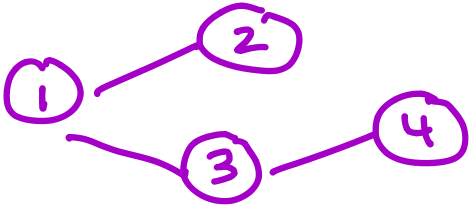
Polynomial systems of graphical models

Elizabeth Gross
University of Hawai'i at Mānoa

ICERM Nonlinear Algebra, Graphical Models Working Group:

Bibhas Adhikari, Alexandros Grosdos, Marc Härkönen,
Cvetelina Hill, Sara Lamboglia, Samantha Sherman,
Elias Tsigaridas, Dane Wilburne

Graphical Models

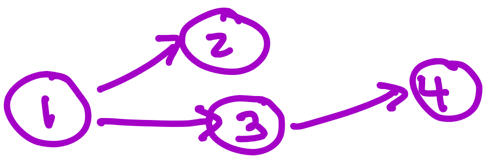


Gaussian graphical models

- Undirected graphs
- Described in detail in "Gaussian Graphical Models: An Algebraic and Geometric Perspective", Uhler (2017)
- Conjecture on ML Degree of cycles in Section 7.4, *Lectures on Algebraic Statistics*, Drton–Sturfmels–Sullivant (2009)

Linear Structural Equation Models

- Directed graphs, mixed graphs
- Overview in "Algebraic Problems in Structural Equation Modeling", Drton (2018)

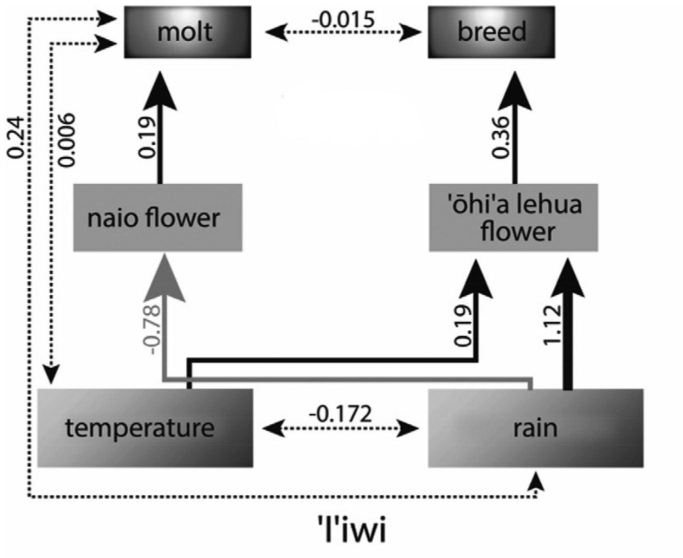


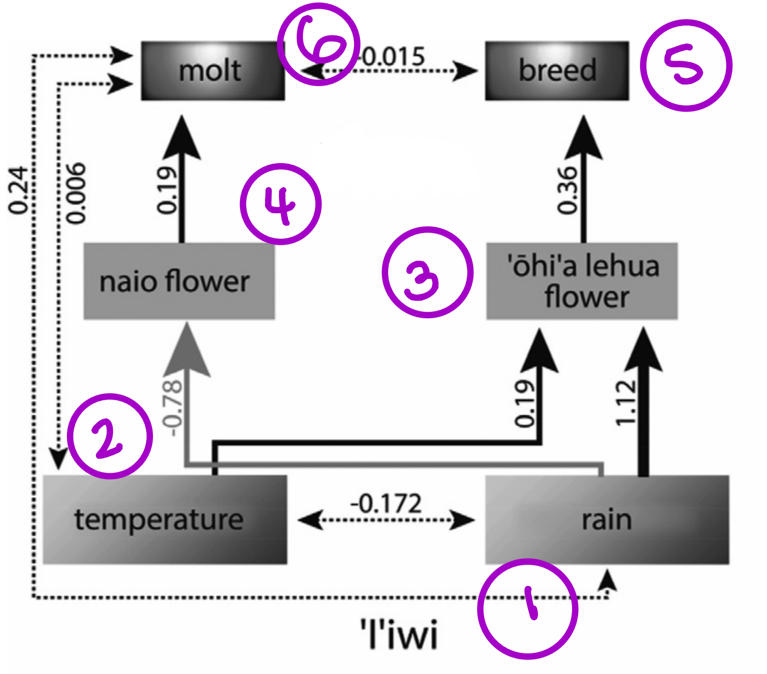
Ecology, 98(11), 2017, pp. 2885–2894

Bottom-up processes influence the demography and life-cycle phenology of Hawaiian bird communities

