Causality and independence in stationary diffusions

Tobias Boege

Department of Mathematics and Statistics
UiT The Arctic University of Norway

Algebraic methods in dynamics and particle physics
MPI-SWS Saarbrücken, October 16

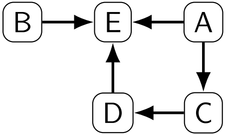
Based on joint works **arXiv:2408.00583** and **arXiv:2510.04985** with Carlos Améndola, Mathias Drton, Benjamin Hollering, Sarah Lumpp, Pratik Misra, and Daniela Schkoda.

Funded by the European Union

Causal modeling with directed graphs

Causal modeling:

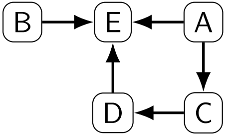
- We want a statistical model which captures the causal structure encoded in a directed graph $\mathcal{G} = (V, E)$.



Causal modeling with directed graphs

Causal modeling:

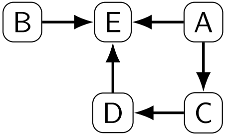
- We want a statistical model which captures the causal structure encoded in a directed graph $\mathcal{G} = (V, E)$.
- Parents of a node are regarded as its direct causes, further-up ancestors are only indirect causes.



Causal modeling with directed graphs

Causal modeling:

- We want a statistical model which captures the causal structure encoded in a directed graph $\mathcal{G} = (V, E)$.
- Parents of a node are regarded as its direct causes, further-up ancestors are only indirect causes.
- The causal diagram may come from expert knowledge or it might be learned from observational data.



Causal modeling in physics

- Hans Reichenbach. *The direction of time*. Edited by Maria Reichenbach. University of California Press, 1971.
- Časlav Brukner. "Quantum causality". In: *Nature Physics* 10.4 (2014). DOI: 10.1038/nphys2930.
- John-Mark A. Allen, Jonathan Barrett, Dominic C. Horsman, Ciarán M. Lee, and Robert W. Spekkens. "Quantum Common Causes and Quantum Causal Models". In: *Phys. Rev. X* 7.3 (2017). DOI: 10.1103/PhysRevX.7.031021.
- Jonathan Barrett, Robin Lorenz, and Ognyan Oreshkov. "Cyclic quantum causal models". In: *Nature Communications* 12.1 (2021). DOI: 10.1038/s41467-020-20456-x.

Linear structural equation models

- Linear structural equation models provide a simple encoding of causal diagrams $\mathcal{G}$. Define a random variable X_j as a linear function of its parents, up to some noise:

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \omega_j).$$

Linear structural equation models

- Linear structural equation models provide a simple encoding of causal diagrams $\mathcal{G}$. Define a random variable X_j as a linear function of its parents, up to some noise:

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \omega_j).$$

- In short: $X = \Lambda X + \varepsilon$ where Λ is $\mathcal{G}$-sparse and $\varepsilon \sim \mathcal{N}(0, \Omega)$ with Ω diagonal.

Linear structural equation models

- Linear structural equation models provide a simple encoding of causal diagrams $\mathcal{G}$. Define a random variable X_j as a linear function of its parents, up to some noise:

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \omega_j).$$

- In short: $X = \Lambda X + \varepsilon$ where Λ is $\mathcal{G}$-sparse and $\varepsilon \sim \mathcal{N}(0, \Omega)$ with Ω diagonal.
- Solutions to this system are multivariate normal distributions with zero mean and covariance matrix Σ satisfying the congruence

$$\boxed{(I - \Lambda)^{\mathsf{T}} \Sigma (I - \Lambda) = \Omega.}$$

LSEMs as statistical models

- The map $\phi(\Lambda, \Omega) := \Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ is rational.

  ☞ The statistical model is an irreducible algebraic subset of PD_V.

LSEMs as statistical models

- The map $\phi(\Lambda, \Omega) := \Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ is rational.

  ☞ The statistical model is an irreducible algebraic subset of PD_V.

