Role of Brain Networks in Imaging Genetics

Suprateek Kundu
Emory University

October 1, 2015

Outline

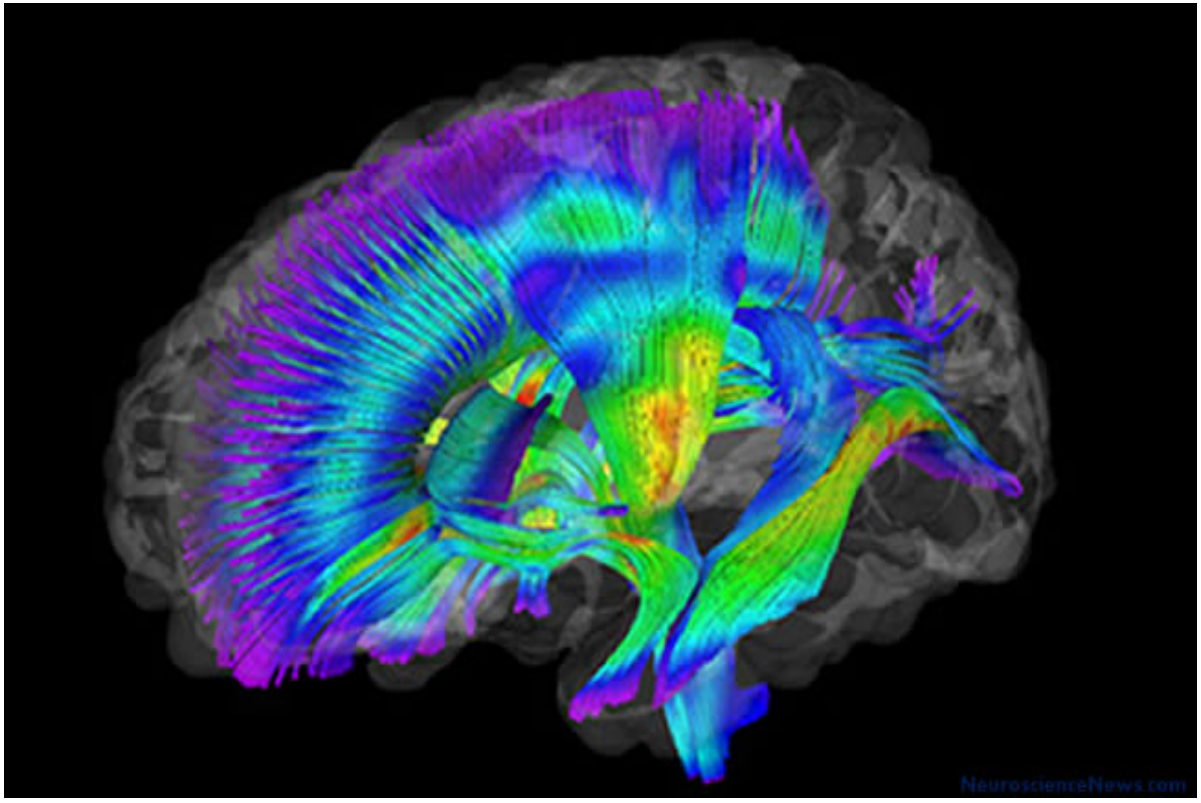
- The nervous system can be described as a network of interconnected neurons
- Modern noninvasive imaging techniques applied to the human brain allow the mapping of anatomical regions
- Brain networks describes the inter-connections between different regions of the brain
- Reasonable to believe that different parts of the brain work together to drive phenotypic outcomes. Understanding brain networks will help us understand the dynamics of the brain.

Outline

- There can be two types of connections - structural and functional
- Structural connectivity focuses on how different regions in the brain are structurally connected (e.g. white matter tracts)
- Functional connectivity describes statistical patterns of dynamic interactions among regions - also called "functional networks"
- Functional networks can be measured when (a) brain is endogenously active (or at rest) as well as in the course of stimulus; & (b) task-evoked perturbations.

Functional Connectivity

Structural Connectivity



Introduction

- Imaging Genetics $\rightarrow$ integrating imaging, genetic & psychological data to answer challenging questions in neurodevelopmental research
- Question of interest $\rightarrow$ (i) how do different brain regions interact in healthy and diseased individuals ? & (ii) what are the important confounders driving such connections ?
- E.g. AD group has reduced functional connectivity compared to MCI group.
- Characterizing brain networks can provide tools for early detection of individuals with a high risk for mental disorders.

