

Alzheimer's Analysis with a Copula-Generated Random Graph Model

Dr. Aaron J. Danielson

Department of Statistics and Actuarial Science
Simon Fraser University



A Network with Directed edges

Nodes are **regions in the brain** corresponding to where resting-state fMRI data was collected.

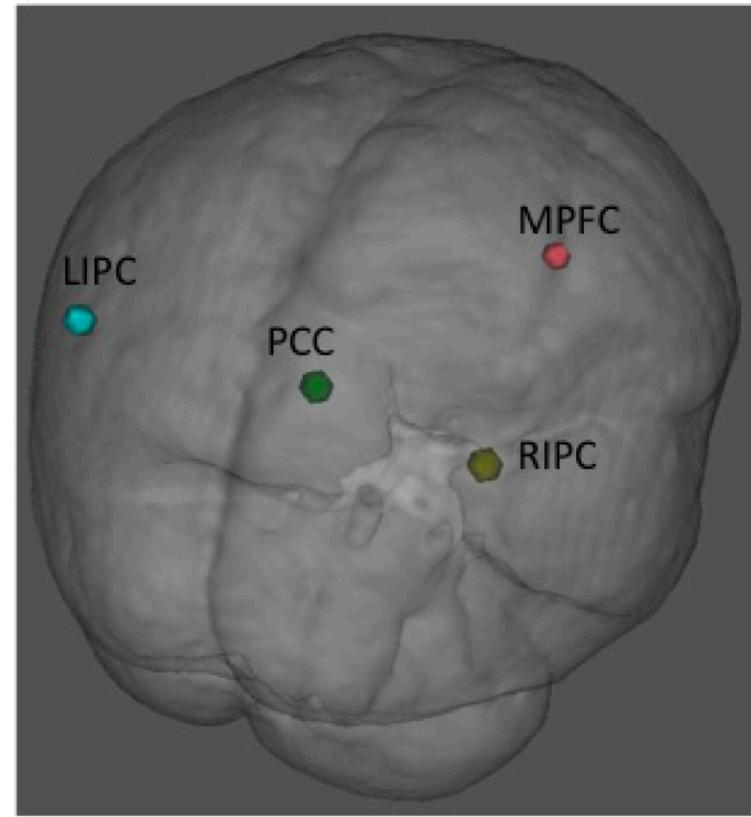
- The edges are real-valued.
- They represent the degree to which one node regulates another node.
- The number indicates the degree to which observation of a signal at a node is followed by observation of a signal at the other node in the dyad.

A Copula-Based Model for the Default Mode Network

- Connectome coined by Dr. Olaf Sporns (Indiana University) and Dr. Patric Hagmann (Lausanne University Hospital) in 2005 to refer to the **map of neural connections in the brain**.
- The connectome contains brain networks defined at different levels of scale:
 - Microscale: Single neurons and synapses.
 - Mesoscale: Populations of neurons.
 - Macroscale: Cortical Areas.
- Our data describes the macro-scale connectome: nodes correspond to regions of interest.
- Edges of the graph are derived from the axons interconnecting those areas.

Mind: A Copula-Based Model for the Connectome

- Examine the Default Mode Network.
- DMN consists of brain regions that tend to be active in a resting-state, i.e., when a subject's mind wanders with no intended task.
- Involved in:
 - Neurological basis for the self.
 - Thinking about others.
 - Memories from the past.
 - Imagination of the future.
- Regions of interest:
 1. LIPC (left intraparietal cortex)
 2. PCC (posterior cingulate cortex)
 3. MPFC (medial prefrontal cortex)
 4. RIPC (right intraparietal cortex)



Mind: A Copula-Based Model for the Connectome

- Data found in the Alzheimer's disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu).
- The data contains:
 - Time series of resting-state fMRI taken at different regions in the brain.
 - Background information on the observational units.
 - A label indicating the disease state for each unit of study: Normal Cognition, Mild Cognitive Impairment, Alzheimer's.
- Previous work transformed the rs-fMRI data into a measure of **effective connectivity** using Dynamic Causal Modeling. (Friston et al., 2017)
- Effective connectivity refers to the directed influence of one brain region on others.
 - Simplest possible circuit diagram that would replicate the observed timing relationships between the recorded neurons.

Mind: A Copula-Based Model for the Connectome

- Coauthor: Dr. Jiguo Cao.
- Inspired by work in *Spectral Dynamic Causal Modeling of Resting-State fMRI* (Nie et al., 2019).
- Model the 4 node network as a Real-Valued **Copula Generated Random Graph** (Danielson, Handcock and Lawrence, in progress).
- The idea is to use a copula to model dependence between the value of particular edges in the network.
- The CGRG framework is convenient as:
 - Requires no intractable normalizing constant unlike Exponential Random Graph Models.
 - Introduces interpretable dependence parameters.
 - Can be estimated in a frequentist or Bayesian manner.
 - Can be adapted to real-valued, count, signed or binary networks.

Copula Generated Random Graph: Directed Edges

- Suppose we observe values of the edges associated with a network $\{A_{ij}^{(s)}\}$ where $i, j \in \{1, \dots, N\}$ index the edges and s indexes the the units of observation (patients).
- Each arc has a marginal distribution

$$p\left(A_{ij}^{(s)} | \delta_i^{(g_s)}, \gamma_j^{(g_s)}, \sigma^{(g_s)}\right) = \mathcal{N}\left(\delta_i^{(g_s)} + \gamma_j^{(g_s)} + \beta^\top x_s, (\sigma^{(g_s)})^2\right).$$

- Since $A_{ij}^{(s)} \in \mathbb{R}$, we choose a Normal distribution.
- The mean contains a **sender** term δ_i and a **receiver** term γ_j .
- The inner product $\beta^\top x_s$ captures the signal from the vector of background information x_s .
- Parameters have a superscript g_s indicating the disease group to which unit s belongs.

