_0 = k_{0n}d_{j0}$, where $c_0$ is the last coefficient on the left hand side of the input output equation (3) and $d_{j0}$ is the last coefficient on the right hand side of (3).

We have that $c_0$ is the sum of products of the edges of $F \in \mathcal{F}_n(\tilde{G})$, and $|\mathcal{F}_n(\tilde{G})| = 1$ for similar reasoning as in the proof of Proposition 4.5. Because $F$ is a spanning incoming forest of $\tilde{G}$ and because of the structure of a directed path model, $k_{0n}$ will be the only outgoing edge of $n$.

Similarly, we have that $d_{j0}$ is the sum of products of the edges of $F \in \mathcal{F}_{n-1}^{jn}(\tilde{G}_n^*)$, and $|\mathcal{F}_{n-1}^{jn}(\tilde{G}_n^*)| = 1$ for the same reasoning as $c_0$. Because of the structure of $\tilde{G}_n^*$, where the outgoing edges of $n$ are removed, all of the edges except $k_{0n}$ in $F \in \mathcal{F}_{n-1}^{jn}(\tilde{G}_n^*)$ will be the same as $F \in \mathcal{F}_n(\tilde{G})$, giving us that $c_0 = k_{0n} d_{j0}$.

Therefore, the leak parameter $k_{0n}$ is globally identifiable. □

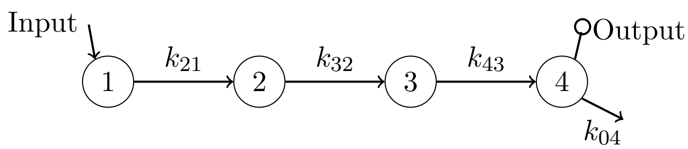



Figure 11: Directed path model with $n = 4$, $In = \{1\}$, and $Out = Leak = \{4\}$.

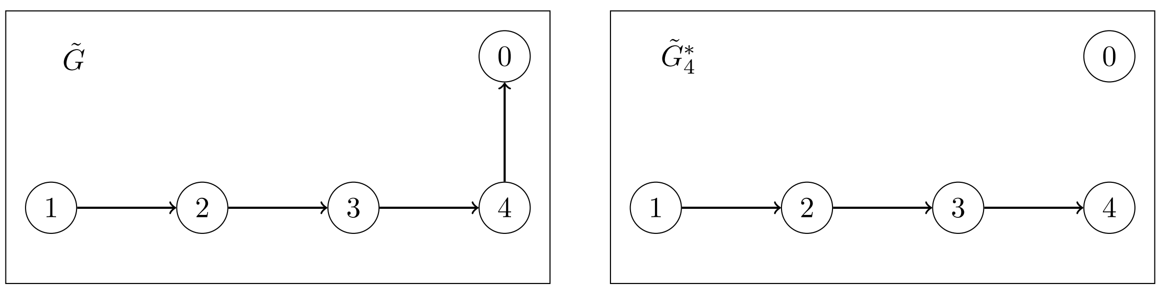



Figure 12: Graphs obtained from the model in Figure 11.

**Example 4.9.** The model in Figure 11 is a directed path model because there is a directed path from input to output, and Proposition 4.8 holds because we have that $In = \{j\} = \{1\}$ and $Out = Leak = \{n\} = \{4\}$.

From this model, we can obtain the graphs $\tilde{G}$ and $\tilde{G}_4^*$, shown in Figure 12. Since $\tilde{G}$ and $\tilde{G}_4^*$ are already the only spanning incoming forest of $\tilde{G}$ and $\tilde{G}_4^*$, we can see that product of the edges in $\mathcal{F}_4(\tilde{G}) = \tilde{G}$ includes all of the same edges as $\mathcal{F}_3^{14}(\tilde{G}_4^*) = \tilde{G}_4^*$ in addition to the leak edge, $4 \to 0$ labeled as $k_{04}$. This gives us $c_0 = k_{04} d_{10}$, so $k_{04}$ is globally identifiable.

**Conjecture 4.10.** *Assume $n \geq 3$. For an n-compartment model, if a leak is introduced at the same compartment as the output, then the identifiability degree of the non-leak parameters does not change.*

**Remark 4.11.** In Conjecture 4.10, the idea of introducing a leak to a model is suggested. This idea, along with other operations and their effects on identifiability, is investigated further in [5].

**Example 4.12.** In Figure 13, $\mathcal{M}'$ is obtained by introducing a leak to $\mathcal{M}$. It is clear to see that the identifiability degree of the non-leak parameters remains the same when the leak is introduced to the same compartment as the output, as we expect from Conjecture 4.10.

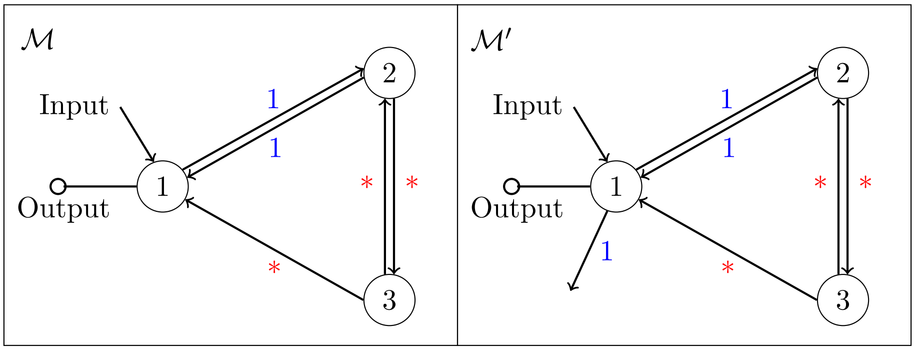



Figure 13: On the left, linear compartmental model $\mathcal{M}$ with $In = Out = \{1\}$ and the identifiability degree of the parameters labeled, with $*$ denoting an unidentifiable parameter. On the right, $\mathcal{M}'$ is $\mathcal{M}$ with the operation of adding a leak in compartment 1 applied.