- If $\mathcal{G}$ is acyclic, then ϕ has a rational inverse (global identifiability).

LSEMs as statistical models

- The map $\phi(\Lambda, \Omega) := \Sigma = (I - \Lambda)^{-\mathsf{T}}\Omega(I - \Lambda)^{-1}$ is rational.
  ☞ The statistical model is an irreducible algebraic subset of PD_V.
- If $\mathcal{G}$ is acyclic, then ϕ has a rational inverse (global identifiability).
- There is an implicit description of the model via conditional independence:

LSEMs as statistical models

- The map $\phi(\Lambda, \Omega) \coloneqq \Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ is rational.

  ☞ The statistical model is an irreducible algebraic subset of PD_V.

- If $\mathcal{G}$ is acyclic, then ϕ has a rational inverse (global identifiability).

- There is an implicit description of the model via conditional independence:
  Σ belongs to the LSEM of $\mathcal{G}$ (acyclic) if and only if

$$[i \perp\!\!\!\perp j \mid \mathrm{pa}(j)] \text{ holds for each } i \to j \notin E(\mathcal{G}) \text{ with } i <_{\mathcal{G}} j.$$

LSEMs as statistical models

- The map $\phi(\Lambda, \Omega) := \Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ is rational.

  ☞ The statistical model is an irreducible algebraic subset of PD_V.

- If $\mathcal{G}$ is acyclic, then ϕ has a rational inverse (global identifiability).

- There is an implicit description of the model via conditional independence:
  Σ belongs to the LSEM of $\mathcal{G}$ (acyclic) if and only if

$$[i \perp\!\!\!\perp j \mid \mathrm{pa}(j)] \text{ holds for each } i \to j \notin E(\mathcal{G}) \text{ with } i <_{\mathcal{G}} j.$$

- Gaußian version of Bayesian networks.

LSEMs as statistical models

- The map $\phi(\Lambda, \Omega) \coloneqq \Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ is rational.

  ☞ The statistical model is an irreducible algebraic subset of PD_V.

- If $\mathcal{G}$ is acyclic, then ϕ has a rational inverse (global identifiability).

- There is an implicit description of the model via conditional independence: Σ belongs to the LSEM of $\mathcal{G}$ (acyclic) if and only if

$$[i \perp\!\!\!\perp j \mid \mathrm{pa}(j)] \text{ holds for each } i \to j \notin E(\mathcal{G}) \text{ with } i <_{\mathcal{G}} j.$$

- Gaußian version of Bayesian networks.

⋝ Interactions between nodes only through the prescribed causal mechanism! ⋜

Two shortcomings of LSEMs

- Identifiability of parameters (Λ, Ω) is lost when the causal structure contains feedback loops (cycles).

Two shortcomings of LSEMs

- Identifiability of parameters (Λ, Ω) is lost when the causal structure contains feedback loops (cycles).
- Different causal structures may define the same model due to Markov equivalence. This stunts any attempts to learn causal structures from data.

Two shortcomings of LSEMs

- Identifiability of parameters (Λ, Ω) is lost when the causal structure contains feedback loops (cycles).
- Different causal structures may define the same model due to Markov equivalence. This stunts any attempts to learn causal structures from data.

Example: The LSEMs of $A \to B \to C$ and $C \to B \to A$ are both defined by the independence $[A \perp\!\!\!\perp C \mid B]$ which is a polynomial equation in the entries of Σ:

$$\sigma_{B,B}\sigma_{A,C} = \sigma_{A,B}\sigma_{B,C}.$$

Two shortcomings of LSEMs

- Identifiability of parameters (Λ, Ω) is lost when the causal structure contains feedback loops (cycles).
- Different causal structures may define the same model due to Markov equivalence. This stunts any attempts to learn causal structures from data.

Example: The LSEMs of $A \to B \to C$ and $C \to B \to A$ are both defined by the independence $[A \perp\!\!\!\perp C \mid B]$ which is a polynomial equation in the entries of Σ:

$$\sigma_{B,B}\sigma_{A,C} = \sigma_{A,B}\sigma_{B,C}.$$

Same model despite **opposite** claims of causation.

