# Blockmodels for Connectome Analysis

Daniel Moyer[ab], Boris Gutman[b], Gautam Prasad[b], Joshua Faskowitz[b], Greg Ver Steeg[c], and Paul Thompson[b]

[a]Dept. of Computer Science, University of Southern California, 840 Childs Way, Los Angeles, CA, United States
[b]USC Stevens Neuroimaging and Informatics Institute, 4676 Admiralty Way, Marina Del Rey, CA, United States
[c]Information Science Institute, 4676 Admiralty Way, Marina Del Rey, CA, United States

## ABSTRACT

In the present work we study a family of generative network models and their applications for modeling the human connectome. We introduce a minor but novel variant of the Mixed Membership Stochastic Blockmodel and apply it and two other related models to two human connectome datasets (ADNI and a Bipolar Disorder dataset) with both control and diseased subjects. We further provide a simple generative classifier that, alongside more discriminating methods, provides evidence that blockmodels accurately summarize tractography count networks with respect to a disease classification task.

**Keywords:** Random Network Models, Connectomics, Diffusion Weighted Imaging

## 1. INTRODUCTION

Due to advances in medical image processing and imaging technologies, recent literature has given much attention to the analysis of connectivity in human brains. In many studies the so-called "connectome" is modeled as a network, and a subsequent network analysis is performed upon it; these analyses range from regressions on simple counting measures (e.g. degree) to spectral analyses and community detection. These studies have been instrumental in shaping the field of connectomics, and have provided both promising clinical applications as well as valuable insight on other scientific discoveries (e.g. the "rich club" phenomena).

Network analysis however has a rich history that is often untapped by current connectome analyses; while intended for a variety of other applications, the methods and models developed therein may be useful for the analysis of brain networks. In particular, we consider a family of generative network models known as Stochastic Blockmodels. Originally defined in the context of social networks,[1] SBMs have since been applied to a variety of network contexts, including internet networks,[2] academic citation networks,[3] protein interaction networks,[4] and structural connectivity networks.[5,6] In particular, the flexibility and general ease of estimation of SBMs has lead to such popularity and widespread usage.

Blockmodels also have two important, related properties that we leverage: first, blockmodels have a basic community structure assumption. We review this in detail in the next section, and leverage this fact to measure "between community" connectivity, which we show to be useful in disease prediction. Second, their estimation is easily extensible to multiple networks. While numerous network measures are simply descriptive, and subsequent statistical analysis must be then undertaken, we may directly analyze multiple networks from within the blockmodel. That is, the community parameters we estimate are estimated from every observed (brain) network at the same time. This is rarely leveraged in most network analyses due to the rarity of data that fits such a paradigm, but is intuitively simple here.

In the present work we investigate the usage of blockmodels for producing efficient representations the human connectome. We introduce a minor but novel variant of a mixed membership blockmodel and apply it and two other related model to two human connectome datasets with both control and diseased subjects. For a simple classification task we demonstrate comparable accuracy of blockmodel representations to the original representation while greatly reducing the number of dimensions. We further provide a simple generative classifier that, alongside more discriminating methods, provides evidence that blockmodels accurately summarize tractography

Further author information: (Send correspondence to Daniel Moyer)
Daniel Moyer: E-mail: dmoyer@ini.usc.edu

count networks. It is important to note that the objective of this paper is not to improve classification (outside of reducing optimization inefficiencies and noise, network reductions are unable to improve classification accuracy due to the data processing inequality) but to generate succinct representations of the networks.

## 2. RELATED WORK

Our work relies upon the strong literature developed in the Blockmodel network analysis community. Among the most influential is the recent work by Karrer and Newton,[2] which introduced the Degree Corrected Stochastic Blockmodel (with Poisson edges). It also provides a short but excellent summary of Poisson distributed edge weights for general blockmodels. Alongside their work is the paper by Airoldi et al[4] which introduced the Mixed Membership variant of blockmodel for binary edges (topological graphs).

Our work also depends highly upon the work of the Connectome community; first and foremost is the work of Sporns et al., whose influence cannot be understated.[7–10] In terms of Bayesian models for connective communities, several recent papers are directly relevant to our work. In particular Hinne et al.[6] and Jbabdi et al.[11] both consider Non-Parametric Bayes approaches to community detection, applying them to similar domains but without a classification task. Their work also highlights the neurological relevance generative network models. Previously in this exact line of research Moyer et al.[12] conducted a study of the same dataset using just the topological blockmodel (no edge weights, no degree correction). This is direct continuation of that work.

