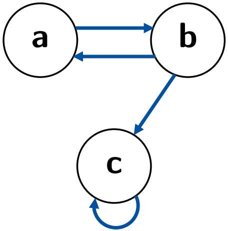
Bayesian estimation and testing for idiographic networks

Björn Siepe & Daniel Heck Philipps-Universität Marburg 12.09.2023

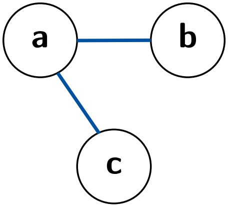
Dynamic Networks

Dynamic Networks

Temporal

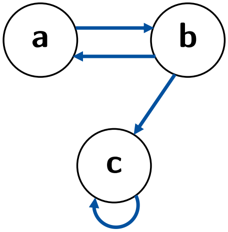


Contemporaneous

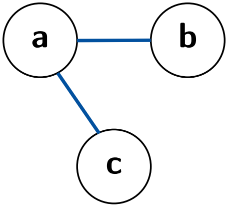


Dynamic Networks

Temporal



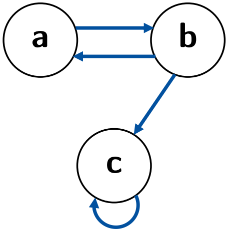
Contemporaneous



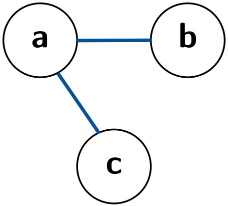
- Issues: Questionable performance in typical psychological data (Hoekstra et al., 2022; Mansueto et al., 2020)

Dynamic Networks

Temporal

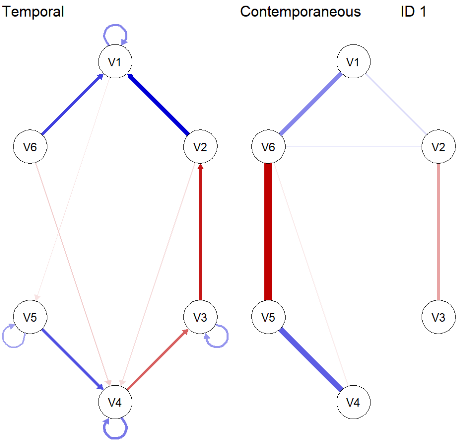


Contemporaneous



- Issues: Questionable performance in typical psychological data (Hoekstra et al., 2022; Mansueto et al., 2020)
- This presentation: New ways to estimate and compare idiographic networks in a Bayesian framework (Siepe and Heck, 2023)

Heterogeneity?



Bayesian gVAR Estimation

Gibbs sampler in R package BGGM (Williams and Mulder, 2021)

Temporal Network

$$\boldsymbol{y}_t = \boldsymbol{B}\boldsymbol{y}_{t-1} + \boldsymbol{\epsilon}_t$$

$$\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Theta}^{-1})$$

Prior:

$$\beta_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, s_\beta)$$

Contemporaneous Network

$$\rho_{ij} = \frac{-\boldsymbol{\Theta}_{ij}}{\sqrt{\boldsymbol{\Theta}_{ii}\boldsymbol{\Theta}_{jj}}}.$$

Prior:

$$\rho \sim \text{Beta}\left(\frac{\delta}{2}, \frac{\delta}{2}\right)$$

Simulation 1: Estimation Performance

- DGP: Real-world examples with six and eight nodes & dense network

Simulation 1: Estimation Performance

- DGP: Real-world examples with six and eight nodes & dense network
- Time series length: $t \in \{50, 100, 200, 400, 1000\}$

Simulation 1: Estimation Performance

- DGP: Real-world examples with six and eight nodes & dense network
- Time series length: $t \in \{50, 100, 200, 400, 1000\}$
- Methods:

Simulation 1: Estimation Performance

- DGP: Real-world examples with six and eight nodes & dense network
- Time series length: $t \in \{50, 100, 200, 400, 1000\}$
- Methods:
  - Bayesian gVAR without thresholding

Simulation 1: Estimation Performance

- DGP: Real-world examples with six and eight nodes & dense network
- Time series length: $t \in \{50, 100, 200, 400, 1000\}$
- Methods:
  - Bayesian gVAR without thresholding
  - Bayesian gVAR with CI-based thresholding