# 5 Parameter Identifiability

We begin by exploring the identifiability degree of parameters in cycle models containing at most one leak, before discussing some mammillary models. Before this, though, we must define the elementary symmetric polynomials on a finite set.

**Definition 5.1.** (Elementary Symmetric Polynomials) Let $A = \{\alpha_1, ..., \alpha_n\}$ be a finite set of variables. Then, for any $i \in \{1, ..., n\}$, the $i^{th}$ *elementary symmetric polynomial* on the elements of $A$ are given by

$$e_i(A) := \sum_{1 \le k_1 < k_2 < ... < k_i \le n} \alpha_{k_1} ... \alpha_{k_i}$$

If $i > n$, then we define $e_i(A) := 0$. Also, we define $e_0(A) = 1$ for any set $A$.

## 5.1 Cycle Models

In this subsection, we will develop a strategy for bounding identifiability degree of models and parameters above using the coefficient map, by proving those values for cycle models with less than two leaks. Before we begin our proofs, we make the following remark on notation:

**Remark 5.2.** In general, in cycle models with exactly one input, we reindex compartments to ensure the input is in compartment 1. This has no effect, however, on the dynamics of our model. In models with multiple inputs, we similarly reorder so that at least one of those inputs is in compartment 1. As such, we assume that there is always an input in compartment 1 in cycle models.

With this, we are equipped to begin proving the identifiability of some cycle models. We start with a lemma on cycle models containing exactly one input and one output, whose output is not in the last compartment.

**Lemma 5.3.** *Consider the $n$-compartment cycle model with $In = \{1\}$, $Out = \{i\}$ for some $i \in \{1, \ldots, n-1\}$, and $Leak = \emptyset$. Let $F = \{k_{21}, k_{32}, \ldots, k_{i,i-1}\}$ be the set of edges between compartment*

*1 and compartment $i$ and $G = \{k_{i+2,i+1}, \dots, k_{1n}\}$ be the set of edges between compartment $i+1$ and compartment 1. Then, edges in $F$ have identifiability degree $(i-1)$, edges in $G$ have identifiability degree $(n-i)$, and $k_{i+1,i}$ is globally identifiable. Moreover, the identifiability degree of the model is $(n-i)!(i-1)!$*

*Proof.* For additional notation, we will define the set $E := \{k_{21}, k_{32}, \dots, k_{1n}\}$ as the set of all edges in our model. Also, we'll say that compartment $n+1$ is the same as compartment 1, for simplicity when iterating through compartments.

It follows routinely that the $(\partial I - A)$ matrix for these models is given by:

$$(\partial I - A) = \begin{pmatrix} \partial + k_{21} & 0 & 0 & \dots & -k_{1n} \\ -k_{21} & \partial + k_{32} & 0 & \dots & 0 \\ 0 & -k_{32} & \partial + k_{43} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \partial + k_{1n} \end{pmatrix}$$

With this, we use Proposition 2.1 determine that the models have the input-output equation

$$\left[\prod_{1 \le \ell \le n} (\partial + k_{\ell+1,\ell}) - k_{21}k_{32}\dots k_{1n}\right] y_i = k_{21}k_{32}\dots k_{i,i-1}\left[\prod_{i+1 \le \ell \le n} (\partial + k_{\ell+1,\ell})\right] u_1$$

Expanding, the products in the input-output equation have coefficients on $\partial$ given by the elementary symmetric polynomials on the sets $E$ and $G$:

$$[\partial^n + e_1(E)\partial^{n-1} + e_2(E)\partial^{n-1} + \dots + e_{n-1}(E)\partial]y_i = k_{21}k_{32}\dots k_{i,i-1}[\partial^{n-i} + e_1(G)\partial^{n-i-1} + \dots + e_{n-i}(G)]u_1$$

So, models of the desired type have a coefficient map which maps

$$\begin{pmatrix} k_{21} \\ k_{32} \\ \vdots \\ k_{1n} \end{pmatrix} \mapsto \begin{pmatrix} e_1(E) \\ e_2(E) \\ \vdots \\ e_{n-1}(E) \\ k_{21}k_{32}\dots k_{i,i-1} \\ [k_{21}k_{32}\dots k_{i,i-1}]e_1(G) \\ \vdots \\ [k_{21}k_{32}\dots k_{i,i-1}]e_{n-i}(G) \end{pmatrix}$$

To determine parameter identifiability, we begin by considering edges in $G$. Note that we may divide $k_{21}k_{32}\dots k_{i,i-1}$ away from each of the last $n-i$ entries in the coefficient map to find exactly the values of the elementary symmetric polynomials on $G$. So, by properties of the elementary symmetric polynomial, we find that each edge in $G$ has at most $n-i$ solutions, and that the set has a total $(n-i)!$ possible permutations of those values. With this, note that we have bounded the identifiability degree of all edges in $G$ by $n-i$.

Now, we will use strong induction to find the identifiability degree of $k_{i+1,i}$. Let us define the set $F' := F \cup \{k_{i+1,i}\}$, so that $F' = E \setminus G$.

As our base case, recall that the coefficient map guarantees us the values of $e_1(E)$ and $e_1(G)$. So, we may find exactly the value of $e_1(E) - e_1(G) = e_1(F')$.

Assume, now, that the first $m$ elementary symmetric polynomials on $F'$ are known (for some $m < i$). Then,

$$e_{m+1}(E) = e_{m+1}(G) + e_m(g)e_1(F') + e_{m-1}(G)e_2(F') + \cdots + e_1(G)e_m(F') + e_{m+1}(F')$$

Using this equation, by induction, every value is known except $e_{m+1}(F')$, so we can solve to determine that value exactly. Hence, by induction, we find $e_\ell(F')$ for every degree $\ell$.