Jared D. Wolfe,[1,2,3,4] C. John Ralph,[1,3] and Andrew Wiegardt[1,2]

[1]*U.S.D.A Forest Service, Pacific Southwest Research Station, Arcata, California 95521 USA*
[2]*Wildlife Department, Humboldt State University, Arcata, California 95521 USA*
[3]*Klamath Bird Observatory, Ashland, Oregon 97520 USA*





Structural Eqns

$$X_1 = \lambda_{01} + \varepsilon_1$$

$$X_2 = \lambda_{02} + \varepsilon_2$$

$$X_3 = \lambda_{03} + \underline{\lambda_{13}} X_1 + \underline{\lambda_{23}} X_2 + \varepsilon_3$$

$$X_4 = \lambda_{04} + \underline{\lambda_{14}} X_1 + \varepsilon_4$$

$$X_5 = \lambda_{05} + \underline{\lambda_{35}} X_3 + \varepsilon_5$$

$$X_6 = \lambda_{06} + \underline{\lambda_{46}} X_4 + \varepsilon$$

directed edges

$$Var[\epsilon] = \begin{bmatrix} \omega_{11} & \omega_{12} & 0 & 0 & 0 & \omega_{16} \\ \omega_{12} & \omega_{22} & 0 & 0 & 0 & \omega_{26} \\ 0 & 0 & \omega_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \omega_{55} & \omega_{56} \\ \omega_{16} & \omega_{26} & 0 & 0 & \omega_{56} & \omega_{66} \end{bmatrix}$$

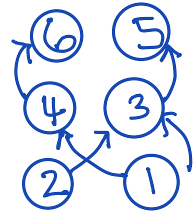
bidirected edges

Structural Equation Models

- $\mathbb{R}^D = \{\Lambda \in \mathbb{R}^{V \times V} : \lambda_{ij} = 0 \text{ if } i \to j \notin D\}$
- $\mathbb{R}^D_{\text{reg}}$ = subset of matrices $\Lambda \in \mathbb{R}^D$ for which $I - \Lambda$ is invertible.
- PD_V = cone of pos def symmetric $V \times V$ matrices.
- $PD(B) = \{\Omega \in PD_V : \omega_{ij} = 0 \text{ if } i \neq j \text{ and } i \leftrightarrow j \notin B\}$.

Example

D:



$$\begin{bmatrix} 0 & 0 & \lambda_{13} & \lambda_{14} & 0 & 0 \\ 0 & 0 & \lambda_{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{35} & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{46} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Definition

The **linear structural equation model** given by a mixed graph $G = (V, D, B)$ on $V = [m]$ is the family of all probability distributions on $\mathbb{R}^m$ with covariance matrix

$$\Sigma = (I - \Lambda)^{-T} \Omega (I - \Lambda)^{-1}$$

for $\Lambda \in \mathbb{R}^D_{\text{reg}}$ and $\Omega \in PD(B)$.

Structural Equation Models

- $\mathbb{R}^D = \{\Lambda \in \mathbb{R}^{V \times V} : \lambda_{ij} = 0 \text{ if } i \to j \notin D\}$
- $\mathbb{R}^D_{\text{reg}}$ = subset of matrices $\Lambda \in \mathbb{R}^D$ for which $I - \Lambda$ is invertible.
- PD_V = cone of pos def symmetric $V \times V$ matrices.
- $PD(B) = \{\Omega \in PD_V : \omega_{ij} = 0 \text{ if } i \neq j \text{ and } i \leftrightarrow j \notin B\}$.

