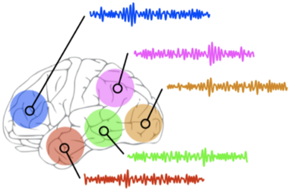
# Graph Topology Identification Based on Covariance Matching

Yongsheng Han, Alberto Natali, Geert Leus

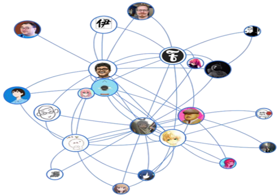
# Graph Topology Identification

## Understanding Hidden Relationships

In many fields, relationships among entities are not directly observable. Graph topology identification helps to infer these hidden structures from nodal data.



Brain Connectivity Network



Social Interaction Network

# Graph Topology Identification (GTI)

## General Process

- GTI relies on the fact that the nodal data is related to the graph.
- Specifically, it follows a distribution determined by the graph

$$\mathbf{x} \sim \mathscr{F}(\mathbf{S}).$$

- This could be the Gaussian distribution:

$$\mathscr{F}(\mathbf{S}) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{S}), \boldsymbol{\Sigma}(\mathbf{S})).$$

- Here, both $\boldsymbol{\mu}(\mathbf{S})$ and $\boldsymbol{\Sigma}(\mathbf{S})$ are functions of the graph structure $\mathbf{S}$.

## Approaches

- Smoothness of data over graph
- Graphical Lasso
- Spectral template
- **Structural equation model (SEM)**

# Structural Equation Model (SEM)

## Setup

- We consider graph signals $\mathbf{x}$ following the model:

$$\mathbf{x} = \mathbf{S}\mathbf{x} + \mathbf{e}, \quad \mathbf{x} = (\mathbf{I} - \mathbf{S})^{-1}\mathbf{e}.$$

- We assume that $\mathbf{S}$ represents the *adjacency matrix* of an *undirected* graph with $N$ nodes, implying

$$\mathbf{S} = \mathbf{S}^\top, \quad \operatorname{diag}(\mathbf{S}) = \mathbf{0}.$$

- We further assume the exogenous variables satisfy

$$\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}).$$

- This overall leads to the following distribution for $\mathbf{x}$:

$$\mathbf{x} \sim \mathcal{N}\Big(\mathbf{0}, \ (\mathbf{I} - \mathbf{S})^{-2}\Big).$$

- Observing $T$ realizations, we collect them into the data matrix

$$\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_T].$$

# Linear Regression

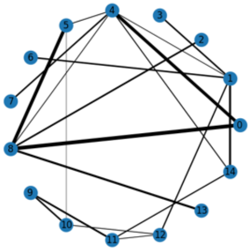
## Optimization Problem

- **SigMatch** seeks $\hat{\mathbf{S}}$ by minimizing:

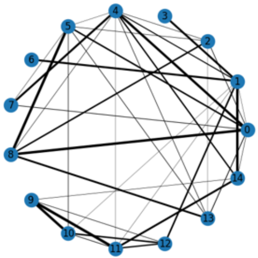
$$\min_{\hat{\mathbf{S}}} \|\mathbf{X} - \hat{\mathbf{S}}\,\mathbf{X}\|_F^2 \quad \text{subject to} \quad \hat{\mathbf{S}} = \hat{\mathbf{S}}^\top,\ \operatorname{diag}(\hat{\mathbf{S}}) = \mathbf{0}.$$

- **Not** the maximum likelihood estimator!
- Only good approach for:
  - A directed acyclic graph (DAG)
  - Deterministic known exogenous variables

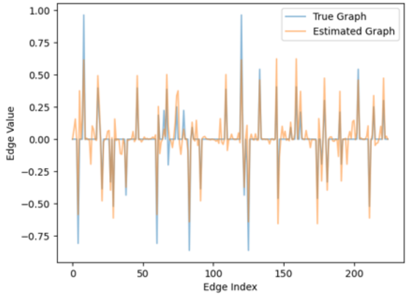
# Analysis of SigMatch




True graph




Estimation graph




## Convergence Issue

We can prove that, for *any* underlying graph shift matrix $\mathbf{S}$, the linear regression method **fails** to converge to the correct graph even for $T \to \infty$.

