A Constrained, Weighted-$\ell 1$ Minimization Approach for Joint Discovery of Heterogeneous Neural Connectivity Graphs

Chandan Singh[1]; Beilun Wang[2]; Yanjun Qi[2]
[1]University of California, Berkeley, [2]University of Virginia

Abstract

Determining functional brain connectivity is crucial to understanding the brain and neural differences underlying disorders, such as autism. Recent studies have used Gaussian graphical models to learn brain connectivity via statistical dependencies across brain regions from neuroimaging. However, previous studies often fail to properly incorporate priors tailored to neuroscience, such as preferring shorter connections. To remedy this problem, the paper here introduces a novel, weighted-$\ell 1$, multi-task graphical model (W-SIMULE).

W-SIMULE elegantly incorporates a flexible prior, along with a parallelizable formulation. Additionally, W-SIMULE extends the often-used Gaussian assumption, leading to considerable performance increases in applications to fMRI data. Due to its elegant domain adaptivity, W-SIMULE can be readily applied to various data types to effectively estimate connectivity.

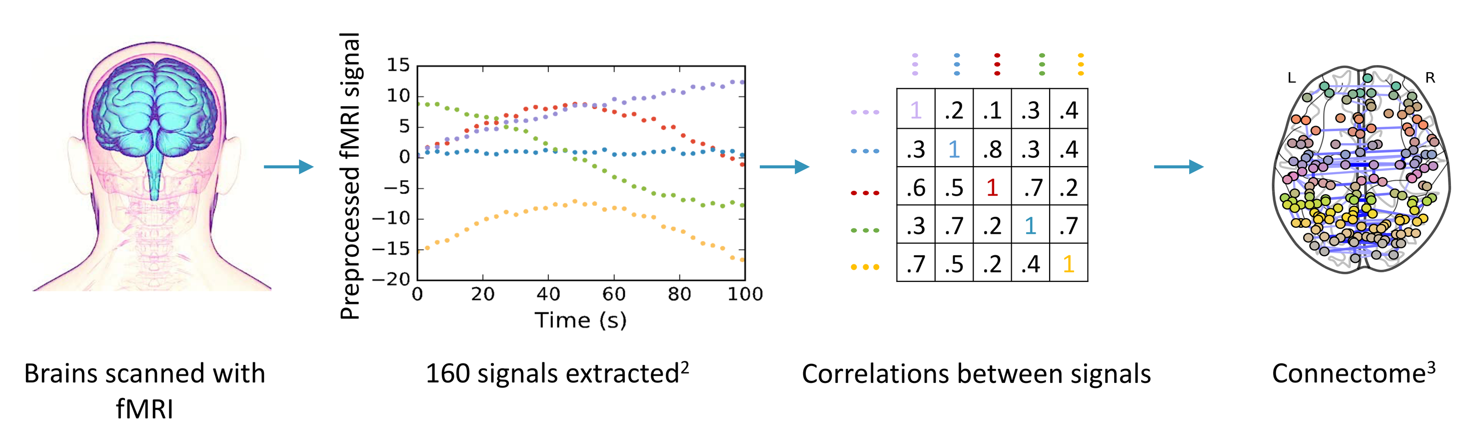


Fig 1. Data pipeline for constructing a functional connectome.[2]

Introduction

Recently, there has been great interest in mapping the interactions between brain regions, a field known as functional connectomics. We focus on the important problem of estimating brain connectivity for more than one group (i.e. a disease group and a control group). This study's main contribution is the novel formulation of W-SIMULE, which arises naturally from brain-imaging data. W-SIMULE is a weighted-ℓ_1, multi-task graphical model which robustly estimates the precision for each group. Its main advantages are:

1. **Effectiveness.** it yields accurate connectivity in terms of log-likelihood and classification accuracy on the ABIDE resting-state fMRI dataset[1]
2. **Domain adaptivity**. it elegantly enforces a prior based on the problem at hand and can overcome the often-incorrect Gaussian assumption by using nonparanormality
3. **Interpretability**. it calculates a connectome for each group which can be tuned to the desired sparsity level and is particularly effective at low sparsity levels
4. **Efficiency.** the formulation is column-wise parallelizable and quickly solvable