Outside of medical imaging, Palovic et al.[5] fit a stochastic block model variant to the *C. elegans* connectome. Their particular domain has data with cellular resolution (the *C. elegans* neural system is only 300 cells total), but their work is very similar in concept to our own, and the work of Hinne et al.[6] and Jbabdi et al.[11]

## 3. BLOCKMODELS

The basic SBM is an immediate extension of the Erdös-Rényi (ER) model. For a graph $\mathcal{G} = (V, E)$ with $N$ fixed nodes but a random number of undirected edges $E$, the ER model generates graphs by considering each possible edge as an independent Bernoulli random variable with probability $b$. The model then adds edges by independent coin flips. Clearly here the edges are either 0 or 1, and one would not expect the same community structures to emerge in multiple observations.

The SBM adds a latent community assignment to this; for each node we first choose a community $z \in \{1, \ldots, K\}$. For each pair of nodes $(i, j)$, the (random) edge between them is then modeled as a Bernoulli random variable with a community interaction parameter $b_{z_i, z_j} \in [0, 1]$, where $z_i$ and $z_j$ are the community assignments of $i$ and $j$ respectively. It can be seen that each individual community's closed subgraph is itself an ER random graph. Furthermore, the bipartite subgraph between any two communities also behaves as an ER graph. For this reason the basic SBM is sometimes referred to as the Erdös Renyi Mixture Model (ERMM) in frequentist contexts;[13] though estimation techniques differ, the underlying principles are the same.

It is helpful to consider a matrix of the edge parameters $\Gamma$ of each model, where $\gamma_{ij}$ is the parameter for the distribution of possible edges between node $i$ and node $j$. In the ER model all parameters $\gamma_{ij}$ are equal, and $\Gamma_{ER} = b\mathbf{1}^T\mathbf{1}$, where $\mathbf{1}$ is a vector of ones. For the basic SBM, we construct a membership matrix $Z \in \{0, 1\}^{K \times N}$, where $Z_{ki} = 1$ if node $i$ is in community $k$ and 0 otherwise. It can be shown easily that $\Gamma_{SBM} = Z^T B Z$, where $B \in [0, 1]^{K \times K}$ the matrix of community interaction parameters. $\Gamma_{SBM}$ is a block matrix up to some permutation of the node indices, hence the name Blockmodel. In this context the least structured model (ER) corresponds with the representation with the lowest rank (1), while increasingly flexible models (SBMs with various numbers of communities) have representations with higher rank.

Often SBMs defined slightly more generally, where $\mathcal{G}$ allowed to have multiple edges between nodes[2] (or, alternatively, where each edge is assigned a non-negative integer weight). In this latter case we model the number of edges between each pair of nodes as a Poisson random variable, with $b_{z_i, z_j} \in [0, \infty)$. The matrix representation of the parameter matrix $\Gamma_{SBM}$ remains almost the same, with the exception that $B \in [0, \infty)^{K \times K}$ has been modified for the support of the parameter of the Poisson distribution. To avoid confusion we use the latter case exclusively from this point forward.

Given community memberships $\{z_i\}$ and interaction matrix B, a random graph $\mathcal{G}$ with weighted adjacency matrix $Adj$ generated from the basic SBM with Poisson edges and without self edges has the following distribution:

$$P(\mathcal{G}|\{z_i\}, B) = \prod_{i<j} \frac{(b_{z_i,z_j})^{Adj_{i,j}}}{A_{i,j}!} \exp(-b_{z_i,z_j})$$

The a posteriori estimation of parameters from an ensemble of observed networks is also well studied; though frequentist approaches exist,[14] we choose to use a collapsed Gibbs Sampling method similar to the one described in Nowikcki et al[15] for the basic model as well as both presented extensions.

After estimating the model parameters we can also calculate an *empirical interaction matrix*. For any realization $Adj^n$ of the random graph we define $\hat{B} = ZAdj^n Z^T$, whose entries $b_{z,z'}$ may be interpreted as the number of edges observed between communities $z$ and $z'$.

