Inference of partial correlations of a multivariate Gaussian time series

Andrew DiLernia

Grand Valley State University
Department of Statistics

October 13, 2025

Table of contents

Motivation

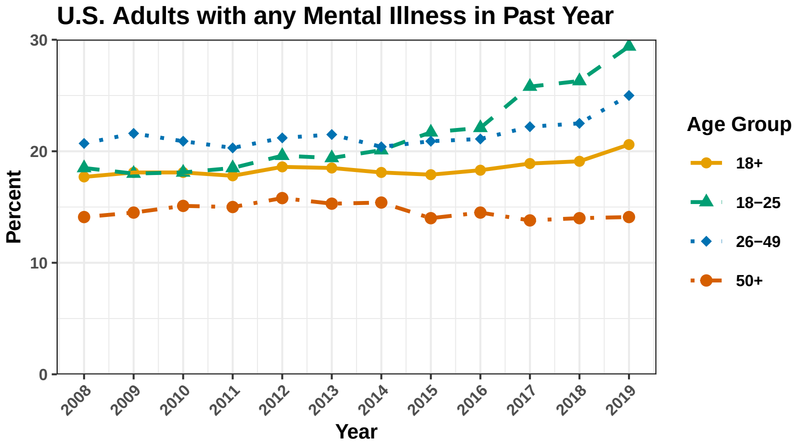


Figure 1: Data from (Substance Abuse and Mental Health Services Administration, 2020)

Motivation

- Over 50 million American adults living with a mental illness e.g., schizophrenia, bipolar disorder, and major depression in 2019 (Substance Abuse and Mental Health Services Administration, 2020).
- Techniques for describing neuronal activity of the brain are essential for improved treatments

Functional MRI (fMRI)

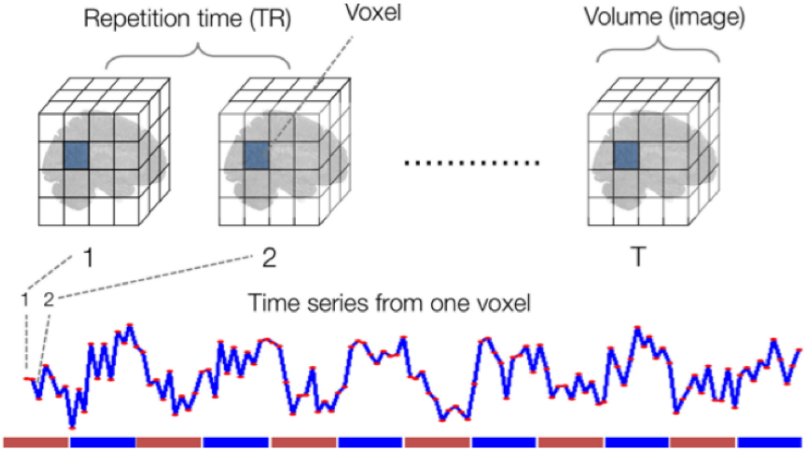
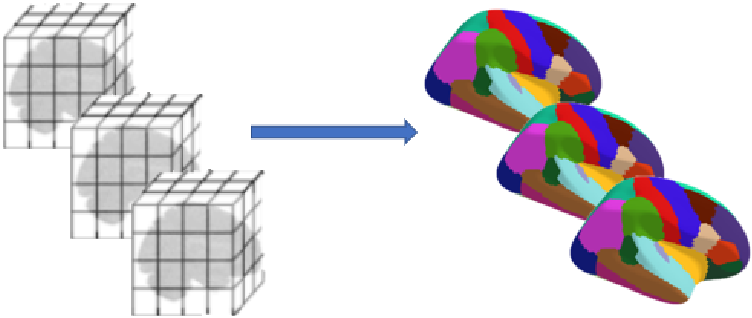


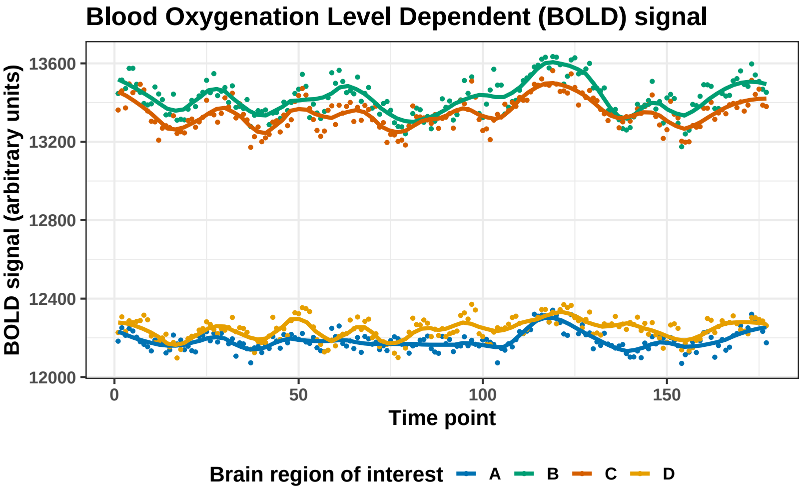
Figure 2: Image from (Wager and Lindquist, 2015).