If D is acyclic then $\det(I - \Lambda) = 1$

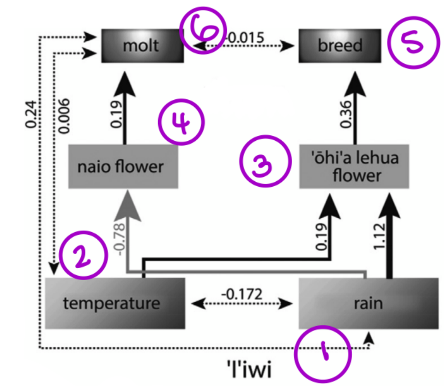
Definition

The **linear structural equation model** given by a mixed graph $G = (V, D, B)$ on $V = [m]$ is the family of all probability distributions on $\mathbb{R}^m$ with covariance matrix

$$\Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$$

for $\Lambda \in \mathbb{R}^D_{\text{reg}}$ and $\Omega \in PD(B)$.

Identifiability



The **covariance parametrization** is

$$\phi_G : \mathbb{R}^D \times PD(B) \to PD_V$$

$$(\Lambda, \Omega) \mapsto (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$$

$$\phi_G(\Lambda, \Omega)_{34} = \omega_{11}\lambda_{13}\lambda_{14} + \omega_{12}\lambda_{23}\lambda_{14}$$

uses trek rule

Definition

The **fiber** of a pair $(\Lambda, \Omega) \in \mathbb{R}^D_{\text{reg}} \times PD(B)$ is

$$\mathcal{F}_G(\Lambda, \Omega) = \{(\Lambda', \Omega') \in \mathbb{R}^D_{\text{reg}} \times PD(B) : \phi_G(\Lambda', \Omega') = \phi_G(\Lambda, \Omega)\}$$

Identifiability

The **covariance parametrization** is

$$\phi_G : \mathbb{R}^D \times PD(B) \to PD_V$$
$$(\Lambda, \Omega) \mapsto (I - \Lambda)^{-T} \Omega (I - \Lambda)^{-1}$$

Definition

The **fiber** of a pair $(\Lambda, \Omega) \in \mathbb{R}^D_{\text{reg}} \times PD(B)$ is

$$\mathcal{F}_G(\Lambda, \Omega) = \{(\Lambda', \Omega') \in \mathbb{R}^D_{\mathrm{reg}} \times PD(B) : \phi_G(\Lambda', \Omega') = \phi_G(\Lambda, \Omega)\}$$

- If the map ϕ_G is injective, then we call the model **global identifiable**.
- If the map ϕ_G is generically injective, then we call the model **generically identifiable**.
- If the map ϕ_G is generically k-to-one, then we call the model **generically locally identifiable**. In this case, we call k the **identifiability degree** of the model.

Identifiability

The **covariance parametrization** is

$$\phi_G : \mathbb{R}^D \times PD(B) \to PD_V$$
$$(\Lambda, \Omega) \mapsto (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$$

Definition

The **fiber** of a pair $(\Lambda, \Omega) \in \mathbb{R}^D_{\text{reg}} \times PD(B)$ is

$$\mathcal{F}_G(\Lambda, \Omega) = \{(\Lambda', \Omega') \in \mathbb{R}^D_{\mathrm{reg}} \times PD(B) : \phi_G(\Lambda', \Omega') = \phi_G(\Lambda, \Omega)\}$$

- If the map ϕ_G is injective, then we call the model **global identifiable**.
- If the map ϕ_G is generically injective, then we call the model **generically identifiable**.
- If the map ϕ_G is generically k-to-one, then we call the model **generically locally identifiable**. In this case, we call k the **identifiability degree** of the model.

care about dim & degree

Another set of equations

Lemma (Foygel–Draisma–Drton (2012), Drton (2018))

Let $G = (V, D, B)$ be a mixed graph, and let $\Sigma = \phi_G(\Lambda_0, \Omega_0)$ for $\Lambda_0 \in \mathbb{R}^D_{reg}$ and $\Omega_0 \in PD(B)$. Then the fiber $\mathcal{F}_G(\Lambda_0, \Omega_0)$ is isomorphic to the set of matrices $\Lambda \in \mathbb{R}^D_{reg}$ that solve the equation system:

$$F_{ij} = [(I - \Lambda)^T \Sigma (I - \Lambda)]_{ij} = 0 \quad i \neq j, i \leftrightarrow j \notin B$$

or more explicitly:

$$F_{ij} = \sigma_{ij} - \sum_{k \to i} \lambda_{ki}\sigma_{kj} - \sum_{l \to j} \lambda_{lj}\sigma_{il} + \sum_{k \to i}\sum_{l \to j} \lambda_{ki}\sigma_{kl}\lambda_{lj} = 0$$

Need to be careful about spurious solutions ($\det(I - \Lambda) = 0$).