### 3.1 Mixed Membership

A recent extension of SBMs is the Mixed Membership Stochastic Blockmodel[4] (MMSB). As the name suggests, the MMSB allows each node to have a distribution of memberships; this distribution is realized in latent community assignments *for each edge*, where we assign labels $z_{ij}$ and $z_{ji}$ to each pair of nodes $(i, j)$. Originally defined for the Bernoulli edge parameters, we provide its natural extension to Poisson edge counts, with the following generative process:

1. For each node $i \in \mathcal{V}$
   a) Draw a membership distribution vector $\pi_i \sim \text{Dirichlet}(\alpha)$
2. For each pair of nodes $(i, j) \in V \times V$, $i \neq j$
   a) Draw a membership indicator $z_{i,j} \sim$ Multinomial$(\pi_i)$
   b) Draw a membership indicator $z_{j,i} \sim$ Multinomial$(\pi_j)$
   c) Sample their interaction $Adj^n_{i,j} \sim$ Poisson$([Z^T BZ]_{ij})$ where $B$ is the $K \times K$ matrix of block interaction probabilities, and where each entry $B_{i,j} \sim$ Gamma$(a, b)$.

Note the use of three prior parameters, $\alpha$, $a$, and,$b$, as well as one hyper parameter $K$. This leads us to the following joint distribution:

$$P(Adj, \{\pi_i\}, \{z_{i,j}, B|\alpha, a, b) = \prod_{i<j} P(Adj_{i,j}|z_{i,j}, B)P(z_{i,j}|\pi_i)P(z_{j,i}|\pi_j) \prod_i P(\pi_i|\alpha) \prod_i P(B|a, b)$$

The joint distribution is very similar in appearance that of the MMSB with Bernoulli edges, and also enjoys the same matrix factorization representation. In terms of the matrix $\Gamma$ of edge distribution parameters, the MMSB can be though of as the relaxation of the community assignment matrix $Z$ from binary $\{0, 1\}$ entries to those on the interval $[0, 1]$, with the constraint that each column sums to 1. While this greatly increases the number of parameters to be learned, it also correspondingly allows for much greater flexibility.

Since community assignment is still $\{0, 1\}$ at the edge level, we cannot write out the interaction parameters exactly as a matrix factorization, but we can write their expectation given the prior: $\mathbb{E}[\Gamma_{MMSB}|\alpha, a, b] = Z^T BZ$. Since the edges are Poisson distributed, this matrix is also the expectation of the edge weights. The empirical interaction matrix $\hat{B} = ZAdjZ^T$ has the same form as that of the basic SBM, but in the mixed membership case $\hat{B}$ will not necessarily have integer valued elements.

It is important to note the original model was defined for directed graphs, and includes outward and inward labels. We here only use one label per ordered pair of nodes, but our estimation method is easily extensible to the directed case.

### 3.2 Degree Correction

Another important extension of SBMs corrects for a tendency for the model to group nodes by degree. The Degree Corrected SBMs (DCSBs) define an additional parameter $\theta_i$ for each node, representing the probability an edge incident on the community of node $i$ is incident on node $i$ itself.[2] This probability is then multiplied onto the usual community interaction parameter to produce the following joint distribution:

$$P(\mathcal{G}|\{z_i\}, B) = \prod_{i<j} \frac{(\theta_i \theta_j b_{z_i,z_j})^{Adj_{i,j}}}{A_{i,j}!} \exp(-\theta_i \theta_j b_{z_i,z_j})$$

Karrer and Newton introduced this model with a maximum likelihood estimator for $\theta_i$ using the degrees $d_i$ of each node: $\hat{\theta}_i = d_i / \sum_{j:z_j=z_i} d_j$

### 3.3 Estimation and Implementation Notes

For all three variants (SBM, MMSB, and DCSB) we fit a posteriori model parameter estimates using a collapsed Gibbs Sampler. This has the advantage of being immediately extensible to multiple networks. We define $E^{z,z'}$ as the set of edges with labels $z$ and $z'$. We define $N_i^z$ as the number of edges from node $i$ with label $z$, and define $N^{z,z'} = |E^{z,z'}| = |\{k, r : z_{k,r} = z, z_{r,k} = z'\}|$ as the number of edges with labels $z$ and $z'$. We further define $C^{z,z'} = \sum_{(k,r)\in E^{z,z'}} Adj_{k,r}$, which is the sum of the weights of edges with labels $z$ and $z'$. The marginal distribution for the basic SBM is exactly:

$$P(z_{i,j} = z|\{z_{k,r}\}_{(k,r)\neq(i,j)}) \propto \frac{1+\alpha}{1+K\alpha} \prod_n \text{Poisson}\left(Adj_{i,j}^n \; ; \; \frac{C^{z,z_{j,i}} + a}{N^{z,z_{j,i}} + b}\right) \tag{1}$$