Two shortcomings of LSEMs

- Identifiability of parameters (Λ, Ω) is lost when the causal structure contains feedback loops (cycles).
- Different causal structures may define the same model due to Markov equivalence. This stunts any attempts to learn causal structures from data.

Example: The LSEMs of $A \to B \to C$ and $C \to B \to A$ are both defined by the independence $[A \perp\!\!\!\perp C \mid B]$ which is a polynomial equation in the entries of Σ:

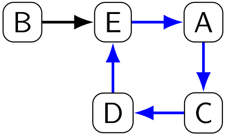
$$\sigma_{B,B}\sigma_{A,C} = \sigma_{A,B}\sigma_{B,C}.$$

Same model despite **opposite** claims of causation.

Some good news: Markov equivalence can be characterized combinatorially!

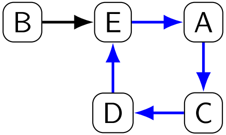
Lyapunov models

- We want a statistical model which captures the causal structure of a directed graph $\mathcal{G} = (V, E)$ with cycles.

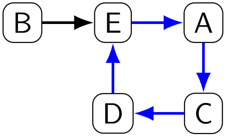


Lyapunov models

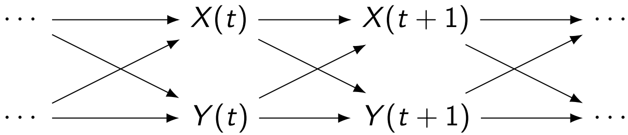
- We want a statistical model which captures the causal structure of a directed graph $\mathcal{G} = (V, E)$ with cycles.
- Idea of [Fit19; VH20]: unroll the feedback loop over time.



Lyapunov models



- We want a statistical model which captures the causal structure of a directed graph $\mathcal{G} = (V, E)$ with cycles.
- Idea of [Fit19; VH20]: unroll the feedback loop over time.



Lyapunov models

- We want a statistical model which captures the causal structure of a directed graph $\mathcal{G} = (V, E)$ with cycles.
- Idea of [Fit19; VH20]: unroll the feedback loop over time.

$$
\begin{tikzcd}
\mathrm{B} \arrow[r] & \mathrm{E} \arrow[r] & \mathrm{A} \arrow[d] \\
 & \mathrm{D} \arrow[u] & \mathrm{C} \arrow[l]
\end{tikzcd}
$$

$$
\begin{tikzcd}
\cdots \arrow[r] \arrow[dr] & X(t) \arrow[r] \arrow[dr] & X(t+1) \arrow[r] \arrow[dr] & \cdots \\
\cdots \arrow[r] \arrow[ur] & Y(t) \arrow[r] \arrow[ur] & Y(t+1) \arrow[r] \arrow[ur] & \cdots
\end{tikzcd}
$$

- Formally, take the stationary distribution of the Ornstein–Uhlenbeck process

$$\mathrm{d}\mathbb{X}(t) = M\mathbb{X}(t)\mathrm{d}t + D\mathrm{d}\mathbb{W}(t),$$

where M is $\mathcal{G}$-sparse, $\mathbb{W}$ a standard Brownian motion, and D diagonal.

Parametrization of Lyapunov models

- The stationary distribution is Gaußian with covariance matrix satisfying the Lyapunov equation $\boxed{M\Sigma + \Sigma M^{\mathsf{T}} + C = 0}$ where $C = DD^{\mathsf{T}}$.

Parametrization of Lyapunov models

- The stationary distribution is Gaußian with covariance matrix satisfying the Lyapunov equation $\boxed{M\Sigma + \Sigma M^{\mathsf{T}} + C = 0}$ where $C = DD^{\mathsf{T}}$.
- If M is stable (all eigenvalues have negative real part) and C is positive definite, then there exists a unique positive definite solution Σ.

