

# Comparing Structural Brain Connectivity by the Infinite Relational Model

Karen S. Ambrosen<sup>\*†</sup>, Tue Herlau<sup>\*</sup>, Tim Dyrby<sup>†</sup>, Mikkel N. Schmidt<sup>\*</sup> and Morten Mørup<sup>\*</sup>

<sup>\*</sup>*Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark*

<sup>†</sup>*Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark*

*Email: kmsa@dtu.dk, tuhe@dtu.dk, timd@drcmr.dk, mnsch@dtu.dk, mmor@dtu.dk*

**Abstract**—The growing focus in neuroimaging on analyzing brain connectivity calls for powerful and reliable statistical modeling tools. We examine the Infinite Relational Model (IRM) as a tool to identify and compare structure in brain connectivity graphs by contrasting its performance on graphs from the same subject versus graphs from different subjects. The inferred structure is most consistent between graphs from the same subject; however, the model is able to predict links in graphs from different subjects on par with results within a subject. The framework proposed can be used as a statistical modeling tool for the identification of structure and quantification of similarity in graphs of brain connectivity in general.

**Keywords**—Neuroimaging, Bayesian Methods, Structural Connectivity, Relational Modelling

## I. INTRODUCTION

The human connectome [1], [2] constitutes a formidable network formed by trillions of connections between billions of neurons [1]. While current technologies cannot measure the full human connectome, functional and diffusion magnetic resonance imaging are key non-invasive techniques for measuring brain connectivity at a spatial resolution in the order of cubic-millimeters. Functional connectivity can be estimated by quantifying similarity between blood oxygen level dependent (BOLD) responses between brain regions [3], [4], whereas structural connectivity between gray-matter regions can be derived from tractography approaches, see also [5].

Using tools from network science, researchers have analyzed graphs of brain connectivity in terms of their functional segregation and integration as quantified by graph measures such as the clustering coefficient and shortest path properties, see also [4], [6]. In [7] structural connectivity graphs were derived between 998 regions of interests (ROI) spanning the whole brain for five subjects (one subject was scanned twice) using tractography based on diffusion spectrum imaging. These graphs were found to include a structural core as well as distinct structural modules [7].

With the growing focus in neuroimaging on modeling graphs of brain connectivity, there is a need for powerful and reliable statistical modeling tools that can identify latent structure. A further challenge is to compare different connectivity graphs, e.g. to assess similarities across different subjects, measuring modalities, etc. The Infinite Relational

Model (IRM) [8]–[10] is a probabilistic generative model of structure in relational data (graphs), in which the nodes of the graph are partitioned into groups with statistically similar connectivity patterns. The number of groups is automatically inferred from data. The IRM can be used to quantify how similar two brain connectivity graphs are either by comparing the group structure estimated for two graphs or by fitting the model on one graph and using the result to predict the other graph, where a low prediction error indicates that the graphs are similar.

In this paper we discuss the following question: *Can the infinite relational model reliably be used to estimate latent group structure and quantify the similarity between brain connectivity graphs?* We address this by contrasting the performance of the IRM on graphs from the same subject versus graphs from different subjects, expecting that similarity should be greater on graphs from the same subject than on graphs from different subjects. As the inference in the IRM is based on Markov chain Monte Carlo (MCMC) we use multiple restarts to assess potential mixing issues of the sampler. To compare graphs we examine normalized mutual information as a measure of consistency of the inferred group structure and predictive log-likelihood and area under curve (AUC) of the receiver operator characteristic to estimate how well a model fitted on one graph can predict another graph. The proposed framework extends to the modeling of other types of brain networks and forms a principled statistical modeling tool for quantifying both the number of functional and structural units in brain networks as well as comparing brain connectivity in general.

## II. METHODS

The infinite relational model is a non-parametric Bayesian generative model for complex networks independently proposed in [8], [9]. The model is an extension of the stochastic block model [11] to include an unbounded number of clusters based on the Chinese Restaurant Process (CRP) (see also [12] for an introduction to the IRM.) The generative model for the IRM is given by

$$\begin{aligned} z &\sim \text{CRP}(\alpha), && \textit{Groups} \\ \eta_{lm} &\sim \text{Beta}(\beta^+, \beta^-), && \textit{Interactions} \\ A_{ij} &\sim \text{Bernoulli}(\eta_{z_i, z_j}), && \textit{Links}. \end{aligned}$$