The marginal distribution of a single edge label $z_{i,j}$ for the MMSB is exactly:

$$P(z_{i,j} = z|\{z_{k,r}\}_{(k,r)\neq(i,j)}) \propto \frac{N_i^z + \alpha}{N+K\alpha} \prod_n \text{Poisson}\left(Adj_{i,j}^n \; ; \; \frac{C^{z,z_{j,i}} + a}{N^{z,z_{j,i}} + b}\right) \tag{2}$$

The marginal distribution of a single edge label $z_{i,j}$ for the DCSB is exactly:

$$P(z_{i,j} = z|\{z_{k,r}\}_{(k,r)\neq(i,j)}) \propto \frac{1+\alpha}{1+K\alpha} \prod_n \text{Poisson}\left(Adj_{i,j}^n \; ; \; \frac{d_i}{\sum_{j:z_j=z} d_j + d_i} \frac{C^{z,z_{j,i}} + a}{N^{z,z_{j,i}} + b}\right) \tag{3}$$

A Gibbs Sampling based estimator for either model requires a prior for the Gamma distribution. Unfortunately, using an uninformative prior would lead us to choose an *improper prior*, i.e. one that is constant on the parameter space and not a proper probability distribution. While this might be allowable in some estimation schemes, for our method an improper prior leads to numerical issues, forcing us to choose and informed prior. However, the choice of a very small informative prior in practice has the same effect.

Finally, when calculating the marginal distributions it is helpful to omit the factorial term in the denominator of the Poisson distribution definition. Since the observed data are constant across each possible value of $z$, the factorial will be canceled in the re-normalization of the marginal.

### 3.4 Simple Bayesian Classifier

Assuming our dataset of networks $\{Adj^n\}$ partitions into two classes, $y = 0$ and $y = 1$ (e.g. diseased and control subjects), and assuming these classes affect network structure or edge weights, we can construct a simple classifier by leveraging Bayes Theorem. Given a training set of networks from both classes, we may train models $\mathcal{M}_0$ and $\mathcal{M}_1$ on each respective class of training data. Our classifier is then:

$$y_n = \arg\max_{y \in \{0,1\}} P(Adj^n | \mathcal{M}_y)$$

This is simply the model that better fits the data in terms of log-likelihood; in practice we instead consider $\gamma P(Adj^n|\mathcal{M}_y)$ and $(1-\gamma)P(Adj^n|\mathcal{M}_{1-y})$, where $\gamma \in [0, 1]$ is a tuning parameter. We set $\gamma$ by further dividing the training set into training and validation folds, and then performing a grid search for each validation fold.

We do not suggest the use of this classifier for prediction. Instead, we describe and use it here in order to validate the use of the blockmodel as a sensitive network description. It is important to note that this classifier is *not* optimized over classification accuracy beyond a multiplicative hyper parameter; moreover, there is no influence by the class on the blockmodels themselves beyond the data upon which each is trained. Thus, assuming the disease affects the connectomes at all, this might be viewed as a qualitative test of blockmodel sensitivity.

## 4. PROCEEDURE

The overall objective of our testing is the validation of blockmodels as accurate network models and not the classification of diseased patients from controls. Both of these datasets are far too small to allow generalizations in terms of clinically applicable classifiers; instead, we use linear SVM classification accuracy as a qualitative measure of the overall goodness-of-fit for each blockmodel, in tandem with the aforementioned simple Bayesian classifier. If the SVM trained on the blockmodel's empirical interaction matrix performs at least as well as the SVM on the full dataset, we conclude that the blockmodel fits the data acceptably well.

While there are other goodness-of-fit measures such as Perplexity and various information criteria, these generally are useful only for comparisons between different models, and do not provide an objective overall-goodness measure.

### 4.1 ADNI

Our data are taken from 96 subjects scanned as part of ADNI-2,[16] a continuation of the ADNI project in which Diffusion Imaging was added to the standard MRI protocol. The same dataset was used in Prasad et al.,[17] in which the following description is featured. It is reproduced here for completeness, but both the data and its description are effectively the same.

