Bayesian Joint Modeling of Multiple Brain Functional Networks

Joshua Lukemire*, Suprateek Kundu*, Giuseppe Pagnoni†, and Ying Guo*

*Department of Biostatistics and Bioinformatics, Emory University

†Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Italy

EMORY ROLLINS SCHOOL OF PUBLIC HEALTH | CBIS Center for Biomedical Imaging Statistics

Abstract

Brain function is organized in coordinated modes of spatio-temporal activity (functional networks) exhibiting an intrinsic baseline structure with variations under different experimental conditions. Existing approaches for uncovering such network structures typically do not explicitly model shared and differential patterns across networks, thus potentially reducing the detection power. We develop an integrative modeling approach for jointly modeling multiple brain networks across experimental conditions. The proposed Bayesian Joint Network Learning approach develops flexible priors on the edge probabilities involving a common intrinsic baseline structure and differential effects specific to individual networks. Conditional on these edge probabilities, connection strengths are modeled under a Bayesian spike and slab prior on the off-diagonal elements of the inverse covariance matrix. The model is fit under a posterior computation scheme based on Markov chain Monte Carlo. An application of the method to fMRI Stroop task data provides unique insights into brain network alterations between cognitive conditions.

Introduction and Goals

- Many researchers are interested in comparing brain networks across cognitive states induced by experimental conditions with the aim of identifying functional connections whose strengths reflect differences or commonalities between these conditions.
- Under a graph-theoretic approach, edges featuring differential strengths correspond to brain connections that are more activated or suppressed during one experimental condition as compared to others.
- The comparison of brain networks across multiple conditions may be performed on a single subject or, as in our case, at a group level. Group level comparisons have the advantage of being able to average out subject-specific idiosyncrasies, potentially providing greater power to detect underlying biological differences and similarities.
- Penalized approaches for the joint estimation of multiple graphical models typically smooth over the strength of connections across networks to enforce shared edges, which is a useful modeling assumption but may not be supported in practical brain network applications.
- In this work, we develop a Bayesian Gaussian graphical modeling approach for estimating multiple networks. The approach, denoted as Bayesian Joint Network Learning (BJNL), is implemented via a fully Gibbs posterior computation scheme which proceeds via Markov chain Monte Carlo (MCMC).
- Our approach models the probability of a connection as a parametric function of a baseline component shared across networks and differential components unique to each network.
- The shared and differential effects are modeled under a Dirichlet process mixture of Gaussians prior (Muller, 1996), and the edge probabilities are estimated by pooling information across experimental conditions, thereby resulting in the joint estimation of multiple brain networks.
- The role of the edge probabilities is twofold - they characterize uncertainty in network estimation and enable direct testing of shared and differential patterns across networks after multiplicity corrections.
- The connection strengths are encapsulated via network specific precision matrices, which are modeled separately for each network under a spike and slab Bayesian graphical lasso prior informed by the above edge probabilities.

Methods

Model

- The pre-whitened fMRI measurements for g-th experimental condition are modeled as $\mathbf{y}_{it}(g) \sim N_p(\mathbf{0}, \Omega_g^{-1})$, $i = 1, \dots, n, t = 1, \dots, T_{ig}, g = 1, \dots, G$, where

$$\pi(\Omega_g) = C_g^{-1} \prod_{k=1}^p E(\omega_{g,kk}; \frac{\alpha}{2}) \left\{ \prod_{k<l} w_{g,kl} N(\omega_{g,kl}; 0, \tau_{g,kl}^{-1}) + (1 - w_{g,kl}) DE(\omega_{g,kl}; \lambda_0) \right\} I(\Omega_g \in M^+),$$
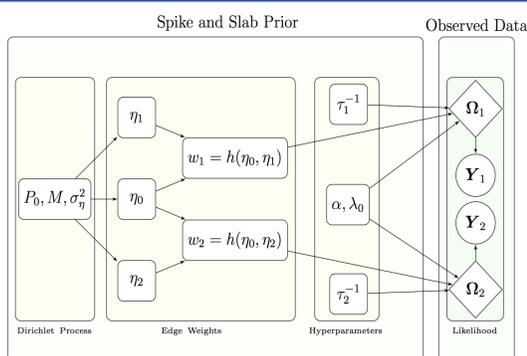
(1)

where $\pi(\cdot)$ denotes the prior distribution, $\omega_{g,kl}$ and $w_{g,kl}$ denote the strength and probability of the functional connection between nodes k and l for network $\mathcal{G}_g$ respectively, M^+ denotes the space of all positive definite matrices, $I(\cdot)$ denotes the indicator function, C_g is the intractable normalizing constant for the prior on the precision matrix, $N_p(\cdot; \mathbf{0}, \Sigma)$ denotes a p-variate Gaussian distribution with mean $\mathbf{0}$ and covariance Σ, and $E(\alpha)$ and $DE(\lambda)$ denote the exponential and double exponential distributions with scale parameters α^{-1} and λ^{-1} respectively.