  Wide & narrow priors

Philipps Universität Marburg

Simulation 1: Estimation Performance

- DGP: Real-world examples with six and eight nodes & dense network
- Time series length: $t \in \{50, 100, 200, 400, 1000\}$
- Methods:
  - Bayesian gVAR without thresholding
  - Bayesian gVAR with CI-based thresholding
  - LASSO gVAR in default setting

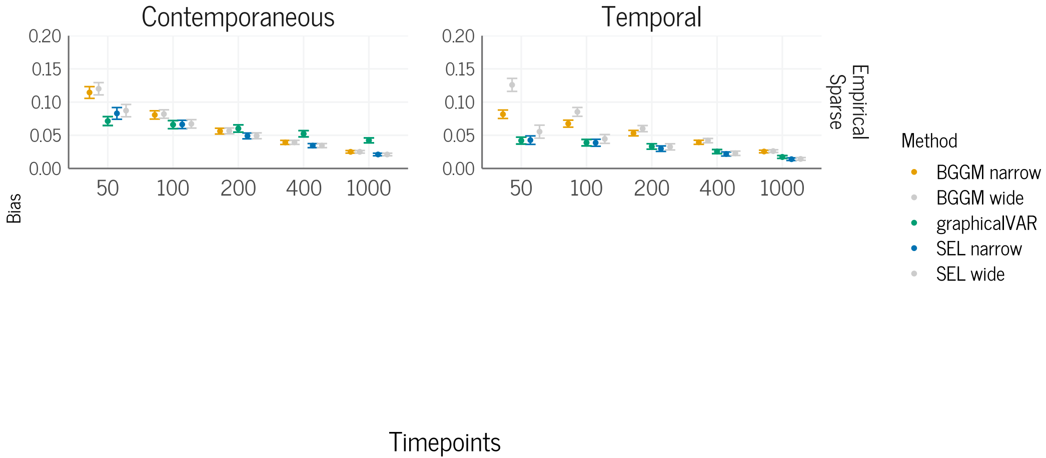
$\left.\begin{array}{l}\text{Bayesian gVAR without thresholding}\\\text{Bayesian gVAR with CI-based thresholding}\end{array}\right\}$ Wide & narrow priors

Simulation 1: Estimation Performance

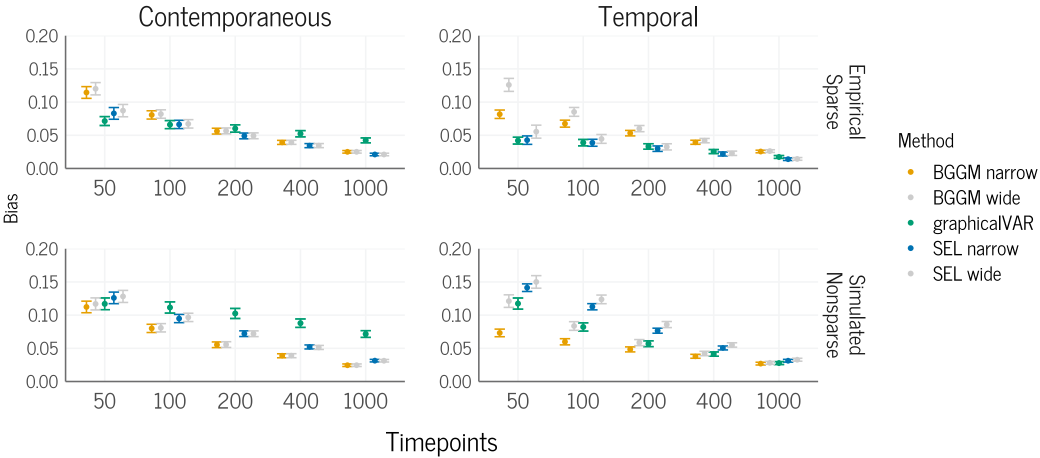
- DGP: Real-world examples with six and eight nodes & dense network
- Time series length: $t \in \{50, 100, 200, 400, 1000\}$
- Methods:
  - Bayesian gVAR without thresholding
  - Bayesian gVAR with CI-based thresholding
  - LASSO gVAR in default setting