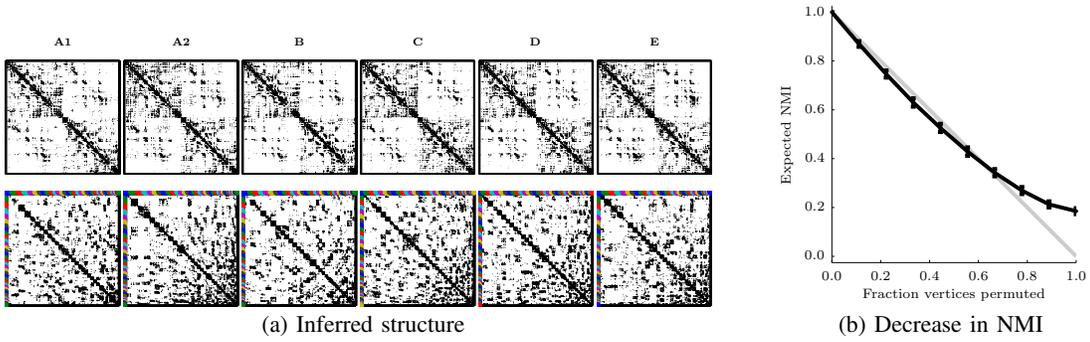


Figure 1: *Left*: Graphs before and after ordering into the maximum likelihood grouping. *Right*: NMI between grouping and itself after a fraction of vertices has been permuted. NMI roughly correspond to the fraction of permuted vertices.

The model partitions the nodes into groups  $\mathbf{z}$  ( $z_i = m$  means node  $i$  is assigned to group  $m$ ). Links are formed between nodes in groups  $l$  and  $m$  independently with probability  $\eta_{lm}$ , and  $\alpha$ ,  $\beta^+$ , and  $\beta^-$  are hyperparameters of the model. As the beta prior on  $\boldsymbol{\eta}$  is conjugate to the Bernoulli likelihood this parameter can be collapsed such that

$$p(\mathbf{A}, \mathbf{z} | \beta^+, \beta^-, \alpha) = \int d\boldsymbol{\eta} p(\mathbf{A} | \mathbf{z}, \boldsymbol{\eta}) p(\mathbf{z}, \boldsymbol{\eta} | \beta^+, \beta^-, \alpha)$$

$$= \left[ \prod_{l \geq m} \frac{B(n_{lm}^+ + \beta^+, n_{lm}^- + \beta^-)}{B(\beta^+, \beta^-)} \right] \left[ \frac{\Gamma(\alpha) \alpha^K}{\Gamma(\alpha + I)} \prod_{k=1}^K \Gamma(n_k) \right],$$

where  $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$  is the normalization constant of the beta distribution,  $n_m$  is the number of nodes in group  $m$ , and  $n_{lm}^+ = \sum_{ij} A_{ij} \delta_{l, z_i} \delta_{m, z_j} / 2^{\delta_{lm}}$  (and similar for  $n_{lm}^-$  with  $A_{ij}$  replaced by  $(1 - A_{ij})$ ) denotes the number of links and non-links between group  $l$  and  $m$ . We use the notation  $\beta = \beta^+ + \beta^-$  and  $n_{lm} = n_{lm}^+ + n_{lm}^-$ . Based on this,  $\mathbf{z}$  can be inferred by MCMC. We use Gibbs sampling in combination with split-merge moves similar to [8].

#### A. Quantifying graph similarity by the IRM

To compare the similarity of graphs based on the IRM, the estimated group structure can be compared directly, or by exploiting that the IRM model is a generative model, a model fitted on one graph can be used to predict other graphs. We compare the following three measures to assess similarity between graphs: normalized mutual information (NMI) between the inferred clustering structure of the graphs, the predictive log-likelihood, and the area under curve (AUC) of the receiver operator characteristic. These three approaches are described below.

1) *Normalized Mutual Information (NMI)*: The NMI between two group structures  $\mathbf{z}$  and  $\mathbf{z}^*$  is given by

$$\text{NMI}(\mathbf{z}, \mathbf{z}^*) = \frac{2\text{I}(\mathbf{z}, \mathbf{z}^*)}{\text{I}(\mathbf{z}, \mathbf{z}) + \text{I}(\mathbf{z}^*, \mathbf{z}^*)},$$

where  $\text{I}(\cdot, \cdot)$  is the mutual information defined by  $\text{I}(\mathbf{z}, \mathbf{z}^*) = \prod_{l, m} p(l, m) \log \left( \frac{p(l, m)}{p(l)p(m)} \right)$  and  $p(l, m)$  is the probability that a node in cluster  $l$  in  $\mathbf{z}$  is in cluster  $m$  in  $\mathbf{z}^*$ . For  $\text{I}(\mathbf{z}, \mathbf{z})$

this reduces to the entropy  $H(\mathbf{z}) = -\sum_m p(m) \log p(m)$ . An important property of mutual information is that it is invariant to permutation of the extracted groups. NMI is bounded by  $[0, 1]$  where 0 indicates that the two group assignments are independent whereas 1 indicates the two groupings are identical up to permutation [13].