The dataset includes diffusion MRI scans from 50 cognitively normal controls, as well as 46 individuals with Alzheimer's Disease. Subjects were scanned on 3-Tesla GE Medical Systems scanners, which acquired both T1-weighted 3D anatomical spoiled gradient echo (SPGR) image volumes as well as diffusion weighted images (DWI).

The T1-weighted images were first cleared of extra-cerebral tissue, then corrected for inhomogeneity and registered to the Colin27 template using FSL FLIRT.[18] DWI images were corrected for head motion and eddy current distortion via FSL's eddy correct tool. Tractographies were generated for each subject through a global probabilistic tractography method based on the Hough transform[19] using 10,000 fibers (note that this number differs from,[17] which used a lookup table accelerated method and 35,000 fibers).

### 4.2 Bipolar Dataset

Our data are comprised of 92 subjects from the Searching for Endophenotypes of Bipolar Disorders Study.

T1 weighted (T1w) and diffusion weighted (DWI) images were taken at three sites with similar protocols. Scanner magnet strengths varied at each site (1.5T  3.0 T). T1w were processed with Freesurfer's recon-all script to obtain cortical parcellation maps.[20] Each T1w image had its bias field inhomogeneity removed with an N4 implementation in the ANTs package.[21] To obtain brain masks, Freesurfers cortical parcellations were binarized, hole filled, and dilated by one voxel. This mask was applied to the N4 corrected T1w to obtain a T1w brain image. The T1w brain images were then linearly aligned to the MNI 152 1mm template using FSLs FLIRT[18] with 6 degrees of freedom, and subsequently down-sampled to 2mm isotropic voxel space. Freesurfers cortical parcellations were also brought into this space, with a nearest neighbor interpolation.

DWIs were first denoised with LPCA, then corrected for motion and eddy distortion by linearly registering with 12 degrees of freedom to a B0 reference image with FSLs eddy correct. Each DWI was brain extracted with FSLs BET. DWI B0s were linearly aligned with 12 degrees of freedom to the subjects corresponding 2mm T1w image in MNI space with FSLs FLIRT. The image similarity metric employed for this alignment was the boundary-based registration using Freesurfer's white matter parcellation to drive the B0 to T1 alignment. The linear registration of the B0 was applied to all volumes of the DWI with bspline interpolation. B-vectors were appropriately adjusted for linearly alignments in both eddy correction and linear registration to the T1w.

Tractography was conducted in the isotropic 2mm MNI 152 space. Probabilistic streamline tractography was performed using Dipys Constrained spherical deconvolution with a spherical harmonics order of 4 was used to model the fiber distribution at each voxel. Tractrography streamlines were seeded 3 times per voxel in a white matter mask generated from Freesurfer's parcellation map, dilated by 1 voxel. Streamline tracking using utilized a probability mass function to determine tracking direction, at 0.5mm steps. Streamlines were terminated in grey matter according to the Dipy implementation of ACT[22]Connectivity matrices were computed by measuring total streamline connections between each Freesurfer cortical region, dilated by 1 voxel. Streamlines shorter than 10mm were discarded. Additionally, streamline tractography was conducted with Dipy's LocalTracking module, with a tensor fit to model water diffusion at each voxel.

### 4.3 Connectome Generation

In order to produce FreeSurfer anatomical regions of interest (ROI) connectivity networks, we segmented each subject's cortex into 34 ROIs per hemisphere. Each region was dilated using an isotropic box kernel to ensure its intersection with white matter. Weighted connectivity networks (connectomes) were then generated by counting for every pair of regions the number of fibers intersecting both regions at any point along the fibers. This network generation proceedure produces weighted adjacency matrices $Adj_n$, which we treat as realizations of random networks.

### 4.4 Cross Validation and Model Fitting

In order to safely validate our models and evaluate their accuracy, we performed a 10-fold stratified cross-validation on each dataset. We then fit each variant of the SBM twice, once to the control subjects in the training folds, and once to disease subjects. We further generated a 3-fold validation step and tuned the hyper parameters for two different SVMs, the first trained on the flattened connectome matrices (weighted adjacencies) and the second trained on the flattened empirical interaction matrices, as measured by the community assignment posterior distributions from the control fit of the SBM. We also select $\gamma$, the hyper parameter of the Bayesian classifier.