In particular, now, $e_i(F') = k_{21}k_{32}\dots k_{i,i-1}k_{i+1,i}$ is known. But, we already know the value of $k_{21}k_{32}\dots k_{i,i-1}$ from the coefficient map. So, we divide this value away from $e_i(F')$ to uniquely identify $k_{i+1,i}$.

It remains to determine the identifiability of edges in $F$. We can use the same induction as earlier to find each of the $i-1$ nontrivial elementary symmetric polynomials on $F$, given the elementary symmetric polynomials on $F'$ and $k_{i+1,i}$, since $F$ and $k_{i+1,i}$ partition $F$.

With this, if each of the elementary symmetric polynomials on the $i-1$ elements of $F$ are known, then we can bound the number of possible solutions for elements of $F$ above by $i-1$, as desired. Then, the total number of possible permutations on those solutions is bounded above by $(i-1)!$.

It remains to bound the identifiability degree of the model. The steps we used to determine the number of solutions of $F$ and $G$ were independent, so we have no information available to strengthen a bound beyond those guaranteed by our solutions for those subsets of $E$. In particular, then, we bound the identifiability degree of the model above by $(n-i)!(i-1)!$.

□

We will make a few remarks about the steps used in this proof and the strength of its result:

**Remark 5.4.** This result is stated as a lemma since its proof method (and result) is generalized by Proposition 5.9. With this, we study this lemma to develop our understanding of the coefficient maps and $(\partial I - A)$ matrices of cycle models.

**Remark 5.5.** At times, we've called the upper bounds on solutions found in this proof the identifiability degree of those solutions (both for parameters and the model itself). While we conjecture that this is exactly the identifiability degree, we have not yet found a condition which generically guarantees a sharp lower bound on identifiability degree. In the following cases, this will remain the case, and we confidently conjecture in each case that the upper bound given is generically the identifiability degree, for the model and each of its edges and leaks.

In this case, the strongest lower bound we may find is $(n-i)!(i-2)!$, where the $(n-i)!$ is from the symmetry of edges in $G$ and the $(i-2)!$ is from the symmetry of edges in $F \setminus \{k_{21}\}$, since we do not yet know the effect of sharing a compartment with an input as happens for $k_{21}$.

The case where a leak is introduced to the above model follows very similarly:

**Lemma 5.6.** *Consider the n-compartment cycle model with $In = \{1\}$, $Out = \{i\}$ $(1 \leq i \leq n-1)$, and $Leak = \{m\}$ for some $m \in \{1, \dots n\}$. Then, identifiability degree of the model and all edges are as in Lemma 5.3, and the identifiability degree of $k_{0m}$ is exactly the identifiability degree of $k_{m+1,m}$.*

*Proof.* If $i = m$, then this case follows directly from Proposition 4.5.

For the other cases, let us maintain our definition $E := \{k_{21}, k_{32}, \dots k_{1n}\}$ and establish a new definition, $E' = \{k_{21}, k_{32}, \dots, k_{m+1,m} + k_{0m}, \dots, k_{1n}\}$.

Now, we consider the case $m < i$. In this case, coefficients obtained from the right-hand side of the input-output equation in Lemma 5.3 remain the same as previously. In fact, one can verify that the only change to the coefficient map is that entries on the left-hand side now are elementary symmetric polynomials on $E'$ rather than $E$, and with this, we gain a new entry $e_n(E') - k_{21}k_{32}\dots k_{1n} = k_{0m}k_{21}\dots \widehat{k_{m+1,m}}\dots k_{1n}$.

So, we can find the identifiability degree of entries in $G$ and $k_{i+1,i}$ using the same methods as in the proof of Lemma 5.3. By reducing once more our elementary symmetric polynomials on $E'$, we obtain the set of nontrivial elementary symmetric polynomials on $\{k_{21}, k_{32}, \dots, k_{m+1,m} + k_{0m}, \dots, k_{i,i-1}\}$ and find each of the set entries' value, up to $i-1$ solutions. Note that this process follows that of Lemma 5.3 exactly, up to the elements of our set of $i-1$ elements.

Now, we have determined the value of $k_{m+1,m} + k_{0m}$ up to $i-1$ solutions, which is dependent on the values taken by each of the elements in $F \setminus \{k_{m+1,m}\}$. But, the coefficient map still contains the product $k_{21}k_{32}\dots k_{i,i-1}$. So, knowing the value of each element in $F \setminus \{k_{m+1,m}\}$ also uniquely determines the value of $k_{m+1,m}$. So, we find a unique value of $k_{m+1,m}$ for each of the $i-1$ total values $(k_{m+1,m} + k_{0m}$ may take, and hence conclude that $k_{0m}$ may itself take $i-1$ values. Since $k_{0m}$ is determined by other edge values, it does not affect model identifiability degree.

Therefore, if $m \leq i$, then the identifiability degree of all edges, and the model itself, are preserved from the case with no leak. Furthermore, in these cases, the identifiability degree of $k_{0m}$ is exactly the identifiability degree of $k_{m+1,m}$.

It remains to prove the case $m > i$. In this case, the product on each side of the input-output equation now contains $k_{m+1,m} + k_{0m}$ in place of $k_{m+1,m}$.

Using the same steps as are contained in the proof of Lemma 5.3, we find the identifiability degree of all edges except $k_{m+1,m}$ to be preserved, and by substitution see that $k_{m+1,m} + k_{0m}$ can be identified up to $n-i$ solutions, which are dependent on the values taken by elements of $G \setminus \{k_{m+1,m}\}$.

Note also that it follows from the steps of proving that number of solutions that we may use induction to uniquely determine the value of $k_{21}k_{32}\dots k_{i+1,i}$. Also, by symmetry, we again find the value of $k_{i+2,i+1} \dots \hat{k_{m+1,m}} \dots k_{1n}$ up to exactly $n-i$ solutions.