2) *Predictive log-likelihood*: To quantify how similar the structure of links are in two graphs we can evaluate how well the IRM inferred on graph  $\mathbf{A}$  predicts the graph  $\mathbf{A}^*$ . The expected predictive log-likelihood for a given group structure  $\mathbf{z}$  is given by

$$\langle \log p(\mathbf{A}^* | \mathbf{z}, \mathbf{A}, \beta^+, \beta^-, \alpha) \rangle_{p(\boldsymbol{\eta} | \mathbf{A}, \mathbf{z})} = \sum_{i > j} A_{ij}^* \langle \log \eta_{z_i, z_j} \rangle_{p(\boldsymbol{\eta} | \mathbf{A}, \mathbf{z})} + (1 - A_{ij}^*) \langle \log(1 - \eta_{z_i, z_j}) \rangle_{p(\boldsymbol{\eta} | \mathbf{A}, \mathbf{z})}$$

where the expectations are  $\langle \log(\eta_{z_i, z_j}) \rangle = \psi(n_{lm}^+ + \beta^+) - \psi(n_{lm} + \beta)$  and  $\langle \log(1 - \eta_{z_i, z_j}) \rangle = \psi(n_{lm}^- + \beta^-) - \psi(n_{lm} + \beta)$  and  $\psi$  is the digamma function  $\psi(x) = \frac{d \log \Gamma(x)}{dx}$ . Averaged over the posterior samples from the MCMC run, this can be used as a predictive similarity measure.

3) *Area Under Curve (AUC)*: An alternative measure for prediction is based on the extend to which the probabilities of generating links inferred by the IRM in graph  $\mathbf{A}$  can be used to separate the class of links and non-links in graph  $\mathbf{A}^*$ . The expected probability of generating a link between node  $i$  and  $j$  is given by

$$\langle p(A_{ij}^* = 1 | \mathbf{z}, \mathbf{A}, \beta^+, \beta^-, \alpha) \rangle_{p(\boldsymbol{\eta} | \mathbf{A}, \mathbf{z})} = \frac{n_{lm}^+ + \beta^+}{n_{lm} + \beta}.$$

This probability can be used as a scoring function ( $s(i, j)$ ) for computing the AUC which is bounded by  $[0, 1]$  where 1 corresponds to a perfect separation of the link and non-links by the scoring function  $s(i, j)$  whereas 0.5 means that the scoring function is no better than chance. A benefit of the AUC is that it is invariant to class-imbalance issues. The AUC is therefore widely used as measure of performance in link-prediction tasks, see also [14].

#### B. Data and experimental setup

The human cortex connectivity dataset [7] available from <http://cmck.org/viewer/datasets> is used. The dataset consists