# Covariance Matching

## Covariance Matching Framework

- **Goal**: Estimate $\mathbf{S}$ so that the theoretical covariance

$$\boldsymbol{\Sigma}_x = (\mathbf{I} - \mathbf{S})^{-2}$$

  is close to the *sample* covariance

$$\mathbf{C}_\mathbf{x} = \frac{1}{T}\mathbf{X}\mathbf{X}^\top.$$

- **Why?** Close to the true maximum likelihood.
- **Optimization problem**:

$$\mathbf{S}^* = \arg\min_{\hat{\mathbf{S}}} \left\|\mathbf{C}_\mathbf{x} - (\mathbf{I} - \hat{\mathbf{S}})^{-2}\right\|_F \quad \text{subject to} \quad \hat{\mathbf{S}} = \hat{\mathbf{S}}^\top, \ \mathrm{diag}(\hat{\mathbf{S}}) = \mathbf{0}.$$

# Solution Approach

## Spectral Template Idea

- Take the **EVD**

$$\mathbf{C}_{\mathbf{x}} = \mathbf{U}_{\mathbf{x}}\ \mathrm{diag}(\boldsymbol{\lambda}_{\mathbf{x}})\ \mathbf{U}_{\mathbf{x}}^{\top},$$

  and model

$$(\mathbf{I} - \hat{\mathbf{S}})^{-1} = \hat{\mathbf{U}}\ \mathrm{diag}(\hat{\boldsymbol{\lambda}})\ \hat{\mathbf{U}}^{\top}$$

  which means

$$(\mathbf{I} - \hat{\mathbf{S}})^{-2} = \hat{\mathbf{U}}\ \mathrm{diag}(\hat{\boldsymbol{\lambda}}^2)\ \hat{\mathbf{U}}^{\top}$$

- **Key Idea**:
  - Rather than solving directly for $\hat{\mathbf{S}}$, we *estimate* $\hat{\mathbf{U}}$ and $\hat{\boldsymbol{\lambda}}$.
  - By choosing $\hat{\mathbf{U}} = \mathbf{U}_{\mathbf{x}}$, this task reduces to fitting $\hat{\boldsymbol{\lambda}}^2$ to $\boldsymbol{\lambda}_{\mathbf{x}}$.
  - The constraint to guarantee is $\mathrm{diag}((\mathbf{I} - \hat{\mathbf{S}})) = \mathrm{diag}\big(\hat{\mathbf{U}}\ \mathrm{diag}(\hat{\boldsymbol{\lambda}}^{-1})\ \hat{\mathbf{U}}^{\top}\big) = \mathbf{1}$.

# Optimization Formulation

## Swapping the objective and the constraint

- The resulting optimization problem becomes:

$$\boldsymbol{\lambda}^* = \arg\min_{\hat{\boldsymbol{\lambda}}} \|\hat{\boldsymbol{\lambda}}^2 - \boldsymbol{\lambda}_{\mathbf{x}}\|_2^2 \quad \text{subject to} \quad \mathrm{diag}(\mathbf{U}_{\mathbf{x}}\, \mathrm{diag}(\hat{\boldsymbol{\lambda}}^{-1})\, \mathbf{U}_{\mathbf{x}}^\top) = \mathbf{1}.$$

- Alternatively, we can swap the objective and constraint
  - We enforce $\hat{\boldsymbol{\lambda}}^2 = \boldsymbol{\lambda}_{\mathbf{x}}$ as a constraint.
  - We fit $\mathrm{diag}(\mathbf{U}_{\mathbf{x}}\, \mathrm{diag}(\hat{\boldsymbol{\lambda}}^{-1})\, \mathbf{U}_{\mathbf{x}}^\top)$ as close as possible to $\mathbf{1}$.
- This results in the following problem

$$\boldsymbol{\lambda}^* = \arg\min_{\hat{\boldsymbol{\lambda}}} \|\mathrm{diag}(\mathbf{U}_{\mathbf{x}}\mathrm{diag}(\hat{\boldsymbol{\lambda}}^{-1})\mathbf{U}_{\mathbf{x}}^\top) - \mathbf{1}\|_2^2 \quad \text{subject to} \quad \hat{\boldsymbol{\lambda}}^2 = \boldsymbol{\lambda}_{\mathbf{x}}.$$

