Graphical continuous Lyapunov models

Tobias Boege

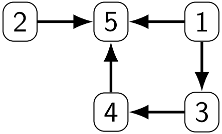
Based on joint works with Carlos Améndola, Mathias Drton,
Benjamin Hollering, Sarah Lumpp, Pratik Misra, Daniela Schkoda and Liam Solus

Department of Mathematics and Statistics
UiT The Arctic University of Norway

Algebra and Discrete Mathematics seminar
Aalto University, 3 December 2024

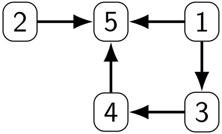
Causal modeling with directed graphs

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



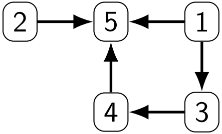
Causal modeling with directed graphs

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



Causal modeling with directed graphs

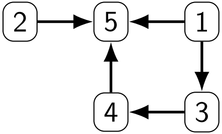
- We want a statistical model which captures the causal structure encoded in a directed graph $G = (V, E)$.
- Parents of node j are regarded as direct causes of j, further-up ancestors are only indirect causes.



- A linear structural equation model defines random variables X recursively via G, parameter matrix Λ and Gaussian noise ε:

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

Causal modeling with directed graphs



- We want a statistical model which captures the causal structure encoded in a directed graph $G = (V, E)$.
- Parents of node j are regarded as direct causes of j, further-up ancestors are only indirect causes.
- A linear structural equation model defines random variables X recursively via G, parameter matrix Λ and Gaussian noise ε:

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

- Solutions to this system are also Gaussian with covariance matrix satisfying the congruence equation $(I - \Lambda)^{\mathsf{T}} \Sigma (I - \Lambda) = \Omega.$

Reasoning with graphical models

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2017WR020412

Key Points:

- We develop a statistical graphical model to characterize the statewide California reservoir system
- We quantify the influence of external physical and economic factors (e.g., statewide PDSI and consumer price index) on the reservoir network
- Further analysis gives a system-wide health diagnosis as a function of PDSI, indicating when heavy management practices may be needed

Supporting Information:

- Supporting Information S1
- Supporting Information S2

A Statistical Graphical Model of the California Reservoir System

A. Taeb[1], J. T. Reager[2], M. Turmon[2], and V. Chandrasekaran[3]

[1]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA, [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [3]Department of Computing and Mathematical Sciences and Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA

Abstract The recent California drought has highlighted the potential vulnerability of the state's water management infrastructure to multiyear dry intervals. Due to the high complexity of the network, dynamic storage changes in California reservoirs on a state-wide scale have previously been difficult to model using either traditional statistical or physical approaches. Indeed, although there is a significant line of research on exploring models for single (or a small number of) reservoirs, these approaches are not amenable to a system-wide modeling of the California reservoir network due to the spatial and hydrological heterogeneities of the system. In this work, we develop a state-wide statistical graphical model to characterize the dependencies among a collection of 55 major California reservoirs across the state; this model is defined with respect to a graph in which the nodes index reservoirs and the edges specify the

Properties of SEMs for *acylic* graphs

If G does **not** contain a directed cycle ("feedback loop"):

- $\mathcal{M}(G)$ is an irreducible algebraic variety and a smooth manifold.

Properties of SEMs for *acylic* graphs

If G does **not** contain a directed cycle ("feedback loop"):

- $\mathcal{M}(G)$ is an irreducible algebraic variety and a smooth manifold.
- The parameters (ω, Λ) are rationally identifiable.