Dimension Reduction

- Data often collected at thousands of voxels yielding high-dimensional data
- Average activation levels within sets of voxels called regions of interest (ROIs) at each time point



Example fMRI Data



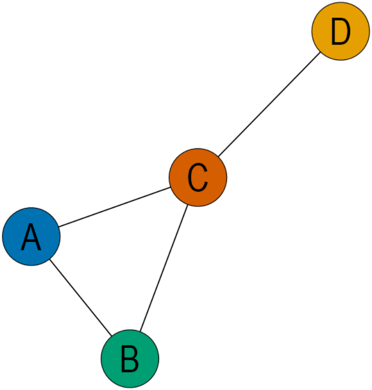
Functional Connectivity (FC)

- **Functional Connectivity (FC)**: temporal dependence of neuronal activity in regions of the brain (Friston et al., 1993)
- Alterations in FC associated with psychiatric disorders and neurodegenerative diseases
- Metrics to describe FC connections: marginal correlation, partial correlation, coherence, and mutual information among others

Graphical Modeling

One FC analysis method is **graphical modeling**, where nodes represent brain regions and edges connect dependent regions.

$$\begin{array}{c c} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{pmatrix} 1.00 & 0.17 & 0.09 & 0.00 \\ 0.17 & 1.00 & 0.18 & 0.00 \\ 0.09 & 0.18 & 1.00 & 0.21 \\ 0.00 & 0.00 & 0.21 & 1.00 \end{pmatrix} \end{array} \rightarrow$$



Partial Correlation Motivation

Why use partial instead of marginal correlation?

- Partial correlations describe linear relationship after removing effect of other variables
- Some assert more closely related to *effective connectivity*, the influence that ROIs exert on one another (Marrelec et al., 2006)

Partial Correlation Coefficient

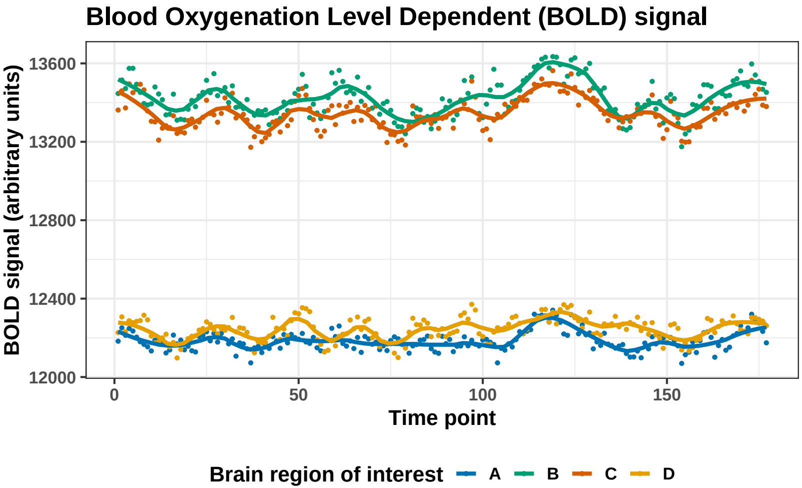
- $\mathbf{x}(t) = \{\mathbf{x}_k\}_{k=1}^{p}$: n-length realization of a p-variate Gaussian process that is second-order stationary and ergodic with $\mathbf{x}_k \in \mathbb{R}^n$
- $\mathbf{e}_i$ and $\mathbf{e}_j$: n-length vectors of contemporaneous OLS residuals from regressing $\mathbf{x}_i$ and $\mathbf{x}_j$ respectively on the other $p-2$ variables $\{\mathbf{x}_k\}_{k\neq i,j}$
- Partial correlation between $\mathbf{x}_i$ and $\mathbf{x}_j$:

$$r_{ij} = f\left(\begin{bmatrix}\mathbf{e}_i \\ \mathbf{e}_j\end{bmatrix}\right) = f(\mathbf{e}_{ij}) = \frac{\mathbf{e}_i^T\mathbf{e}_j}{\sqrt{\mathbf{e}_i^T\mathbf{e}_i\mathbf{e}_j^T\mathbf{e}_j}}$$

Inference of Partial Correlations

- How to describe uncertainty of the estimated partial correlations?
- For the naïve approach assuming normal independent observations, the estimated standard error for r_{ij} is $\sqrt{\frac{1-r_{ij}^2}{n-p}}$ (Cramer, 1974).
- Is this assumption reasonable for fMRI data?

Autocorrelation and fMRI



Motivation

- Need to conduct inference of partial correlations for multivariate time series data
- Inferential methods have been provided under various assumptions (e.g., independence, normality, the population partial correlations being 0)
- Important to provide flexible inferential methods with more reasonable assumptions for fMRI data

Derived Asymptotic Distribution

We derived an asymptotic distribution for the partial correlations of a multivariate time series with mild regularity conditions, providing the following:

- Estimator for the covariance matrix of partial correlations of a multivariate time series
- Inferential methods using this distribution for the partial correlations of a multivariate time series

How did we derive this distribution?

