# IDENTIFYING MODULAR RELATIONS IN COMPLEX BRAIN NETWORKS

*Kasper Winther Andersen*[a,b], *Morten Mørup*[a], *Hartwig Siebner*[b], *Kristoffer H Madsen*[b], *Lars Kai Hansen*[a]

[a]DTU Informatics, Technical University of Denmark.
[b]Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre

## ABSTRACT

We evaluate the infinite relational model (IRM) against two simpler alternative nonparametric Bayesian models for identifying structures in multi subject brain networks. The models are evaluated for their ability to predict new data and infer reproducible structures. Prediction and reproducibility are measured within the data driven NPAIRS split-half framework. Using synthetic data drawn from each of the generative models we show that the IRM model outperforms the two competing models when data contain relational structure. For data drawn from the other two simpler models the IRM does not overfit and obtains comparable reproducibility and predictability. For resting state functional magnetic resonance imaging data from 30 healthy controls the IRM model is also superior to the two simpler alternatives, suggesting that brain networks indeed exhibit universal complex relational structure in the population.

***Index Terms***— Infinite Relational Model, Complex Networks, fMRI

## 1. INTRODUCTION

The brain is composed of about $10^{11}$ neurons connected by more than $10^{14}$ synapses forming an almost unfathomable complex network [1]. Neuroimaging, e.g., by functional magnetic resonance imaging (fMRI), has the potential to identify structures in the brain network. fMRI allows one to measure the blood oxygen level and thereby an indirect measure of the neuronal activity in a relatively fine spatial resolution. As the signal-to-noise ratio in fMRI is rather poor we are in many cases forced to average within a population of subjects to obtain reproducible results. This averaging further limits the spatial resolution of our models as considerable individual variability exist in both anatomy and function [2].

While the dominant paradigm in functional neuroimage analysis has been to identify local components of the network based on the functional segregation hypothesis, see e.g. [3], there is a current move towards modeling more global properties of the brain network including both functional and structural aspects, for reviews see, e.g., [4, 5]. Functional brain network models typically summarize temporal relations between regions of the brain whereas structural networks are obtained by, e.g., using tractography methods to track the white matter bundles connecting different brain regions. One of the challenges in brain imaging is to understand how brain modules identified under the functional segregation approach work together to perform important information processing tasks such as perception, cognition, and decision making. We currently pursue a strategy which is designed to mediate between functional localization and global description [6]. The main idea is to search for so-called 'communities' of nodes characterized as groups of nodes that share certain connectivity properties, see e.g. [7] for background and references on community detection.

In neuroimaging models we typically face two equally important objectives, namely on one hand to identify predictive models, i.e., models that generalize to new data, and on the other hand that the structures we learn are trustworthy, e.g., being reproducible across different subject groups. The two objectives are quantified in the so-called NPAIRS split-half resampling framework [8]. This framework has been used for evaluation of Bayesian models before, e.g., in [9] in which generalizability and reproducibility of the parameter posterior distribution were of interest, however, the framework has not been used before for evaluation of community detection models. Here we for the first time investigate the predictability and reproducibility of three Bayesian nonparametric models which are able to infer community structure in complex networks.

The models are the infinite relational model (IRM) [10, 11], the infinite diagonal model (IDM), and the model proposed by [12] which will be termed infinite Hofman-Wiggins (IHW) model, the term infinite referring to the fact the complexity in terms of the number of communities is formally unlimited. The three models vary in expressiveness, the IRM allows different link characteristics within and between components and are thereby able to infer general relations between communities, the IDM is parametrized with different link probabilities within communities but have a single common between community link probability, while the link probabilities in IHW are described with only two parameters; one within and one between component link probability.

As mentioned we are interested in detecting community

structure shared by a population of fMRI subjects. Thus while we estimate the functional networks in each individual independently, all three models represent all subjects by a common community structure and common link probabilities. An additional benefit of this approach, compared to individual models investigated [6] is that it become possible to compute test likelihood for unseen data (subjects). Also, having common link probabilities the latent variables of the model do not scale with the number of subjects, hence, allow for analysis of a large number of subjects.