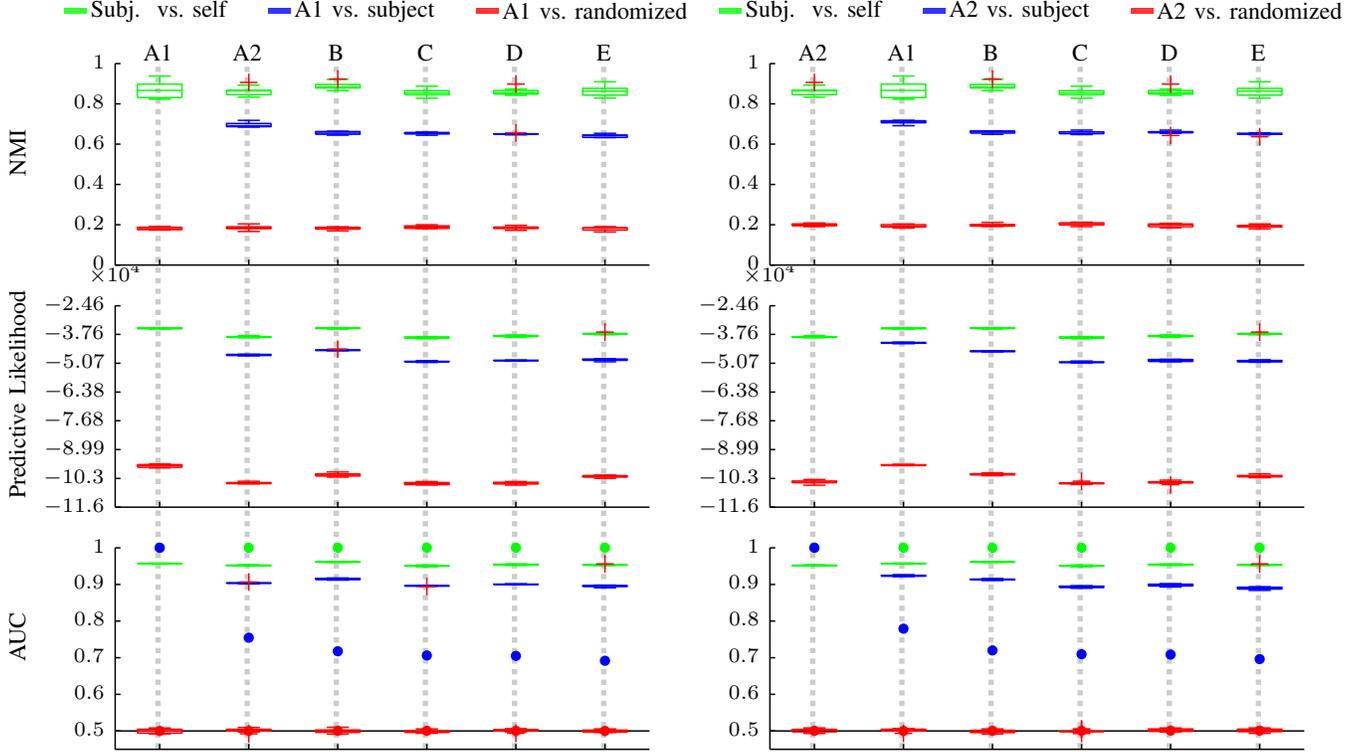


Figure 2: Each column indicates similarity between scan A1 (or A2) and the other scans according to NMI, predictive likelihood and AUC. The green boxes indicate variability across the 10 MCMC restarts, the blue boxes how well subject A1 (or A2) generalize across the other scans and the red is a baseline given by random permutations. The red crosses are outliers. The dots in the lower row correspond to naively predicting  $A^* = A$ .

of six structural connectivity graphs obtained from tractography based on diffusion spectrum imaging (DSI) from five subjects [7]. Graph A1 and A2 are obtained from two different scans of the same subject. The graphs have  $J = 998$  nodes and were symmetrized and binarized before our analysis. The number of links and the graph densities are listed in Table I. The MCMC inference is initialized at random and is run for 50 000 iterations. Every 25th sample is saved, resulting in 2 000 samples. 10 MCMC restarts are made for each graph. The priors are selected as  $\beta^+ = \beta^- = 1$  and  $\alpha = \log(J)$ . The number of clusters is initialized uniformly at random between 1 and 200.

### III. RESULTS

The similarity of the groupings of nodes is found by calculating the NMI between the assignment matrices from MCMC restart 1 and 2, 2 and 3, ..., 10 and 1. The NMI for each subject versus itself is shown in Figure 2 as green boxes indicating an upper bound on the similarity. Instead of averaging NMI over the posterior distribution, we use the single posterior sample with the highest likelihood, thus the NMI for a graph versus itself should in theory be equal to one. The blue boxes are the NMI for MCMC restart 1 between subject A1 (A2 in second column) and all the other subjects—this indicates the estimated similarity

Subject	No. of links	Graph density [%]
A1	27 040	2.71
A2	29 730	2.98
B	28 444	2.86
C	29 866	3.00
D	29 702	2.98
E	28 744	2.89

Table I: The number of links and density of the 6 graphs.

between subjects. The red boxes show NMI between A1 (A2) and a random permutation of each subject indicating a lower bound on the similarity. The NMI within a subject is around 0.85 corresponding to a fraction of 10% of the nodes are permuted, as shown in Figure 1b. The reason why the NMI within a subject is less than 1 can be attributed to lack of mixing of the MCMC sampler making it unable to identify the same highest likelihood solution. This indicates that the MCMC sampler should either be run for a much larger number of iterations, which may be impractical, or that more efficient inference procedures should be devised. Nonetheless, the results gives an indication of the magnitude of error due to lack of mixing. The NMI between subject A1 and A2 is slightly higher than NMI between any other combinations. This indicates that the graph structure is more similar within a subject across scans than between subjects, but further investigations are needed to confirm the result.