As mentioned above, the product $k_{0m}k_{21} \dots \hat{k_{m+1,m}} \dots k_{1n}$ is now given in the coefficient map of the model. So, we find the product $k_{21}k_{32} \dots \hat{k_{m+1,m}} \dots k_{1n}$ up to $n-i$ solutions, and hence find $k_{0m}$ up to $n-i$ solutions, each determined by the values of edges in $G \setminus \{k_{m+1,m}\}$.

So, since those $n-i$ sets of distinct edge values in $G \setminus \{k_{m+1,m}\}$ determine $k_{0m}$ and $k_{m+1,m} + k_{0m}$, we can find a unique value of $k_{m+1,m}$ for each $k_{0m}$, and hence conclude that identifiability degree of $k_{m+1,m}$ is preserved from the case with no leaks, as desired.

Therefore, the introduction of a leak to a model of the type discussed in Lemma 5.3 preserves identifiability degree of the model and of all edges, and the leak itself has identifiability degree equal to that of the edge leaving its compartment.

□

Now, we will quickly discuss the case where the output is in compartment $n$, before moving on to the cases with many inputs and/or outputs.

**Proposition 5.7.** Consider the n-compartment cycle model with $In = \{1\}$, $Out = \{n\}$, and $Leak = \emptyset$. In this model, the identifiability degree of $k_{1n}$ is $n-1$, the identifiability degree of all other edges is $(n-1)^2$, and the model identifiability degree is $(n-1)(n-1)!$

*Proof.* It is routine to determine that the coefficient map for models of this class is given by

$$c : \begin{pmatrix} k_{21} \\ k_{32} \\ \vdots \\ k_{1n} \end{pmatrix} \mapsto \begin{pmatrix} e_1(E) \\ e_2(E) \\ \vdots \\ e_{n-1}(E) \\ k_{21}k_{32} \dots k_{n,n-1} \end{pmatrix}$$

where $E := \{k_{21}, \dots, k_{1n}\}$ is the set of all edges.

Also, let $F := \{k_{21}, k_{32}, \dots, k_{n,n-1}\}$ be the set of all edges except $k_{1n}$. Then, $F$ and $\{k_{1n}\}$ partition $E$, so we gain the relation

$$e_i(E) = e_i(F) + k_{1n}e_{i-1}(F) \tag{4}$$

, for every value $i$ (assuming $e_{-k} = 0$ for any set, where $-k < 0$).

Note as well that $e_{n-1}(F)$ is given as the last entry in the coefficient map. So, we may use the relation given in Equation 4 – with the previous entry in the coefficient map, $e_{n-1}(E)$, to determine $k_{1n}e_{n-2}(F)$.

We can continue for smaller degrees in the same fashion, determining $k_{1n}^{k-1}e_{n-k}(F)$ for each degree $k \in \{1, \dots, n\}$. In particular, we can find $k_{1n}^{n-1}e_0(F) = k_{1n}^{n-1}$, and hence can find $n-1$ solutions for $k_{1n}$.

While finding those values for $k_{1n}$, we extracted each of the elementary symmetric polynomials on $F$, dependent on $k_{1n}$. That is, given $k_{1n}$, we can find the elementary symmetric polynomials on $F$, and hence can determine the entries of $F$ up to $n-1$ solutions for each edge, with a total solution set size of $(n-1)!$.

Since those values are dependent on the $k_{1n}$ value chosen, we conclude that there are $(n-1)^2$ values for each edge in $F$, and a total of $(n-1)(n-1)!$ possible edge combinations for the model. □

It remains to introduce a leak to this system.

**Proposition 5.8.** Consider the n-compartment cycle model described in Proposition 5.7. If a leak is introduced in the $n^{th}$ compartment, then it is globally identifiable, and identifiability degree of the model and its edges are preserved. If a leak is introduced in any other compartment, then the identifiability degree of the model becomes $n!$, $k_{1n}$ has $n$ solutions, and all other parameters (including the leak) have identifiability degree $n(n-1)$.

*Proof.* The introduction of a leak in compartment n follows from Proposition 4.5. That identifiability degree is preserved in this case is a consequence of the procedure used to prove Proposition 4.5, as well as that in the proof of 5.7.

Suppose, then, that we introduce a single leak in compartment $m \in \{1, \ldots, n-1\}$. Then, our coefficient map is given by

$$c : \begin{pmatrix} k_{0m} \\ k_{21} \\ \vdots \\ k_{1n} \end{pmatrix} \mapsto \begin{pmatrix} e_1(E') \\ e_2(E') \\ \vdots \\ e_{n-1}(E') \\ k_{0m}k_{21}k_{32}\ldots \widehat{k_{m+1,m}} \ldots k_{1n} \\ k_{21}k_{32}\ldots k_{n,n-1} \end{pmatrix}$$

where $E'$ is the set of all edges, substituting $(k_{m+1,m} + k_{0m})$ for $k_{m+1,m}$.

With this, let $a := k_{21}k_{32}\ldots k_{n,n-1}$ be the last entry in the coefficient map, and $b := k_{0m}k_{21}\ldots \widehat{k_{m+1,m}} \ldots k_{1n}$ be the previous entry in the coefficient map. Then,

$$k_{1n}a + b = e_n(E') = k_{1n}e_{n-1}(F') \tag{5}$$

where $F' = E' \setminus k_{1n}$.

From here, we reduce the degree of the elementary symmetric polynomial on $F'$ as in the proof of Proposition 5.7, this time determining $k_{1n}$ up to $n$ solutions. Then, we use the same procedure to find $n-1$ solutions for each entry in $F'$. In particular, given values for all edges except $k_{m+1,m}$, we can find $n(n-1)$ solutions for $(k_{m+1,m} + k_{0m})$.