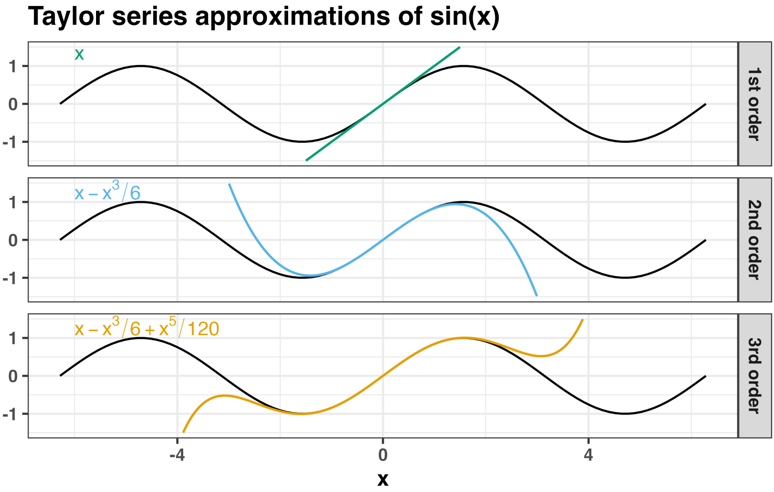
Taylor Series

The Taylor series of a univariate function $f(x)$ about a point a is

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$
$$= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

where $f^{(n)}(a)$ is the nth derivative of $f(\cdot)$ at a.

Taylor Series Approximation



Estimating Moments with Taylor Series

Using Taylor series expansion about μ_x, we can estimate the moments of a function, $f(\cdot)$, of a random variable, x:

First Moment of $f(x)$:

$$\mathbb{E}[f(x)] \approx f(\mu_x) + \frac{1}{2} f''(\mu_x) \sigma_x^2$$

where $\mu_x = \mathbb{E}[x]$ and $\mathrm{Var}(x) = \sigma_x^2$.

Second Moment of $f(x)$:

$$\mathbb{E}[f(x)^2] \approx f(\mu_x)^2 + f(\mu_x) f''(\mu_x) \sigma_x^2$$

Multivariate Taylor Series

The Taylor series of a scalar-valued multivariate function $f(\mathbf{x})$, $f : \mathbb{R}^k \to \mathbb{R}$, about a vector $\mathbf{a}$ is

$$f(\mathbf{x}) = f(\mathbf{a}) + (\mathbf{x} - \mathbf{a})^T \nabla f(\mathbf{a}) + \frac{1}{2}(\mathbf{x} - \mathbf{a})^T H_f(\mathbf{a})(\mathbf{x} - \mathbf{a}) + \cdots$$

where

- $\nabla f(\mathbf{a}) \in \mathbb{R}^k$ is the gradient vector of $f(\cdot)$ at $\mathbf{a}$
- $H_f(\mathbf{a}) \in \mathbb{R}^{k \times k}$ is the Hessian matrix of $f(\cdot)$ at $\mathbf{a}$

Multivariate Taylor Series Approximations of Moments

If $\mathbf{x}$ is a random vector, the approximations for the mean and variance of $f(\mathbf{x})$ using an expansion around μ_x are given by:

$$\mathbb{E}[f(\mathbf{x})] = f(\mu_x) + \frac{1}{2}\,\mathrm{trace}(H_f(\mu_x)\Sigma_x)$$

$$\mathrm{Var}[f(\mathbf{x})] = \nabla f(\mu_x)^T \Sigma_x \nabla f(\mu_x) + \frac{1}{2}\,\mathrm{trace}\left(H_f(\mu_x)\Sigma_x H_f(\mu_x)\Sigma_x\right)$$

where

- ∇f and H_f denote the gradient and the Hessian matrix respectively
- $\mathbb{E}[\mathbf{x}] = \mu_x$
- Σ_x is the covariance matrix of $\mathbf{x}$

Taylor Series Approximation

We approximate the function, $r_{ij} = f(e_{ij})$ using a second-order Taylor series expansion around $\varepsilon_{ij} = \left[\varepsilon_i^T, \quad \varepsilon_j^T\right]^T$, the vector such that $\rho_{ij} = f(\varepsilon_{ij})$, as

$$f(e_{ij}) \approx f(\varepsilon_{ij}) + (e_{ij} - \varepsilon_{ij})^T \nabla f(\varepsilon_{ij}) + \frac{1}{2}(e_{ij} - \varepsilon_{ij})^T H\{f(\varepsilon_{ij})\}(e_{ij} - \varepsilon_{ij}),$$

where

- $\nabla f(\varepsilon_{ij}) \in \mathbb{R}^{2n}$ is the gradient vector of $f(\varepsilon_{ij})$
- $H\{f(\varepsilon_{ij})\} \in \mathbb{R}^{2n \times 2n}$ is the Hessian matrix of $f(\varepsilon_{ij})$

Asymptotic Variance

Theorem 1.