Copula Generated Random Graph: Frank Copula

- Edges from the same dyad depend upon one another via a copula

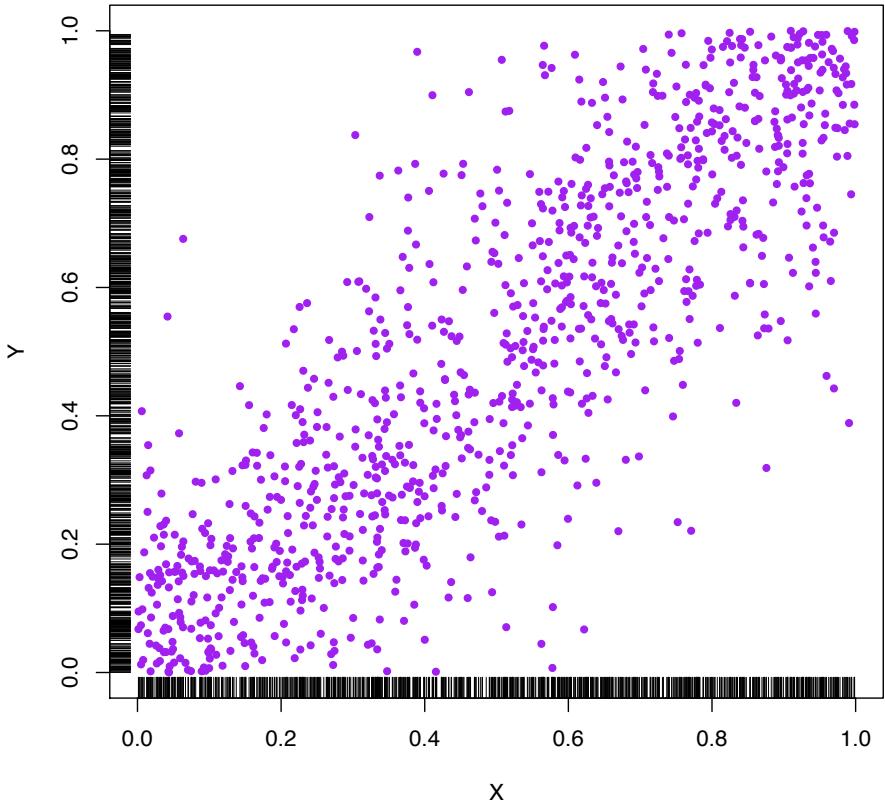
$$p\left(A_{ij}^{(s)}, A_{ji}^{(s)} | \rho^{(g_s)}, \delta_i^{(g_s)}, \delta_j^{(g_s)}, \gamma_i^{(g_s)}, \gamma_j^{(g_s)}, \sigma^{(g_s)}\right) = \frac{\partial}{\partial A_{ij}^{(s)}} \frac{\partial}{\partial A_{ji}^{(s)}} C_\rho \left(F\left(A_{ij}^{(s)}\right), F\left(A_{ji}^{(s)}\right)\right).$$

- The dependence parameter ρ measures association between the edge values.
- In social networks, this is called **reciprocity**.
- We to choose the Frank copula because:
 - $\rho \in \mathbb{R}/\{0\}$ implies the model can represent negative and positive reciprocity.
 - the function is simple to compute

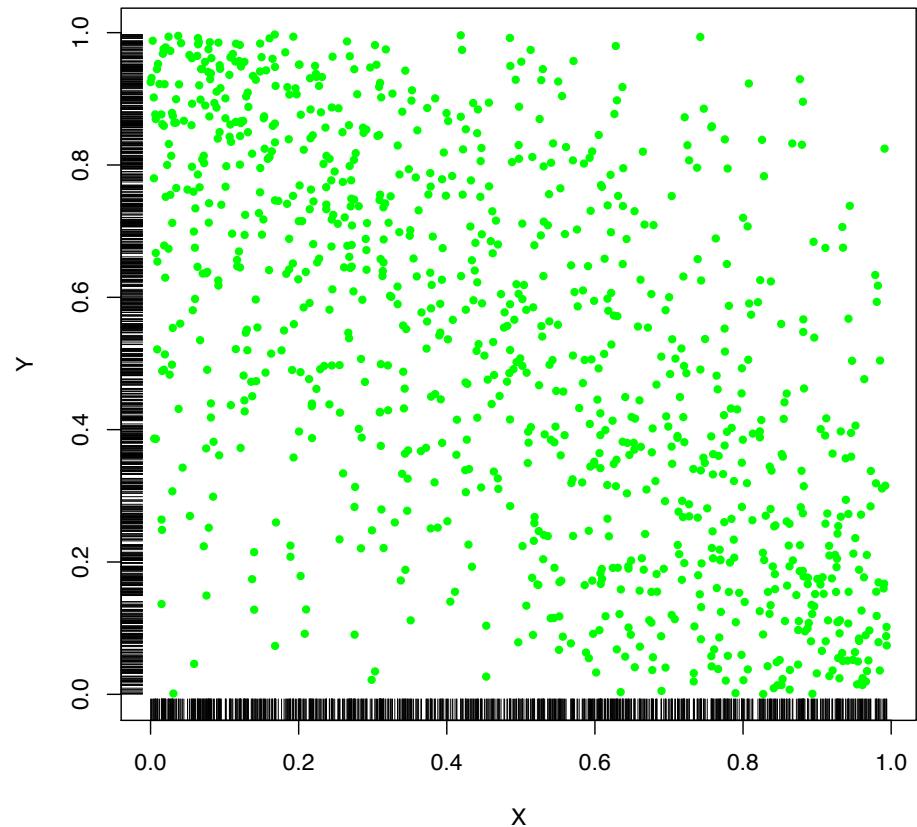
$$C_\rho(F(A_{ij}), F(A_{ji})) = -\frac{1}{\rho} \log \left[1 + \frac{\left(\exp\left\{-\rho F(A_{ij})\right\} - 1\right) \left(\exp\left\{-\rho F(A_{ji})\right\} - 1\right)}{\exp\left\{-\rho\right\} - 1} \right]$$

Frank Copula with Normal Marginals: Some Pictures

Bivariate Sample from Frank Copula (Positive Dependence)



Bivariate Sample from Frank Copula (Negative Dependence)



CGRG for Brain Networks: The Likelihood Function

- The joint pdf of the arc values is

$$\frac{\partial}{\partial F(A_{ij}^{(s)})} \frac{\partial}{\partial F(A_{ji}^{(s)})} C_\rho \left(F(A_{ij}^{(s)}), F(A_{ji}^{(s)}) \right) f \left(A_{ij}^{(s)} \right) f \left(A_{ji}^{(s)} \right).$$

- Then, given the assumption that dyads are conditionally independent, the likelihood is

$$L(\boldsymbol{\rho}, \boldsymbol{\delta}, \boldsymbol{\gamma}, \boldsymbol{\sigma} | \mathbf{A}^{(1)}, \dots, \mathbf{A}^{(S)}) = \prod_{s=1}^S \prod_{i=1}^n \prod_{j < i} p \left(A_{ij}^{(s)}, A_{ji}^{(s)} | \rho^{(g_s)}, \delta_i^{(g_s)}, \delta_j^{(g_s)}, \gamma_i^{(g_s)}, \gamma_j^{(g_s)}, \sigma^{(g_s)} \right).$$

- The different group-specific parameters share a common prior.