Parametrization of Lyapunov models

- The stationary distribution is Gaußian with covariance matrix satisfying the Lyapunov equation $\boxed{M\Sigma + \Sigma M^{\mathsf{T}} + C = 0}$ where $C = DD^{\mathsf{T}}$.
- If M is stable (all eigenvalues have negative real part) and C is positive definite, then there exists a unique positive definite solution Σ.
- The Lyapunov equation is a linear matrix equation in Σ, so it can be rewritten via vectorization and Kronecker products:

$$(I \otimes M + M \otimes I)\,\mathrm{vec}\,\Sigma = -\,\mathrm{vec}\,C.$$

Parametrization of Lyapunov models

- The stationary distribution is Gaußian with covariance matrix satisfying the Lyapunov equation $\boxed{M\Sigma + \Sigma M^{\mathsf{T}} + C = 0}$ where $C = DD^{\mathsf{T}}$.
- If M is stable (all eigenvalues have negative real part) and C is positive definite, then there exists a unique positive definite solution Σ.
- The Lyapunov equation is a linear matrix equation in Σ, so it can be rewritten via vectorization and Kronecker products:

$$(I \otimes M + M \otimes I)\operatorname{vec}\Sigma = -\operatorname{vec} C.$$

- The unique solution is obtained via Cramer's rule.

Parametrization of Lyapunov models

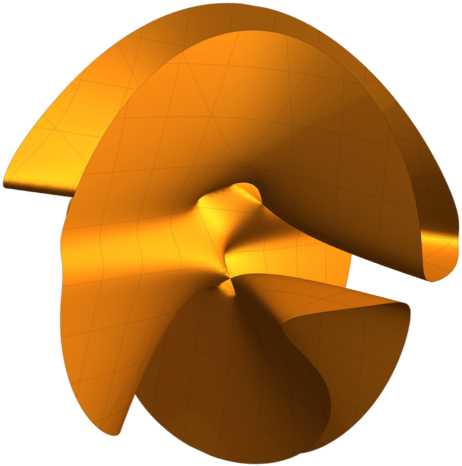
- The stationary distribution is Gaußian with covariance matrix satisfying the Lyapunov equation $\boxed{M\Sigma + \Sigma M^{\mathsf{T}} + C = 0}$ where $C = DD^{\mathsf{T}}$.
- If M is stable (all eigenvalues have negative real part) and C is positive definite, then there exists a unique positive definite solution Σ.
- The Lyapunov equation is a linear matrix equation in Σ, so it can be rewritten via vectorization and Kronecker products:

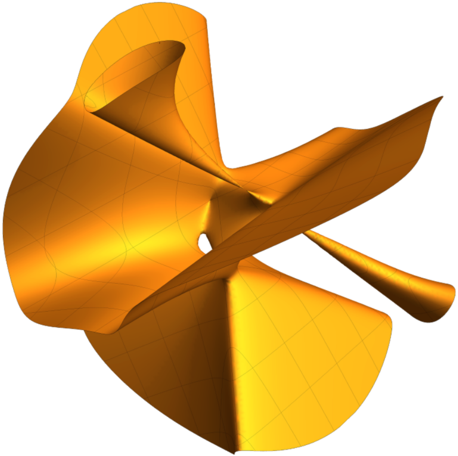
$$(I \otimes M + M \otimes I) \operatorname{vec} \Sigma = -\operatorname{vec} C.$$

- The unique solution is obtained via Cramer's rule.

$\geqslant$ The Lyapunov model is an irreducible algebraic variety! $\leqslant$

Beauty is in the eye of the beholder...





Parameter identifiability

$$M\Sigma + \Sigma M^{\mathsf{T}} + C = 0.$$

- The Lyapunov equation is **also** a linear matrix equation in M, equivalent to:

$$(\Sigma \otimes I + (I \otimes \Sigma)K_n)\operatorname{vec} M = -\operatorname{vec} C,$$

where K_n is the commutation matrix satisfying $K_n \operatorname{vec} M = \operatorname{vec}(M^{\mathsf{T}})$.