# Addressing Sign Ambiguity and Final Optimization

## Sign Ambiguity

- Note that $\hat{\boldsymbol{\lambda}}^2 = \boldsymbol{\lambda}_\mathbf{x}$, introduces a sign ambiguity for $\hat{\boldsymbol{\lambda}}$.
- Denoting the sign vector as $\hat{\mathbf{q}} \in \{-1, 1\}^N$, the constraint $\hat{\boldsymbol{\lambda}}^2 = \boldsymbol{\lambda}_\mathbf{x}$ can be rewritten as

$$\hat{\boldsymbol{\lambda}} = \operatorname{diag}(\hat{\mathbf{q}})\, \boldsymbol{\lambda}_\mathbf{x}^{1/2}.$$

- The objective function can then be written as

$$\| \operatorname{diag}(\mathbf{U}_\mathbf{x} \operatorname{diag}(\hat{\mathbf{q}})\operatorname{diag}(\boldsymbol{\lambda}_\mathbf{x}^{-1/2})\mathbf{U}_\mathbf{x}^\top) - \mathbf{1}\|_2^2 = \|\left(\mathbf{U}_\mathbf{x} \odot \mathbf{U}_\mathbf{x}\right) \operatorname{diag}\!\left(\boldsymbol{\lambda}_\mathbf{x}^{-1/2}\right)\hat{\mathbf{q}} - \mathbf{1}\|_2^2$$

## Final Problem

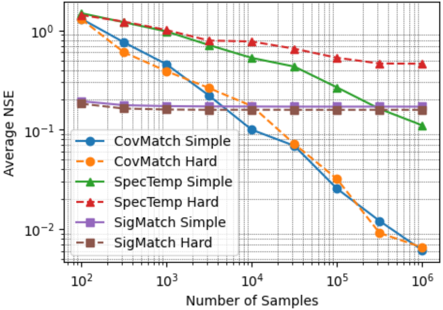
- Putting this all together and defining $\mathbf{W} = \left(\mathbf{U}_\mathbf{x} \odot \mathbf{U}_\mathbf{x}\right) \operatorname{diag}\!\left(\boldsymbol{\lambda}_\mathbf{x}^{-1/2}\right)$, the final problem becomes a binary quadratic programming problem

$$\mathbf{q}^* = \arg \min_{\hat{\mathbf{q}} \in \{-1,1\}^N} \left\|\mathbf{W}\, \hat{\mathbf{q}} - \mathbf{1}\right\|_2^2.$$

# Results and Observations

## Settings

- Graph with 20 nodes and 20 edges
- 100 graph realizations
- Simple: $\mathbf{U} \odot \mathbf{U}$ high rank
- Hard: $\mathbf{U} \odot \mathbf{U}$ low rank




## Key Observations

- **SpecTemp** does not exploit SEM structure and has problems with rank loss
- **SigMatch** never converges, but shows better performance with fewer samples.
- **CovMatch** maintains robust performance across both simple and hard scenarios.

# Polynomial Model

## Motivation and Model

- The SEM model can be seen as a special case of a more general *polynomial model* with

$$h(\mathbf{S}) = (\mathbf{I} - \mathbf{S})^{-1}.$$

- We consider graph signals:

$$\mathbf{x} \;=\; h(\mathbf{S})\,\mathbf{e}, \qquad h(\mathbf{S}) \;=\; \sum_{\ell=0}^{L-1} h_\ell\,\mathbf{S}^\ell.$$