  $\}$ Wide & narrow priors
- 1000 Monte Carlo replications

Simulation 1: Results



Simulation 1: Results



Philipps Universität Marburg

Idea of the Test

Idea of the Test

- Enough evidence that differences between networks are more than estimation uncertainty? (Williams et al., 2020)

$$
\begin{array}{cccc}
B_{a} & B_{b} & D & ||D||_F \\
\begin{pmatrix} 0.3 & 0.5 & 0.5 \\ 0.5 & 0.3 & 1 \\ 0.3 & 0.1 & 0.3 \end{pmatrix} - & \begin{pmatrix} 0.1 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.3 \\ 0.3 & 0.1 & 0.1 \end{pmatrix} = & \begin{pmatrix} 0.2 & 0 & 0.3 \\ 0 & 0.1 & 0.7 \\ 0.3 & 0 & 0.2 \end{pmatrix} \longrightarrow & 0.87
\end{array}
$$

Idea of the Test

- Enough evidence that differences between networks are more than estimation uncertainty? (Williams et al., 2020)

$$
\begin{array}{ccccccc}
B_{\color{blue}{a}} & & B_{\color{red}{b}} & & D & & ||D||_F \\
\begin{pmatrix} 0.3 & 0.5 & 0.5 \\ 0.5 & 0.3 & 1 \\ 0.3 & 0.1 & 0.3 \end{pmatrix} & - & \begin{pmatrix} 0.1 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.3 \\ 0.3 & 0.1 & 0.1 \end{pmatrix} & = & \begin{pmatrix} 0.2 & 0 & 0.3 \\ 0 & 0.1 & 0.7 \\ 0.3 & 0 & 0.2 \end{pmatrix} & \longrightarrow & 0.87
\end{array}
$$

- Randomly draw matrix pairs from posterior of two networks

Philipps Universität Marburg

Idea of the Test

- Enough evidence that differences between networks are more than estimation uncertainty? (Williams et al., 2020)

$$
\begin{array}{ccccccc}
B_{a} & & B_{b} & & D & & ||D||_F \\
\begin{pmatrix} 0.3 & 0.5 & 0.5 \\ 0.5 & 0.3 & 1 \\ 0.3 & 0.1 & 0.3 \end{pmatrix} & - & \begin{pmatrix} 0.1 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.3 \\ 0.3 & 0.1 & 0.1 \end{pmatrix} & = & \begin{pmatrix} 0.2 & 0 & 0.3 \\ 0 & 0.1 & 0.7 \\ 0.3 & 0 & 0.2 \end{pmatrix} & \longrightarrow & 0.87
\end{array}
$$

- Randomly draw matrix pairs from posterior of two networks
  ➥ Obtain reference distribution of uncertainty

Idea of the Test

- Enough evidence that differences between networks are more than estimation uncertainty? (Williams et al., 2020)

$$
\begin{array}{ccccccc}
B_{\color{blue}a} & & B_{\color{red}b} & & D & & ||D||_F \\
\begin{pmatrix} 0.3 & 0.5 & 0.5 \\ 0.5 & 0.3 & 1 \\ 0.3 & 0.1 & 0.3 \end{pmatrix} & - & \begin{pmatrix} 0.1 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.3 \\ 0.3 & 0.1 & 0.1 \end{pmatrix} & = & \begin{pmatrix} 0.2 & 0 & 0.3 \\ 0 & 0.1 & 0.7 \\ 0.3 & 0 & 0.2 \end{pmatrix} & \longrightarrow & 0.87
\end{array}
$$

- Randomly draw matrix pairs from posterior of two networks
  ➥ Obtain reference distribution of uncertainty
- Comparison of empirical norm with reference distribution for temporal and contemporaneous network

Idea of the Test

- Enough evidence that differences between networks are more than estimation uncertainty? (Williams et al., 2020)

$$
\begin{array}{ccccccc}
B_a & & B_b & & D & & ||D||_F \\
\begin{pmatrix} 0.3 & 0.5 & 0.5 \\ 0.5 & 0.3 & 1 \\ 0.3 & 0.1 & 0.3 \end{pmatrix} & - & \begin{pmatrix} 0.1 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.3 \\ 0.3 & 0.1 & 0.1 \end{pmatrix} & = & \begin{pmatrix} 0.2 & 0 & 0.3 \\ 0 & 0.1 & 0.7 \\ 0.3 & 0 & 0.2 \end{pmatrix} & \longrightarrow & 0.87
\end{array}
$$

- Randomly draw matrix pairs from posterior of two networks
  - ➥ Obtain reference distribution of uncertainty
- Comparison of empirical norm with reference distribution for temporal and contemporaneous network
- Decision: Heuristic rule

Performance of the Test

Performance of the Test

- DGPs & time series length:
  Same as in first simulation

Performance of the Test

- DGPs & time series length: Same as in first simulation
- Induce differences between matrices