We measure classification performance in terms of accuracy, precision, and recall. This is repeated for each of the 10-folds. In each test fold, the folds are held as close to even class distribution as possible. The naive classification accuracy is approximately 0.5. In a few cases the SVMs did not converge within 100,000 gradient descent steps, thus naive accuracy was not achieved.

Table 1: Classifier performance on ADNI data.

| Type | K | Accuracy | Precision | Recall |
|---|---|---|---|---|
| Baseline SVM | – | 0.80 | 0.84 | 0.73 |
| SBM-SVM | K = 10 | 0.80 | 0.79 | 0.78 |
| SBM-Bayes | K = 6 | 0.71 | 0.73 | 0.67 |
| MMSB-SVM | K = 20 | 0.82 | 0.82 | 0.80 |
| MMSB-Bayes | K = 10 | 0.71 | 0.67 | 0.80 |
| DCSB-SVM | K = 20 | 0.78 | 0.83 | 0.70 |
| DCSB-Bayes | K = 6 | 0.54 | – | – |

Table 2: Classifier performance on Bipolar data.

| Type | K | Accuracy | Precision | Recall |
|---|---|---|---|---|
| Baseline SVM | – | 0.58 | 0.58 | 0.57 |
| SBM-SVM | K = 18 | 0.56 | 0.62 | 0.59 |
| SBM-Bayes | K = 16 | 0.60 | 0.62 | 0.62 |
| MMSB-SVM | K = 16 | 0.61 | 0.69 | 0.53 |
| MMSB-Bayes | K = 14 | 0.56 | 0.60 | 0.86 |
| DCSB-SVM | K = 18 | 0.60 | 0.63 | 0.64 |
| DCSB-Bayes | K = 24 | 0.50 | 0.45 | 0.90 |

## 5. RESULTS

We ran the cross-validation procedure for each value of $K \in \{6, 8, \ldots, 24\}$. We here show the best accuracy for both the Bayesian Classifier and the empirical interaction matrix SVMs, as well as the baseline SVM. Overall, the ADNI dataset is a much easier classification task than the Bipolar Disorder dataset.

As can be seen in Table 1, each Blockmodel-SVM perform approximately as well as the Baseline SVM. Furthermore, in two of three cases the Bayesian classifier also performed moderately well; this is qualitatively significant considering the fact they are not explicitly optimized for classification accuracy. The Degree Corrected Blockmodel's Bayesian classifier, shown on the last row of the table, classifies some folds as entirely one class, invalidating Precision and Recall scores.

As can be seen in Table 2, a similar result was produced for the Bipolar dataset as was produced for the ADNI dataset, albeit with much lower accuracy overall. This may be due to data set noise (the data were collected on different MRI machines, and age of the bipolar cohort has much higher variance), but also may be attributed to the disease itself. While Alzheimer's disease is clearly structurally degenerative, Bipolar disorder is not nearly as well understood.

### 5.1 Discussion

It is interesting to note that in almost every case the Bayesian classifier performed better with *fewer* blocks. Especially clear in the Alzheimer's Disease dataset results, there is a trend toward small $K$ values for the Bayesian classifier, while the SVM has optimal performance at higher values of $K$.

Unfortunately we should not expect the baseline SVM to be greatly outperformed by any of the Blockmodel-SVM methods. Since the empirical interaction matrices are essentially linear combinations of the rows and then the columns of the original weight matrices, the set of possible classifiers from the Blockmodel-SVM method is a subset of those from the baseline SVM. In this context, the Blockmodel-SVM could be thought of as having an enforced coefficient structure in the original weight matrix space; this clearly aids in classification somewhat, but does not lead to drastic increases in accuracy.

## 6. CONCLUSION AND FUTURE WORK

In this paper we have defined a novel variant of the Stochastic blockmodel. We have provided evidence empirically that it and several related variants may be useful tools in the analysis of brain networks, also known as connectomes.

This being said, however, there are many things left open. In particular the selection of the number of blocks remains an open problem; while information criteria may be occasionally useful, it is not clear how to treat the degree correction term, nor is it clear which criterion is best. Beyond this, the analysis of the discovered communities may provide scientific insight; for larger datasets the block groupings should be explored and analyzed as groupings of neurological regions.

## ACKNOWLEDGMENTS

This work was supported in part by NIH Grant U54 EB020403: ENIGMA Center for Worldwide Medicine, Imaging & Genomics, as well as the Rose Hills Fellowship at the University of Southern California. The authors would further like to thank the helpful comments of the anonymous reviewers, as well as Neda Jahanshad, Talia Nir, and Julio Villalon for multiple helpful conversations.

