arXiv
bit.ly/4dD21Nt

SpARCD: A Spectral Graph Framework for Revealing Differential Functional Connectivity in fMRI Data

מעצבים את העתיד מאז 1925
האוניברסיטה העברית בירושלים

Shira Yoffe[1], Ziv Ben-Zion[3], Guy Gurevitch[6], Talma Hendler[4,5,6,7], Malka Gorfine[2], and Ariel Jaffe[1]

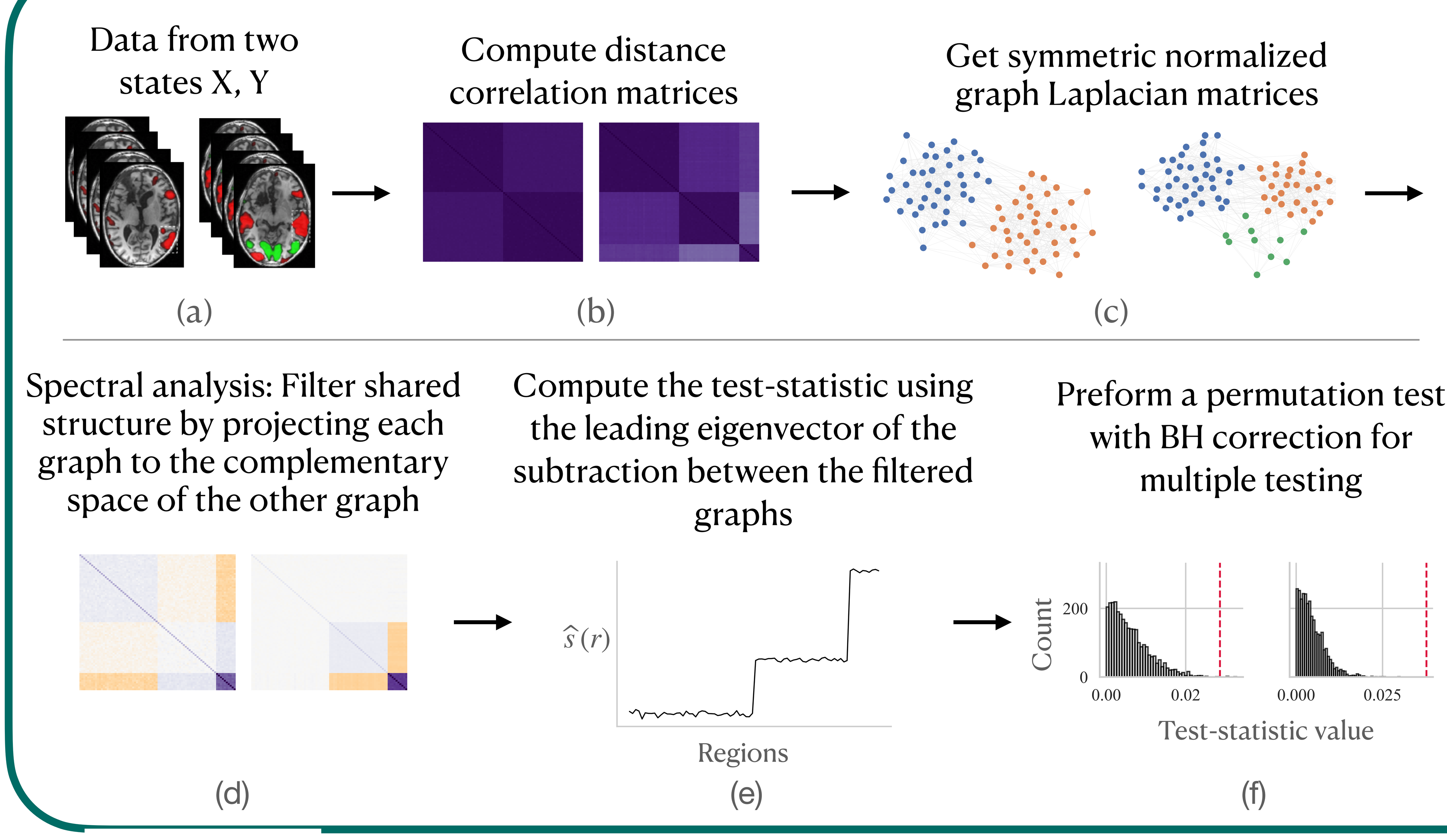
[1] Department of Statistics and Data Science, Hebrew University of Jerusalem [2] Department of Statistics and Operations Research, Tel Aviv University [3] School of Public Health, Faculty of Social Welfare and Health Sciences, University of Haifa [4] Sagol School of Neuroscience, Tel Aviv University [5] Sagol Brain Institute, Tel Aviv Sourasky Medical Center [6] Gray Faculty of Medical and Health Sciences, Tel Aviv University [7] School of Psychological Sciences, Faculty of Social Sciences, Tel Aviv Universit

Problem and Motivation

Functional connectivity (FC) networks quantify statistical dependence between brain regions using fMRI BOLD time series.
Objective: Given two experimental or clinical conditions, identify brain regions whose connectivity patterns differ between conditions.
Challenges: (i) noisy fMRI measurements, (ii) high-dimensional network structure, (iii) severe multiple-comparison burden, and (iv) potentially nonlinear dependencies.Existing approaches often have limited statistical power or fail to capture distributed connectivity changes [1].