- Information is pooled across experimental conditions to estimate the edge weights $w_{g,kl}$, $k \neq l, k, l = 1, \dots, p$, leading to joint estimation of multiple networks.
- By pooling information to model the edge probabilities instead of the edge strengths, we are able to jointly model multiple brain networks without constraining the edge strengths in separate networks to be similar.
- The prior weights represent the unknown probabilities of having functional connections, and are modeled via a parametric link function comprising unknown shared and differential effects as:

$$w_{g,kl} = h(\eta_{0,kl}, \eta_{g,kl}), \eta_{0,kl} \sim f, \eta_{g,kl} \sim f, f \sim P, P = DP(MP_0),$$

(2)

for $k \neq l, k, l = 1, \dots, p, g = 1, \dots, G$, where $h(\cdot)$ is the parametric link function relating the probability for edge (k, l) in network $\mathcal{G}_g$ to the network specific differential effect ($\eta_{g,kl}$) and common effect ($\eta_{0,kl}$) across all networks, and $DP(MP_0)$ denotes a Dirichlet process mixture prior defined by the precision parameter M and base measure $P_0 \equiv N(0, \sigma_\eta^2)$.

Parameter Diagram

Spike and Slab Prior | Observed Data

Dirichlet Process | Edge Weights | Hyperparameters | Likelihood

Acknowledgements

Reprints of this poster presentation are available at http://web1.sph.emory.edu/bios/CBIS

Simulation Study

Simulation Setup

We conducted a series of simulations to compare group level network estimation between BJNL and two popular penalized network estimation methods: the graphical lasso (GL) (Friedman, 2008), which estimates networks separately, and the Joint Graphical Lasso (JGL) (Danaher, 2014) under a fused lasso penalty, which pools information across graphs to jointly estimate multiple networks. We considered all combinations of the following simulation settings:

- Erdos-Renyi random graphs, small-world networks, scale-free networks
- 40 nodes, 100 nodes
- 25% similarity, 50% similarity, 75% similarity between graphs

For each combination of simulation settings, we generated data from 100 subjects with 300 time points each.
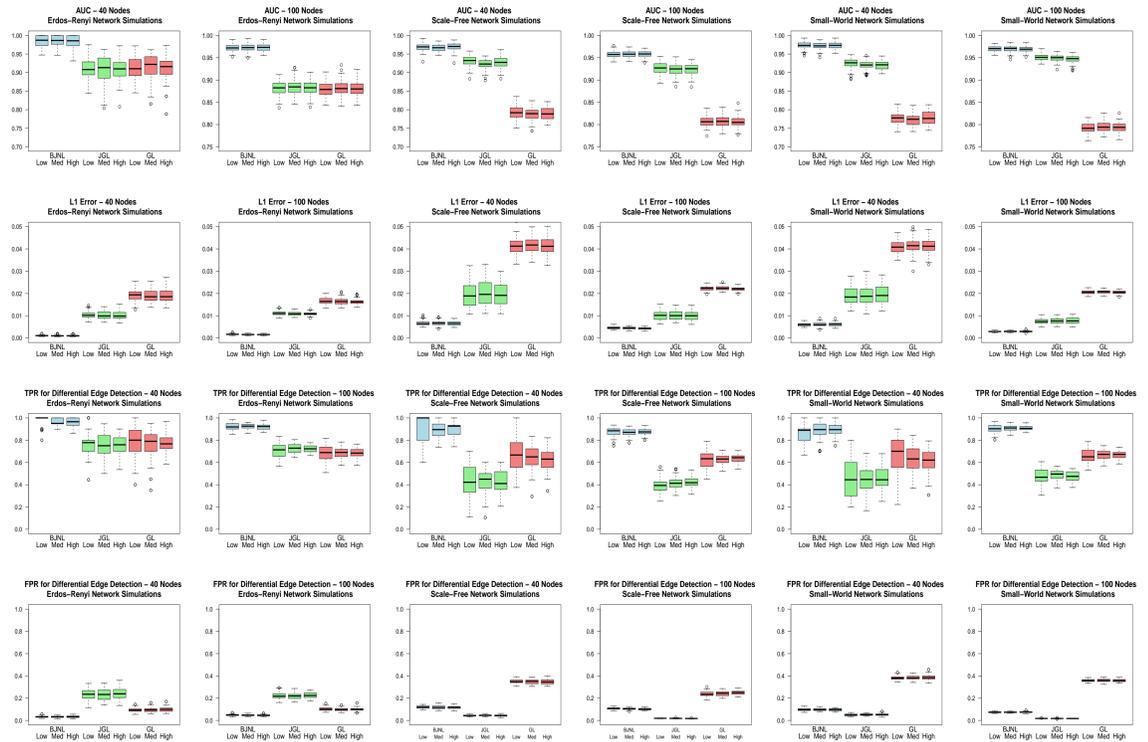
Simulation Results