Performance of the Test

- DGPs & time series length: Same as in first simulation
- Induce differences between matrices
- Matrix Norms

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \xrightarrow{\text{Largest} \times \{1.4,\, 1.6\}} \begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & \color{red}{1.4} \\ 0.3 & 0 & 0.3 \end{pmatrix}$$

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \xrightarrow{\text{All} \pm \{0.05,\, 0.10,\, 0.15\}} \begin{pmatrix} \color{red}{0.2} & \color{red}{0.6} & \color{red}{-0.1} \\ \color{red}{-0.4} & \color{red}{0.4} & \color{red}{1.1} \\ \color{red}{0.1} & \color{red}{0.1} & \color{red}{0.2} \end{pmatrix}$$

$$\begin{pmatrix} \color{blue}{0.3} & \color{red}{0.5} & \color{gray}{0} \\ \color{blue}{-0.5} & \color{red}{0.3} & \color{gray}{1} \\ \color{blue}{0.3} & \color{red}{0} & \color{gray}{0.3} \end{pmatrix} \xrightarrow{\text{Permute Columns}} \begin{pmatrix} \color{gray}{0} & \color{blue}{0.3} & \color{red}{0.5} \\ \color{gray}{1} & \color{blue}{-0.5} & \color{red}{0.3} \\ \color{gray}{0.3} & \color{blue}{0.3} & \color{red}{0} \end{pmatrix}$$

Performance of the Test

- DGPs & time series length: Same as in first simulation
- Induce differences between matrices
- Matrix Norms
  - Absolute value (ℓ_1) norm

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \xrightarrow{\text{Largest} \times \{1.4,\ 1.6\}} \begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1.4 \\ 0.3 & 0 & 0.3 \end{pmatrix}$$

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \xrightarrow{\text{All} \pm \{0.05,\ 0.10,\ 0.15\}} \begin{pmatrix} 0.2 & 0.6 & -0.1 \\ -0.4 & 0.4 & 1.1 \\ 0.1 & 0.1 & 0.2 \end{pmatrix}$$

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \xrightarrow{\text{Permute Columns}} \begin{pmatrix} 0 & 0.3 & 0.5 \\ 1 & -0.5 & 0.3 \\ 0.3 & 0.3 & 0 \end{pmatrix}$$

Performance of the Test

- DGPs & time series length: Same as in first simulation
- Induce differences between matrices
- Matrix Norms
  - Absolute value (ℓ_1) norm
  - Frobenius (ℓ_2) norm

Largest × $\{1.4, 1.6\}$

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1.4 \\ 0.3 & 0 & 0.3 \end{pmatrix}$$

All $\pm \{0.05, 0.10, 0.15\}$

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.2 & 0.6 & -0.1 \\ -0.4 & 0.4 & 1.1 \\ 0.1 & 0.1 & 0.2 \end{pmatrix}$$

Permute Columns

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 0.3 & 0.5 \\ 1 & -0.5 & 0.3 \\ 0.3 & 0.3 & 0 \end{pmatrix}$$

Performance of the Test

- DGPs & time series length: Same as in first simulation
- Induce differences between matrices
- Matrix Norms
  - Absolute value (ℓ_1) norm
  - Frobenius (ℓ_2) norm
  - Maximum (ℓ_∞) norm

Largest × {1.4, 1.6}

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1.4 \\ 0.3 & 0 & 0.3 \end{pmatrix}$$

All ± {0.05, 0.10, 0.15}

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.2 & 0.6 & -0.1 \\ -0.4 & 0.4 & 1.1 \\ 0.1 & 0.1 & 0.2 \end{pmatrix}$$

Permute Columns

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 0.3 & 0.5 \\ 1 & -0.5 & 0.3 \\ 0.3 & 0.3 & 0 \end{pmatrix}$$

Performance of the Test

- DGPs & time series length: Same as in first simulation
- Induce differences between matrices
- Matrix Norms
  - Absolute value (ℓ_1) norm
  - Frobenius (ℓ_2) norm
  - Maximum (ℓ_∞) norm

Largest × {1.4, 1.6}

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1.4 \\ 0.3 & 0 & 0.3 \end{pmatrix}$$

