Benchmarking Brain Connectivity Graph Inference: A Novel Validation Approach [1]

Alice Chevaux[†], Ali Fakhar[†], Kévin Polisano[†], Irène Gannaz[*], Sophie Achard[†]

[†]Univ. Grenoble Alpes, CNRS, Inria, LJK, F-38000 Grenoble, France
[*]Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France

EUSIPCO 2025

[1]This work was supported by the Agence Nationale de la Recherche under the France 2030 programme, reference ANR-23-IACL-0006.

Motivations for fMRI data

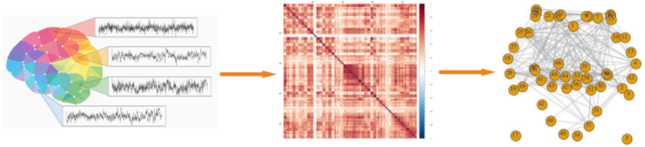


Figure 1: Illustration of the usual inference of graph for fMRI data

Objective: Recover the adjacency matrix A of a matrix Σ a positive semi-definite matrix

- Σ represent the connectivity (resp. correlation or precision matrix)
- Observations are i.i.d realisations of $\boldsymbol{X} = (X_1, \ldots, X_p)^\top \longrightarrow \mathcal{N}(0, \Sigma)$ or resp. $\boldsymbol{X} \longrightarrow \mathcal{N}(0, \Sigma^{-1})$
- number of observations: T , dimension of the matrix: $p \times p$

Problem Statement

Shortcomings of Current Approaches:

- For statistical methods
  1. Few mathematical guarantee with realistic settings (T low, p high)
  2. Difficulty to simulate matrices

Problem Statement

Shortcomings of Current Approaches:

- For statistical methods
  1. Few mathematical guarantee with realistic settings (T low, p high)
  2. Difficulty to simulate matrices

- For benchmark papers
  1. Rely on real datasets [3, 6]
  2. Simulate under the assumption of sparsity [5]

Problem Statement

Shortcomings of Current Approaches:

- For statistical methods
  1. Few mathematical guarantee with realistic settings (T low, p high)
  2. Difficulty to simulate matrices
- For benchmark papers
  1. Rely on real datasets [3, 6]
  2. Simulate under the assumption of sparsity [5]

Objectives:

- Simulate PSD matrices according to parameters that we choose
- Propose a pipeline to measure the performance of a method

Parameters of interest

- **Graph density** (d): proportion of edges in the adjacency matrix

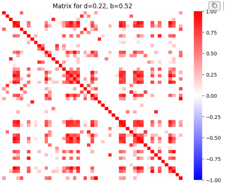


Figure 2: PSD matrix with $b = 0.52$ and $d = 0.22$

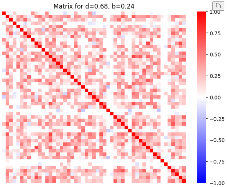


Figure 3: PSD matrix with $b = 0.24$ and $d = 0.68$

Parameters of interest

- **Graph density** (d): proportion of edges in the adjacency matrix
- **Sample size** (T)

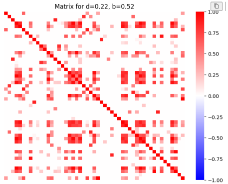


Figure 2: PSD matrix with $b = 0.52$ and $d = 0.22$

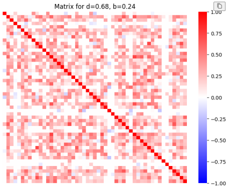


Figure 3: PSD matrix with $b = 0.24$ and $d = 0.68$

Parameters of interest

- **Graph density** (d): proportion of edges in the adjacency matrix
- **Sample size** (T)
- **Signal-to-noise level** (b): the mean value of the nonzero coefficients in Σ