We propose **SpARCD**, a spectral graph framework for identifying condition-specific brain regions.

Methods



- The *distance correlation* (Fig. 1 (b)):

$$\widehat{W} = \widehat{\mathscr{D}}(X_r, X_{r'}) = \frac{\langle \tilde{D}_r, \tilde{D}_{r'} \rangle}{\sqrt{\langle \tilde{D}_r, \tilde{D}_r \rangle \langle \tilde{D}_{r'}, \tilde{D}_{r'} \rangle}}$$

- Filter shared spectral structure between graphs (Fig. 1 (d)):

$$\widehat{Q}_X = I_R - \sum_{k=1}^{K} \hat{v}_X^{(k)} \hat{v}_X^{(k)\top} \quad \rightarrow \quad \widehat{L}_Y^{\text{filt}} = \widehat{Q}_X (I_R - \widehat{L}_Y) \widehat{Q}_X$$

Main Novelty:
(i) SpARCD uses distance correlation (d-Corr) to build FC networks, capturing linear and nonlinear dependencies.
(ii) It applies spectral filtering to construct a differential connectivity operator that captures condition-specific changes.
(iii) A permutation-based testing framework provides interpretable region-level significance maps for robust detection of altered connectivity.

Results

Simulation Studies: Evaluation across five connectivity scenarios demonstrated the robustness of SpARCD under diverse patterns of network change.
Full results are reported in the paper.

Real Data: Data: We applied SpARCD to resting-state fMRI data comparing PTSD and control groups. Results: The method identified eight significant regions (Fig. 3), primarily within visual processing areas, suggesting altered connectivity organization in visual cortical systems in PTSD. Additional Real-Data Analyses: Consistent connectivity-alteration patterns were also observed across two EFMT datasets and in EFMT vs. resting-state comparisons, supporting the robustness of the proposed framework.