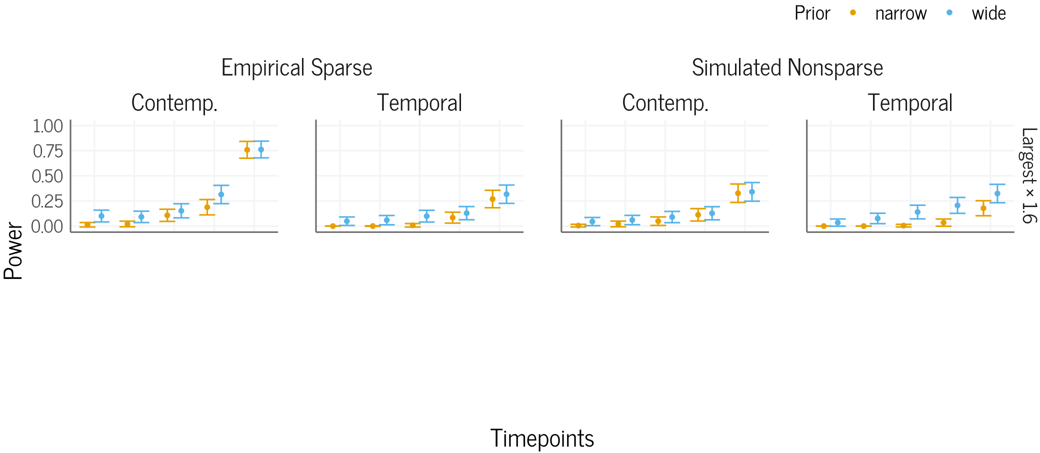
All $\pm$ {0.05, 0.10, 0.15}

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.2 & 0.6 & -0.1 \\ -0.4 & 0.4 & 1.1 \\ 0.1 & 0.1 & 0.2 \end{pmatrix}$$

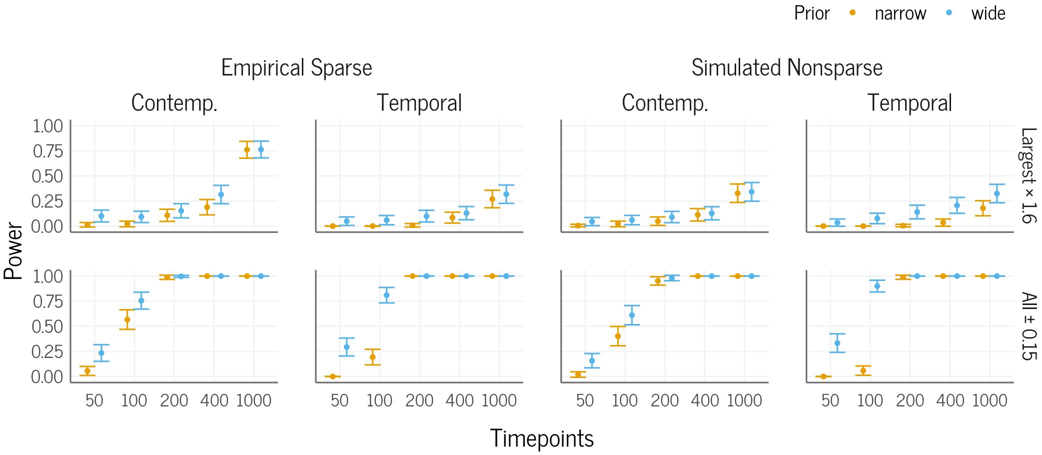
Permute Columns

$$\begin{pmatrix} 0.3 & 0.5 & 0 \\ -0.5 & 0.3 & 1 \\ 0.3 & 0 & 0.3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 0.3 & 0.5 \\ 1 & -0.5 & 0.3 \\ 0.3 & 0.3 & 0 \end{pmatrix}$$