- $\rho^{(g_s)} \sim \pi(\rho | \rho^0)$ and $\rho^0 \sim \pi(\rho^0 | a_0, b_0)$
- $\delta_i^{(g_s)} \sim \pi(\delta_i | \delta_i^0)$ and $\delta_i^0 \sim \pi(\delta_i^0 | c_0, d_0)$
- $\gamma_i^{(g_s)} \sim \pi(\gamma_i | \gamma_i^0)$ and $\gamma_i^0 \sim \pi(\gamma_i^0 | e_0, f_0)$

Posterior Sampling

- The model parameters (ρ, δ, γ) are estimated hierarchically.
- The posterior distributions for the parameters satisfy:

$$p(\rho^{(g)} | \delta, \gamma, A^{(1)}, \dots, A^{(S)}) \propto \pi(\rho^{(g)} | \rho^0) \prod_{s: g_s=g} \prod_{i=1}^n \prod_{j:j>i} p(A_{ij}^{(s)}, A_{ji}^{(s)} | \rho^{(g)}, \delta_i^{(g)}, \delta_j^{(g)}, \gamma_i^{(g)}, \gamma_j^{(g)}),$$

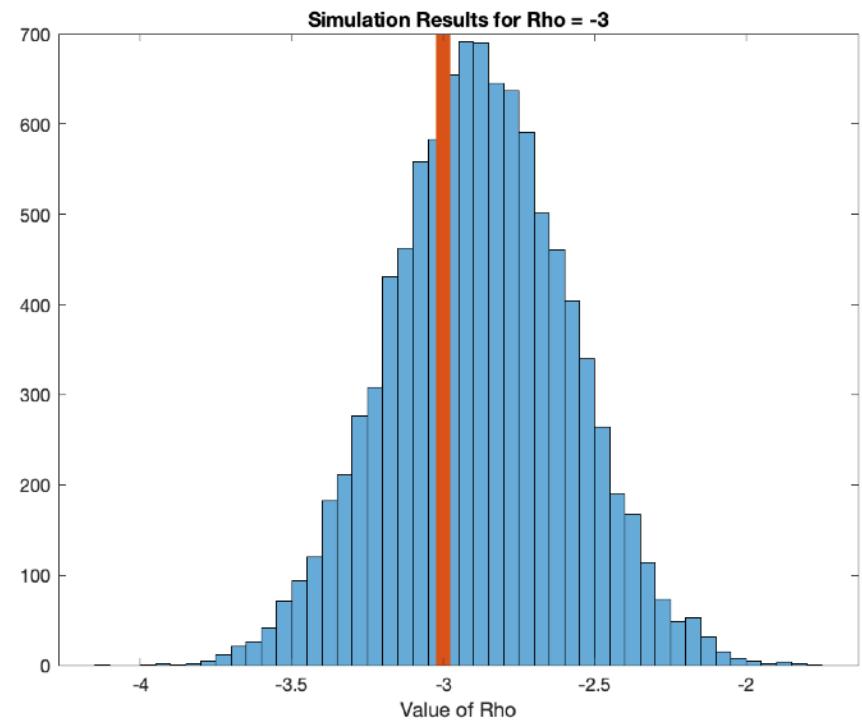
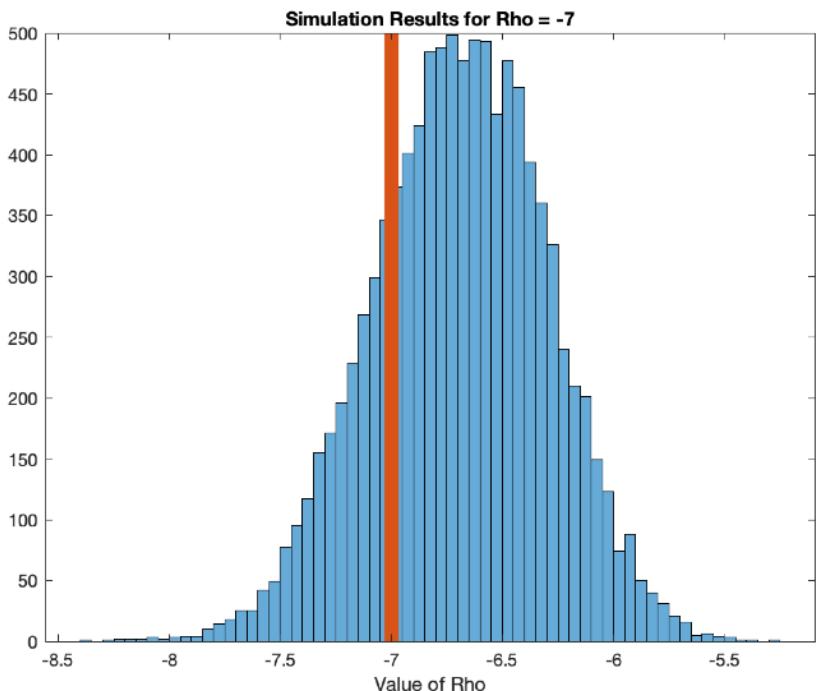
$$p(\delta_i^{(g)} | \delta_{-i}, \gamma, A^{(1)}, \dots, A^{(S)}) \propto \pi(\delta_i^{(g)} | \delta_i^0) \prod_{s: g_s=g} p(A_{ij}^{(s)}, A_{ji}^{(s)} | \rho^{(g)}, \delta_i^{(g)}, \delta_j^{(g)}, \gamma_i^{(g)}, \gamma_j^{(g)}),$$

$$p(\gamma_j^{(g)} | \delta, \gamma_{-j}, A^{(1)}, \dots, A^{(S)}) \propto \pi(\gamma_j^{(g)} | \gamma_j^0) \prod_{s: g_s=g} p(A_{ij}^{(s)}, A_{ji}^{(s)} | \rho^{(g)}, \delta_i^{(g)}, \delta_j^{(g)}, \gamma_i^{(g)}, \gamma_j^{(g)}).$$