Aims

- I will focus on conditional graphical models which :
  (1) estimates the brain network after accounting for covariates;
  (2) infers significant covariates which influence the brain network & the imaging outcome;
  (3) discovers groups of ROIs working together to drive different brain functions, & the interactions among them

- Methods applicable to diverse imaging outcomes (e.g. PET, MRI, brain vol) & scalable to high resolution images

- Our approach reports population brain networks, and not individual specific networks

ROI = region of interest

Brain Networks & Graphical Models

- Brain networks can be estimated using graphical modeling approaches based on partial corr (e.g. SICE, BICE)
- Provides an efficient way to estimate the full set of direct functional connections between all pairs of nodes
- Recognized as one of the most successful approaches for modeling brain functional connectivity (Smith et. al, 2011)
- Graphical modeling is one of the important ongoing research areas in statistics

Role of Confounders

- Promising to integrate supplementary information to reliably estimate the brain network & build better predictive models.
- Makes comparisons more meaningful across disease groups
- Can help us infer which confounders influence which ROIs & the connections between these ROIs
- Calls for new graphical modeling tools integrating covariate information

Graphs

- A graph is defined by the vertex set $V = \{1, \ldots, p\}$ indexing the set of random variables $(X_1, \ldots, X_p)$, and the edge set E.

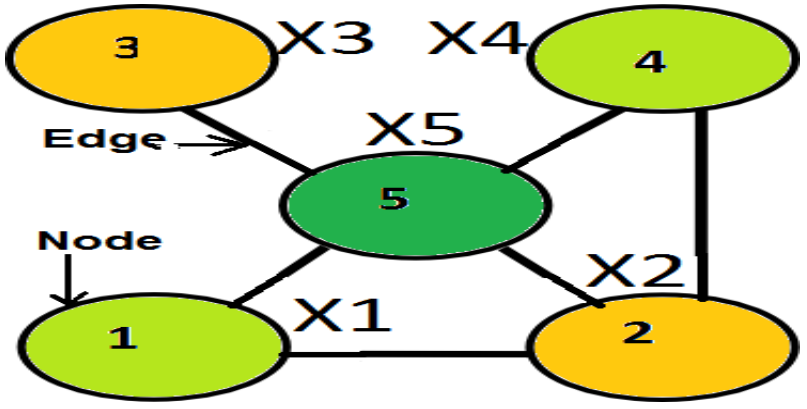


Figure: Visual depiction of an undirected graph. Circles represent nodes and straight lines represent edges.

- Directed graphs/ Bayesian networks & undirected graphs
- An undirected graph $G = (V, E)$ encapsulates a set of conditional dependency relationships dictated by edge set E

Gaussian Graphical Model (GGM)

- Goal of a graphical model is to propose a probabilistic framework for estimating significant edges, and also to possibly estimate the strength of connection via partial corr.
- Definition: A GGM refers to a set of continuous random variables specified by a Gaussian distribution, defined conditional on the graph $G = (V, E)$, with missing edges in E corresponding to zeros on the inverse covariance matrix.
- That is, GGMs specify

$$\boxed{\mathbf{x} \sim N(0, \Sigma_G), \text{ with } \Sigma_G^{-1}(i,j) = 0 \text{ for all pairs } (i,j) \notin E}.$$

- Both Σ, G unknown and need to be estimated. Sometimes Σ treated as a nuisance parameters, and only G is estimated.

Gaussian Graphical Model (GGM)

- For our example, $\Sigma^{-1} = \begin{pmatrix} \omega_{11} & \omega_{12} & 0 & 0 & \omega_{15} \\ \omega_{21} & \omega_{22} & 0 & \omega_{24} & \omega_{25} \\ 0 & 0 & \omega_{33} & 0 & \omega_{35} \\ 0 & \omega_{42} & 0 & \omega_{44} & \omega_{45} \\ \omega_{51} & \omega_{52} & \omega_{53} & \omega_{54} & \omega_{55} \end{pmatrix}$.
- Note that $\Sigma_G \in M^+(G) =$ set of all symmetric P.D. matrices having $\Sigma_G^{-1}(i,j) = 0$ for all $(i,j) \notin E$.
- Graph estimation $\equiv$ determination of structural 0s in Σ^{-1}.

Edges

- What do edges in the graph imply ?
- Absence of an edge between a and b means that X_a is conditionally independent of X_b given all other variables.
- Presence of an edge between a and b implies a significant partial corr between X_a & X_b.

Frequentist Approaches for GGMs

- Typically rely on neighborhood selection techniques (Meinshausen, 2006; Peng et. al, 2009) or penalized maximum likelihood approaches (Yuan and Lin, 2007).
- Advantageous in being computationally feasible for high dimensional graphs, and often have theoretical justifications.
- However limitations include
  (i) frequentist point estimates can be unstable in high dimensions;
  (ii) difficult to report finite sample measures of uncertainty.
- Measures of uncertainty important in imaging studies, to address inherent heterogeneity.

Bayesian Approaches for GGMs

- Bayesian approaches specify a prior on unknown quantities and use posterior distributions to draw inferences.

$$\underline{\text{Model:}}\ \mathbf{x}_i \sim N(0, \Sigma_G),\ \underline{\text{Prior:}}\ \Sigma_G \sim \pi(\Sigma \mid G), G \sim \pi(G)$$

- Uses MCMC to sequentially draw samples from the conditional distributions: $\Sigma_G \sim \pi(\Sigma_G \mid G, X)$ & $G \sim \pi(G \mid X, \Sigma_G)$ over several iterations.