Assume $x(t)$ is a multivariate Gaussian time series satisfying mild regularity conditions. Then the asymptotic variance of r_{ij} is $\tilde{\gamma}_{ij} = 1/2\left(\textit{trace}\left[\mathbf{H}\{f(\varepsilon_{ij})\}\Sigma_{ij}\mathbf{H}\{f(\varepsilon_{ij})\}\Sigma_{ij}\right]\right)$ where $\Sigma_{ij} = \textit{Cov}(\mathbf{e}_{ij})$.

where

$$\Sigma_{ij} = \mathsf{Cov}(\mathbf{e}_{ij}) = \begin{bmatrix} \mathsf{Cov}(\mathbf{e}_i) & \mathsf{Cov}(\mathbf{e}_i, \mathbf{e}_j) \\ \mathsf{Cov}(\mathbf{e}_j, \mathbf{e}_i) & \mathsf{Cov}(\mathbf{e}_j) \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$$

Proposed Variance Estimator

Proposed estimator for this asymptotic variance:

$$\hat{\gamma}_{ij} = 1/2 \left(\text{trace} \left[\mathbf{H}\{f(\mathbf{e}_{ij})\} \hat{\Sigma}_{ij} \mathbf{H}\{f(\mathbf{e}_{ij})\} \hat{\Sigma}_{ij} \right] \right)$$

- Tapered covariance estimators for the blocks of the covariance matrix $\hat{\Sigma}_{ij}$ (McMurry and Politis, 2010)
- Method of moments estimators for the Hessian matrix using the residual vector $\mathbf{e}_{ij}$

Inferential Methods

- Wald confidence intervals for ρ_{ij}:

$$r_{ij} \pm Z_{\alpha/2} \times \sqrt{\widehat{\gamma}_{ij}}$$

- $Z_{\alpha/2}$ is the $\alpha/2$ quantile of the standard normal distribution
- $\sqrt{\widehat{\gamma}_{ij}}$ is the estimated standard error of r_{ij}

Simulation Settings

- Simulated 1,000 data sets from first-order vector autoregressive, VAR(1), model for $n = 100$ or 500 time points, $p = 5$, 10, or 15 variables with autocorrelation parameter ϕ
- Three different amounts of autocorrelation: $\phi \in \{0, 0.40, 0.80\}$
- Generated partial correlations from the set $\{-0.30, 0, 0.30\}$ with equal probability

Comparison Methods

Naive approach: assumes independent and normally distributed observations

- $r_{ij} \pm t^*_{(n-p)}(1 - r^2_{ij})^{1/2}(n-p)^{-1/2}$ (Cramer, 1974)

where $t^*_{(n-p);\alpha/2}$ is the $\alpha/2$ quantile of a t-distribution with $n-p$ degrees of freedom.

Comparison Methods

Fisher-transformation approach:

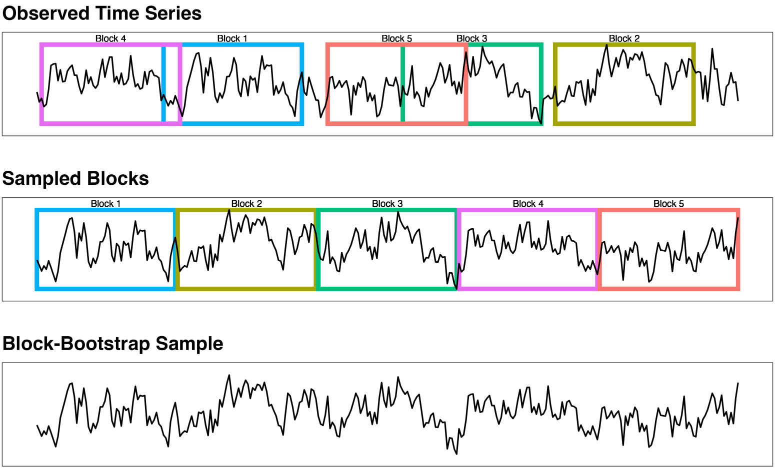
- $\tanh^{-1}(r_{ij}) = 1/2 \log\{(1 + r_{ij})/(1 - r_{ij})\}$ (Fisher, 1915)
- Construct confidence intervals for $\tanh^{-1}(\rho_{ij})$ centered around $\tanh^{-1}(r_{ij})$ using an estimated standard error of $(n - p - 1)^{-1/2}$ (Cramer, 1974)

Comparison Methods

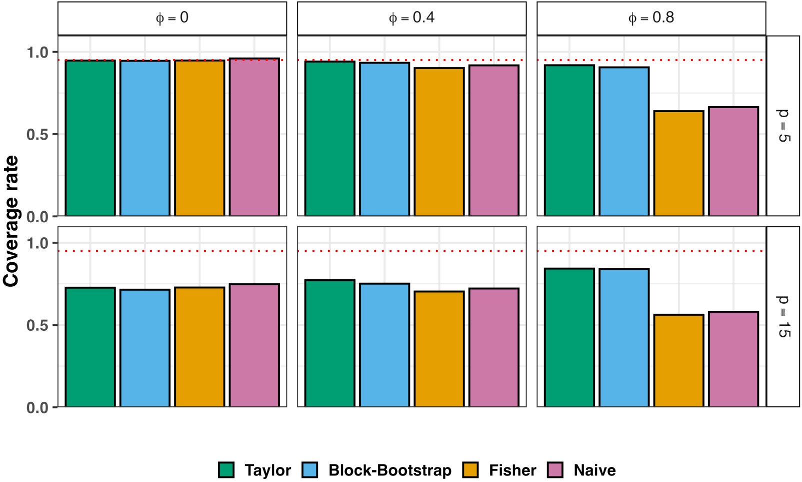
Block-bootstrap: nonparametric, assumes stationary time series