Finally, we can use $a$ to find $n(n-1)$ solutions for $k_{m+1,m}$. Pairing this with the values found for $(k_{m+1,m} + k_{0m})$, we find $n(n-1)$ solutions for all parameters except $k_{1n}$ and conclude that the model has an overall identifiability degree of $n(n-1)! = n!$. □

The next case to observe is the general case with up to one leak, and many inputs and/or outputs.

**Proposition 5.9.** Consider an n-compartment cycle model with $In \cup Out = \{i_1 < \cdots < i_k\}$, $Leak = \emptyset$, $In = \{i_{1_1} < \cdots < i_{1_j}\}$, and $Out = \{i_{2_1} < \cdots < i_{2_\ell}\}$. Also, assume that each edge has a distinct value. Then, edges leaving output compartments are globally identifiable. Furthermore, edges between compartments $i_m$ and $i_{m+1}$ have identifiability degree $D_m := (i_{m+1} - i_m)$ if there is no output in compartment $i_m$, and $D_m := (i_{m+1} - i_m - 1)$ otherwise. Then, the model has identifiability degree $\prod_{m=1}^{k} (D_m)!$

*Proof.* We give an outline of the proof method for this statement:

Given a set of inputs, we can consider the system formed by only considering one input at a time, with each output – the same process used in determining the input-output equations with many inputs and outputs. Then, recall that to prove the identifiability of the cycle model's parameters, we found the elementary symmetric polynomials on subsets of parameters determined by the location of the output relative to the input.

We can repeat this process for each distinct input-output pair, finding the elementary symmetric polynomials on every such subset of edges. In this process, we determine all edges leaving output compartments globally.

Then, note that the set of points in the domain of the coefficient map which do not have distinct parameters have Lebesgue measure zero, and hence excluding those points will keep any result generically.

Using this property, we can consider the overlap in edge values between adjacent edge sets to determine the values of the parameters on the paths of minimal length between input and output locations, the length of which is $(i_{m+1} - i_m)$, if $i_m$ is not an output compartment, and which is $(i_{m+1} - i_m - 1)$ if compartment $i_m$ contains and output.

With this, we have found the values (up to symmetry) of edges between compartments $i_m$ and $i_{m+1}$, for any $m \in 1, \dots, k$. We now use the same argument as earlier to conclude that the model's overall identifiability degree is the product of each smaller path identifiability degree, given by $\prod_{m=1}^{k} (D_m)!$ as defined in the proposition statement. □

**Remark 5.10.** While in the single input and output case, we had to worry about the location of our output, upon the introduction of another input or output, this problem becomes obsolete. Indeed, if another input is introduced, then the new input compartment is not preceded by the output, and hence the edge leaving the output compartment will behave as expected. Alternatively, in the case with another output and a single input, the set of parameters between outputs is determined up to symmetry, as are those preceding the new output. Also, though, the set of parameters between the new output and the input are determined up to symmetry, so since parameters are distinct, we can uniquely determine the value of $k_{1n}$.

**Corollary 5.11.** *In a cycle model with multiple inputs or outputs and exactly one leak, the identifiability degree of the model and its parameters are the same as in the case with no leak. In addition, the identifiability degree of the leak is exactly the identifiability degree of the edge leaving its compartment.*

*Proof.* This result follows from the procedure in the proof outline of Proposition 5.9 and the discussion of adding a leak to models with exactly one input and output. The same symmetry used in the proof of Proposition 5.9 holds, and as discussed in the remark, we no longer are concerned with the location of outputs relative to inputs. So, we can use the same logic as in the proof of Lemma 5.6 to find the identifiability degree of the leak. □

## 5.2 Mammillary Models

**Definition 5.12.** A linear compartmental model $\mathcal{M}$ is $\mathcal{M}_n(2,2)$ if $\mathcal{M} = (G, In, Out, Leak)$, where $G$ is the n-compartment mammillary graph with central compartment 1, $In = Out = \{2\}$, and $Leak = \emptyset$. The general model $\mathcal{M}_n(2,2)$ is shown in Figure 14.

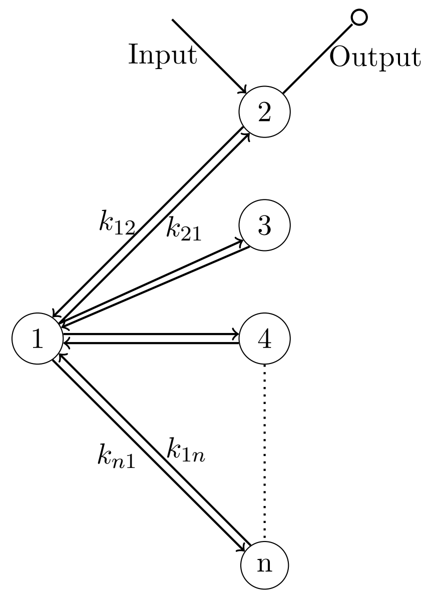



Figure 14: n-compartment mammillary model in $\mathcal{M}_n(2,2)$ with $In = Out = \{2\}$ and without leaks

We will find and prove the identifiability degree of $\mathcal{M}_n(2,2)$ models, as well as that of each of the parameters in those models, using the coefficient map. For our result, we require the following lemma:

**Lemma 5.13.** *Let $n \in \mathbb{N}$. The set* $A = \left\{ \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} e_{1,1} \\ e_{1,2} \\ \vdots \\ e_{1,n} \end{pmatrix}, ..., \begin{pmatrix} e_{n-1,1} \\ e_{n-1,2} \\ \vdots \\ e_{n-1,n} \end{pmatrix} \right\}$ *is linearly independent over $\mathbb{R}[\alpha_1, ..., \alpha_n]$, where each $\alpha_i$ is a distinct nonzero variable and $e_{i,j}$ denotes the $i^{th}$ elementary symmetric polynomial on the set $A_j = \{\alpha_1, ..., \hat{\alpha_j}, ..., \alpha_n\}$ of all variables except $\alpha_j$.*

*Proof.* We begin with the $n = 1$ case, before using induction on the dimension of $A$ to prove other cases. If $n = 1$, then $A$ contains only the vector $\begin{pmatrix} 1 \end{pmatrix}$, and therefore is linearly independent over $\mathbb{R}[\alpha_1, \ldots, \alpha_n]$.