Method: W-SIMULE

The problem of determining functional brain connectivity concerns using the covariance matrix (Σ) to calculate the precision matrix Ω), which represents conditional correlations between brain areas. To do this, W-SIMULE uses four properties:

1. Sparsity
2. Multi-task learning with K groups
3. A prior matrix of positive weights W
4. A nonparanormal assumption

Combining these elements yields the novel formulation of W-SIMULE :

$$\widehat{\Omega}_I^{(1)}, \dots, \widehat{\Omega}_I^{(K)}, \widehat{\Omega}_S = \sum_i \operatorname{argmin} \left\| W \cdot \Omega_I^{(i)} \right\|_1 + \epsilon K \left\| W \cdot \Omega_S \right\|_1$$

$$\text{Subject to: } \left\| \Sigma_N^{(i)} \left(\Omega_I^{(i)} + \Omega_S \right) - I \right\|_\infty \le \lambda, i = 1{:}K$$

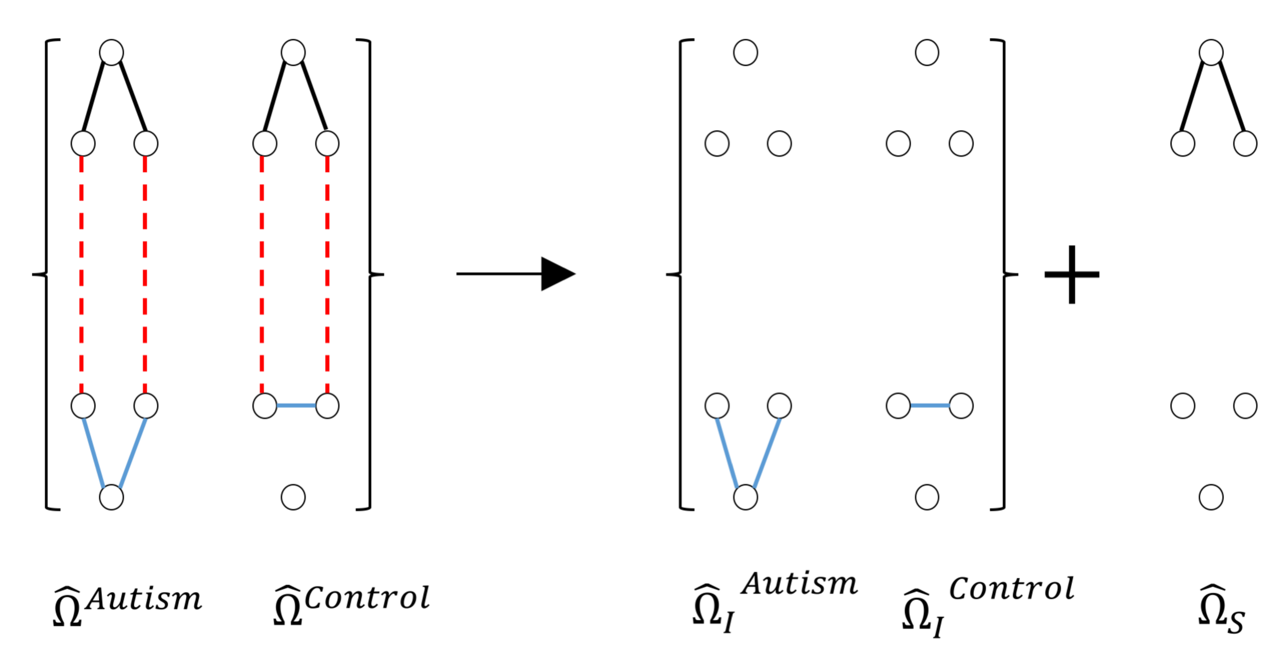


Fig 2. Toy example depicting W-SIMULE . Left shows potential edges present in the data and right shows learned edges. Long edges (red) are spatially penalized and discarded, edges that differ between groups (blue) are learned individually, and edges shared between groups (black) are learned in $\widehat{\boldsymbol{\Omega}}_{\boldsymbol{S}}$.