- Used the $\alpha/2$ and $1 - (\alpha/2)$ quantiles calculated from 1,000 bootstrap samples
- Selected the block-length using an automatic selection algorithm for stationary multivariate time series data (Politis and White, 2004)

Block-Bootstrap

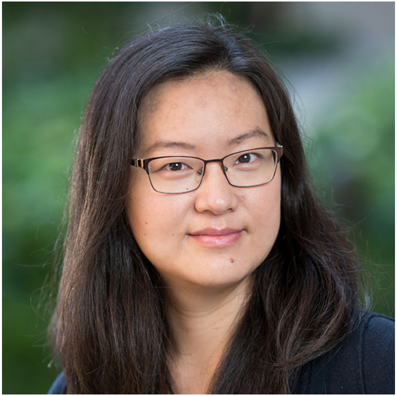
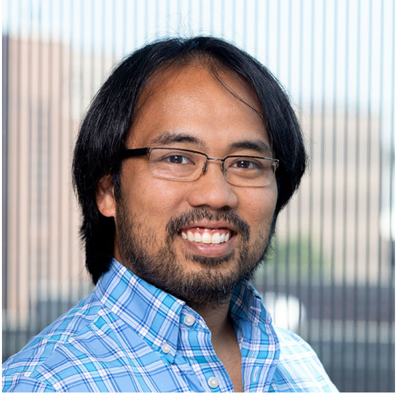


Confidence Interval Coverage Rates: $n = 500$



Collaborators

Thank you to my advisors for their guidance and support.

Dr. Lin Zhang	Dr. Mark Fiecas
	

Thank you

- A corresponding manuscript published in *Biometrika* in 2024, supplementary material, and an R package implementing the proposed covariance estimator are available on my website: https://www.andrewdilernia.com/publication/pccov/
- Questions?

Proposed Wald Test

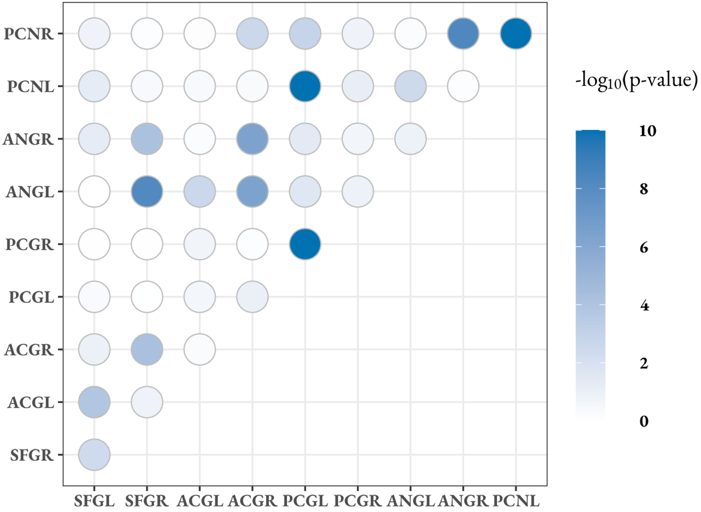
Wald test for testing whether individual partial correlations were 0 or not i.e., H_0: $\rho_{ij} = 0$ vs. H_A: $\rho_{ij} \neq 0$:

$$T_{w;ij} = r_{ij}^2 / \hat{\gamma}_{ij}$$

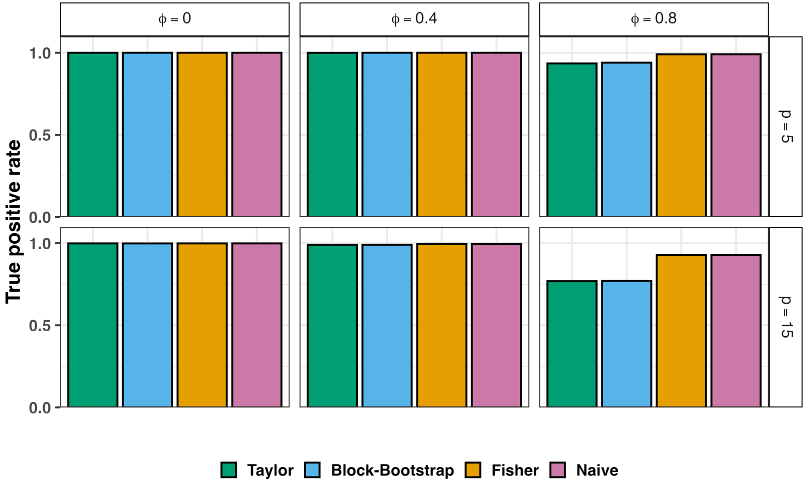
Case Study

- Used proposed methods to analyze data from the Autism Brain Imaging Data Exchange (ABIDE) initiative (Craddock et al., 2013)
- Data consisted of resting-state fMRI data with $n = 175$ volumes for $p = 10$ regions of interest in the Default Mode Network from the Automated Anatomical Labeling (AAL) atlas
- Most aligns with the $n = 100$ observations, $p = 10$ variables, and $\phi = 0.4$ simulation setting