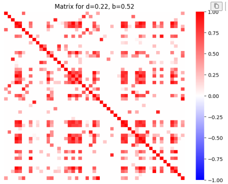


Figure 2: PSD matrix with $b = 0.52$ and $d = 0.22$

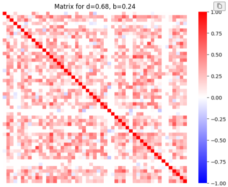


Figure 3: PSD matrix with $b = 0.24$ and $d = 0.68$

Convex Optimization

Objectives

- Find correlation matrix matching adjacency matrix A (with a number of edges n_A)
- Control signal-to-noise ratio
- Choose a target matrix $\bar{\Sigma}$ (initialisation value)

Optimization Problem

$$\Sigma \succcurlyeq 0, \Sigma_{ii} = 1, \quad A_{ij} = 0 \Longrightarrow \Sigma_{ij} = 0. \tag{1}$$

$$\frac{1}{2|n_A|} \sum_{i \neq j} \Sigma_{ij} \geq b. \tag{2}$$

$$\begin{aligned} \underset{C}{\text{minimize}} \quad & \frac{1}{2} \|\Sigma - \bar{\Sigma}\|_F^2, \\ \text{subject to} \quad & \text{constraints (1) and (2),} \end{aligned} \tag{3}$$

Simulation of a set of matrices

- Pipeline of simulation:
  1. simulate A according to a type of graph for different graph densities d
  2. sample b between 0 and 1
  3. sample $\bar{\Sigma}$
- Chordal graph simulation offer a larger range of b for every density

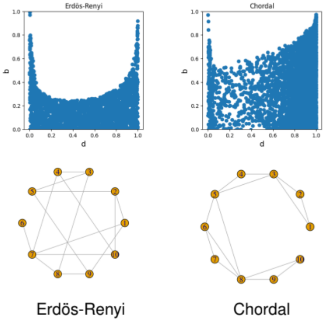


Figure 4: Representation of the set of matrices we were able to simulate with respect to the mean value of the non-zeros coefficients b and the proportion of edges d, with a chordal graph structure.

Methods to compare

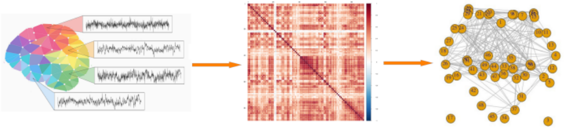


Figure 5: Illustration of the usual inference of graph for fMRI data

Methods with an arbitrary threshold:

- Proportional thresholding
- Hard-thresholding

Statistical methods to choose a threshold:

- Multiple testing with correction [7],[1],[4]
- Percolation-threshold [2]
- Threshold based on a mixture-model

Sparse Gaussian Graphical Model

- Graphical Lasso

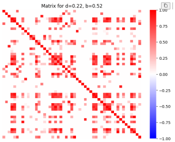
Is there an optimal threshold ?

- Hard-thresholding consists in applying a threshold τ between 0 and 1 on the empirical correlation matrix $\hat{\Sigma}$ to obtain $\hat{A}(\tau) = (\hat{\Sigma} > \tau)$
- **Limit cases** :

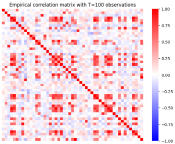
$$\hat{A}(\tau) = \begin{cases} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} & \text{if } \tau = 0 \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \text{if } \tau = 1 \end{cases}$$

- To compare $\hat{A}$ and A we use: **Accuracy**:

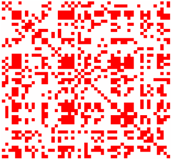
$$\frac{TP + TN}{TP + TN + FP + FN}$$



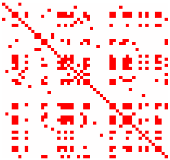
(a) Ground-truth Σ



(b) Empirical correlation $\hat{\Sigma}$



(c) $\hat{A}(0.1)$



(d) $\hat{A}(0.4)$

Figure 6: Differences between the ground-truth correlation matrix Σ, the empirical correlation matrix obtained with $T = 100$ observations and the $\hat{A}(\tau)$ adjacency matrix estimated using $\tau = 0.1$ and $\tau = 0.4$

Is there an optimal threshold ?