*In conclusion: The main contributions of this paper are i) a demonstration that a data driven predictability and reproducibility framework for evaluation of neuroimaging models can be adapted to Bayesian community detection and ii) use of the framework to test whether the expressiveness of the full IRM is needed to describe the functional networks of the resting state as measured by fMRI.*

## 2. METHODS

Data in this paper consists of $N$ undirected graphs, one per subject, defined on a common set of nodes and described by individual adjacency matrices $\mathbf{A}^{(n)}$ for subject $n$. $A^{(n)}_{i,j} = 1$ if nodes $i$ and $j$ have similar hemodynamic responses and zero otherwise.

### 2.1. Models

Formally we write the three generative models as:

$$\begin{aligned} \mathbf{Z}|\alpha &\sim \mathrm{CRP}(\alpha) \\ \rho_{a,b}|\boldsymbol{\beta}^+,\boldsymbol{\beta}^- &\sim \mathrm{Beta}(\beta^+_{a,b},\beta^-_{a,b}) \\ A^{(n)}_{i,j}|\mathbf{Z},\boldsymbol{\rho} &\sim \mathrm{Bernoulli}(\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}^\top_{j_r}). \end{aligned}$$

$\mathbf{Z}$ is the $[J \times D]$ binary matrix indicating group membership for each node and have a prior distribution given by the Chinese Restaurant Process (CRP) with the hyperprior parameter $\alpha$. $\mathbf{z}_{i_r}$ is the $i^{\text{th}}$ row vector of $\mathbf{Z}$. $\boldsymbol{\rho}$ is a symmetric matrix indicating the probability of links between each pair of components. Here $\boldsymbol{\rho}$ is shared across graphs. The $\boldsymbol{\rho}$ structure define the difference between the models compared in this paper, which will be described below. The link probabilities have a prior given by the Beta function with the two hyperprior parameters $\beta^+_{a,b}$ and $\beta^-_{a,b}$ giving the pseudocount of links and nonlinks between components $a$ and $b$ respectively. The adjacency matrix $\mathbf{A}^{(n)}$ for graph $n$ has the prior given by a Bernoulli trial biased by the link probability of links between the components in which the two nodes participate. Three different structures for the link probability matrix are considered. The full IRM model having a full upper triangular matrix, the infinite diagonal model (IDM) having unique elements in the diagonal and identical off-diagonal elements, and the model proposed in [12] having two parameters describing the within and between component link probabilities, respectively.

$$\boldsymbol{\rho}_{\mathrm{IRM}} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1} & \rho_{m2} & \cdots & \rho_{mm} \end{bmatrix} \tag{1}$$

$$\boldsymbol{\rho}_{\mathrm{IDM}} = \begin{bmatrix} \rho_1 & \rho_0 & \cdots & \rho_0 \\ \rho_0 & \rho_2 & \cdots & \rho_0 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_0 & \rho_0 & \cdots & \rho_m \end{bmatrix} \tag{2}$$

$$\boldsymbol{\rho}_{\mathrm{IHW}} = \begin{bmatrix} \rho_c & \rho_0 & \cdots & \rho_0 \\ \rho_0 & \rho_c & \cdots & \rho_0 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_0 & \rho_0 & \cdots & \rho_c \end{bmatrix} \tag{3}$$

The joint likelihood over graphs can be written as

$$\begin{aligned} &P(\mathbf{A}^{(1)},...,\mathbf{A}^{(N)}|\mathbf{Z},\boldsymbol{\rho}) \\ &= \prod_n \prod_{j>i} (\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}^\top_{j_r})^{A^{(n)}_{i,j}} (1-\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}^\top_{j_r})^{\left(1-A^{(n)}_{i,j}\right)} \\ &= \prod_{j\geq i} (\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}^\top_{j_r})^{\left(\sum_n A^{(n)}_{i,j}\right)} (1-\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}^\top_{j_r})^{\left(N-\sum_n A^{(n)}_{i,j}\right)} \end{aligned} \tag{4}$$

We note that the joint likelihood can be written efficiently using the aggregate adjacency matrix, $\mathbf{A}^{\mathrm{tot}} = \sum_n \mathbf{A}^{(n)}$.