Parameter identifiability

$$M\Sigma + \Sigma M^{\mathsf{T}} + C = 0.$$

- The Lyapunov equation is **also** a linear matrix equation in M, equivalent to:

  $$(\Sigma \otimes I + (I \otimes \Sigma)K_n)\operatorname{vec} M = -\operatorname{vec} C,$$

  where K_n is the commutation matrix satisfying $K_n \operatorname{vec} M = \operatorname{vec}(M^{\mathsf{T}})$.
- It has redundant rows since Σ and C are symmetric, but:

Parameter identifiability

$$M\Sigma + \Sigma M^{\mathsf{T}} + C = 0.$$

- The Lyapunov equation is **also** a linear matrix equation in M, equivalent to:

$$(\Sigma \otimes I + (I \otimes \Sigma)K_n)\operatorname{vec} M = -\operatorname{vec} C,$$

where K_n is the commutation matrix satisfying $K_n \operatorname{vec} M = \operatorname{vec}(M^{\mathsf{T}})$.

- It has redundant rows since Σ and C are symmetric, but:

Theorem ([DHADH23])

Let $\mathcal{G}$ be a simple directed graph (i.e., having no directed 2-cycles). Given Σ in the Lyapunov model of $\mathcal{G}$ with fixed diffusion matrix C, the parameter matrix M is uniquely recoverable as a rational function of Σ.

Parameter identifiability

$$M\Sigma + \Sigma M^{\mathsf{T}} + C = 0.$$

- The Lyapunov equation is **also** a linear matrix equation in M, equivalent to:

$$(\Sigma \otimes I + (I \otimes \Sigma)K_n)\operatorname{vec} M = -\operatorname{vec} C,$$

where K_n is the commutation matrix satisfying $K_n \operatorname{vec} M = \operatorname{vec}(M^{\mathsf{T}})$.

- It has redundant rows since Σ and C are symmetric, but:

Theorem ([DHADH23])

Let $\mathcal{G}$ be a simple directed graph (i.e., having no directed 2-cycles). Given Σ in the Lyapunov model of $\mathcal{G}$ with fixed diffusion matrix C, the parameter matrix M is uniquely recoverable as a rational function of Σ.

Lyapunov models are more often globally identifiable than LSEMs!

Independence structure

Fundamental question for a dynamical system / stochastic process:

At the stationary points, which coordinates end up correlated, and how?

Independence structure

Fundamental question for a dynamical system / stochastic process:

At the stationary points, which coordinates end up correlated, and how?

The marginal independence graph $\widehat{\mathcal{G}}$ of $\mathcal{G} = (V, E)$ is the undirected graph on vertices V in which $i - j \in \widehat{E}$ if and only if i and j have a common ancestor in $\mathcal{G}$.

Independence structure

Fundamental question for a dynamical system / stochastic process:

At the stationary points, which coordinates end up correlated, and how?

The marginal independence graph $\widehat{\mathcal{G}}$ of $\mathcal{G} = (V, E)$ is the undirected graph on vertices V in which $i - j \in \widehat{E}$ if and only if i and j have a common ancestor in $\mathcal{G}$.

Theorem ([BDHLMS25])

The Lyapunov model of $\mathcal{G}$ is Markov-perfect to $\widehat{\mathcal{G}}$:

Independence structure

Fundamental question for a dynamical system / stochastic process:

At the stationary points, which coordinates end up correlated, and how?

The marginal independence graph $\widehat{\mathcal{G}}$ of $\mathcal{G} = (V, E)$ is the undirected graph on vertices V in which $i - j \in \widehat{E}$ if and only if i and j have a common ancestor in $\mathcal{G}$.