Properties of SEMs for *acylic* graphs

If G does **not** contain a directed cycle ("feedback loop"):

- $\mathcal{M}(G)$ is an irreducible algebraic variety and a smooth manifold.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov property of the DAG, e.g.,

$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : X_i \perp\!\!\!\perp X_j \mid X_{\mathrm{pa}(j)} \text{ whenever } ij \notin E\}.$$

Properties of SEMs for *acylic* graphs

If G does **not** contain a directed cycle ("feedback loop"):

- $\mathcal{M}(G)$ is an irreducible algebraic variety and a smooth manifold.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov property of the DAG, e.g.,

$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : X_i \perp\!\!\!\perp X_j \mid X_{\mathrm{pa}(j)} \text{ whenever } ij \notin E\}.$$

⋛ Interactions between nodes only through the prescribed causal mechanism ⋚

Properties of SEMs for *acylic* graphs

If G does **not** contain a directed cycle ("feedback loop"):

- $\mathcal{M}(G)$ is an irreducible algebraic variety and a smooth manifold.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov property of the DAG, e.g.,

$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : X_i \perp\!\!\!\perp X_j \mid X_{\mathrm{pa}(j)} \text{ whenever } ij \notin E\}.$$

⋟ **Interactions between nodes only through the prescribed causal mechanism** ⋞

- Model equivalence $\mathcal{M}(G) = \mathcal{M}(H)$ is combinatorially characterized: if and only if G and H have the same skeleton and v-structures.

Properties of SEMs for *acylic* graphs

If G does **not** contain a directed cycle ("feedback loop"):

- $\mathcal{M}(G)$ is an irreducible algebraic variety and a smooth manifold.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov property of the DAG, e.g.,

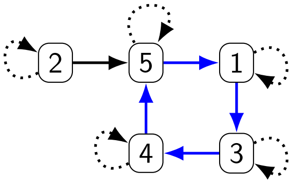
$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : X_i \perp\!\!\!\perp X_j \mid X_{\mathrm{pa}(j)} \text{ whenever } ij \notin E\}.$$

⋝ Interactions between nodes only through the prescribed causal mechanism ⋜

- Model equivalence $\mathcal{M}(G) = \mathcal{M}(H)$ is combinatorially characterized: if and only if G and H have the same skeleton and v-structures.
  - Markov equivalence = ambiguity about the direction of causality.

Lyapunov models

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

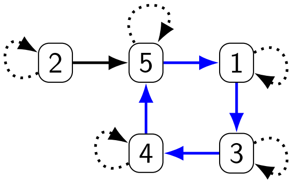


Lyapunov models

- We want a statistical model which captures the causal structure of a directed graph $G = (V, E)$ with cycles.
- Consider stationary distributions of the Ornstein–Uhlenbeck processes satisfying

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

where $m_{ji} = 0$ if $ij \notin E$ and $\mathbb{W}$ is a Brownian motion.

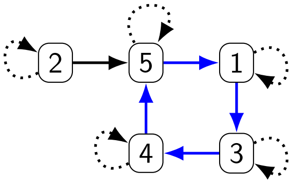


Lyapunov models

- We want a statistical model which captures the causal structure of a directed graph $G = (V, E)$ with cycles.
- Consider stationary distributions of the Ornstein–Uhlenbeck processes satisfying

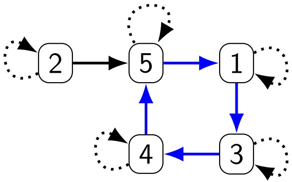
$$\mathrm{d}\mathbb{X}(t) = M(\mathbb{X}(t) - \mu)\mathrm{d}t + D\mathrm{d}\mathbb{W}(t),$$

  where $m_{ji} = 0$ if $ij \notin E$ and $\mathbb{W}$ is a Brownian motion.
- Temporal perspective of stochastic process accommodates feedback loops.



Lyapunov models

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



- Consider stationary distributions of the Ornstein–Uhlenbeck processes satisfying

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

  where $m_{ji} = 0$ if $ij \notin E$ and $\mathbb{W}$ is a Brownian motion.

- Temporal perspective of stochastic process accommodates feedback loops.

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

Parametrization

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

- 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

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

- 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

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

- 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

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

- 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:

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

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

⋛ The Lyapunov model is an irreducible algebraic variety!* ⋚

*Actually a semialgebraic set with irreducible Zariski closure.