## REFERENCES

[1] Wang, Y. J. and Wong, G. Y., "Stochastic blockmodels for directed graphs," *Journal of the American Statistical Association* **82**(397), 8–19 (1987).

[2] Karrer, B. and Newman, M. E., "Stochastic blockmodels and community structure in networks," *Physical Review E* **83**(1), 016107 (2011).

[3] Gopalan, P. K. and Blei, D. M., "Efficient discovery of overlapping communities in massive networks," *Proceedings of the National Academy of Sciences* **110**(36), 14534–14539 (2013).

[4] Airoldi, E. M., Blei, D. M., Fienberg, S. E., and Xing, E. P., "Mixed membership stochastic blockmodels," in [*Advances in Neural Information Processing Systems*], 33–40 (2009).

[5] Pavlovic, D. M., Vértes, P. E., Bullmore, E. T., Schafer, W. R., and Nichols, T. E., "Stochastic blockmodeling of the modules and core of the caenorhabditis elegans connectome," *PloS one* **9**(7), e97584 (2014).

[6] Hinne, M., Ekman, M., Janssen, R. J., Heskes, T., and van Gerven, M., "Probabilistic clustering of the human connectome identifies communities and hubs," *PloS one* **10**, e0117179 (2015).

[7] Sporns, O., Tononi, G., and Kötter, R., "The human connectome: a structural description of the human brain," *PLoS computational biology* **1**(4), e42 (2005).

[8] Rubinov, M. and Sporns, O., "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage* **52**(3), 1059–1069 (2010).

[9] Bullmore, E. and Sporns, O., "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience* **10**(3), 186–198 (2009).

[10] Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., and Sporns, O., "Mapping the structural core of human cerebral cortex," *PLoS biology* **6**(7), e159 (2008).

[11] Jbabdi, S., Woolrich, M. W., and Behrens, T. E. J., "Multiple-subjects connectivity-based parcellation using hierarchical dirichlet process mixture models," *NeuroImage* **44**(2), 373–384 (2009).

[12] Moyer, D., Gutman, B., Prasad, G., Ver Steeg, G., and Thompson, P., "Mixed membership stochastic blockmodels for the human connectome," in [*2nd International Workshop on Bayesian and Grahical Models for Biomedical Imaging–BAMBI*],

[13] Daudin, J.-J., Picard, F., and Robin, S., "A mixture model for random graphs," *Statistics and computing* **18**(2), 173–183 (2008).

[14] Wasserman, S. and Anderson, C., "Stochastic a posteriori blockmodels: Construction and assessment," *Social Networks* **9**(1), 1–36 (1987).

[15] Nowicki, K. and Snijders, T. A. B., "Estimation and prediction for stochastic blockstructures," *Journal of the American Statistical Association* **96**(455), 1077–1087 (2001).

[16] Trojanowski, J. Q., Vandeerstichele, H., Korecka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., Blennow, K., Soares, H., Simon, A., Lewczuk, P., et al., "Update on the biomarker core of the alzheimer's disease neuroimaging initiative subjects," *Alzheimer's & Dementia* **6**(3), 230–238 (2010).

[17] Prasad, G., Joshi, S. H., and Thompson, P. M., "Optimizing brain connectivity networks for disease classification using epic," in [*Proceedings/IEEE International Symposium on Biomedical Imaging: from nano to macro. IEEE International Symposium on Biomedical Imaging*], **2014**, 834, NIH Public Access (2014).

[18] Jenkinson, M., Bannister, P., Brady, M., and Smith, S., "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage* **17**(2), 825–841 (2002).

[19] Aganj, I., Lenglet, C., Jahanshad, N., Yacoub, E., Harel, N., Thompson, P. M., and Sapiro, G., "A hough transform global probabilistic approach to multiple-subject diffusion mri tractography," *Medical image analysis* **15**(4), 414–425 (2011).

[20] Fischl, B., "Freesurfer," *Neuroimage* **62**(2), 774–781 (2012).

[21] Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C., "A reproducible evaluation of ants similarity metric performance in brain image registration," *Neuroimage* **54**(3), 2033–2044 (2011).

[22] Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., Nimmo-Smith, I., and Contributors, D., "Dipy, a library for the analysis of diffusion mri data," *Frontiers in neuroinformatics* **8** (2014).