If $n = 2$, then $A = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \alpha_2 \\ \alpha_1 \end{pmatrix} \right\}$. Now, suppose that there exist $b_0, b_1 \in \mathbb{R}[\alpha_1, \ldots, \alpha_n]$ such that $b_0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + b_1 \begin{pmatrix} \alpha_2 \\ \alpha_1 \end{pmatrix} = 0$. Then, $b_0 + b_1\alpha_2 = b_0 + b_1\alpha_1$, so $b_1(\alpha_2 - \alpha_1) = 0$. Since $\alpha_1 \neq \alpha_2$ (by assumption), we require $b_1 = 0$. So, our equations reduce to $b_0 = 0$ and we conclude that $A$ is linearly independent over $\mathbb{R}[\alpha_1, \ldots, \alpha_n]$.

Before our induction, it is important to introduce the following notation. In this proof, we will use $v_j$ to refer to the element of $A$ whose entries have degree $j$. That is, $v_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ and in general,

$$v_j = \begin{pmatrix} e_{j,1} \\ e_{j,2} \\ \vdots \\ e_{j,n} \end{pmatrix}.$$

Now, for induction, suppose that $A = \left\{ \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} e_{1,1} \\ e_{1,2} \\ \vdots \\ e_{1,n-1} \end{pmatrix}, \ldots, \begin{pmatrix} e_{n-2,1} \\ e_{n-2,2} \\ \vdots \\ e_{n-2,n-1} \end{pmatrix} \right\}$ is linearly independent over $\mathbb{R}[\alpha_1, \ldots, \alpha_{n-1}]$, for some natural number $n \geq 3$.

To prove the $n+1$ case, suppose that there exist polynomials $b_0, \ldots, b_{n-1} \in \mathbb{R}[\alpha_1, \ldots, \alpha_n]$ such that $\sum_{k=0}^{n-1} b_k v_k = 0$. Then, in particular, for each $j \in \{1, \ldots, n\}$, we obtain the equation $\sum_{k=0}^{n-1} b_k e_{k,j} = 0$ (where $e_{0,j} = 1$ for every index $j$).

Now, consider the system of $n - 1$ equations given by $\sum_{k=0}^{n-1} b_k(e_{k,\ell} - e_{k,n}) = 0$, for each $\ell \in \{1, \ldots, n-1\}$. Note that for each $\ell$, these equations simplify to $(\alpha_n - \alpha_\ell) \sum_{k=1}^{n-1} b_k e_{k-1,\{\ell,n\}} = 0$, where $e_{k,\{\ell,n\}}$ denotes the $k^{th}$ elementary symmetric polynomial on the set $\{\alpha_1, \ldots, \hat{\alpha_\ell}, \ldots, \alpha_{n-1}\}$.

Since $\alpha_\ell$ and $\alpha_n$ are distinct for any $\ell \in \{1, \ldots, n-1\}$, we divide to obtain the set of equations $\sum_{k=1}^{n-1} b_k e_{k-1,\{\ell,n\}}$. Observe that $\alpha_n$ does not appear in any polynomial of the form $e_{k,\{\ell,n\}}$. Then, any linear dependence relation over $\mathbb{R}[\alpha_1, \ldots, \alpha_n]$ of those polynomials may be reduced not to contain $\alpha_n$ (that is, may be rewritten with coefficients in $\mathbb{R}[\alpha_1, \ldots, \alpha_{n-1}]$). So, to prove linear independence of these terms over $\mathbb{R}[\alpha_1, \ldots, \alpha_n]$, it suffices to show linear independence over $\mathbb{R}[\alpha_1, \ldots, \alpha_{n-1}]$.

But, this is exactly the set of equations generated by expanding the vectors in the $n^{th}$ case, up to coefficient reindexing. So, by induction, $b_1 = \cdots = b_{n-1} = 0$, and by returning to the original linear combination we conclude that $b_0 = 0$, also.

So, for any $n \in \mathbb{N}$, $A$ is linearly independent over $\mathbb{R}[\alpha_1, \ldots, \alpha_n]$ as desired.

□

With this, we may find the identifiability degree of $\mathcal{M}_n(2,2)$ mammillary models, as well as the identifiability degrees of the models' parameters. This builds on the work done in [2, Proposition 3.13].

**Proposition 5.14.** The identifiability degree of $\mathcal{M}_n(2,2)$ is $(n-2)!$ Furthermore, $k_{12}$ and $k_{21}$ are globally identifiable, and each other parameter has identifiability degree $n-2$.

*Proof.* Identifiability degree of $k_{12}$ and $k_{21}$ follows from [2, Proposition 3.13]. The identifiability degree of the model is also given in this theorem, but we'll develop an alternative method to determine it below as we find the identifiability degree of all parameters.