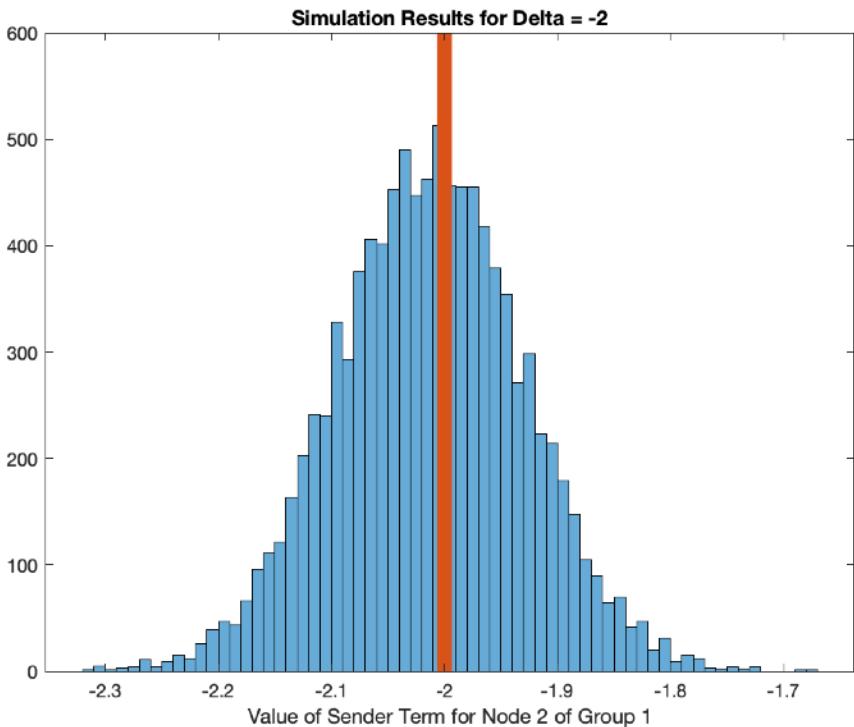
CGRG: Simulation

- Consider a two-group model for a 5 node network
 - $\rho^{(1)} = -7$ and $\rho^{(2)} = -3$
 - $\delta^{(1)} = (2, -2,.25, -.5, 1)$ and $\delta^{(2)} = (-1,.33, 8, -5, 9)$
 - $\gamma^{(1)} = (15,.5, -2, 2, -1)$ and $\gamma^{(2)} = (3, -12, 0, 4, -3)$
 - $\sigma = (1,1)$
- The model recovers the parameters.
- But, it can't recover them all without setting one of the sender or receiver effects to be the baseline term.
- Observe that these effects cannot be identified.
- Let $\tilde{\delta} = \delta + \mathbf{k}$ and $\tilde{\gamma} = \gamma - \mathbf{k}$ with $\mathbf{k} = (k, \dots, k)$.
- Then the likelihoods based on (δ, γ) and $(\tilde{\delta}, \tilde{\gamma})$ are observationally equivalent.

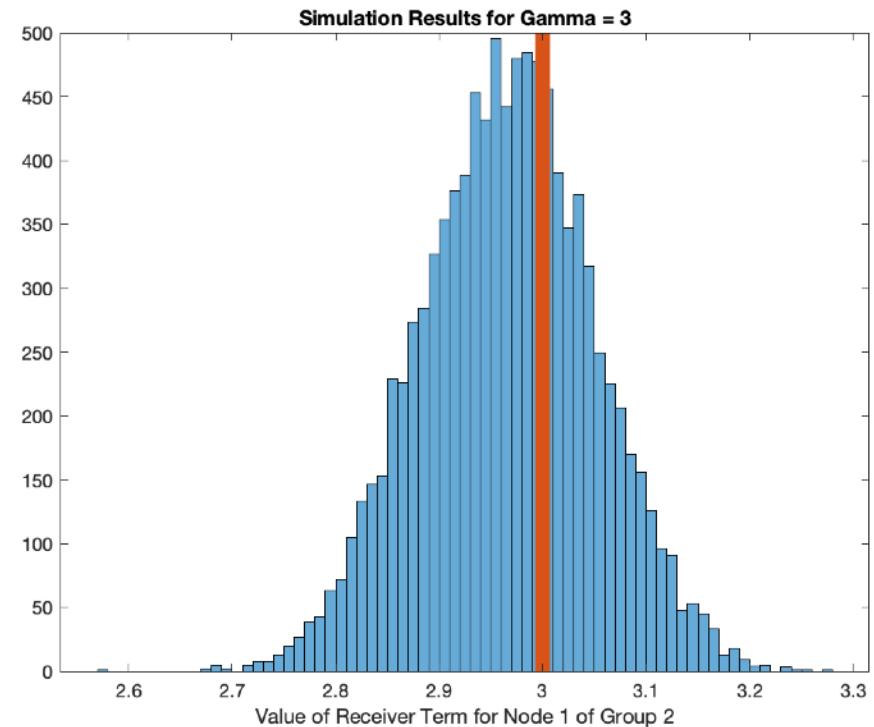
CGRG: Simulation Results for Reciprocity Parameter



CGRG: Simulation for Sender and Receiver Terms



Negative sender term indicates a lower propensity to regulate other nodes.