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:

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

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

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

Theorem ([Det+23])

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

Vanishing ideal

- Parameter identifiability can be used to implicitize the parametrization map.

Vanishing ideal

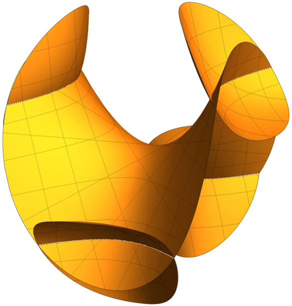
- Parameter identifiability can be used to implicitize the parametrization map.
- [BS24] gives generators of the vanishing ideal up to saturation.

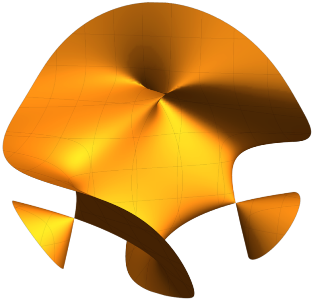
Vanishing ideal

- Parameter identifiability can be used to implicitize the parametrization map.
- [BS24] gives generators of the vanishing ideal up to saturation.
- Lyapunov models generally have a positive-dimensional singular locus but so far no model with singularities in the positive definite cone is known.

Vanishing ideal

- Parameter identifiability can be used to implicitize the parametrization map.
- [BS24] gives generators of the vanishing ideal up to saturation.
- Lyapunov models generally have a positive-dimensional singular locus but so far no model with singularities in the positive definite cone is known.





Restricted trek rule

Lemma ([Boe+24])

Let $G = (V, E)$ be a DAG, fix $C = 2I$ and $m_{ii} = -1/2\zeta$ for all $i \in V$. Then the following trek rule holds:

$$\sigma_{ij} = \sum_{\substack{T:(\ell,r)\text{-trek} \\ \text{from } i \text{ to } j}} 2\zeta^{\ell+r+1} \binom{\ell+r}{\ell} m_T, \tag{1}$$

where the trek monomial m_T associated to a trek T is given by $m_T = \prod_{e \in T} m_e$, i.e., the product over all the edges e in that trek.

Restricted trek rule

Lemma ([Boe+24])

Let $G = (V, E)$ be a DAG, fix $C = 2I$ and $m_{ii} = -1/2\zeta$ for all $i \in V$. Then the following trek rule holds:

$$\sigma_{ij} = \sum_{\substack{T:(\ell,r)\text{-trek} \\ \text{from } i \text{ to } j}} 2\zeta^{\ell+r+1} \binom{\ell+r}{\ell} m_T, \tag{1}$$

where the trek monomial m_T associated to a trek T is given by $m_T = \prod_{e \in T} m_e$, i.e., the product over all the edges e in that trek.

- For acyclic linear SEMs, a similar trek rule holds but without the blue term.

Restricted trek rule

Lemma ([Boe+24])

Let $G = (V, E)$ be a DAG, fix $C = 2I$ and $m_{ii} = -1/2\zeta$ for all $i \in V$. Then the following trek rule holds:

$$\sigma_{ij} = \sum_{\substack{T:(\ell,r)\text{-trek} \\ \text{from } i \text{ to } j}} 2\zeta^{\ell+r+1}\binom{\ell+r}{\ell} m_T, \tag{1}$$

where the trek monomial m_T associated to a trek T is given by $m_T = \prod_{e \in T} m_e$, i.e., the product over all the edges e in that trek.

- For acyclic linear SEMs, a similar trek rule holds but without the blue term.
- This disturbs the conditional independence structure familiar from linear SEMs.

Conditional independence

The marginal independence graph $\widehat{G}$ of $G = (V, E)$ is the undirected graph on vertices V in which $ij \in \widehat{E}$ if and only if there exists a trek between i and j in G.

Conditional independence

The marginal independence graph $\widehat{G}$ of $G = (V, E)$ is the undirected graph on vertices V in which $ij \in \widehat{E}$ if and only if there exists a trek between i and j in G.