- Goal is to approx the posterior distributions $\pi(\Sigma \mid X)$ & $\pi(G \mid X)$ & obtain posterior summaries using MCMC samples.

- Bayesian approaches provide measures of uncertainty (e.g. credible intervals, posterior prob), which are important in addressing heterogeneity in imaging studies.

Conditional Graphical Models

- Conditional graphical models estimate associations after accounting for possible covariate effects.

$$\mathbf{x}_i = \mathbf{z}_i B + \boldsymbol{\epsilon}_i, \boldsymbol{\epsilon} \sim N(0, \Sigma_G),\ \Sigma_G \sim \pi(\Sigma \mid G), G \sim \pi(G), B \sim \pi(B)$$

- $\mathbf{x}_i$ is $1 \times p$ dim outcome, $\mathbf{z}_i$ is $1 \times q$ dim covariate, and B is $q \times p$ dim matrix of regression coeff.

- (k,l)-th element of B quantifies the effect of the k-th covariate on the l-th outcome measurement.

- Challenging to accurately estimate the large number of parameters in B and G simultaneously.

Novelties of Proposed Approach

- We propose a new conditional graphical model which uses patterns in coeff matrix B to simplify graph estimation.
- One of the first approaches in imaging genetics, for identifying true associations in the brain after accounting for confounders.
- Specifies a mixture prior on the columns of B to identify clusters of nodes related to covariates by similar magnitudes.
- Clusters represent functional modules $=$ inter-connected groups of ROIs working together to drive brain functions.
- Uses these clusters to develop novel graphical priors suitable for large scale problems involving high res imaging outcomes.