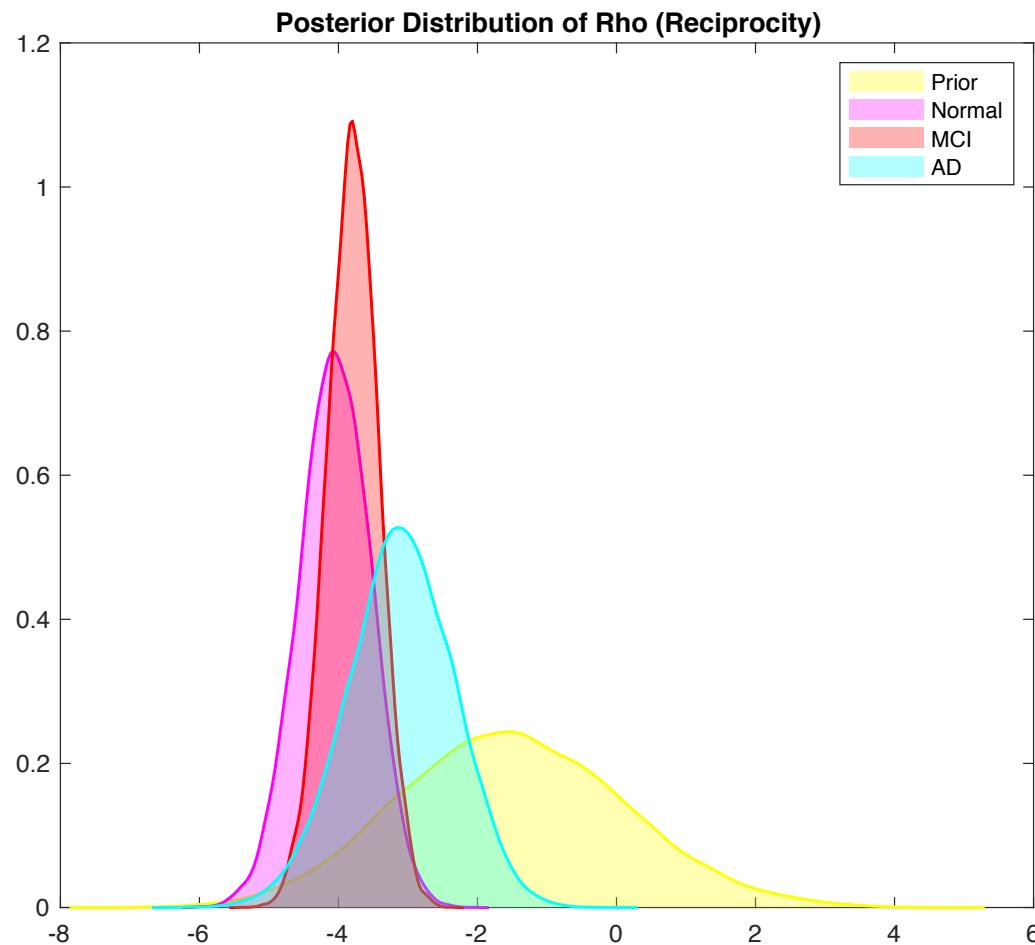


Positive receiver term indicates a higher propensity to be regulated by other nodes.

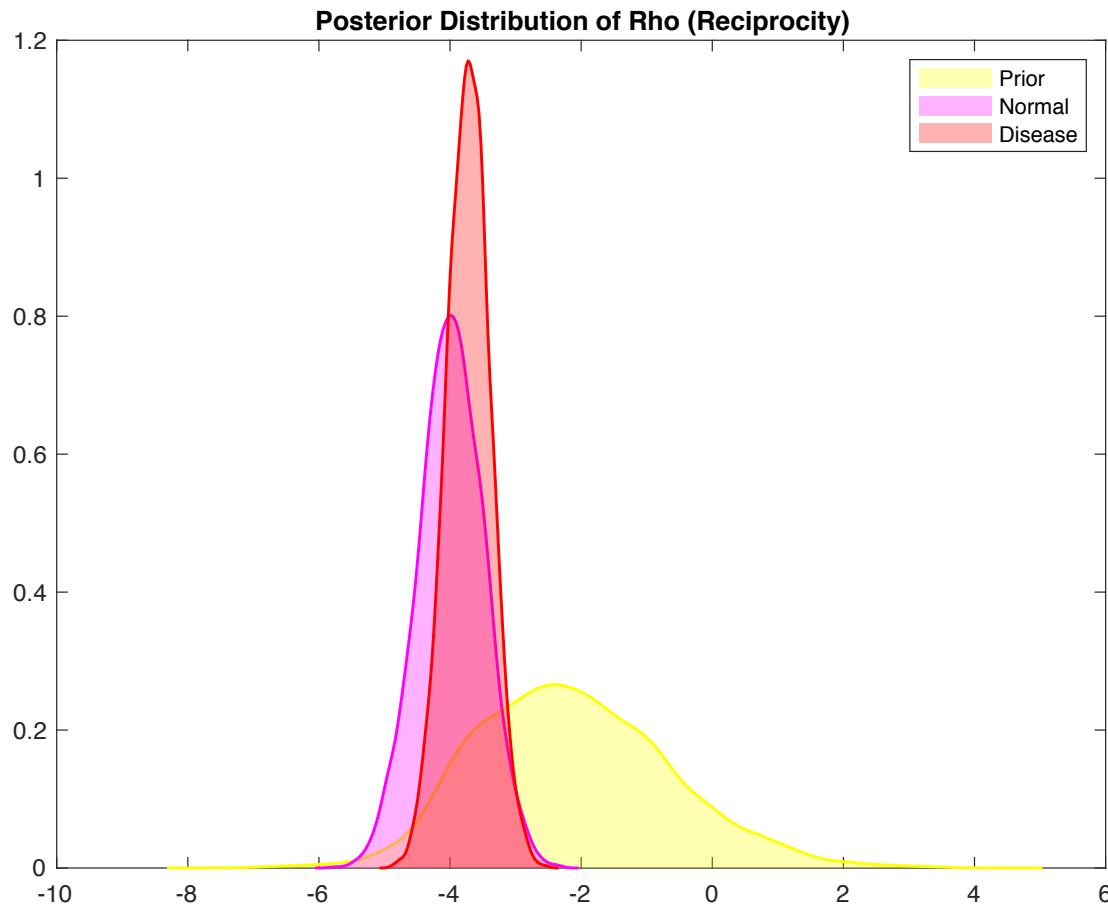
CGRG for the Connectome: Results for Reciprocity

- No matter whether two (disease or normal) or three (AD,MCI,NL) groups are used in the hierarchical model, the reciprocity parameters are negative.
- Interpretation: if $A_{ij} > 0$, then it is likely that $A_{ji} < 0$.
- If a signal at node i leads to a signal at node j , then a signal at j is unlikely to lead to a signal at node i .
- In networks a lack of reciprocity suggests a hierarchical structure.
- There is evidence that the reciprocity parameter in the normal group is more negative than the reciprocity parameters associated with the various diseased groups considered in our study.
- This corroborates the notion that healthy brains are hierarchically-organized complex systems.