W-SIMULE has three hyperparameters:

1. W – enforces a different prior or changes how strictly it is enforced
2. λ - controls the total sparsity of the resulting precision matrices
3. ε – controls how strictly the group penalty is imposed

W-SIMULE	CLIME[4]	GLASSO[5]	SIMULE[6]	JGL (fused)[7]	SIMONE[8]
58.62	46.55	53.71	57.96	56.90	53.71

Table 1. Classification accuracy obtained on the ABIDE dataset using various methods.

Results

Enforcing the prior.[9] W-SIMULE effectively enforces the prior (Fig 3A). As the *dist* prior is raised to a higher power (thereby increasing the spread of the prior weights), the prior is more strictly enforced, resulting in a lower average edge length at every sparsity level. The "optimal" line shows the lowest possible average edge length as a function of graph sparsity.

Log-likelihood. Connectomes are generated for various sparsity levels and their resulting log-likelihoods are plotted in Fig 3B. W-SIMULE (*dist² prior)* outperforms all of the relevant baselines, especially for very sparse connectomes, which are the most interpretable.

Classification Accuracy. Table 1 displays the maximum accuracy achieved for each baseline, after sweeping over hyperparameters. W-SIMULE (*W=anatomical*) yields a classification accuracy of 58.62%.

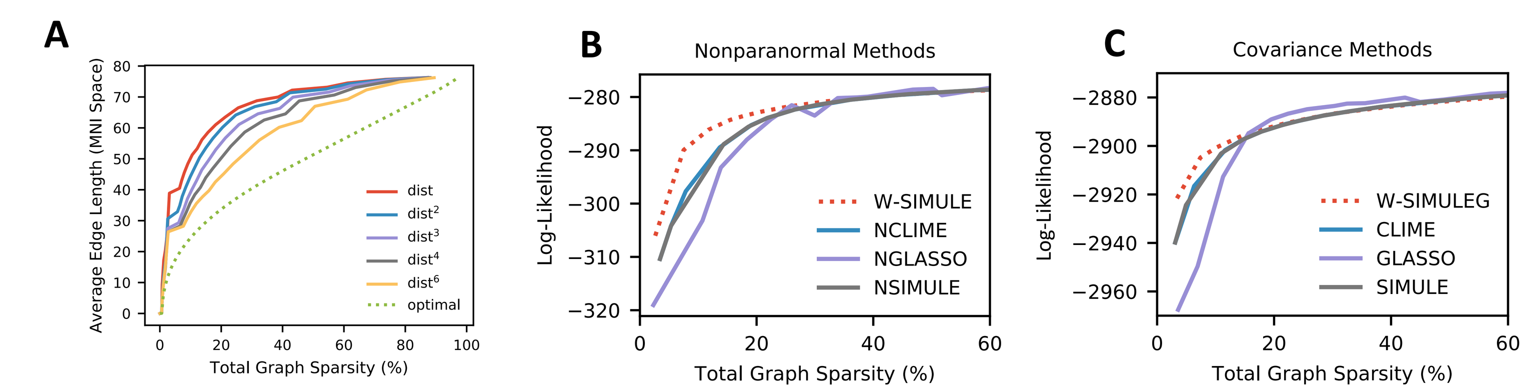
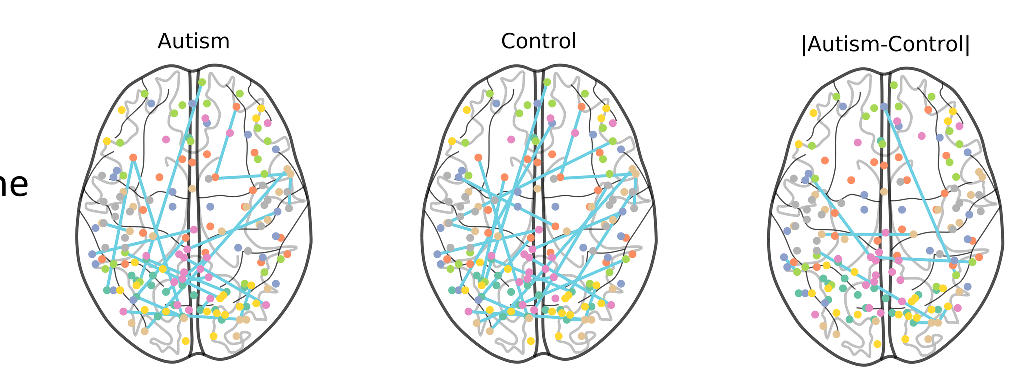