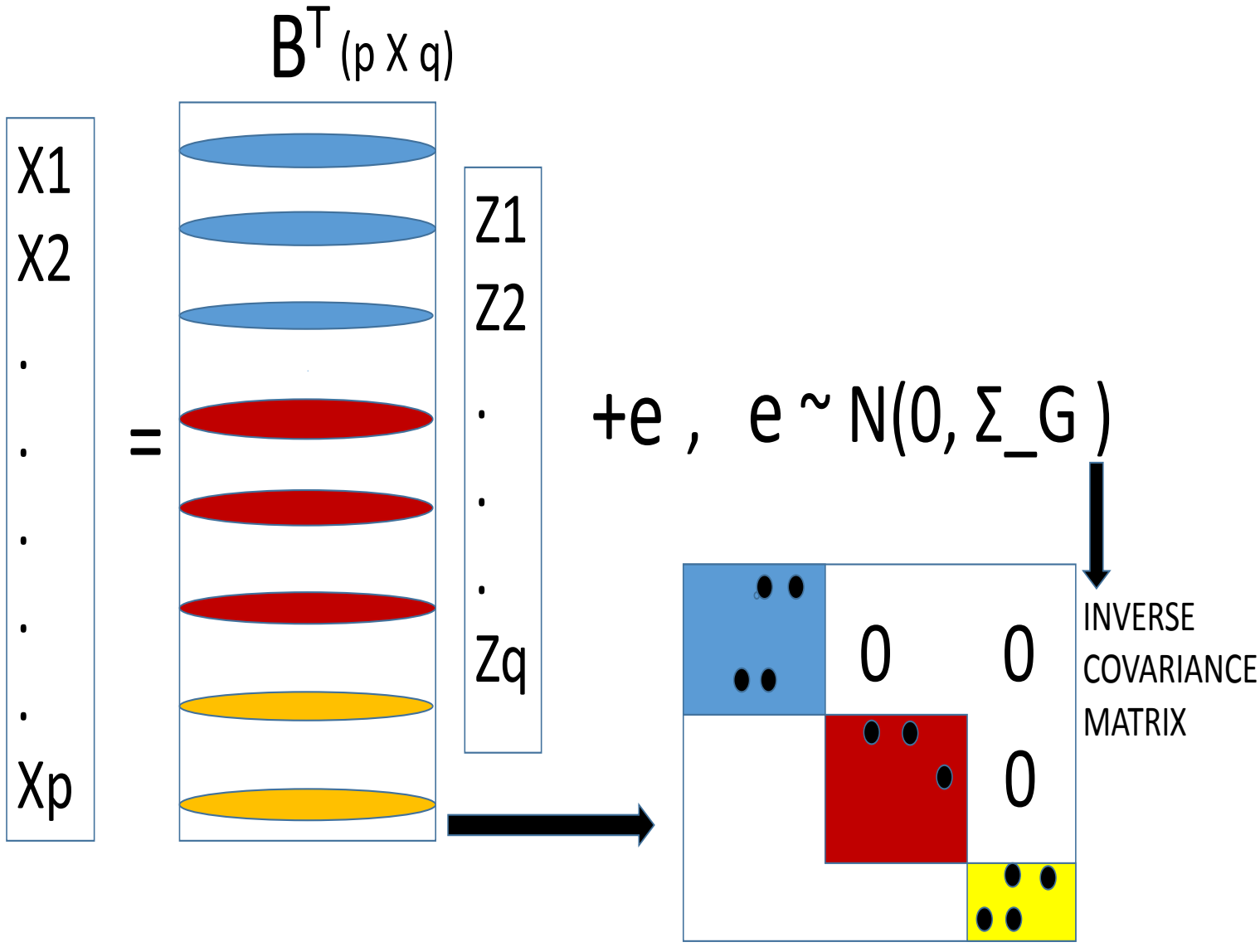


Figure: Schematic representation of the proposed model.

Model Features

- *Assumption 1:* X_a, X_b belonging to the same cluster are related to each covariate by the same magnitude. X_a & X_b may or may not be connected
- *Assumption 2:* $X_a \perp X_b$ given covariate info where a & b belong to distinct clusters
- In other words, the association between distinct clusters of nodes can be fully explained by covariate information
- The number of clusters & their sizes are random
- Also provides a regression set-up involving covariates which can be used for variable selection & prediction

Formal Notations

- Model

$$\mathbf{x}_i = \mathbf{z}_i B + \boldsymbol{\epsilon}_i, \boldsymbol{\epsilon} \sim N(0, \Sigma_G),\ \Sigma_G \sim \pi(\Sigma \mid G), G \sim \pi(G), B \sim \pi(B)$$

- The columns of $B = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p)$ are clustered under the prior:

$$\boldsymbol{\beta}_k \sim \sum_{j=1}^{\infty} w_j \delta_{\boldsymbol{\eta}_j},\ \boldsymbol{\eta}_j \sim \prod_{l=1}^{q} \text{Laplace}(\eta_{jl}; \lambda_j),$$

- Blk diag structure on Σ via a novel graphical prior:

$$e(k,l) \sim \text{Ber}(\omega) 1(\cup_{h=1}^{H}(k \in S_h, l \in S_h)) + \delta_0 1(k \in S_h, l \in S_{h'}, h \neq h'),$$

- Here $e(k,l)$ is the binary edge inclusion indicator, δ_0 denotes a point mass at 0, & $S_1, \ldots, S_H$ denote clusters
- Given G, $\Sigma_G \sim \pi(\Sigma \mid G)$ is specified using hyper inverse-Wishart
- Posterior computation proceeds via a combination of fully Gibbs & reversible jump moves

Analysis of ADNI Data

- ADNI 1 collected longitudinal FDG-PET scans at multiple times points across different imaging sites
- We look at the PET scans at baseline for 50 AD & 121 MCI patients, & 71 healthy controls (HC)
- We fit our model separately to the three groups & compare the functional brain networks after accounting for strongly associated covariates

Goals

- Identify important connections in the functional brain network after accounting for age, gender, handedness, weight, and genes
- Identify communities or collections of ROIs in the brain which work together to drive brain functions
- Identify important genes influencing the imaging phenotype and the sub-networks for these communities
- Compare the above findings for AD, MCI & HC groups