Theorem ([Boe+24])

Then the Lyapunov model of G is Markov-perfect to $\widehat{G}$:

$$[i \perp\!\!\!\perp j \mid K] \text{ holds if and only if } V \setminus (\{i, j\} \cup K) \text{ separates } i \text{ and } j \text{ in } \widehat{G}.$$

Conditional independence

The marginal independence graph $\widehat{G}$ of $G = (V, E)$ is the undirected graph on vertices V in which $ij \in \widehat{E}$ if and only if there exists a trek between i and j in G.

Theorem ([Boe+24])

Then the Lyapunov model of G is Markov-perfect to $\widehat{G}$:

$$[i \perp\!\!\!\perp j \mid K] \text{ holds if and only if } V \setminus (\{i, j\} \cup K) \text{ separates } i \text{ and } j \text{ in } \widehat{G}.$$

- All conditional independence statements are implied by absences of treks.
- Lyapunov models are **not** defined by conditional independence relations.

Conditional independence and geometry

- Linear SEMs are defined by conditional independence relations, so their geometric properties are preserved under restriction to correlation matrices.

Conditional independence and geometry

- Linear SEMs are defined by conditional independence relations, so their geometric properties are preserved under restriction to correlation matrices.
- The Lyapunov model on $V = \{1, 2, 3\}$ with missing edge $1 \not\to 3$ is cut out by the following irreducible quintic form:

$$\sigma_{11}\sigma_{12}^2\sigma_{13}\sigma_{22} - \sigma_{11}^2\sigma_{13}\sigma_{22}^2 - \sigma_{11}\sigma_{12}^3\sigma_{23} + \sigma_{11}\sigma_{12}\sigma_{13}^2\sigma_{23} + \sigma_{11}^2\sigma_{12}\sigma_{22}\sigma_{23} + \sigma_{12}\sigma_{13}^2\sigma_{22}\sigma_{23} - \sigma_{11}^2\sigma_{13}\sigma_{23}^2 - 2\sigma_{12}^2\sigma_{13}\sigma_{23}^2 + \sigma_{11}\sigma_{13}\sigma_{22}\sigma_{23}^2 - \sigma_{11}\sigma_{12}^2\sigma_{13}\sigma_{33} - \sigma_{11}\sigma_{13}\sigma_{22}^2\sigma_{33} + \sigma_{11}^2\sigma_{12}\sigma_{23}\sigma_{33} + \sigma_{12}^3\sigma_{23}\sigma_{33} = 0$$

Conditional independence and geometry

- Linear SEMs are defined by conditional independence relations, so their geometric properties are preserved under restriction to correlation matrices.
- The Lyapunov model on $V = \{1, 2, 3\}$ with missing edge $1 \not\to 3$ is cut out by the following irreducible quintic form:

$$\sigma_{11}\sigma_{12}^2\sigma_{13}\sigma_{22} - \sigma_{11}^2\sigma_{13}\sigma_{22}^2 - \sigma_{11}\sigma_{12}^3\sigma_{23} + \sigma_{11}\sigma_{12}\sigma_{13}^2\sigma_{23} + \sigma_{11}^2\sigma_{12}\sigma_{22}\sigma_{23} + \sigma_{12}\sigma_{13}^2\sigma_{22}\sigma_{23} - \sigma_{11}^2\sigma_{13}\sigma_{23}^2 - 2\sigma_{12}^2\sigma_{13}\sigma_{23}^2 + \sigma_{11}\sigma_{13}\sigma_{22}\sigma_{23}^2 - \sigma_{11}\sigma_{12}^2\sigma_{13}\sigma_{33} - \sigma_{11}\sigma_{13}\sigma_{22}^2\sigma_{33} + \sigma_{11}^2\sigma_{12}\sigma_{23}\sigma_{33} + \sigma_{12}^3\sigma_{23}\sigma_{33} = 0$$

- On the space of correlation matrices $\sigma_{11} = \sigma_{22} = \sigma_{33} = 1$, this turns reducible

$$(\sigma_{13} - \sigma_{12}\sigma_{23}) \cdot (1 - \sigma_{12}\sigma_{13}\sigma_{23}) = 0$$

and implies the CI relation $[1 \perp\!\!\!\perp 3 \mid 2]$ which does not hold on the entire model.