CGRG for the Connectome: Results (Three Groups)



CGRG for the Connectome: Results (Two Groups)

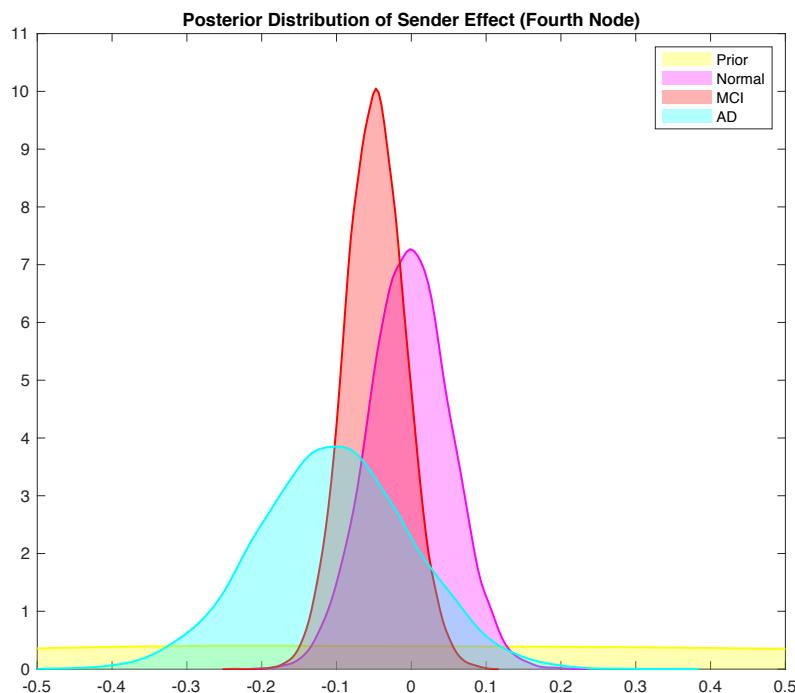


CGRG for the Connectome: Sender/Receiver Terms

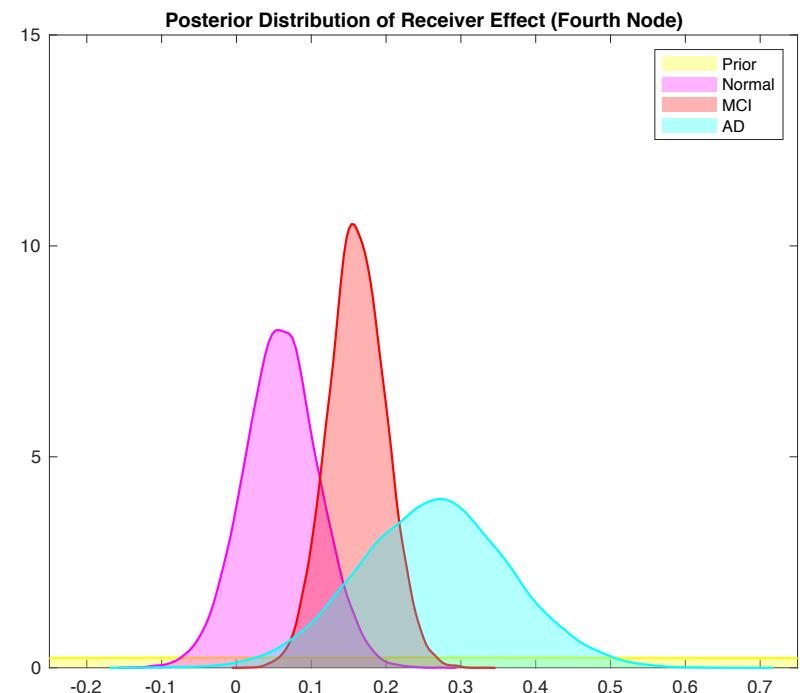
- We observe some differences in the terms comprising the mean function in the marginal distribution.
- The degree to which certain nodes activate others differs across disease states.
- If structural terms vary across diseases states, then the model could be used to detect disease.

CGRG for the Connectome: Sender/Receiver Terms

Sender



Receiver



CGRG: Disease State Detection

- Now that we have a sample from the posterior distribution of the model parameters, we can investigate the model's ability to recognize disease states. The posterior probability that unit s belongs to group g is