Another set of equations

Lemma (Foygel–Draisma–Drton (2012), Drton (2018))

Let $G = (V, D, B)$ be a mixed graph, and let $\Sigma = \phi_G(\Lambda_0, \Omega_0)$ for $\Lambda_0 \in \mathbb{R}^D_{reg}$ and $\Omega_0 \in PD(B)$. Then the fiber $\mathcal{F}_G(\Lambda_0, \Omega_0)$ is isomorphic to the set of matrices $\Lambda \in \mathbb{R}^D_{reg}$ that solve the equation system:

$$F_{ij} = [(I - \Lambda)^T \Sigma (I - \Lambda)]_{ij} = 0 \quad i \neq j, i \leftrightarrow j \notin B$$

or more explicitly:

variables

$$F_{ij} = \sigma_{ij} - \sum_{k \to i} \lambda_{ki}\sigma_{kj} - \sum_{l \to j} \lambda_{lj}\sigma_{il} + \sum_{k \to i}\sum_{l \to j} \lambda_{ki}\sigma_{kl}\lambda_{lj} = 0$$

Need to be careful about spurious solutions ($\det(I - \Lambda) = 0$).

Another set of equations

Lemma (Foygel–Draisma–Drton (2012), Drton (2018))

Let $G = (V, D, B)$ be a mixed graph, and let $\Sigma = \phi_G(\Lambda_0, \Omega_0)$ for $\Lambda_0 \in \mathbb{R}^D_{reg}$ and $\Omega_0 \in PD(B)$. Then the fiber $\mathcal{F}_G(\Lambda_0, \Omega_0)$ is isomorphic to the set of matrices $\Lambda \in \mathbb{R}^D_{reg}$ that solve the equation system:

$$F_{ij} = [(I - \Lambda)^T \Sigma (I - \Lambda)]_{ij} = 0 \quad i \neq j, i \leftrightarrow j \notin B$$

or more explicitly:

parameters

$$F_{ij} = \sigma_{ij} - \sum_{k \to i} \lambda_{ki} \sigma_{kj} - \sum_{l \to j} \lambda_{lj} \sigma_{il} + \sum_{k \to i} \sum_{l \to j} \lambda_{ki} \sigma_{kl} \lambda_{lj} = 0$$

Need to be careful about spurious solutions ($\det(I - \Lambda) = 0$).

Another set of equations

Lemma (Foygel–Draisma–Drton (2012), Drton (2018))

Let $G = (V, D, B)$ be a mixed graph, and let $\Sigma = \phi_G(\Lambda_0, \Omega_0)$ for $\Lambda_0 \in \mathbb{R}^D_{reg}$ and $\Omega_0 \in PD(B)$. Then the fiber $\mathcal{F}_G(\Lambda_0, \Omega_0)$ is isomorphic to the set of matrices $\Lambda \in \mathbb{R}^D_{reg}$ that solve the equation system:

$$F_{ij} = [(I - \Lambda)^T \Sigma (I - \Lambda)]_{ij} = 0 \quad i \neq j, i \leftrightarrow j \notin B$$

or more explicitly:

Equations are <u>bilinear</u>

$$F_{ij} = \sigma_{ij} - \sum_{k \to i} \lambda_{ki}\sigma_{kj} - \sum_{l \to j} \lambda_{lj}\sigma_{il} + \sum_{k \to i}\sum_{l \to j} \lambda_{ki}\sigma_{kl}\lambda_{lj} = 0$$

Need to be careful about spurious solutions ($\det(I - \Lambda) = 0$).

Identifiability Results

Theorem

Acyclic and no bidirected edges $\Rightarrow$ globally identifiable.

Theorem (Drton–Foygel–Sullivant)

G does not contain a subgraph whose B is connected and D has a unique sink $\Rightarrow$ globally identifiable.

Theorem (Brito–Pearl)

Acyclic and simple $\Rightarrow$ generically identifiable

Theorem (Brito–Pearl)

G criterion $\Rightarrow$ generically identifiable.

Theorem (Foygel–Draisma–Drton)

Half-trek criterion $\Rightarrow$ generically identifiable.