Theorem ([BDHLMS25])

The Lyapunov model of $\mathcal{G}$ is Markov-perfect to $\widehat{\mathcal{G}}$:

- *Every Σ in the model has $\sigma_{ij} = 0$ (i.e., $[i \perp\!\!\!\perp j]$) if and only if $i - j \notin \widehat{E}$.*

Independence structure

Fundamental question for a dynamical system / stochastic process:

At the stationary points, which coordinates end up correlated, and how?

The marginal independence graph $\widehat{\mathcal{G}}$ of $\mathcal{G} = (V, E)$ is the undirected graph on vertices V in which $i - j \in \widehat{E}$ if and only if i and j have a common ancestor in $\mathcal{G}$.

Theorem ([BDHLMS25])

The Lyapunov model of $\mathcal{G}$ is Markov-perfect to $\widehat{\mathcal{G}}$:

- *Every Σ in the model has $\sigma_{ij} = 0$ (i.e., $[i \perp\!\!\!\perp j]$) if and only if $i - j \notin \widehat{E}$.*
- *All valid conditional independencies on the model are implied by these zeros.*

Independence structure

Fundamental question for a dynamical system / stochastic process:

At the stationary points, which coordinates end up correlated, and how?

The marginal independence graph $\widehat{\mathcal{G}}$ of $\mathcal{G} = (V, E)$ is the undirected graph on vertices V in which $i - j \in \widehat{E}$ if and only if i and j have a common ancestor in $\mathcal{G}$.

Theorem ([BDHLMS25])

The Lyapunov model of $\mathcal{G}$ is Markov-perfect to $\widehat{\mathcal{G}}$:

- *Every Σ in the model has $\sigma_{ij} = 0$ (i.e., $[i \perp\!\!\!\perp j]$) if and only if $i - j \notin \widehat{E}$.*
- *All valid conditional independencies on the model are implied by these zeros.*

Strange: Lyapunov models are **not** defined by conditional independence relations.

Causal structure for acyclic graphs

The big surprise: Lyapunov models encode a lot of causal information.

Causal structure for acyclic graphs

The big surprise: Lyapunov models encode a lot of causal information.

Theorem ([ABHM25])

If $\mathcal{G}$ and $\mathcal{H}$ are acyclic then they define the same Lyapunov model if and only if they have the same (undirected) adjacencies and all induced subgraphs of size four have the same model.

Causal structure for acyclic graphs

The big surprise: Lyapunov models encode a lot of causal information.

Theorem ([ABHM25])

If $\mathcal{G}$ and $\mathcal{H}$ are acyclic then they define the same Lyapunov model if and only if they have the same (undirected) adjacencies and all induced subgraphs of size four have the same model. In this case, $\mathcal{G}$ can be transformed into $\mathcal{H}$ by a series of distinct structured edge reversals without leaving the equivalence class.

Causal structure for acyclic graphs

The big surprise: Lyapunov models encode a lot of causal information.

Theorem ([ABHM25])

If $\mathcal{G}$ and $\mathcal{H}$ are acyclic then they define the same Lyapunov model if and only if they have the same (undirected) adjacencies and all induced subgraphs of size four have the same model. In this case, $\mathcal{G}$ can be transformed into $\mathcal{H}$ by a series of distinct structured edge reversals without leaving the equivalence class.

☞ Bayesian networks require all 3-induced subgraphs to be equivalent.

Causal structure for acyclic graphs

The big surprise: Lyapunov models encode a lot of causal information.

Theorem ([ABHM25])

If $\mathcal{G}$ and $\mathcal{H}$ are acyclic then they define the same Lyapunov model if and only if they have the same (undirected) adjacencies and all induced subgraphs of size four have the same model. In this case, $\mathcal{G}$ can be transformed into $\mathcal{H}$ by a series of distinct structured edge reversals without leaving the equivalence class.

☞ Bayesian networks require all 3-induced subgraphs to be equivalent.

Corollary

Model equivalence can be checked in polynomial time. Model equivalence classes can be enumerated in an output-sensitive manner.