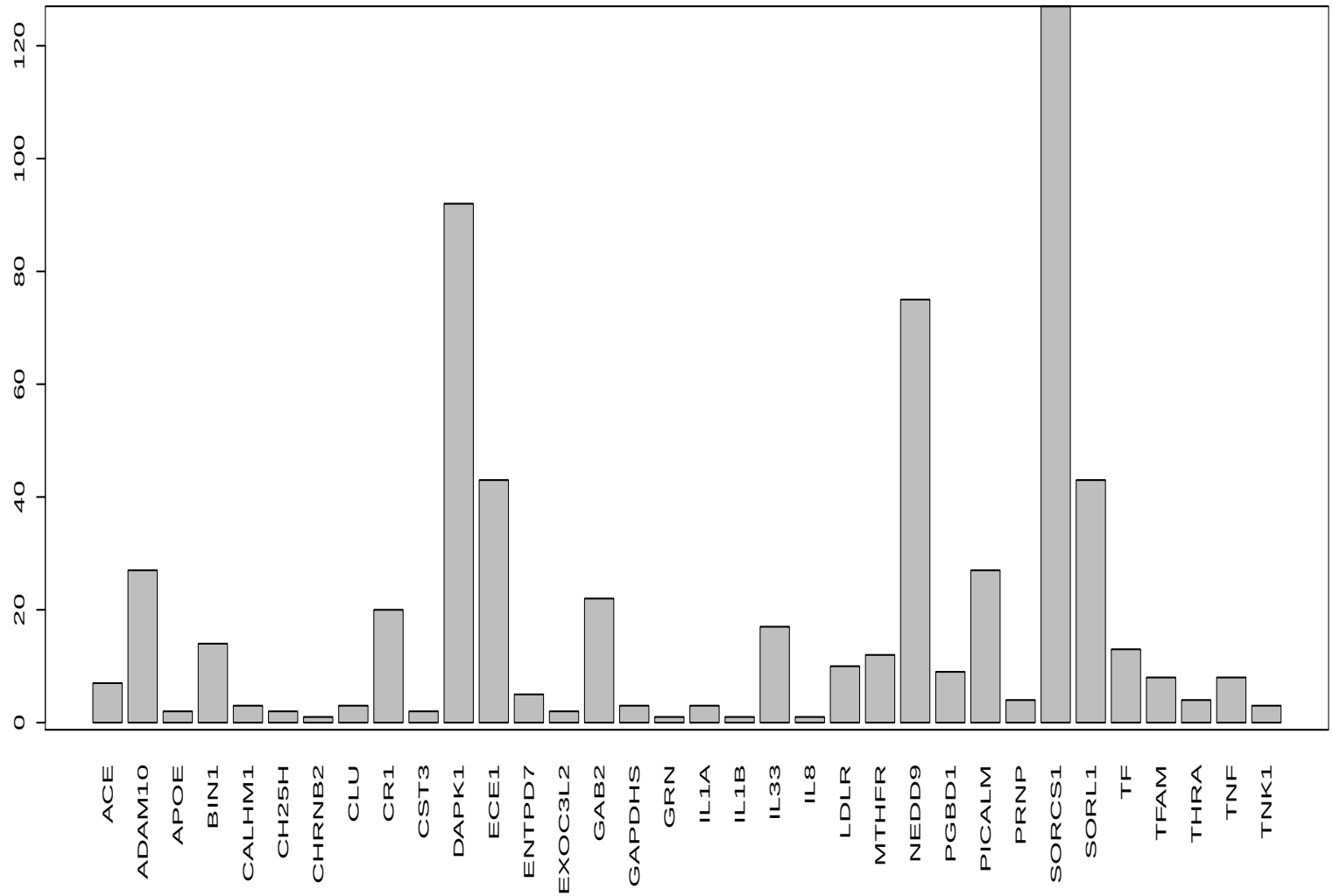
Pre-processing Steps - PET

- Standard pre-processing steps including co-registration, normalization and spatial smoothing (8 mm FWHM) were applied to the PET dataset
- We consider 90 brain regions that are defined according to the automated anatomical labeling (AAL) system.
- Computed the PET regional summaries using the first principal component scores over all voxels with each region (Bowman et al, 2012)
- Although we work with the AAL atlas involving 90 ROIs, we can accommodate higher resolution images as well

Pre-processing Steps - SNPs

- SNP data in the ADNI study were genotyped using the Human 610-Quad BeadChip
- Only focused on SNPs that belongs to top 40 candidate genes reported in the AlzGene database (www.alzgene.org)
- Removed the SNPs (a) with more than 1% missing values; (b) minor allele frequency (MAF) less than 5% and (c) the Hardy-Weinberg Equilibrium (HWE) p-value less than 10^{-6}
- The final dataset includes 614 SNPs(0/1) on 37 genes. We also included age, gender, handedness & weight

Top 37 genes & the number of SNPs per gene.