Extended by: Chen (2015), Drton–Weihs (2016), and Weihs–Robinson–Dufresne–Kenkel–Kubjas–McGee–Nguyen–Robeva (2018)

Identifiability Degree Results

Theorem (ICERM Group)

Simple $\Rightarrow$ generically locally identifiable (finite-to-one).

Theorem (Drton–Foygel–Sullivant)

The identifiability degree of a cycle with length ≥ 3 is 2.

Theorem (Foygel–Draisma–Drton)

Analysis of the identifiability degree of mixed graphs with up to five nodes. The maximum identifiability degree observed was 10.

Identifiability Degree Results

Theorem (ICERM Group)

Simple ⇒ generically locally identifiable (finite-to-one).

uses combinatorial matrix theory, gives us an expression for $\phi(\Lambda,\Omega)_{ij}$

Theorem (Drton–Foygel–Sullivant)

The identifiability degree of a cycle with length ≥ 3 is 2.

Theorem (Foygel–Draisma–Drton)

Analysis of the identifiability degree of mixed graphs with up to five nodes. The maximum identifiability degree observed was 10.

Identifiability Degree Results

Theorem (ICERM Group)

Simple $\Rightarrow$ generically locally identifiable (finite-to-one).

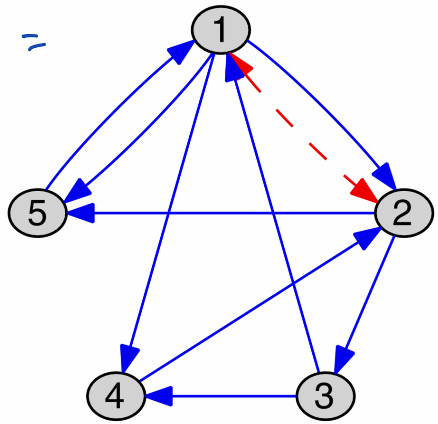
Theorem (Drton–Foygel–Sullivant)

The identifiability degree of a cycle with length ≥ 3 is 2.

Theorem (Foygel–Draisma–Drton)

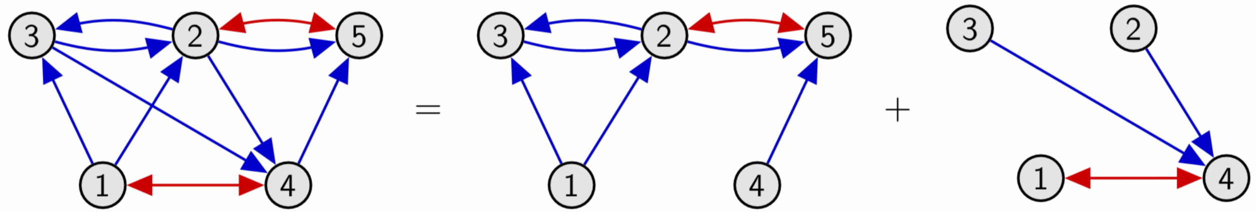
Analysis of the identifiability degree of mixed graphs with up to five nodes. The maximum identifiability degree observed was 10.

$G =$



Identifiability Degree of G = 10

Tian's Decomposition



Theorem (Tian (2005))

The degree of identifiability of a mixed graph G is the product of the degrees of identifiability of its mixed components $G[C]$, $C \in \mathcal{C}(G)$. In particular, ϕ_G is (generically) injective if and only if each $\phi_{G[C]}$ is so, for $C \in \mathcal{C}(G)$.

Identifiability Degree Results

Theorem (ICERM Group)

The identifiability degree of a cycle plus an incoming edge is 1.

Theorem (ICERM Group)

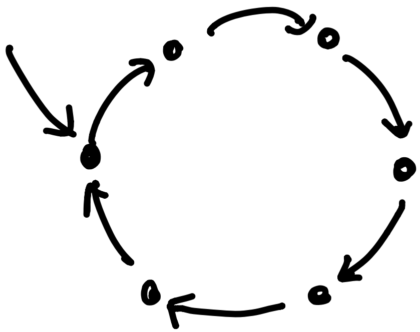
The degree of a cycle plus an outgoing edge is 2.

Theorem (ICERM Group)

Let $C_0, \dots, C_n$ be cycles and let v_{C_i} be a vertex in C_i for every i. Consider the graph G obtained by adding the edge $v_{C_i} \to v_{C_{i+1}}$ for every $i = 0, \dots, n-1$. Then the identifiability degree is 2.