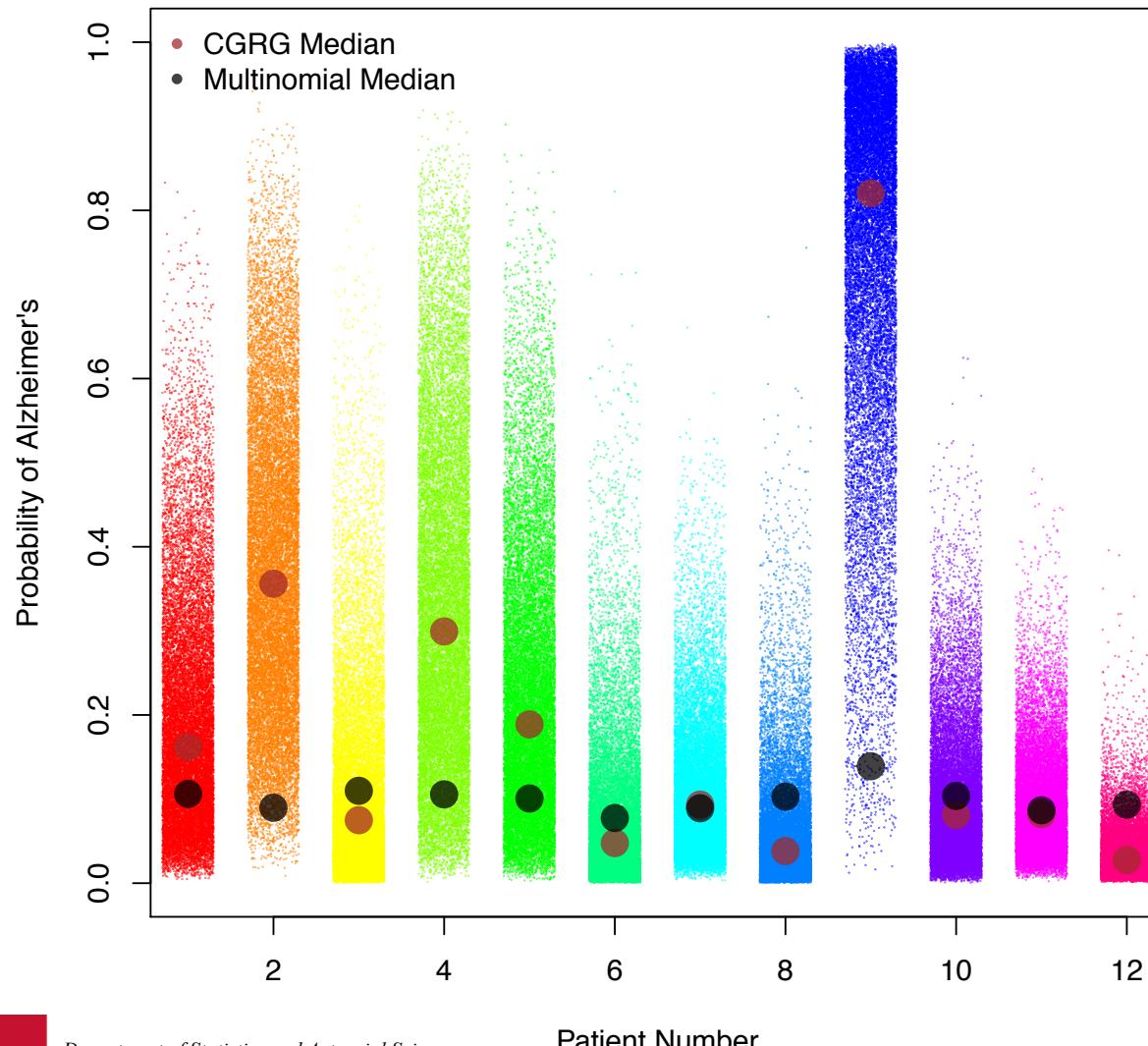
$$p(g_s = g | A^s, \rho, \delta, \gamma, X_s) = \frac{p(A^s | g_s = g, \rho, \delta, \gamma, X_s)p(g_s = g | \theta, X_s)}{\sum_{g=1}^G p(A^s | g_s = g, \rho, \delta, \gamma, X_s)p(g_s = g | \theta, X_s)}.$$

- Model the probability of group membership as the multinomial distribution

$$p(g_s = g | \theta, X_s) = \frac{\exp \left\{ \theta_g^\top X_s \right\}}{\sum_{g=1}^G \exp \left\{ \theta_g^\top X_s \right\}}.$$

CGRG: Disease State Detection (Three Group Model)

Alzheimer's Detection and CGRG



Leave-One-Out Approximation

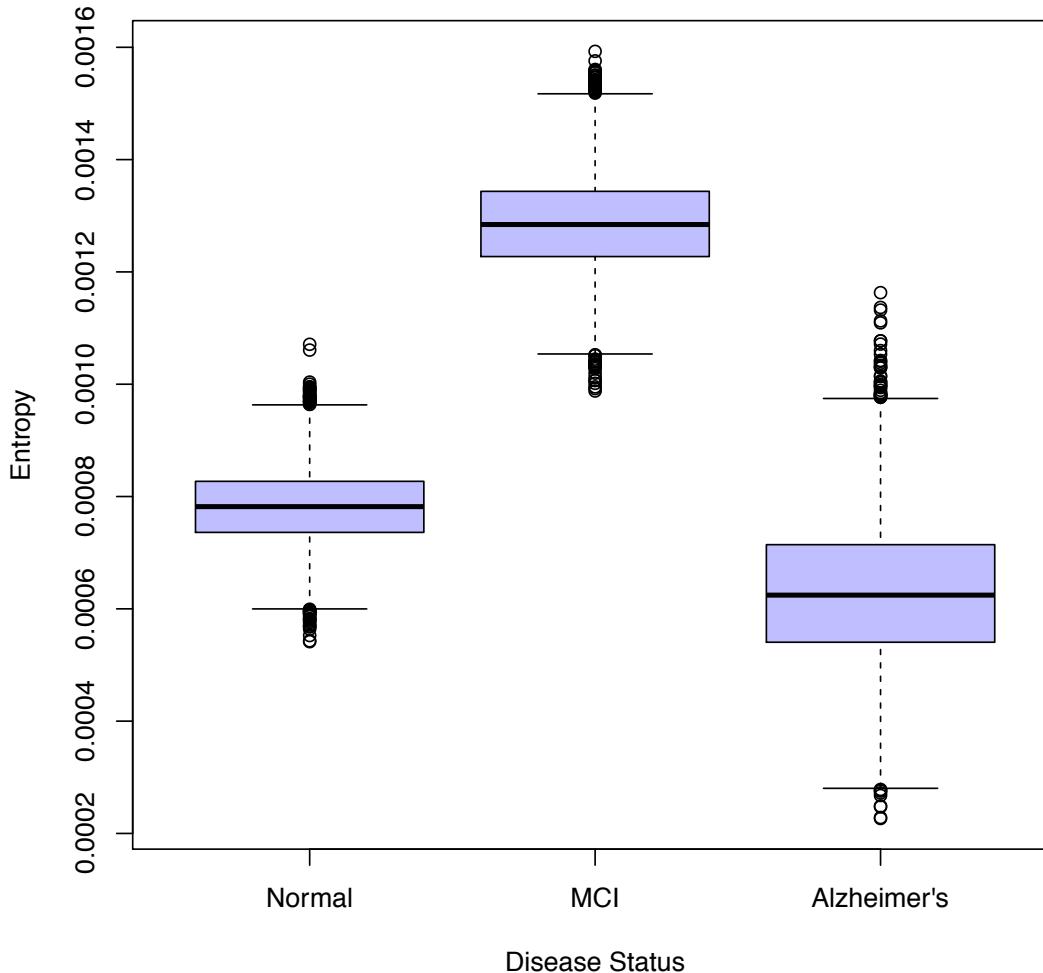
- The full model with group specific reciprocity, sender and receiver terms likely overfits the data.
- Number of units in the AD group is small (12).
- To simplify, estimate disease specific terms for just reciprocity.
- Then importance sampling methods can be used to generate approximate leave-one-out diagnostics (Vehtari and Gelman, 2015).
- Results are inconclusive.
- Need a larger dataset.