Causal discovery

Theorem ([ABHM25])

*An acyclic causal structure $\mathcal{G}$ is uniquely identifiable from data if and only if it does **not** contain an edge $i \to j$ with the following properties:*

Causal discovery

Theorem ([ABHM25])

An acyclic causal structure $\mathcal{G}$ is uniquely identifiable from data if and only if it does ***not*** *contain an edge $i \to j$ with the following properties:*

- $\mathrm{pa}(i) \cup \{i\} = \mathrm{pa}(j)$ *and* $\mathrm{ch}(i) = \mathrm{ch}(j) \cup \{j\}$,

Causal discovery

Theorem ([ABHM25])

An acyclic causal structure $\mathcal{G}$ is uniquely identifiable from data if and only if it does ***not*** *contain an edge $i \to j$ with the following properties:*

- $\mathrm{pa}(i) \cup \{i\} = \mathrm{pa}(j)$ *and* $\mathrm{ch}(i) = \mathrm{ch}(j) \cup \{j\}$,
- *for all* $k \in \mathrm{pa}(j)$ *and* $l \in \mathrm{ch}(i)$ *the edge* $k \to l$ *exists in* $\mathcal{G}$,

Causal discovery

Theorem ([ABHM25])

An acyclic causal structure $\mathcal{G}$ is uniquely identifiable from data if and only if it does ***not*** *contain an edge $i \to j$ with the following properties:*

- $\mathrm{pa}(i) \cup \{i\} = \mathrm{pa}(j)$ *and* $\mathrm{ch}(i) = \mathrm{ch}(j) \cup \{j\}$,
- *for all* $k \in \mathrm{pa}(j)$ *and* $l \in \mathrm{ch}(i)$ *the edge* $k \to l$ *exists in* $\mathcal{G}$,
- *if* $k - i$ *or* $k - j$ *in* $\widehat{\mathcal{G}}$ *then* $i \to k$ *or* $k \to j$ *in* $\mathcal{G}$.

Causal discovery

Theorem ([ABHM25])

An acyclic causal structure $\mathcal{G}$ is uniquely identifiable from data if and only if it does ***not*** *contain an edge $i \to j$ with the following properties:*

- $\mathrm{pa}(i) \cup \{i\} = \mathrm{pa}(j)$ *and* $\mathrm{ch}(i) = \mathrm{ch}(j) \cup \{j\}$,
- *for all* $k \in \mathrm{pa}(j)$ *and* $l \in \mathrm{ch}(i)$ *the edge* $k \to l$ *exists in* $\mathcal{G}$,
- *if* $k - i$ *or* $k - j$ *in* $\widehat{\mathcal{G}}$ *then* $i \to k$ *or* $k \to j$ *in* $\mathcal{G}$.

☞ $A \to B \to C$ and $C \to B \to A$ define different Lyapunov models!

Causal discovery

Theorem ([ABHM25])

An acyclic causal structure $\mathcal{G}$ is uniquely identifiable from data if and only if it does ***not*** *contain an edge $i \to j$ with the following properties:*

- $\mathrm{pa}(i) \cup \{i\} = \mathrm{pa}(j)$ *and* $\mathrm{ch}(i) = \mathrm{ch}(j) \cup \{j\}$,
- *for all* $k \in \mathrm{pa}(j)$ *and* $l \in \mathrm{ch}(i)$ *the edge* $k \to l$ *exists in* $\mathcal{G}$,
- *if* $k - i$ *or* $k - j$ *in* $\widehat{\mathcal{G}}$ *then* $i \to k$ *or* $k \to j$ *in* $\mathcal{G}$.

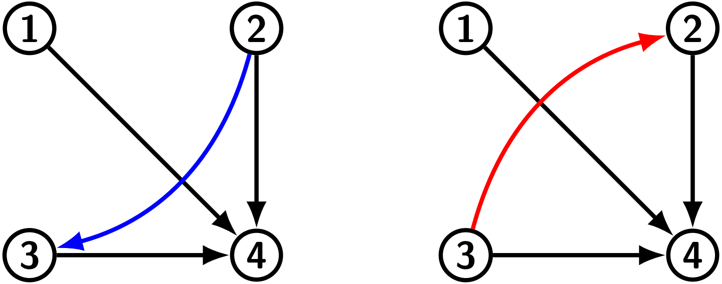
☞ $A \to B \to C$ and $C \to B \to A$ define different Lyapunov models!