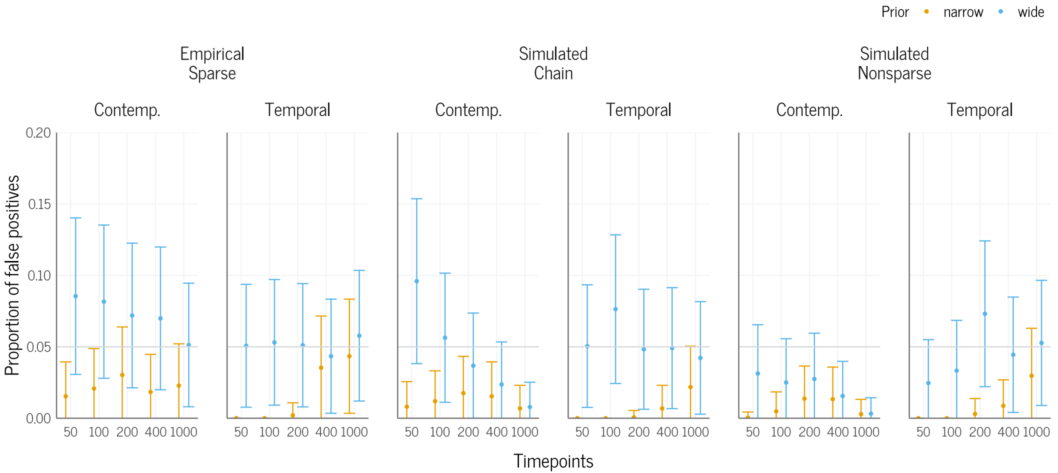
Performance of the Test



Performance of the Test



False Positive Rate



Limitations

Limitations

- Typical limitations of idiographic network models still apply

Limitations

- Typical limitations of idiographic network models still apply
    - Assumption of stationarity

Limitations

- Typical limitations of idiographic network models still apply
  - Assumption of stationarity
  - Studying heterogeneity likely works better with multilevel models

Limitations

- Typical limitations of idiographic network models still apply
  - Assumption of stationarity
  - Studying heterogeneity likely works better with multilevel models
- No ‘proper’ Bayesian edge selection

Limitations

- Typical limitations of idiographic network models still apply
  - Assumption of stationarity
  - Studying heterogeneity likely works better with multilevel models
- No ‘proper’ Bayesian edge selection
- Limited freedom in estimation method

Implications

Implications

- Bayesian Estimation can be a viable alternative to LASSO for idiographic networks

Implications

- Bayesian Estimation can be a viable alternative to LASSO for idiographic networks
  - ➥ Choice of estimation method depends on sparsity assumption

Implications

- Bayesian Estimation can be a viable alternative to LASSO for idiographic networks
  ➥ Choice of estimation method depends on sparsity assumption
- The test may guard against wrong interpretations of heterogeneity

Implications

- Bayesian Estimation can be a viable alternative to LASSO for idiographic networks
  - ➥ Choice of estimation method depends on sparsity assumption
- The test may guard against wrong interpretations of heterogeneity
  - Implemented in R package `tsnet` (https://github.com/bsiepe/tsnet)

Implications

- Bayesian Estimation can be a viable alternative to LASSO for idiographic networks
  - ➥ Choice of estimation method depends on sparsity assumption
- The test may guard against wrong interpretations of heterogeneity
  - Implemented in R package `tsnet` (https://github.com/bsiepe/tsnet)
  - Another very recent invariance test in Hoekstra et al., 2023

Implications

- Bayesian Estimation can be a viable alternative to LASSO for idiographic networks
  - ➥ Choice of estimation method depends on sparsity assumption
- The test may guard against wrong interpretations of heterogeneity
  - Implemented in R package `tsnet` (https://github.com/bsiepe/tsnet)
  - Another very recent invariance test in Hoekstra et al., 2023
  - Potential use in intra-individual comparisons

References

Hoekstra, R. H. A., Epskamp, S., & Borsboom, D. (2022). Heterogeneity in individual network analysis: Reality or illusion? *Multivariate Behavioral Research*, *0*(0), 1–25. https://doi.org/10.1080/00273171.2022.2128020

Hoekstra, R. H. A., Epskamp, S., Nierenberg, A. A., Borsboom, D., & McNally, R. J. (2023). Testing similarity in longitudinal networks: The Individual Network Invariance Test (INIT). https://doi.org/10.31234/osf.io/ugs2r

Mansueto, A. C., Wiers, R., van Weert, J. C. M., Schouten, B. C., & Epskamp, S. (2020). Investigating the Feasibility of Idiographic Network Models. https://doi.org/10.31234/osf.io/hgcz6

Siepe, B. S., & Heck, D. W. (2023). Bayesian Estimation and Comparison of Idiographic Network Models. https://doi.org/10.31234/osf.io/uwfjc

Williams, D. R., & Mulder, J. (2021). BGGM: Bayesian Gaussian Graphical Models.

Williams, D. R., Rast, P., Pericchi, L. R., & Mulder, J. (2020). Comparing gaussian graphical models with the posterior predictive distribution and bayesian model selection. *Psychological methods*, *25*(5), 653–672. https://doi.org/10.1037/met0000254

Get in Touch

bjoern.siepe@uni-marburg.de
bsiepe.github.io (incl. slides)
psyarxiv.com/uwfjc/