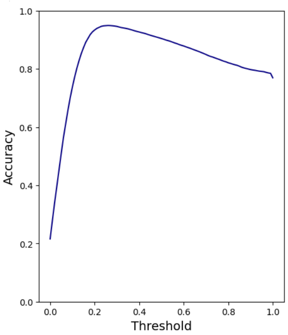
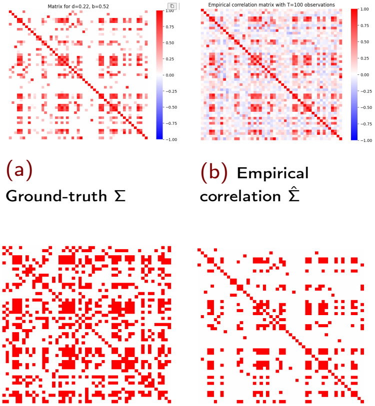


Figure 6: Accuracy obtained when applying different thresholds on an empirical correlation matrix for a number of observations $T = 100$



(a) Ground-truth Σ

(b) Empirical correlation $\hat{\Sigma}$

(c) $\hat{A}(0.1)$

(d) $\hat{A}(0.4)$

Figure 7: Differences between the ground-truth correlation matrix Σ, the empirical correlation matrix obtained with $T = 100$ observations and the $\hat{A}(\tau)$ adjacency matrix estimated using $\tau = 0.1$ and $\tau = 0.4$

Is there an optimal threshold ?

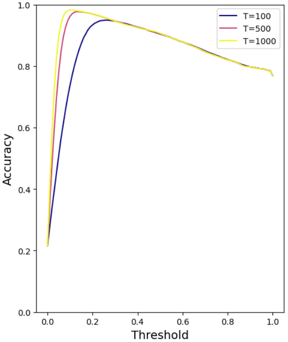
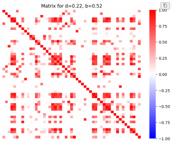
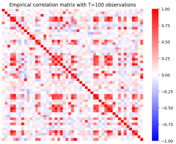


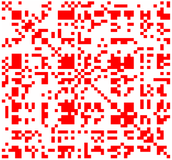
Figure 6: Accuracy obtained when applying different thresholds on an empirical correlation matrix depending on the number of observations T



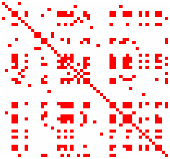
(a) Ground-truth Σ



(b) Empirical correlation $\hat{\Sigma}$



(c) $\hat{A}(0.1)$



(d) $\hat{A}(0.4)$

Figure 7: Differences between the ground-truth correlation matrix Σ, the empirical correlation matrix obtained with $T = 100$ observations and the $\hat{A}(\tau)$ adjacency matrix estimated using $\tau = 0.1$ and $\tau = 0.4$

Effects of parameters on the optimal threshold

- How does the parameters d and b affect the optimal threshold we hope to find ?

- Accuracy itself is not enough to evaluate a method due to d
- The threshold choice should depend on b and T

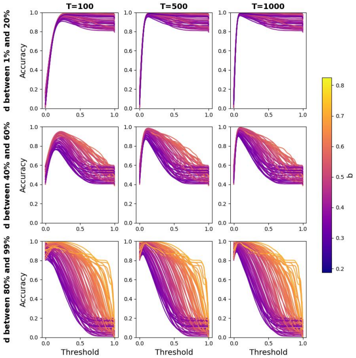


Figure 8: Accuracy obtained when applying different thresholds on an empirical correlation matrix, depending on the graph density d, the number of observations T, and the mean value of the non-zeros coefficients b.