It is routine to show that the matrix $(\partial I - A)$ is given by

$$\partial I - A = \begin{pmatrix} \partial + k_{21} + k_{31} + ... + k_{n1} & -k_{12} & -k_{13} & \dots & -k_{1n} \\ -k_{21} & \partial + k_{12} & 0 & \dots & 0 \\ -k_{31} & 0 & \partial + k_{13} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -k_{n1} & 0 & 0 & ... & \partial + k_{1n} \end{pmatrix}$$

Then, by Proposition 2.1, the entries in the coefficient map for $\mathcal{M}_n$ are given by the coefficients of $\partial$ in $det(\partial I - A)$ and $det((\partial I - A)_{22})$. We calculate

$$\begin{aligned} det(\partial I - A) &= -k_{21} det \begin{pmatrix} -k_{12} & -k_{13} & ... & -k_{1n} \\ 0 & \partial + k_{13} & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \partial + k_{1n} \end{pmatrix} - (\partial + k_{12}) det((\partial I - A)_{22}) \\ &= k_{12}k_{21} \prod_{3 \le i \le n} (\partial + k_{1i}) - (\partial + k_{12}) det((\partial I - A)_{22}) \end{aligned}$$

Then, the $\ell^{th}$ degree coefficient of $\partial$ in $det(\partial I - A)$ is given by

$$k_{12}k_{21}e_{n-2-\ell}(E_{in}) - [det((\partial I - A)_{22})]_{\ell-1} - k_{12}[det((\partial I - A)_{22})]_{\ell}$$

where $E_{in} := \{k_{13}, k_{14}, \dots, k_{1n}\}$ is the set of edges from compartment 1 into any compartment except 2, and $[det((\partial I - A)_{22}]_k$ is the $k^{th}$ degree coefficient of $\partial$ in $det((\partial I - A)_{22})$.

Since $k_{12}$ and $k_{21}$ are globally identifiable and coefficients of $det((\partial I - A)_{22})$ are given in the coefficient map, we conclude that we can find exactly the values of each of the elementary symmetric polynomials on $E_{in}$. So, we can find each inward edge up to $n-2$ symmetric solutions, and determine that there are $(n-2)!$ distinct solutions for the parameter values taken by those edges.

It remains to determine the behavior of the edges in $E_{out} := \{k_{31}, k_{41}, \dots, k_{n1}\}$. To do so, we will determine the coefficients of $\partial$ in $det((\partial I - A)_{22})$.

One can determine that $det((\partial I - A)_{22})$ is given by

$$det((\partial I - A)_{22}) = (\partial + k_{21} + k_{31} + \cdots + k_{n1}) \prod_{i \in \{3,\ldots,n\}} (\partial + k_{1i}) - k_{13}k_{31} \prod_{i \in \{\hat{3},4,\ldots,n\}} (\partial + k_{1i}) - \cdots - k_{1n}k_{n1} \prod_{i \in \{3,\ldots,n-1,\hat{n}\}} (\partial + k_{1i}) \quad (6)$$

Note that the product $\prod_{i \in \{3,\ldots,n\}} (\partial + k_{1i})$ has exactly the elementary symmetric polynomials on $E_{in}$ as its coefficients on $\partial$. Recall that the values of these polynomials are known. So, since $k_{21}$ is also known, we subtract away $(\partial + k_{21}) \prod_{i \in \{3,\ldots,n\}} (\partial + k_{1i})$ from $det((\partial I - A)_{22})$ to find a set of known coefficients in terms of only edges in $E_{in}$ and $E_{out}$.

Here, our coefficients simplify to $k_{31}e_m(E_{in} \setminus \{k_{13}\}) + \cdots + k_{n1}e_m(E_{in} \setminus \{k_{1n}\})$, for each $m \in \{0, \ldots, n-1\}$. Since the coefficients on $k_{31}, k_{41}, \ldots, kn1$ are linearly independent (by Lemma 5.13), we have a system of $n-2$ equations in $n-2$ unknown values, and hence determine the values of edges in $E_{out}$ up to $n-2$ solutions (dependent on the values of edges in $E_{in}$).

With this, the identifiability degree of the model is $(n-2)!$, $k_{12}$ and $k_{21}$ are globally identifiable, and all other edges have identifiability degree $n-2$. □

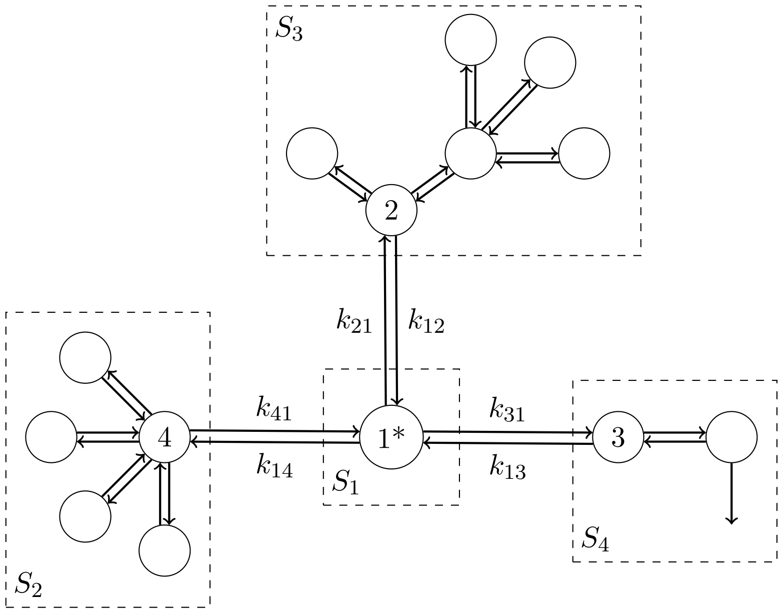
Similar methods to those above are promising in the proofs of identifiability degree of other model classes. While this marks the end of the models considered in this paper, other mammillary, cycle, and other models have promising coefficient maps, waiting to be broken down to determine the identifiability of those models.

# 6 Identifiability Degree of Tree Models

In this section we will look at *tree models* and their identifiability degree, as discussed in [3].