Corollary

Lyapunov models refine the Markov equivalence classes of Bayesian networks.

A sparsest (connected) example of non-identifiability

- $\mathrm{pa}(i) \cup \{i\} = \mathrm{pa}(j)$ and $\mathrm{ch}(i) = \mathrm{ch}(j) \cup \{j\}$,
- for all $k \in \mathrm{pa}(j)$ and $l \in \mathrm{ch}(i)$ the edge $k \to l$ exists in $\mathcal{G}$,
- if $k - i$ or $k - j$ in $\widehat{\mathcal{G}}$ then $i \to k$ or $k \to j$ in $\mathcal{G}$.



Summary on Lyapunov models

- Represent a new paradigm in causal modeling via stochastic processes.

Summary on Lyapunov models

- Represent a new paradigm in causal modeling via stochastic processes.
- Temporal perspective accommodates feedback loops.

Summary on Lyapunov models

- Represent a new paradigm in causal modeling via stochastic processes.
- Temporal perspective accommodates feedback loops.
- Parameters are globally identifiable more often than for LSEMs.

Summary on Lyapunov models

- Represent a new paradigm in causal modeling via stochastic processes.
- Temporal perspective accommodates feedback loops.
- Parameters are globally identifiable more often than for LSEMs.
- Conditional independence structure is very coarse.

Summary on Lyapunov models

- Represent a new paradigm in causal modeling via stochastic processes.
- Temporal perspective accommodates feedback loops.
- Parameters are globally identifiable more often than for LSEMs.
- Conditional independence structure is very coarse.
- Despite that, they encode strictly more causal information than Bayesian networks.

Summary on Lyapunov models

- Represent a new paradigm in causal modeling via stochastic processes.
- Temporal perspective accommodates feedback loops.
- Parameters are globally identifiable more often than for LSEMs.
- Conditional independence structure is very coarse.
- Despite that, they encode strictly more causal information than Bayesian networks.

		Lyapunov DAG models		Bayesian network models	
n	DAGs	Distinct models	Identifiable	Distinct models	Identifiable
3	25	17	13	11	4
4	543	461	423	185	59
5	29 281	27 697	26 761	8 782	2 616
6	3 781 503	3 715 745	3 665 673	1 067 825	306 117

References

[ABHM25] Carlos Améndola, Tobias Boege, Benjamin Hollering, and Pratik Misra. *Structural Identifiability of Graphical Continuous Lyapunov Models*. 2025. arXiv: 2510.04985 [math.ST].

[BDHLMS25] Tobias Boege, Mathias Drton, Benjamin Hollering, Sarah Lumpp, Pratik Misra, and Daniela Schkoda. "Conditional independence in stationary diffusions". In: *Stochastic Processes and their Applications* (2025). DOI: 10.1016/j.spa.2025.104604.

[DHADH23] Philipp Dettling, Roser Homs, Carlos Améndola, Mathias Drton, and Niels Richard Hansen. "Identifiability in Continuous Lyapunov Models". In: *SIAM Journal on Matrix Analysis and Applications* 44.4 (2023), pp. 1799–1821.

[Fit19] Katherine E. Fitch. *Learning Directed Graphical Models from Gaussian Data*. Preprint, arXiv:1906.08050 [cs.LG]. 2019. arXiv: 1906.08050. URL: https://arxiv.org/abs/1906.08050.

[VH20] Gherardo Varando and Niels Richard Hansen. "Graphical continuous Lyapunov models". In: *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*. Ed. by Jonas Peters and David Sontag. Vol. 124. Proceedings of Machine Learning Research. PMLR, 2020, pp. 989–998.