Estimated brian network

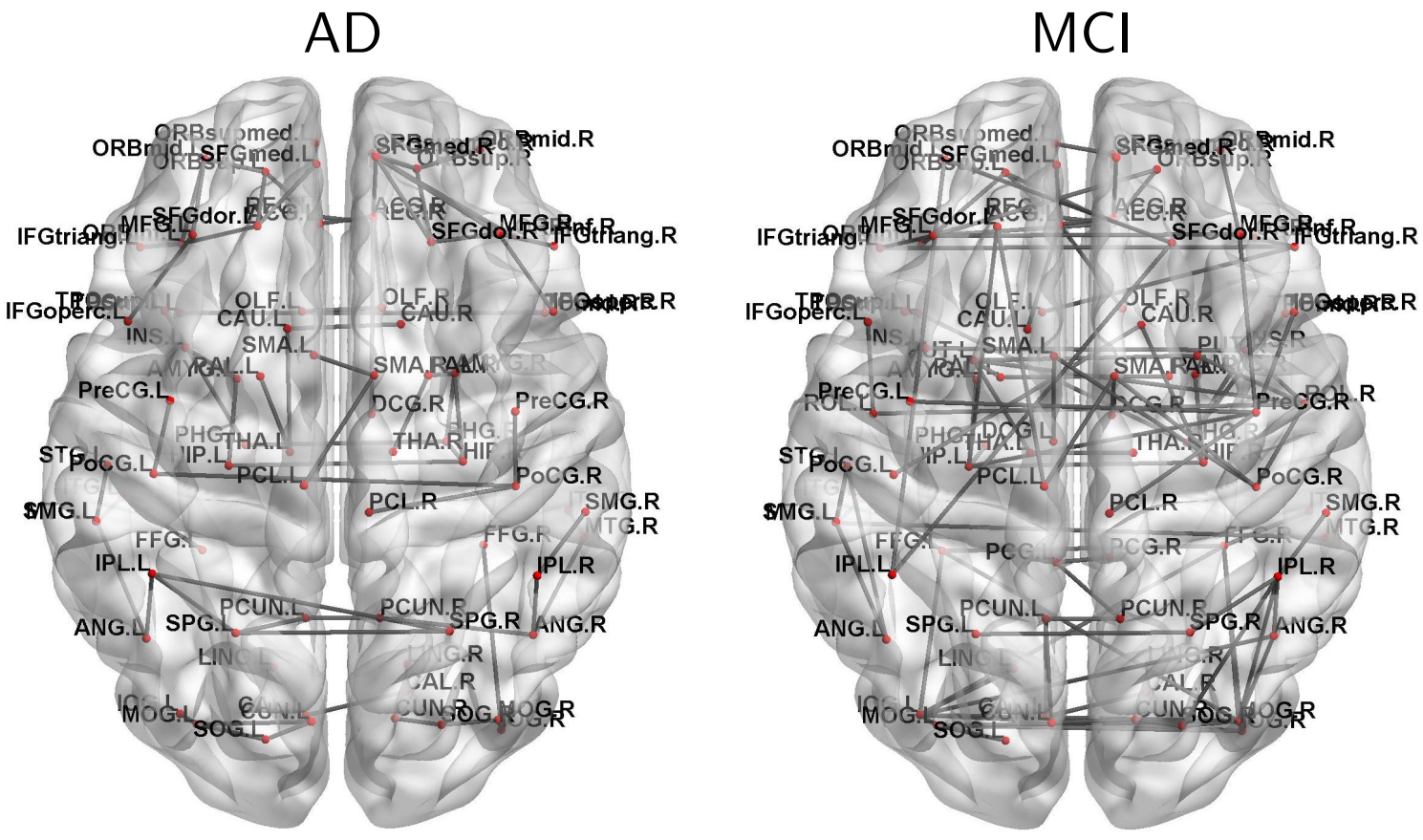


Figure: Functional brain network estimation for AD, MCI groups

Estimated brian network

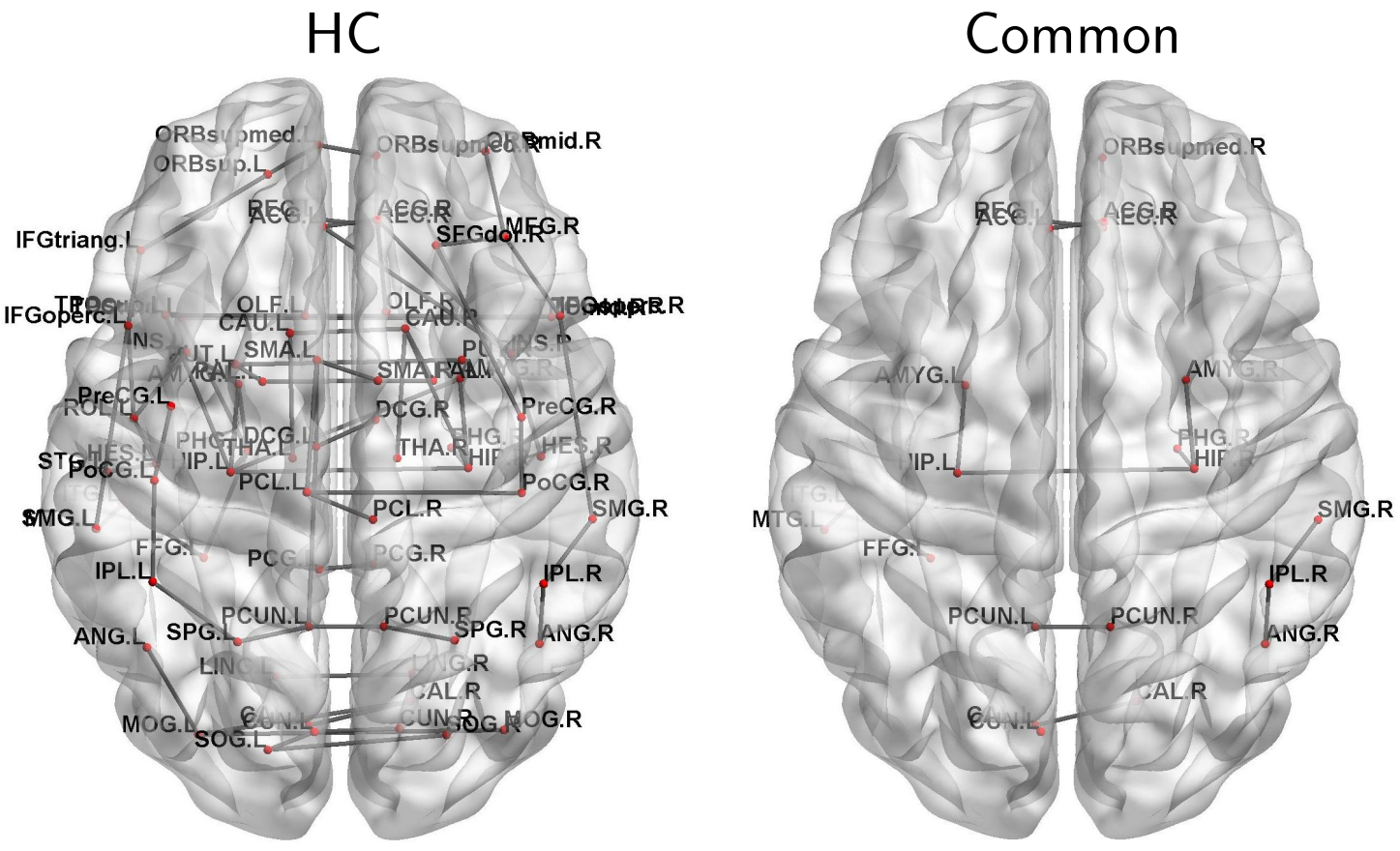


Figure: Functional brain network estimation for the HC group and the common edges shared by the three networks