Figure: Box plots of the AUC, L1 Error, and TPR/FPR for differential edge detection for the simulations. Within each approach, the box plots are organized as: low difference, medium difference, and high difference in edges between experimental conditions, in that order.

Stroop Task

Data

- We applied the proposed BJNL to a fMRI Stroop task study to investigate similarities and differences in the brain network under exertion and relaxed task performance.
- Data were in the AAL 90 node atlas.
- Functional module partition into nine resting state networks (Smith, 2009).
- 45 subjects, 100 time points per subject.
- Subjects were instructed to perform the Stroop task "with maximum effort" (EXR) and "as relaxed as possible" (RLX).
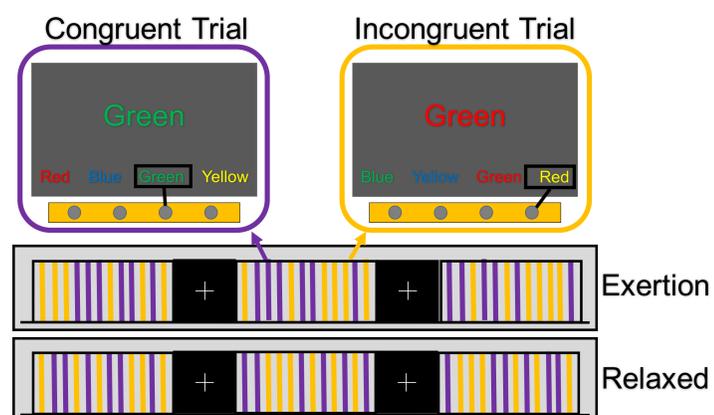

Figure: An illustration of the Stroop task involving task blocks of congruent and incongruent trials, indicated by purple bars and yellow bars respectively, and fixation blocks denoted by a centrally fixated cross.

Results

- T-tests ($p < 0.01$, FDR-corrected) of the Fisher's Z-transformed partial correlation differences at each MCMC iteration revealed 247 significantly different edges between the EXR and RLX conditions, 226 of which lay within our functional module partition.
- Circle plots of the sum of the strengths of the significant connections between modules are provided below separately for positive and negative connections

EXR - Positive | EXR – Negative | RLX– Positive | RLX– Negative
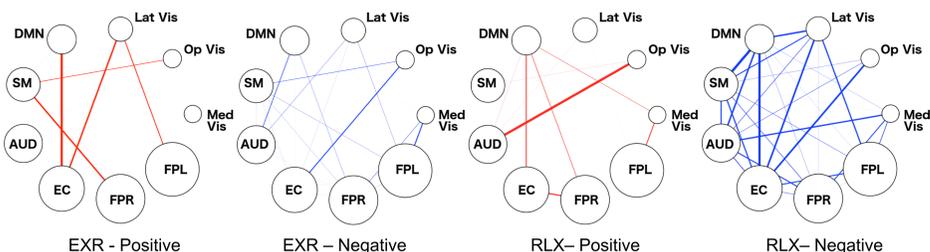
Interpretation

- BJNL revealed differences in the executive control and left frontal parietal networks.
- These networks are associated with high level cognitive function.
- Relaxed task performance featured significantly more negative connections between regions.

References

1. Muller, P., Erkanli, A., and West, M. (1996). Bayesian curve fitting using multivariate normal mixtures. Biometrika, 83(1):6779.
2. Danaher, P., Wang, P., and Witten, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 76(2):373397.
3. Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics, 9(3):432441.
4. Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., et al. (2009). Correspondence of the brains functional architecture during activation and rest. Proceedings of the National Academy of Sciences, 106(31):1304013045.