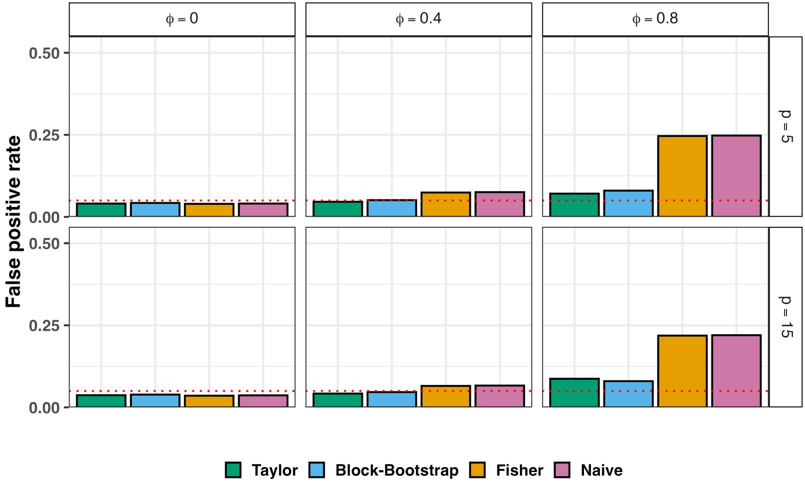
Case Study


PCNR
PCNL
ANGR
ANGL
PCGR
PCGL
ACGR
ACGL
SFGR
SFGL SFGR ACGL ACGR PCGL PCGR ANGL ANGR PCNL
-log10(p-value)
10
8
6
4
2
0

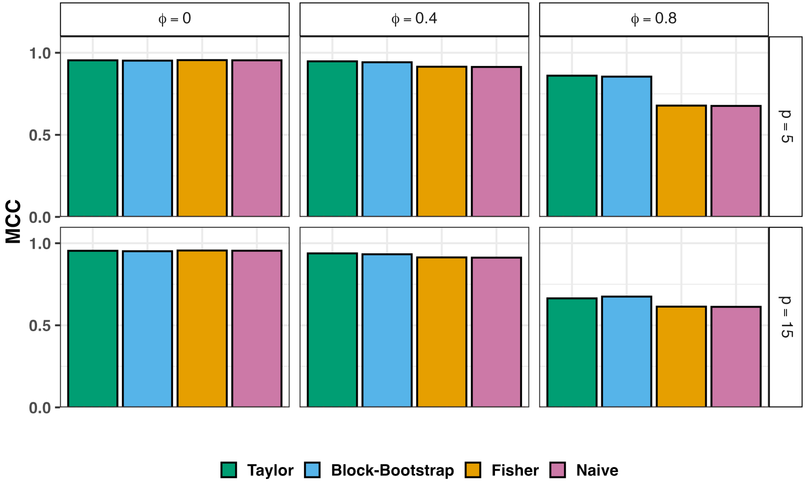
True Positive Rates: $n = 500$



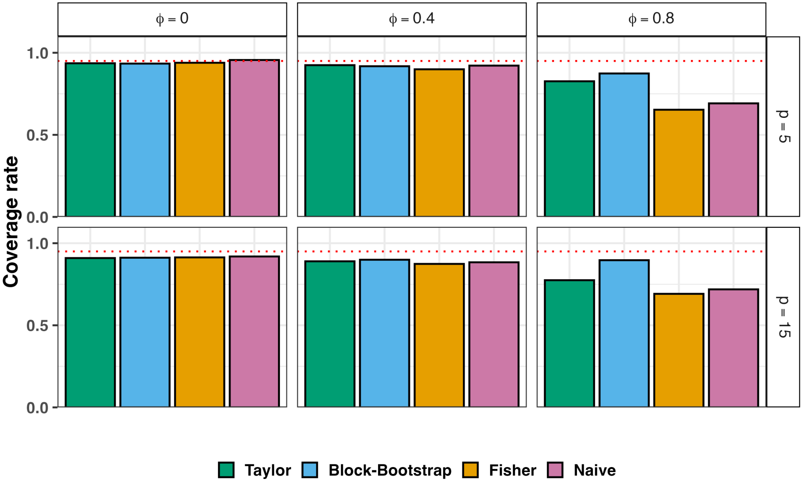
False Positive Rates: $n = 500$



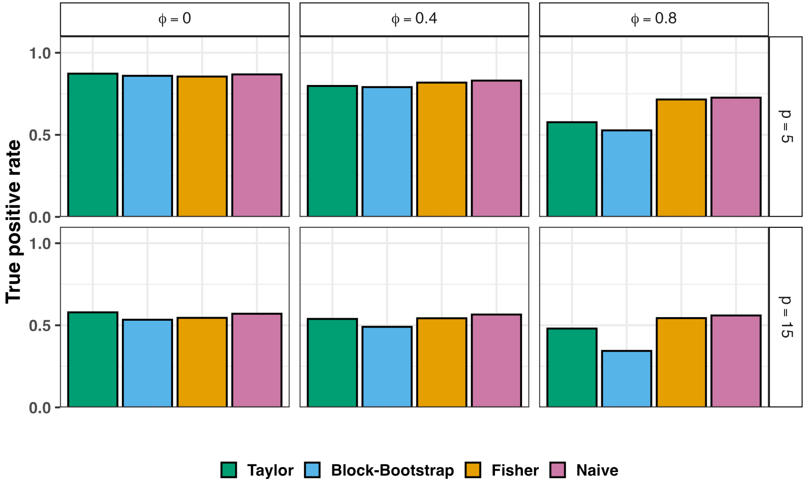
Matthews correlation coefficient (MCC): $n = 500$