Global Performances of calibrated methods (1)

To compare $\hat{A}$ and A we use:

- **Accuracy**: $\dfrac{TP+TN}{TP+TN+FP+FN}$
- **True Positive Rate (TPR)**: $\dfrac{TP}{TP+FN}$
- **False Positive Rate (FPR)**: $\dfrac{FP}{FP+TN}$

where TP (True Positives) are correctly detected edges ($\hat{A}_{ij}=1, A^*_{ij}=1$), TN (True Negatives) are correctly absent edges ($\hat{A}_{ij}=0, A^*_{ij}=0$), FP (False Positives) are incorrectly added edges ($\hat{A}_{ij}=1, A^*_{ij}=0$), and FN (False Negatives) are missed edges ($\hat{A}_{ij}=0, A^*_{ij}=1$).

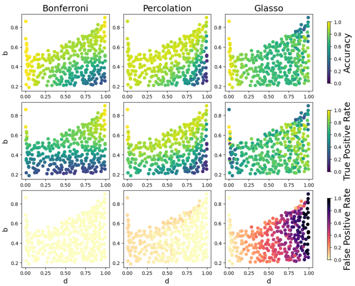
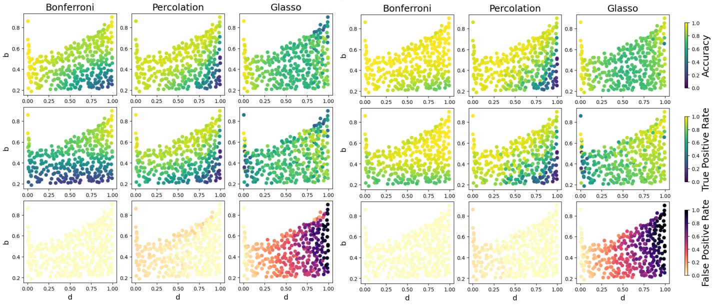


Figure 9: Accuracy , False Positive Rate and True positive Rate of 3 methods (Multiple testing with Bonferonni, Percolation thresholding and Graphical Lasso) for differents PSD matrices depending on b and d for $T=100$

Global Performances of calibrated methods (2)

- What are the parameters that affect the performances of the different methods ?



(a) T=100 (b) T=1000

Figure 10: Comparison of the Bonferonni procedure, the Graphical Lasso and the Percolation threshold methods (from left to right) using several metrics (from top to bottom): Accuracy, False Positive Rate (FPR) and True positive Rate (TPR).

Conclusion

Contributions:

- Method to simulate PSD matrices according to parameters
- Pipeline to evaluate a method
- Meaningful comparisons for users to have a better understanding of the limitations and particularities of well-known methods

Perspectives:

- Include new statistical methods and new metrics of performance
- Propose a ready-to-use package for users to confront their own method

code is available at : https://gricad-gitlab.univ-grenoble-alpes.fr/users/polisank/projects

Bibliography

[1] Carlo Bonferroni. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R istituto superiore di scienze economiche e commericiali di firenze*, 8:3–62, 1936.

[2] Cécile Bordier, Carlo Nicolini, and Angelo Bifone. Graph analysis and modularity of brain functional connectivity networks: searching for the optimal threshold. *Frontiers in neuroscience*, 11:441, 2017.

[3] Kamalaker Dadi, Mehdi Rahim, Alexandre Abraham, Darya Chyzhyk, Michael Milham, Bertrand Thirion, Gaël Varoquaux, Alzheimer's Disease Neuroimaging Initiative, et al. Benchmarking functional connectome-based predictive models for resting-state fMRI. *NeuroImage*, 192:115–134, 2019.

[4] Jelle J Goeman and Aldo Solari. Multiple hypothesis testing in genomics. *Statistics in medicine*, 33(11):1946–1978, 2014.

[5] Ginette Lafit, Francis Tuerlinckx, Inez Myin-Germeys, and Eva Ceulemans. A partial correlation screening approach for controlling the false positive rate in sparse Gaussian graphical models. *Scientific reports*, 9(1):17759, 2019.

[6] František Váša and Bratislav Mišić. Null models in network neuroscience. *Nature Reviews Neuroscience*, 23(8):493–504, 2022.

[7] Daniel Yekutieli and Yoav Benjamini. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference*, 82(1-2):171–196, 1999.