#### *2.1.1. IRM Model Inference*

In this section we derive the posterior likelihood of a node's assignment given the assignments of the remaining nodes, for the IRM model, which is needed in the model inference. For the two other models please refer to [7]. By integrating out $\boldsymbol{\rho}$ from eq (4) the posterior can be written as

$$\begin{aligned} &P(\mathbf{A}^{\mathrm{tot}}|\mathbf{Z},\boldsymbol{\beta}^+,\boldsymbol{\beta}^-) \\ &= \int P(\mathbf{A}^{\mathrm{tot}}|\mathbf{Z},\boldsymbol{\rho})P(\boldsymbol{\rho}|\boldsymbol{\beta}^+,\boldsymbol{\beta}^-)\mathrm{d}\boldsymbol{\rho} \\ &= \prod_{a\geq b} \frac{\mathrm{Beta}(M^+_{a,b}+\beta^+_{a,b}, M^-_{a,b}+\beta^-_{a,b})}{\mathrm{Beta}(\beta^+_{a,b},\beta^-_{a,b})}, \end{aligned}$$

where $M^+_{a,b}$ is the total number of links and $M^+_{a,b}$ is the total number of non-links between the groups $a$ and $b$ across all graphs. Using Bayes' theorem the posterior likelihood can be found as

$$\begin{aligned} &P(\mathbf{Z}|\mathbf{A}^{(1)},...,\mathbf{A}^{(N)},\boldsymbol{\beta}^+,\boldsymbol{\beta}^-,\alpha) \propto \\ &P(\mathbf{A}^{\mathrm{tot}}|\mathbf{Z},\boldsymbol{\beta}^+,\boldsymbol{\beta}^-)P(\mathbf{Z}|\alpha) = \\ &\prod_{a\geq b} \frac{\mathrm{Beta}(M^+_{a,b}+\beta^+_{a,b}, M^-_{a,b}+\beta^-_{a,b})}{\mathrm{Beta}(\beta^+_{a,b},\beta^-_{a,b})} \\ &\times \left[\alpha^D \frac{\Gamma(\alpha)}{\Gamma(J+\alpha)} \prod_a \Gamma(n_a)\right], \end{aligned}$$

where $n_a$ is the number of nodes assigned to component $a$. The posterior likelihood for a node's assignment given the assignment of the remaining nodes is given by

$$P(Z_{i,a}=1|\mathbf{Z}\backslash\mathbf{z}_{i_r},\mathbf{A}^{(1)},...,\mathbf{A}^{(N)}) \propto \begin{cases} m_a \prod_b \frac{\text{Beta}(M^+_{a,b}+\beta^+_{a,b},M^-_{a,b}+\beta^-_{a,b})}{\text{Beta}(\beta^+_{a,b},\beta^-_{a,b})} \text{ if } m_a > 0 \\ \alpha \prod_b \frac{\text{Beta}(M^+_{a,b}+\beta^-_{a,b},M^-_{a,b}+\beta^-_{a,b})}{\text{Beta}(\beta^+_{a,b},\beta^-_{a,b})} \text{ otherwise.} \end{cases} \quad (5)$$

$m_a = \sum_{j\neq i} \mathbf{Z}(j,a)$ is the size of the $a^{th}$ functional group disregarding the assignment of the $i^{th}$ node. This posterior likelihood can be evaluated efficiently since we only need $\mathbf{M}^+$ and $\mathbf{M}^-$ and further to evaluate the Beta function for entries affected by the considered assignment change.

### 2.2. NPAIRS Evaluation Criteria

The models' predictability is evaluated using test log likelihood. The data is randomly split in 2 equally sized sets (S1 and S2) and inference is made separately on each set. The node assignment and link probabilities from the MAP solution are used to calculate the test log likelihood of the other (unseen) split. The average test log likelihood of the two splits is used as the predictability measure. The log likelihood of set S2 for the model inferred using S1 is given as

$$\log L(\mathbf{Z},\boldsymbol{\rho}|\mathbf{A}^{\text{S2},(1)},...,\mathbf{A}^{\text{S2},(N)}) = \frac{1}{N}\sum_{n=1}^{N}\sum_{j>i}\Big[A^{\text{S2},(n)}_{i,j}\log(\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}^\top_{j_r}) + (1-A^{\text{S2},(n)}_{i,j})\log(1-\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}^\top_{j_r})\Big]$$