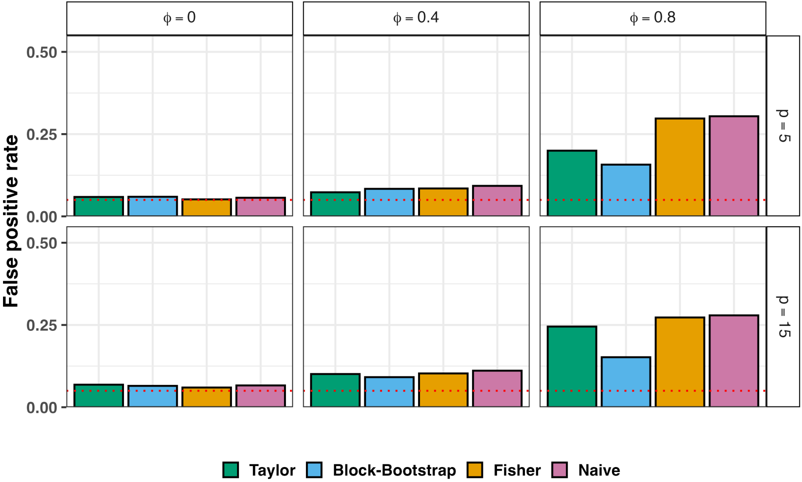
Confidence Interval Coverage Rates: $n = 100$