- We again assume $\mathbf{S} = \mathbf{S}^\top$, $\mathrm{diag}(\mathbf{S}) = \mathbf{0}$, and $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
- Then the covariance of $\mathbf{x}$ is

$$\mathbf{x} \sim \mathcal{N}\big(\mathbf{0},\; h^2(\mathbf{S})\big).$$

# Covariance Matching for Polynomial Models

## Key Steps

- We wish to match
$$\mathbf{C}_{\mathbf{x}} \approx h^2(\hat{\mathbf{S}}).$$
- Assuming now that $\hat{\mathbf{S}} = \hat{\mathbf{U}}\ \mathrm{diag}(\hat{\boldsymbol{\lambda}})\ \hat{\mathbf{U}}^\top$, we have
$$h(\hat{\mathbf{S}}) = \hat{\mathbf{U}}\ \mathrm{diag}(h(\hat{\boldsymbol{\lambda}}))\ \hat{\mathbf{U}}^\top.$$
- Setting $\hat{\mathbf{U}} = \mathbf{U}_{\mathbf{x}}$ (from the EVD of $\mathbf{C}_{\mathbf{x}}$) as before, we need to match $h^2(\hat{\boldsymbol{\lambda}})$ to $\boldsymbol{\lambda}_{\mathbf{x}}$ (from the EVD of $\mathbf{C}_{\mathbf{x}}$).
- This matching is turned into a constraint, leading to a set of polynomial equations
$$h^2(\hat{\lambda}_i) = \lambda_{i,\mathbf{x}}, \quad i = 1, \ldots, N.$$
- The constraint $\mathrm{diag}(\hat{\mathbf{S}}) = \mathbf{0}$ is on the other hand turned into the objective
$$\| \mathrm{diag}(\mathbf{U}_{\mathbf{x}}\ \mathrm{diag}(\hat{\boldsymbol{\lambda}})\ \mathbf{U}_{\mathbf{x}}^\top)\| = \left\|(\mathbf{U}_{\mathbf{x}} \odot \mathbf{U}_{\mathbf{x}})\ \hat{\boldsymbol{\lambda}}\right\|_2^2.$$

# Optimization Formulation

## Final Form

- Let $\lambda_{i,\mathbf{x}}$ be the $i$-th eigenvalue of $\mathbf{C}_{\mathbf{x}}$.
- Then each scalar constraint $h^2(\hat{\lambda}_i) = \lambda_{i,\mathbf{x}}$ has a (finite) set of roots $\{\tilde{c}_i^1, \ldots, \tilde{c}_i^{p_i}\}$.
- All this put together leads to the following problem

$$\min_{\hat{\lambda}} \left\| (\mathbf{U}_{\mathbf{x}} \odot \mathbf{U}_{\mathbf{x}}) \, \hat{\lambda} \right\|_2^2 \quad \text{subject to} \quad \hat{\lambda}_i \in \{\tilde{c}_i^1, \ldots, \tilde{c}_i^{p_i}\}.$$

- This is a discrete optimization problem.

# SEM with Colored Exogenous Variables

## Setup

- So far, we assumed $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
- In many settings, $\mathbf{e}$ can be Gaussian with a *known*, *non*-identity covariance:

$$\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{e}}), \quad \boldsymbol{\Sigma}_{\mathbf{e}} \neq \mathbf{I}.$$

- Since the SEM is given by $\mathbf{x} = (\mathbf{I} - \mathbf{S})^{-1}\,\mathbf{e}$, this leads to

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\mathbf{x}^\top\} = (\mathbf{I} - \mathbf{S})^{-1}\,\boldsymbol{\Sigma}_{\mathbf{e}}\,(\mathbf{I} - \mathbf{S})^{-\top}.$$

- Defining $\mathbf{H} = (\mathbf{I} - \mathbf{S})^{-1}$, then results in the following matching problem

$$\hat{\mathbf{H}}\,\boldsymbol{\Sigma}_{\mathbf{e}}\,\hat{\mathbf{H}}^\top \approx \mathbf{C}_{\mathbf{x}}.$$

- **Key Idea**: Instead of matching $\hat{\mathbf{H}}\,\boldsymbol{\Sigma}_{\mathbf{e}}\,\hat{\mathbf{H}}^\top \approx \mathbf{C}_{\mathbf{x}}$, we match

$$\left(\hat{\mathbf{H}}\,\boldsymbol{\Sigma}_{\mathbf{e}}\right)^2 \approx \mathbf{C}_{\mathbf{x}}\,\boldsymbol{\Sigma}_{\mathbf{e}}.$$

- Results in a structure where an unknown square has to be matched to something we know
- This now involves the EVD of $\mathbf{C}_{\mathbf{x}}\,\boldsymbol{\Sigma}_{\mathbf{e}}$.