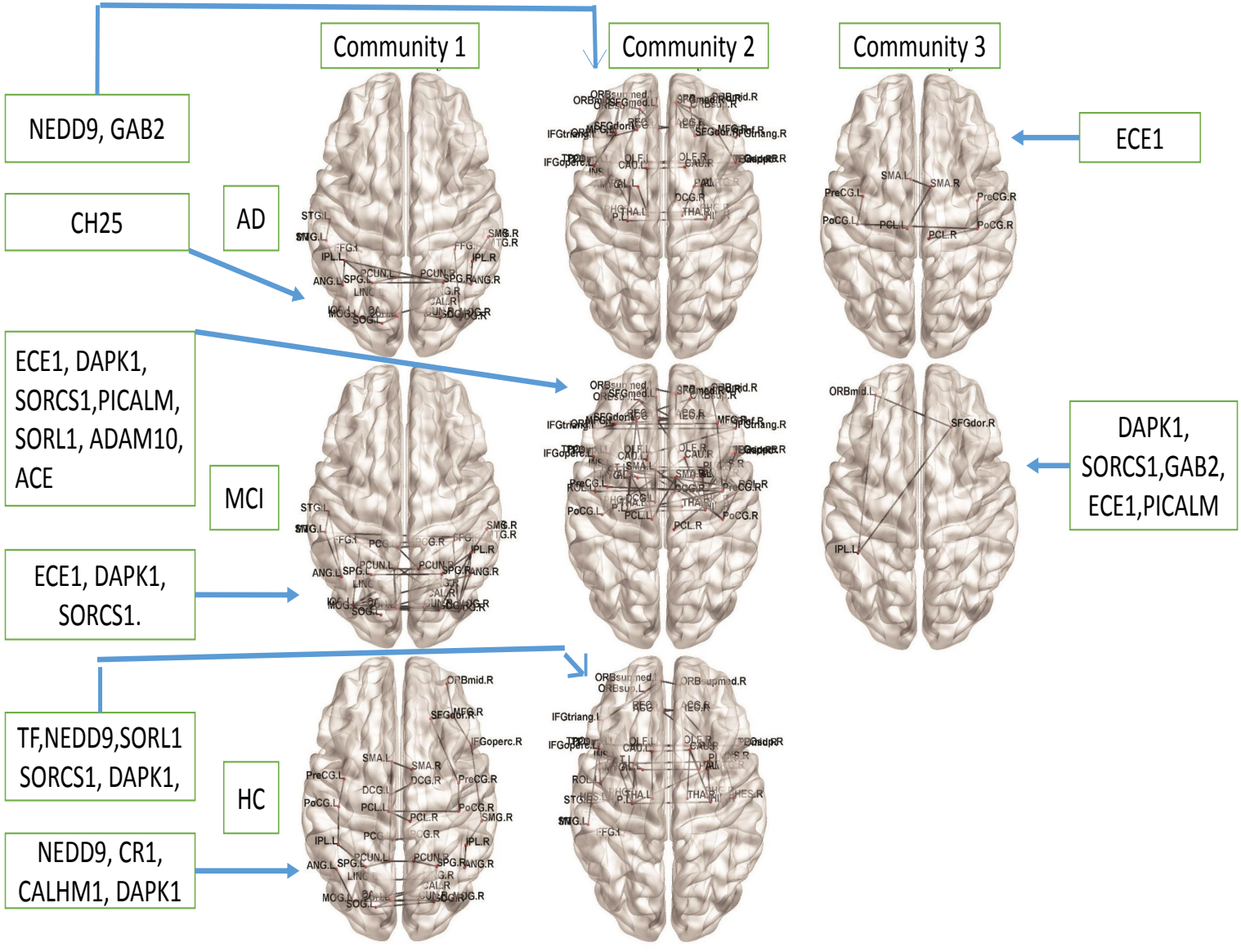


Figure: Sub-network community detection for AD, MCI & HC groups

Findings - brain networks

- AD, MCI and HC networks have 79, 102 and 73 important edges, respectively. AD has fewer connections within and between the two hemispheres compared to the MCI
- There are 14 edges shared by all the three groups
- Functional connections between left & right Precuneus (self-consciousness) appear in all 3 networks. Implies that AD & MCI subjects have the similar functional activities between the two Precuneus region as the HC subjects.
- All 3 networks contain the edges between the left and right Hippocampus (HIP)
- HIP has only 2 connections in AD compared to 6 in MCI, possibly pointing towards a known damage in HIP in AD patients

Findings - brain networks

- We identify 3, 3 & 2 sub network communities in the AD, MCI & HC group resp
- AD and MCI have two similar communities, while the HC sub network communities are quite different
- Postcentral gyrus (PoCG), Precentral gyrus (PreCG), Paracentral lobule (PCL) & Supplementary motor area (SMA), in community 3 have fewer connections in AD
- These ROIs are mainly related to the motor skill & sense of touch - potentially implies reduced motor skills and sense of touch for AD
- Increased connectivity in HC subjects

Findings - genetic effects

- SNP "rs2018334" on NEDD9 is significantly associated with the sub-network community 2 in AD
- We conjecture that it might be an important factor that separates regions HIP & PHC away from others
- GAB2 is an important gene for both AD & MCI - prior evidence of it conferring increased risk for sporadic AD
- CH25H was associated with AD, but not MCI or HC. Known to be associated with amyloid-β metabolism

Findings - genetic effects

- Genes which promote MCI but not related to HC include
  (a) ECE1 which is associated with cognitive ability in elderly individuals and disease risk,
  (b) ADAM-10 & PICALM, which regulate amyloid-β, with the latter being associated with late onset AD
- SORL1 related to MCI & HC - known to be a potential tool for identifying MCI subjects at high risk of conversion to AD
- Age is also found to be negatively related to community 1 in MCI

Simulation

- We fit our model to the PET data for 121 individuals with mild cognitive impairment (MCI) obtained from the ADNI dataset, and then use the fitted model to simulate data.