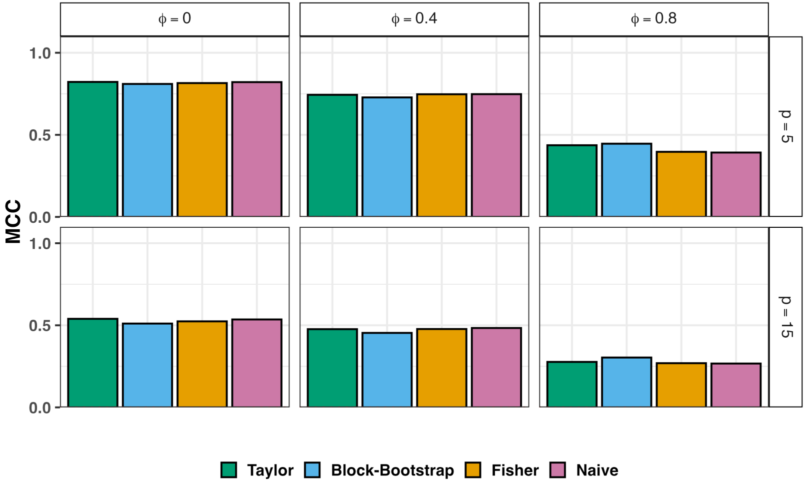
True Positive Rates: $n = 100$



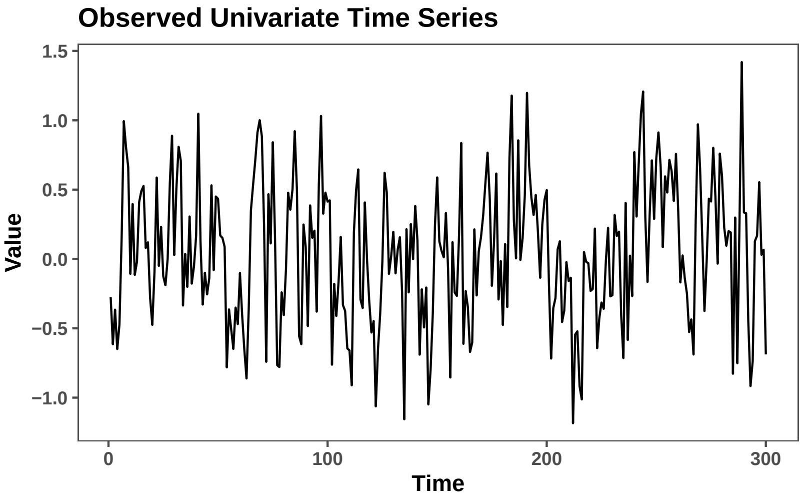
False Positive Rates: $n = 100$



Matthews correlation coefficient (MCC): $n = 100$



Univariate Time Series

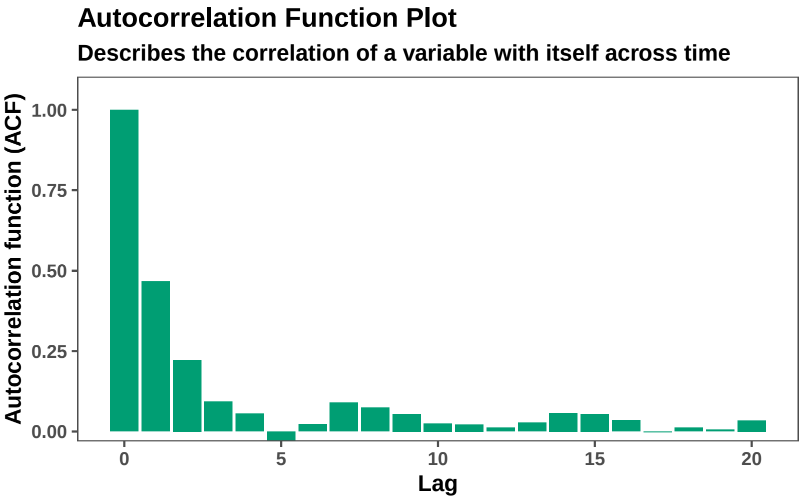


Autocovariance Matrix

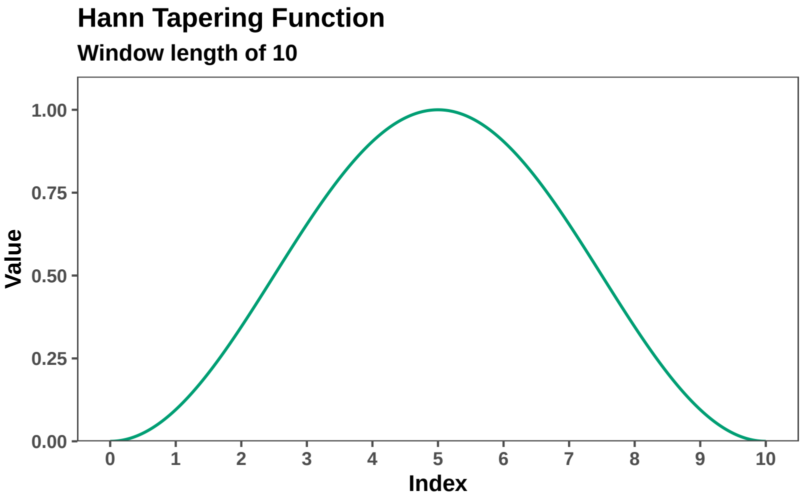
The autocovariance matrix of a univariate time series describes the covariance across time as a function of the lag.

$$\mathsf{Cov}(\mathbf{e}_i) = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \cdots & \gamma(2n-1) \\ \gamma(1) & \gamma(0) & \cdots & \cdots & \gamma(2n-2) \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \gamma(2n-1) & \gamma(2n-2) & \cdots & \cdots & \gamma(0) \end{bmatrix}$$

Temporal Autocorrelation



Tapering Function



Vector Autoregressive Model

A p-variate, first-order vector autoregressive model, denoted VAR(1), is

$$\mathbf{y}_t = A_0 + A_1 \mathbf{y}_{t-1} + \varepsilon_t$$

where

- $\mathbf{y}_t \in \mathbb{R}^p$
- $A_1 \in \mathbb{R}^{p \times p}$
- Every error term has a mean of zero: $\mathbb{E}[\varepsilon_t] = 0$
- The contemporaneous covariance matrix of error terms is a $k \times k$ positive-semidefinite matrix: $\mathbb{E}[\varepsilon_t \varepsilon_t'] = \Omega$
- No serial correlation in individual error terms: $\mathbb{E}[\varepsilon_t \varepsilon_{t-k}'] = 0 \quad$ for any non-zero k

References

Cameron Craddock, Yassine Benhajali, Carlton Chu, François Chouinard, Alan Evans, Andras Jakab, Budhachandra S Khundrakpam, Jamie D Lewis, Qingyang Li, Michael Milham, Chao-Gan Yan, and Pierre Bellec. The neuro bureau preprocessing initiative: Open sharing of preprocessed neuroimaging data and derivatives. *Neuroinformatics*, 2013.

E. M. Cramer. Brief report: The distribution of partial correlations and generalizations. *Multivariate Behavioral Research*, 9(1):119–122, 1974. doi: 10.1207/s15327906mbr0901_9. URL https://doi.org/10.1207/s15327906mbr0901_9. PMID: 26828735.

R. A. Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521, 1915. ISSN 0006-3444.

K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak. Functional connectivity: The principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow & Metabolism*, 13(1):5–14, 1993. ISSN 0271-678X.

G. Marrelec, A. Krainik, H. Duffau, M. Pélégrini-Issac, S. Lehéricy, J. Doyon, and H. Benali. Partial correlation for functional brain interactivity investigation in functional MRI. *NeuroImage*, 32(1):228–237, 2006. ISSN 1053-8119.

References (cont.)

Thomas L. McMurry and Dimitris N. Politis. Banded and tapered estimates for autocovariance matrices and the linear process bootstrap. *Journal of Time Series Analysis*, 31(6):471–482, 2010.

D. N. Politis and H. White. Automatic block-length selection for the dependent bootstrap. *Econometric reviews*, 23(1):53–70, 2004. ISSN 1532-4168.

Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the united states: Results from the 2019 national survey on drug use and health. Technical Report HHS Publication No. PEP20-07-01-001, NSDUH Series H-55, Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, 5600 Fishers Lane, Rockville, MD 20857, 2020. URL https://www.samhsa.gov/data/report/2019-nsduh-annual-national-report.

T. D. Wager and M. A. Lindquist. *Principles of fMRI*, chapter 6. Leanpub, 2015.