Conjecture (ICERM Group)

Gluing over a vertex or along an edge a chain of cycles has identifiability degree 1.



Identifiability Degree Results

Theorem (ICERM Group)

The identifiability degree of a cycle plus an incoming edge is 1.

Theorem (ICERM Group)

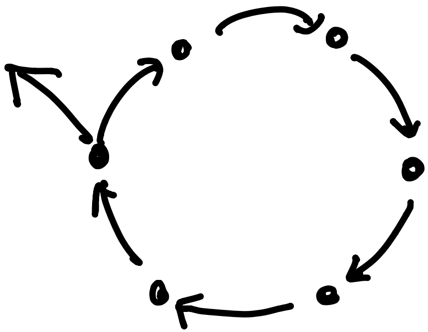
The degree of a cycle plus an outgoing edge is 2.

Theorem (ICERM Group)

Let $C_0, \ldots, C_n$ be cycles and let v_{C_i} be a vertex in C_i for every i. Consider the graph G obtained by adding the edge $v_{C_i} \to v_{C_{i+1}}$ for every $i = 0, \ldots, n-1$. Then the identifiability degree is 2.

Conjecture (ICERM Group)

Gluing over a vertex or along an edge a chain of cycles has identifiability degree 1.



Theorem (ICERM Group)

The identifiability degree of a cycle plus an incoming edge is 1.

Theorem (ICERM Group)

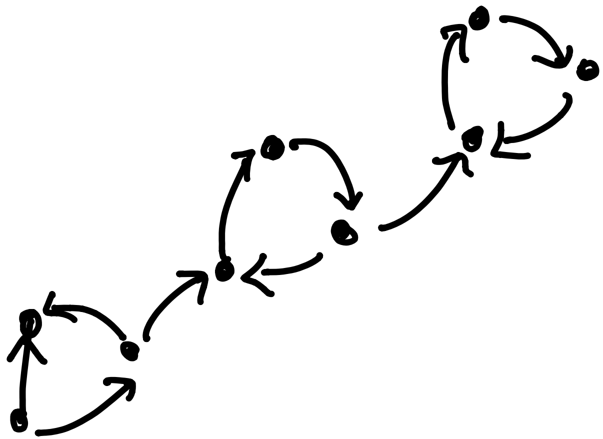
The degree of a cycle plus an outgoing edge is 2.

Theorem (ICERM Group)

Let $C_0, \ldots, C_n$ be cycles and let v_{C_i} be a vertex in C_i for every i. Consider the graph G obtained by adding the edge $v_{C_i} \to v_{C_{i+1}}$ for every $i = 0, \ldots, n-1$. Then the identifiability degree is 2.

Conjecture (ICERM Group)

Gluing over a vertex or along an edge a chain of cycles has identifiability degree 1.



Identifiability Degree Results

Theorem (ICERM Group)

The identifiability degree of a cycle plus an incoming edge is 1.

Theorem (ICERM Group)

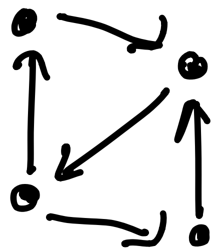
The degree of a cycle plus an outgoing edge is 2.

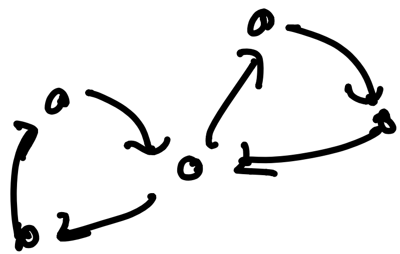
Theorem (ICERM Group)

Let $C_0, \ldots, C_n$ be cycles and let v_{C_i} be a vertex in C_i for every i. Consider the graph G obtained by adding the edge $v_{C_i} \to v_{C_{i+1}}$ for every $i = 0, \ldots, n-1$. Then the identifiability degree is 2.

Conjecture (ICERM Group)

Gluing over a vertex or along an edge a chain of cycles has identifiability degree 1.





Questions

- Can the identifiability degree of a mixed graph be arbitrarily large?
- What is the relationship between the number of vertices n of a mixed graph G and the range of possible identifiability degrees for G?
- For a fixed n is there a way to build a mixed graph G with maximum identifiability degree?

We need your help!

Mahalo! Thank you!