Fig 3. Connectivity results. **A** shows that W-SIMULE effectively enforces priors that penalizes long edges. **B & C** show that W-SIMULE can maximize the log-likelihood more effectively than other methods..

Fig 4. Sparse connectome generated by W-SIMULE with the anatomical prior. The autism graph, control graph, and their difference are shown.



Conclusions / Future Work

W-SIMULE great potential for future applications. As brain-imaging datasets become more complex and include more structural data coupled with functional data, W-SIMULE will become increasingly important to neuroscience. This is especially true for studies with small sample sizes, such as task-specific studies, which require strong priors and multi-task learning in order to robustly determine connectivity. As the spatial resolution of fMRI increases, spatial penalization will become more important in constructing accurate ROIs and brain connections.

Many problems outside of neuroscience can benefit from W-SIMULE; it can utilize diverse priors to find conditional independence between nodes in any multi-task setting. Thus, W-SIMULE can be readily applied to gene-network estimation, image processing (where physical distance could be used as a prior in images), and many other problems that currently utilize Gaussian graphical models.

References

1. Di Martino, A.; Yan, C.-G.; Li, Q.; Denio, E.; Castellanos, F. X.; Alaerts, K.; Anderson, J. S.; Assaf, M.; Bookheimer, S. Y.; Dapretto, M.; et al. 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry* 19(6):659–667.
2. Preprocessing conducted by ABIDE preprocessed connectomes project: Craddock, C. 2014. Preprocessed connectomes project: open sharing of preprocessed neuroimaging data and derivatives. In *61st Annual Meeting*. AACAP.
3. Connectomes visualized with nilearn: Abraham, A.; Pedregosa, F.; Eickenberg, M.; Gervais, P.; Muller, A.; Kossaifi, J.; Gramfort, A.; Thirion, B.; and Varoquaux, G. 2014. Machine learning for neuroimaging with scikit-learn. *arXiv preprint arXiv:1412.3919*.
4. Cai, T.; Liu, W.; and Luo, X. 2011. A constrained 1 minimiza- tion approach to sparse precision matrix estimation. *Journal of the American Statistical Association* 106(494):594–607.
5. Friedman, J.; Hastie, T.; and Tibshirani, R. 2008. Sparse in- verse covariance estimation with the graphical lasso. *Biostatistics* 9(3):432–441.
6. Wang, B.; Singh, R.; and Qi, Y. 2016. A constrained l1 minimiza- tion approach for estimating multiple sparse gaussian or nonparanormal graphical models. *arXiv preprint arXiv:1605.03468*.
7. 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 (Statisti- cal Methodology)* 76(2):373–397.
8. Chiquet, J.; Grandvalet, Y.; and Ambroise, C. 2011. Inferring mul- tiple graphical structures. *Statistics and Computing* 21(4):537–553.
9. Dosenbach, N. U.; Nardos, B.; Cohen, A. L.; Fair, D. A.; Power, J. D.; Church, J. A.; Nelson, S. M.; Wig, G. S.; Vogel, A. C.; Lessov-Schlaggar, C. N.; et al. 2010. Prediction of individual brain maturity using fmri.