CGRG: Entropy and the Connectome

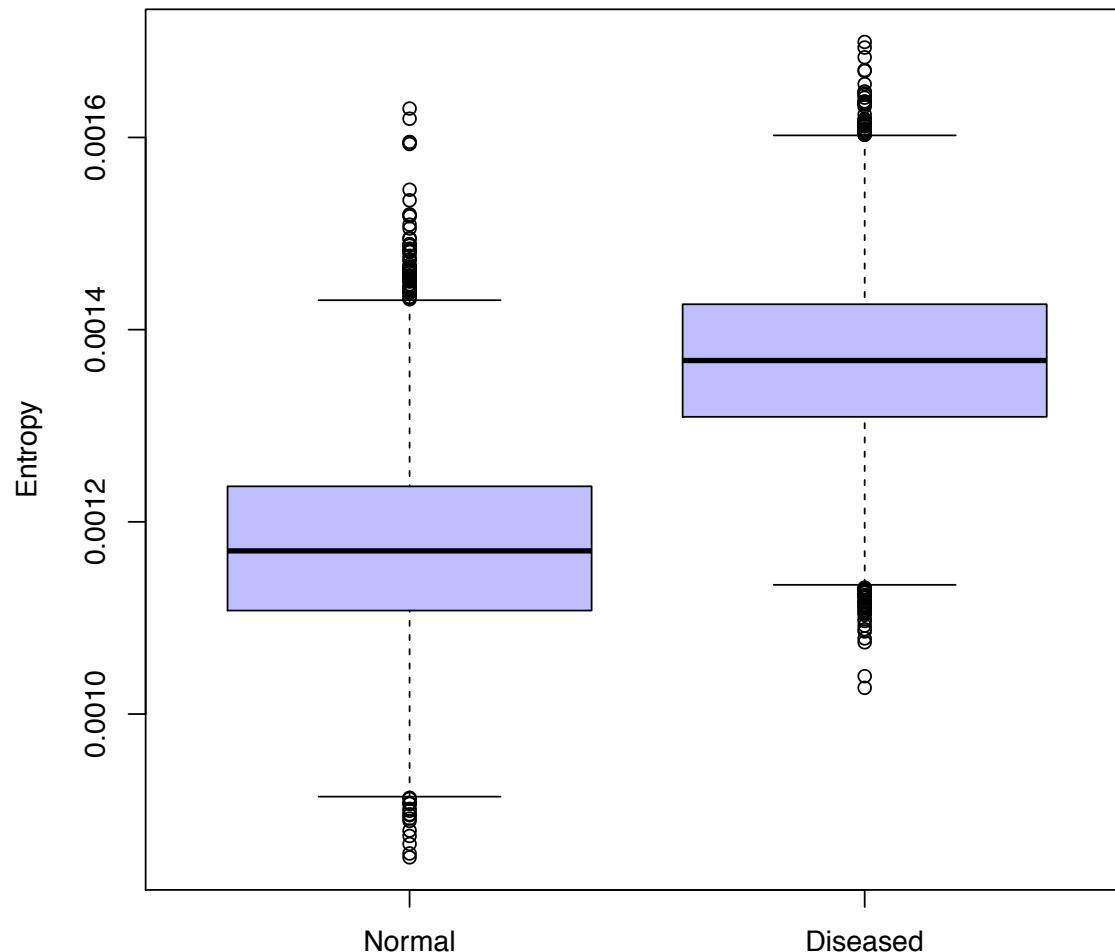
- Theory suggests healthy neural systems feature lower entropy configurations.
- A lower entropy brain exhibits a more predictable regulation network. That is, it is clear which node regulates other nodes.
- If regulatory ties are more reciprocal, then there is greater uncertainty as to which node regulates which.
- Sample R networks. For each disease group compute the Monte Carlo approximation of the entropy

$$-\int_A f(A) \ln(f(A)) dA \approx -\frac{1}{R} \sum_{r=1}^R p(A^{(r)}) \ln p(A^{(r)}).$$

CGRG: Entropy in the DMN (Three Group Model)



CGRG: Entropy in the DMN (Two Group Model)



CGRGs and the Brain: (Tentative) Conclusions

- Structural differences exist across disease groups.
- Can fit the network data better using structural terms.
- Better fits for the disease state better using structural terms.
- But, need more observations to draw strong conclusions about detection of new data points.
- Also, increasing the number of nodes measured in the brain could lead to the discovery of important structural differences.
- We chose interpretable structural parameters inherited from social network theory.
- But, perhaps brain function does not adhere to these intuitive concepts.
- Probability distribution induced by the MCI parameters exhibits higher entropy than the normal and Alzheimer's groups.

Next Steps for this Project

- How to choose the best model for disease detection?
- Which parameters should have group-specific values?
- Consider other methods for detection and prediction of disease states.
- Try a frequentist estimation approach.
 - Suppose the prior means are viewed as nuisance parameters.
 - Choose group specific parameters($\rho^{(1)}, \dots, \rho^{(G)}, \delta^{(1)}, \dots, \delta^{(G)}, \gamma^{(1)}, \dots, \gamma^{(G)}$) to maximize the likelihood marginalized over the prior means

$$\int \int \int p(A^{(1)}, \dots, A^{(S)} | \rho, \delta, \gamma, \sigma) \pi(\rho | \rho^0) \pi(\rho^0) \pi(\delta | \delta^0) \pi(\delta^0) \pi(\gamma | \gamma^0) \pi(\gamma^0) \pi(\sigma) d\rho^0 d\delta^0 d\gamma^0$$

- Then asymptotic results hold for the parameter estimates.
- This is similar to an EM approach in which we average over the values of the prior mean.
- Consider node and group specific reciprocity parameters.