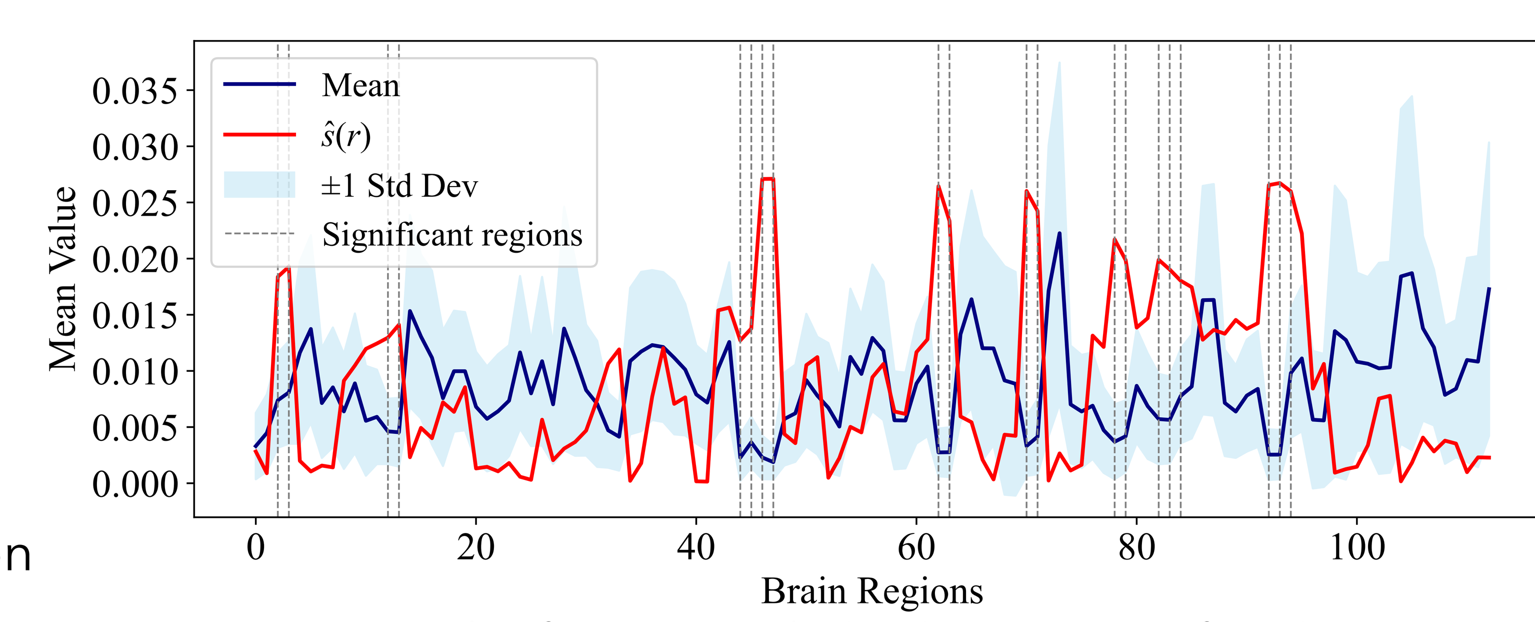


Figure 3: Results of SpARCD applied to the resting-state fMRI dataset. The observed test statistic s(r) (red) is shown together with the mean (blue) and standard deviation (light blue) of the permutation-based null distribution.

Theoretical Guarentees

Theorem 1 (Finite-sample validity of the block permutation test)
Fix a brain region r, and let $\hat{s}(r)$ denote the observed test statistic computed from block data $\mathscr{B}$ and condition labels A. Let $\mathscr{G}$ denote the set of allowable block permutations, and let $\hat{s}^{(\pi)}(r)$ denote the statistic computed from the permuted labels πA, for $\pi \in \mathscr{G}$. Assume that under the null hypothesis, the condition labels are exchangeable over $\mathscr{G}$ conditional on the unlabeled block structure; that is,
$(\mathscr{B}, A) \mid \mathcal{O} \overset{d}{=} (\mathscr{B}, \pi A) \mid \mathcal{O}$, for all $\pi \in \mathscr{G}$, where $\mathcal{O} = \{(\mathscr{B}, \pi A) : \pi \in \mathscr{G}\}$denotes the permutation orbit. If the statistic is computed using the same deterministic procedure for both observed and permuted data, and if all permutations in $\mathscr{G}$ are enumerated, then the exact permutation p-value $p_r^* = \frac{1}{|\mathscr{G}|} \sum_{\pi \in \mathscr{G}} \mathbb{I}\{\hat{s}^{(\pi)}(r) \geq \hat{s}(r)\}$ satisfies $\Pr_{H_0}(p_r^* \leq \alpha \mid \mathcal{O}) \leq \alpha$ for every $\alpha \in [0,1]$. Consequently, $\Pr_{H_0}(p_r^* \leq \alpha) \leq \alpha$.

Theorem 2 (Stability of eigenvector-based region ranking) Let R denote the number of brain regions, treated as fixed while $n \to \infty$. Let L_d be the population contrast matrix, and let $\widehat{L}_d$ denote its empirical counterpart. Assume $\Delta_n = \|\widehat{L}_d - L_d\|_2 = o_p(1)$. Let $\lambda_1, \ldots, \lambda_R$ denote the eigenvalues of L_d, ordered so that $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_R|$, and assume that the leading absolute eigenvalue is separated from the rest of the spectrum. Specifically, define $\rho = |\lambda_1| - \max_{j\geq 2}|\lambda_j| > 0$, and the eigengap is defined by $\varphi = \min_{j\geq 2}|\lambda_1 - \lambda_j| > 0$. Let v_d be a unit-norm eigenvector associated with λ_1, and define $s(r) = \frac{|v_d(r)|}{\|v_d\|_1}$, $r = 1,\ldots,R$. Then, as $n \to \infty$, the following statements hold with probability tending to one:

- Eigenvector stability: There exists a sign $\eta \in \{-1,1\}$ such that $\|\hat{v}_d - \eta v_d\|_2 \leq C_{\mathrm{DK}} \frac{\Delta_n}{\varphi}$, where C_{DK} is a universal Davis--Kahan constant.
- Score stability: $\sup_{1\leq r\leq R} |\hat{s}(r) - s(r)| = O_p\left(\frac{\Delta_n}{\varphi}\right)$.
- Ranking stability: If $s(r_1) - s(r_2) > 2 \sup_{1\leq r\leq R} |\hat{s}(r) - s(r)|$, then $\hat{s}(r_1) > \hat{s}(r_2)$.
- Top-K stability: Let S_K denote the set of the K regions with largest population scores, and define $\Delta_K = s_{(K)} - s_{(K+1)}$, where $s_{(1)} \geq \cdots \geq s_{(R)}$ are the ordered population scores. If $\Delta_K > 2 \sup_{1\leq r\leq R} |\hat{s}(r) - s(r)|$, then the estimated top-K set equals S_K.

References:
[1] Zalesky, A., Fornito, A., Bullmore, E., 2010. Network-based statistic: Identifying differences in brain networks. NeuroImage 53, 1197–1207.