**Definition 6.1.** (Tree Model) A *tree model*, $\mathcal{M}_t$, is a system such that, given any compartment, the rest of the system may be subdivided into subsystems, $S_i$, in which every subsystem is connected to the *experimental compartment* (the compartment containing both the input and output) bidirectionally and there are no connections between compartments of different subsystems.

**Example 6.2.** The following is a tree model with $1^*$ denoting the experimental compartment.




The graph is separated into the subsystems $S_2$, $S_3$, and $S_4$ off of compartment 1.

**Conjecture 6.3.** *For a tree model $\mathcal{M}_t$ the identifiability degree $D$ is given by*

$$D = (n-1)! \prod_{i=2}^{N} \frac{q_i!}{n_i!}$$

*where $n$ is the total number of compartments, $n_i$ is the number of compartments in a subsystem $S_i$, $N$ is the number of subsystems, and $q_i$ is the number of bidirected edges with symmetry (edges that can be switched without changing the system) within a subsystem $S_i$.*

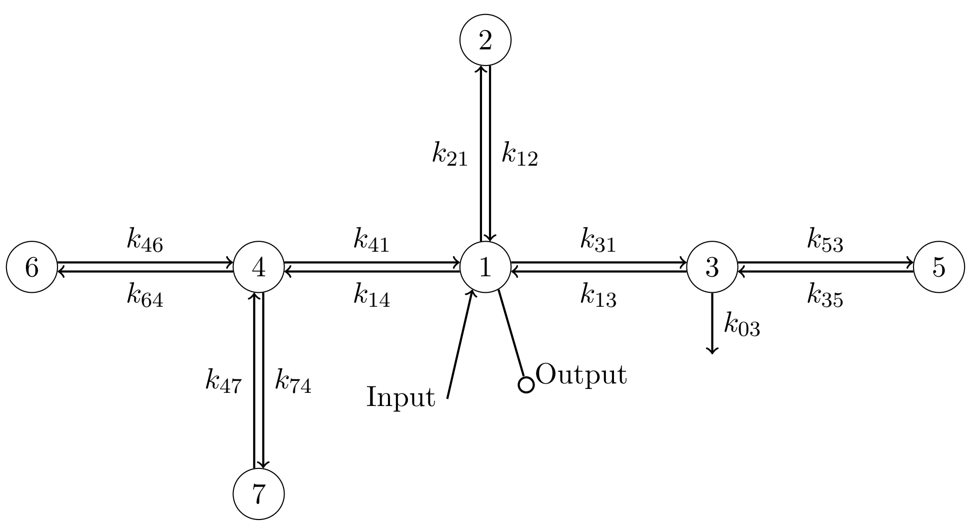



Figure 15: Tree model with $S_2 = \{4,6,7\}$, $S_3 = \{2\}$, $S_4 = \{3,5\}$, and 1 set of symmetric edges, $\{k_{46}, k_{64}, k_{47}, k_{74}\}$.

**Example 6.4.** Using the model in Figure 15, the number of different parametrization vectors (identifiability degree) is predicted in [3] to be

$$\frac{(n-1)!}{n_2!n_3!n_4!}2! = \frac{6!}{1!2!3!}2! = 120$$

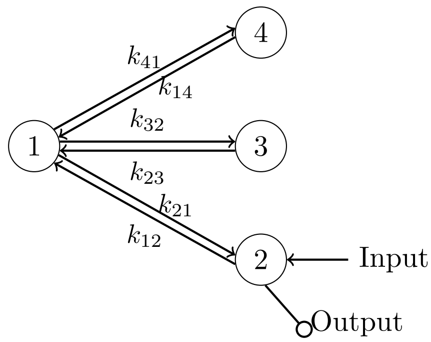



Figure 16: Tree model with experimental compartment 2 and $S_2 = \{1, 3, 4\}$ and the set of symmetric edges $\{k_{23}, k_{32}, k_{14}, k_{41}\}$.

**Example 6.5.** Using the model in Figure 16, the number of different parametrization vectors (identifiability degree) is predicted to be

$$(n-1)!\prod_{i=2}^{2}\frac{q_i!}{n_i!} = 3!\frac{2!}{3!} = 2$$

which we can check is true using the coefficient map.

## Acknowledgements

This work is carried out at the *Texas A&M Mathematics REU 2025* under the thematic program *Algebraic Methods in Computational Biology* funded by the NSF REU site grant DMS-2150094.

A special thanks to Dr. Anne Shiu and Zeytoon Kazemimoghaddam for their mentorship throughout the research and writing of this paper.


## Additional Resources

Code was written to generate and determine properties of linear compartmental models given graph structure and input/output/leak locations. This is available at https://github.com/tylerhuneke/linear-compartmental-models

## References


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

[2] Katherine Clemens, Jonathan Martinez, Anne Shiu, Michaela Thompson, and Benjamin Warren. Parameter identifiability of linear-compartmental mammillary models, 2025.

[3] Claudio Cobelli, Alfredo Lepschy, and Gianni Romanin Jacur. Identifiability results on some constrained compartmental systems. *Mathematical Biosciences*, 47:173–195, 1979.

[4] Alexis Edozie, Odalys Garcia-Lopez, Tyler Huneke, Tegan Keen, and Viridiana Neri. Catenary/cyclic/mammillary model database. https://docs.google.com/spreadsheets/d/19a_QLtWyYvlpDDb2qQTdlFpStz3EadHi4pboNTcGK1E/edit?usp=sharing, 2025. [Online; accessed 17-July-2025].

[5] Elizabeth Gross, Heather Harrington, Nicolette Meshkat, and Anne Shiu. Linear compartmental models: Input-output equations and operations that preserve identifiability. *SIAM Journal on Applied Mathematics*, 79(4):1423–1447, 2019.

[6] Elizabeth Gross, Nicolette Meshkat, and Anne Shiu. Identifiability of linear compartmental models: The singular locus. *Advances in Applied Mathematics*, 2022.