- The dataset in question contains PET measurements recorded from $p = 42$ regions of interest (ROIs) in the brain, with supplementary data on $q = 546$ SNPs having levels 0,1,2.
- These SNPs were chosen by ranking p-values obtained via univariate analyses, and then dichotomized.
- This fitted model has 670 non-zero regression coefficients and 140 edges.
- Corresponds to a high dimensional setting, making inferences challenging.

Simulation

- Compared our approach with (a) Bayesian lasso (BLASSO) applied independently for p univariate regressions; (b) frequentist Lasso, applied independently for p univariate regressions; (c) sparse seemingly unrelated regressions or SSUR, which is an alternate conditional graphical modeling approach; (d) hyper inverse-Wishart (HIW) which is a graphical modeling approach without incorporating covariates.

- Comparison criteria (a) out of sample mean square error (MSE);
  (b) L_2 error in estimating the true regression coefficients ($||\hat{\beta}||^2_{L_2}$);
  (c) sensitivity corresponding to a false discovery rate of 1%;
  (d) area under the ROC curve for variable selection;
  (e) L_1 error in estimating the true precision matrix ($||\hat{\Omega}||_{L_1}$); &
  (f) BIC.

Simulation

Table: Simulation results.

Method	MSE	$\|\|\hat{\beta}\|\|_{L_2}^2$	SE_G(1% FDR)	AUC(var)	$\|\|\hat{\Omega}\|\|_{L_1}$	BIC
spHIW	0.01	0.0015	0.47	0.94	0.17	9.1
BLASSO	0.02	0.0028	NA	0.51	NA	51.61
SSUR	0.02	0.0031	0.45	0.78	0.97	7.92
LASSO	0.04	0.0006	NA	0.59	NA	6.13
HIW	NA	NA	0.47	NA	0.58	NA

Simulation - Conclusions

- spHIW does comparably to BLASSO in terms of variable selection, however BLASSO has a much higher AIC, BIC;
- does a better job in estimating the true inverse covariance matrix compared to HIW, which compensates for lack of covariate information by reporting large conditional variances
- SSUR often performs poorly in terms of variable selection by reporting overly large models (AUC ≈ 0.5), which also results in poor precision matrix estimation
- accurate variable selection is an important factor in the estimation of conditional associations
- spHIW maintains a desired balance between our dual goals of variable selection & graphical model estimation, while also preserving parsimony (low BIC)

Thank you !