# Covariance Matching for Colored Exogenous Variables

## Final Optimization

- As discussed we need to match $(\hat{\mathbf{H}}\,\boldsymbol{\Sigma}_{\mathbf{e}})^2 \approx \mathbf{C}_{\mathbf{x}}\,\boldsymbol{\Sigma}_{\mathbf{e}}$.
- For the right hand side we obtain the EVD

$$\mathbf{C}_{\mathbf{x}}\,\boldsymbol{\Sigma}_{\mathbf{e}} = \mathbf{U}_{\mathbf{xe}}\ \mathrm{diag}\big(\boldsymbol{\lambda}_{\mathbf{xe}}\big)\,\mathbf{U}_{\mathbf{xe}}^{-1}.$$

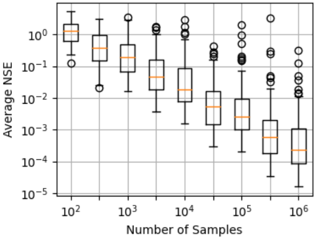
- For the left hand side, we use the model

$$\hat{\mathbf{H}}\,\boldsymbol{\Sigma}_{\mathbf{e}} = \hat{\mathbf{U}}\ \mathrm{diag}\big(\hat{\boldsymbol{\lambda}}\big)\,\hat{\mathbf{U}}^{-1}, \quad \big(\hat{\mathbf{H}}\,\boldsymbol{\Sigma}_{\mathbf{e}}\big)^2 = \hat{\mathbf{U}}\ \mathrm{diag}\big(\hat{\boldsymbol{\lambda}}^2\big)\,\hat{\mathbf{U}}^{-1}.$$

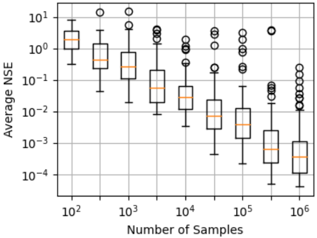
- We set $\hat{\mathbf{U}} = \mathbf{U}_{\mathbf{xe}}$ and define a sign vector $\hat{\mathbf{q}} \in \{-1, 1\}^N$ satisfying $\hat{\boldsymbol{\lambda}} = \mathrm{diag}(\hat{\mathbf{q}})\boldsymbol{\lambda}_{\mathbf{xe}}^{1/2}$.
- The hollow constraint on $\hat{\mathbf{S}}$ again leads to a binary quadratic programming problem:

$$\min_{\hat{\mathbf{q}}}\Big\|\big((\boldsymbol{\Sigma}_{\mathbf{e}}\,\mathbf{U}_{\mathbf{xe}}) \odot \mathbf{U}_{\mathbf{xe}}^{-\top}\big)\,\mathrm{diag}\big(\boldsymbol{\lambda}_{\mathbf{xe}}^{-1/2}\big)\,\hat{\mathbf{q}} - \mathbf{1}\Big\|_2^2.$$

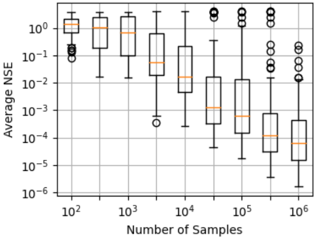
# Experiments




SEM sparse graph




SEM fully connected graph




Polynomial model sparse graph

## Settings

- Positive and negative edge weights
- For SEM, colored exogenous variables
  - Sparse: 20 nodes and 20 edges with rank loss
  - Fully connected: 20 nodes with all edges; no particular attention to rank
- For polynomial model, white exogenous variables and filter order 3

# Conclusions and Future Work

## Conclusions

- We introduced a **covariance matching** approach for a SEM on undirected graphs in white noise.
- **Key advantages:**
  - Close to maximum likelihood estimation.
  - No scale ambiguities and convergence issues.
- We extended the approach to **polynomial models** and a **colored** exogenous variables.

## Future Work

- Incorporating **sparsity priors** or advanced regularizers for large-scale graphs.
- Exploring this method for **directed acyclic graphs**
- Exploring **dynamic** or **time-varying** topologies with evolving network structures.

# Thank you for your attention!