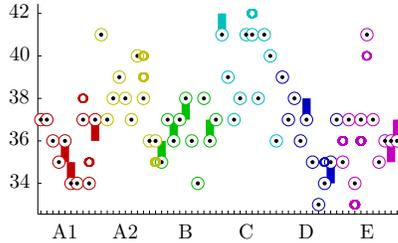


Figure 3: Boxplot of the number of groups in each of the 10 MCMC restarts per subject.

The NMI between subjects is well above random, suggesting a common latent structure between subjects.

The predictive log-likelihood and the AUC between subject A1 (A2 in second column) and the other subjects are found using every 25th sample of the last 25 000 samples using subject A1 (A2) as training data. The predictive log-likelihood and AUC is shown as blue boxes, the green boxes indicate how a subject predicts its own graph, and the red boxes show how A1 (A2) predicts the other graphs randomly permuted. For reference, the dots in the AUC plot indicate baseline results when naively predicting that graphs are equal. The predicted log-likelihood and AUC between subjects is well above random, again supporting a common latent structure; however, when training the model on subject A1 (left column), subject B has the highest predictive log-likelihood and AUC. This might be due to differences in graph density since A1 has the lowest graph density and graph B the second lowest density, but it also shows that predictive performance should be used with caution to assess graph similarity. When training the model on subject A2 (right column), subject A1 has the highest predictive log-likelihood and AUC: Here, as expected, the predictive performance is best within a subject.

Figure 3 shows a boxplot of the number of clusters in each MCMC restart where the color indicates the different subjects. Each box shows the distribution of the number of clusters from every 25th sample of the last 25 000 samples; however, most are centered on a single number again indicating that the MCMC sampler does not mix properly. Nevertheless, both within each subject and across subjects the number of components is fairly consistent.

#### IV. CONCLUSION

We proposed a framework for comparing graphs of brain connectivity based on the structure inferred by the infinite relational model. We tested the framework on six benchmark structural connectivity graphs derived from diffusion spectrum imaging and found that all the networks were consistent both within and between subjects. The inferred structure appeared to be slightly more consistent as quantified by NMI within a subject than from this subject to the other four subjects. However, it was observed that the inferred models predict structural connectivity equally well within a subject

as across subjects. In particular, the structure inferred were significantly more consistent than would be expected by random and also more consistent than predicting on the raw graphs. We believe the proposed framework has promising applications for identifying structure and comparing brain connectivity data in general.

#### V. ACKNOWLEDGEMENT

This work is funded by the Lundbeck Foundation.

#### REFERENCES

- [1] O. Sporns, G. Tononi, and R. Kötter, "The human connectome: a structural description of the human brain," *PLoS computational biology*, vol. 1, no. 4, p. e42, 2005.
- [2] P. Hagmann, "From diffusion mri to brain connectomics," Ph.D. dissertation, 2005.
- [3] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
- [4] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.
- [5] A. Daducci, S. Gerhard, A. Griffa, A. Lemkaddem, L. Cammoun, X. Gigandet, R. Meuli, P. Hagmann, and J.-P. Thiran, "The connectome mapper: An open-source processing pipeline to map connectomes with mri," *PLoS ONE*, vol. 7, 12 2012.
- [6] D. S. Bassett and E. Bullmore, "Small-world brain networks," *The neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.
- [7] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *PLoS biology*, vol. 6, no. 7, p. e159, 2008.
- [8] C. Kemp, J. B. Tenenbaum, T. L. Griffiths, T. Yamada, and N. Ueda, "Learning systems of concepts with an infinite relational model," in *Proceedings of the national conference on artificial intelligence*, vol. 21, no. 1. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006, p. 381.
- [9] Z. Xu, V. Tresp, K. Yu, and H.-P. Kriegel, "Learning infinite hidden relational models," in *Proceedings of the 22nd International Conference on Uncertainty in Artificial Intelligence*, 2006.
- [10] K. W. Andersen, M. Morup, H. Siebner, K. H. Madsen, and L. K. Hansen, "Identifying modular relations in complex brain networks," in *Machine Learning for Signal Processing (MLSP), 2012 IEEE International Workshop on*. IEEE, 2012, pp. 1–6.
- [11] K. Nowicki and T. A. B. Snijders, "Estimation and prediction for stochastic blockstructures," *Journal of the American Statistical Association*, vol. 96, no. 455, pp. 1077–1087, 2001.
- [12] M. N. Schmidt and M. Mørup, "Non-parametric bayesian modeling of complex networks," *Signal Processing Magazine, IEEE*, vol. 30, no. 3, pp. 110–128, 2013.
- [13] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *The Journal of Machine Learning Research*, vol. 11, pp. 2837–2854, 2010.
- [14] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2011.