Likewise, the reproducibility of the identified community structures is measured by the mutual information (MI) between the node assignments $\mathbf{Z}^{(S1)}$ and $\mathbf{Z}^{(S2)}$ of the MAP solution for each split,

$$\text{MI}(\mathbf{Z}^{(S1)},\mathbf{Z}^{(S2)}) = \sum_{i=1}^{D1}\sum_{j=1}^{D2} p(\mathbf{z}^{(S1)}_{i_r},\mathbf{z}^{(S2)}_{j_r})\log\left(\frac{p(\mathbf{z}^{(S1)}_{i_r},\mathbf{z}^{(S2)}_{j_r})}{p(\mathbf{z}^{(S1)}_{i_r})p(\mathbf{z}^{(S2)}_{j_r})}\right).$$

## 3. EXPERIMENTS AND RESULTS

In this section we present results for two different dataset. The first data set is synthetic data generated from each of the three different models. The second dataset is a dataset of resting state fMRI data from 30 healthy controls.

Model inference is based on a Gibbs sampling scheme in combination with split-merge Metropolis-Hastings updates [6, 10, 13]. Each node is initially assigned to one of 50 groups at random. The algorithm runs for 500 iterations, where at each iteration a Gibbs sampling scan is followed by a Split-Merge step. In the model inference $\alpha = \log(J)$ and $\boldsymbol{\beta}^+ = \beta^- = 1$.

### 3.1. Synthetic Data

Data from each of the three different generative models, IRM, IDM, and IHW, were generated. Assignments for $J = 100$ nodes was drawn from the CRP with $\alpha = 5$. Then the link probabilities between components was drawn from the Beta function using

$$\beta^+_{a,b} = \begin{cases} 2 \text{ if a=b} \\ 1 \text{ otherwise} \end{cases} \text{ and } \beta^-_{a,b} = \begin{cases} 3 \text{ if a=b} \\ 5 \text{ otherwise.} \end{cases}$$

constrained according to equations (1-3) for the three models. Then, using the component assignments and link probabilities 20 different adjacency matrices (graphs) per dataset was drawn using the Bernoulli function. For testing the models' predictability and reproducibility 500 different half-splits per dataset was generated. For each split each of the models was inferred and the predictability and reproducibility was evaluated as described in section 2.2.

Reproducibility-predictability plots are shown in Fig. 1 and the histogram of the number of estimated components are shown in Fig. 2. From Fig. 1 it is evident that for the IRM data with complex relations the IRM is superior in estimating the complex component relations as seen by having better reproducibility and predictability. The two other models produce comparable results with the IDM model having slightly better reproducibility. The IRM model is close in estimating the true number of components while the two other models underestimates the number of component with a factor 2.

For the IDM generated data the IDM model has better predictability and reproducibility than the IRM model, while the IHW model obtains slightly better reproducibility than the IDM model but with worse predictability. The better reproducibility of the IHW model should be seen in the light of overestimated number of components which increases the mutual information if the assignments of the between splits agree. The IRM underestimate the true number of components, which also cause the lower reproducibility.

For the IHW generated data all three models have comparable results with the IHW model having slightly better reproducibility than the IDM which is turn is slightly better than the IRM model. The IDM and IHW estimates the true number of components well, while the IRM model underestimates the true number of components.

### 3.2. Resting state fMRI data

Resting state functional magnetic resonance imaging (rs-fMRI) data from $N = 30$ healthy controls was recorded for 20 min (482 volumes) per subject. The first two volumes were discarded to account for T1 equilibrium effects, the remaining 480 volumes were realigned to the time-series mean