What we know so far

		Rat. param.	Rat. ident.	Smooth	CI Markov prop.	Struct. ident.
LSEM	Acyclic	✓	✓	✓	✓	✗
	Simple	✓	✗	?	✓	✗
Lyap.	Acyclic	✓	✓	?	✗	✗[a]
	Simple	✓	✓	?	✗	✗

[a] But appears better than SEMs

What we know so far

		Rat. param.	Rat. ident.	Smooth	CI Markov prop.	Struct. ident.
LSEM	Acyclic	✓	✓	✓	✓	✗
	Simple	✓	✗	?	✓	✗
Lyap.	Acyclic	✓	✓	?	✗	✗[a]
	Simple	✓	✓	?	✗	✗

[a]But appears better than SEMs

Further directions:

What we know so far

		Rat. param.	Rat. ident.	Smooth	CI Markov prop.	Struct. ident.
LSEM	Acyclic	✓	✓	✓	✓	✗
	Simple	✓	✗	?	✓	✗
Lyap.	Acyclic	✓	✓	?	✗	✗[a]
	Simple	✓	✓	?	✗	✗

[a]But appears better than SEMs

Further directions:

- Is there an easier formula for the parametrization than Cramer's rule?

What we know so far

		Rat. param.	Rat. ident.	Smooth	CI Markov prop.	Struct. ident.
LSEM	Acyclic	✓	✓	✓	✓	✗
	Simple	✓	✗	?	✓	✗
Lyap.	Acyclic	✓	✓	?	✗	✗[a]
	Simple	✓	✓	?	✗	✗

[a]But appears better than SEMs

Further directions:

- Is there an easier formula for the parametrization than Cramer's rule?
- Are Lyapunov models smooth?

What we know so far

		Rat. param.	Rat. ident.	Smooth	CI Markov prop.	Struct. ident.
LSEM	Acyclic	✓	✓	✓	✓	✗
	Simple	✓	✗	?	✓	✗
Lyap.	Acyclic	✓	✓	?	✗	✗[a]
	Simple	✓	✓	?	✗	✗

[a] But appears better than SEMs

Further directions:

- Is there an easier formula for the parametrization than Cramer's rule?
- Are Lyapunov models smooth?
- Are the irreducible factors $|A(\Sigma)|$ positive on the set of stable matrices?

What we know so far

		Rat. param.	Rat. ident.	Smooth	CI Markov prop.	Struct. ident.
LSEM	Acyclic	✓	✓	✓	✓	✗
	Simple	✓	✗	?	✓	✗
Lyap.	Acyclic	✓	✓	?	✗	✗[a]
	Simple	✓	✓	?	✗	✗

[a] But appears better than SEMs

Further directions:

- Is there an easier formula for the parametrization than Cramer's rule?
- Are Lyapunov models smooth?
- Are the irreducible factors $|A(\Sigma)|$ positive on the set of stable matrices?
- Relation of a graph's linear SEM and Lyapunov model for correlation matrices?

References

[ABHM25] Carlos Améndola, Tobias Boege, Benjamin Hollering, and Pratik Misra. The algebraic geometry of graphical continuous Lyapunov models. 2025+.

[Boe+24] Tobias Boege, Mathias Drton, Benjamin Hollering, Sarah Lumpp, Pratik Misra, and Daniela Schkoda. Conditional Independence in Stationary Diffusions. 2024. arXiv: 2408.00583 [math.ST].

[BS24] Tobias Boege and Liam Solus. Real birational implicitization for statistical models. 2024. arXiv: 2410.23102 [math.ST].

[Det+23] 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.

[VR20] 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.