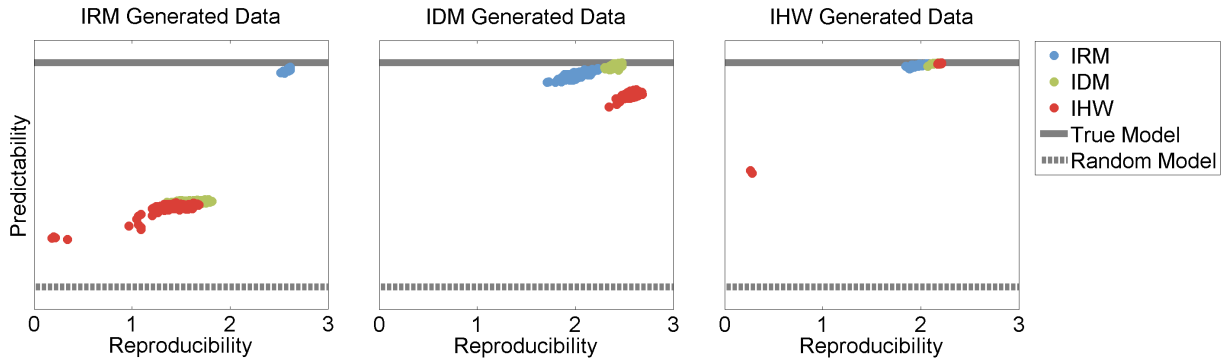


**Fig. 1**. Predictability (test log likelihood) as function of reproducibility (mutual information) for data generated with the IRM, IDM, and IHW models respectively. Solid lines indicate the log likelihood for the true model (which generated the data). Dotted lines indicate log likelihood for a random model in which all elements in the $\rho$ matrix are identical. For the IRM generated data the IRM model is superior compared to the two other models in both the predictability and reproducibility measures with the two other models having comparable results. For the IDM generated data the IDM is slightly better better than the IRM model and has a better predictability than the IHW, though with the IHW reproducing slightly better. For the IHW generated data all models are almost comparable though with slightly better results in favor of the simpler models.

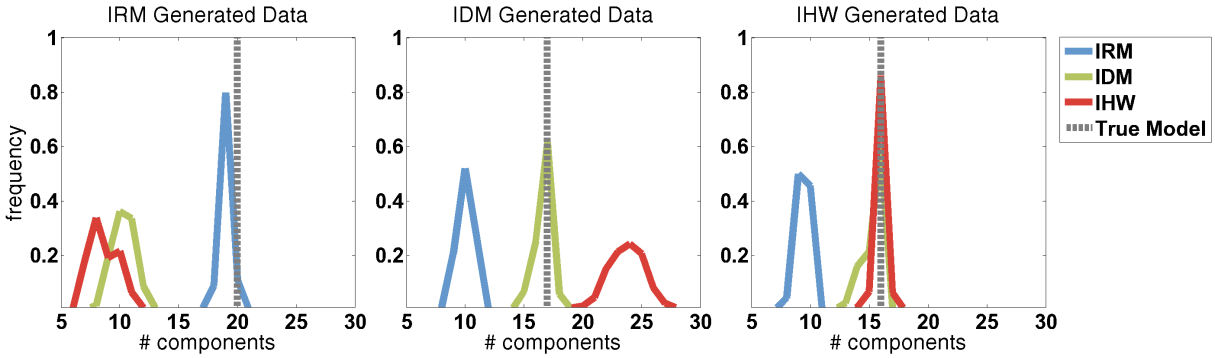


**Fig. 2**. Histogram of number of components estimated using the three models. Data was generated using either the IRM, IDM, or IHW generative model. For the IRM generated data the IRM model is close in estimating the true number of components (indicated with the vertical dotted line) while the two other models underestimate the number of components. For the IDM generated data the IDM model closely finds the right number of components while the IRM model underestimate and IHW overestimate the number of components. For the IHW generated data both the IDM and IHW finds the true number of components while the IRM underestimate.

and spatially normalized to the MNI template using SPM. Nuisance effects related to residual movement or physiological effects were removed using a linear filter comprised of 24 motion related and a total of 64 physiological effects including cardiac, respiratory, respiration volume over time, and time series from left and right hemispheres CSF and white matter voxels.

The mean signal in each of the 116 regions covered in the AAL database [14] was extracted and the $[116 \times 116]$ correlation matrix was created for each subject. These matrices were thresholded to include the top 1000 links and thereby formed the adjacency matrices. 200 different half-splits were generated and the reproducibility and predictability were calculated between the half-splits for each of the three models.

For visualization of the estimated communities a co-occurrence matrix of the count of each pair of nodes were in the same community were made. Based on this co-occurrence matrix we performed agglomerative hierarchical clustering using average linkage. This clustering was thresholded to include the number of communities corresponding to the median number found across the 400 half-splits. Fig. 3 shows the reproducibility-predictability plot and Fig. 4 plots the number of components estimated using each of the three models. The more complex IRM model interestingly yields better repro-

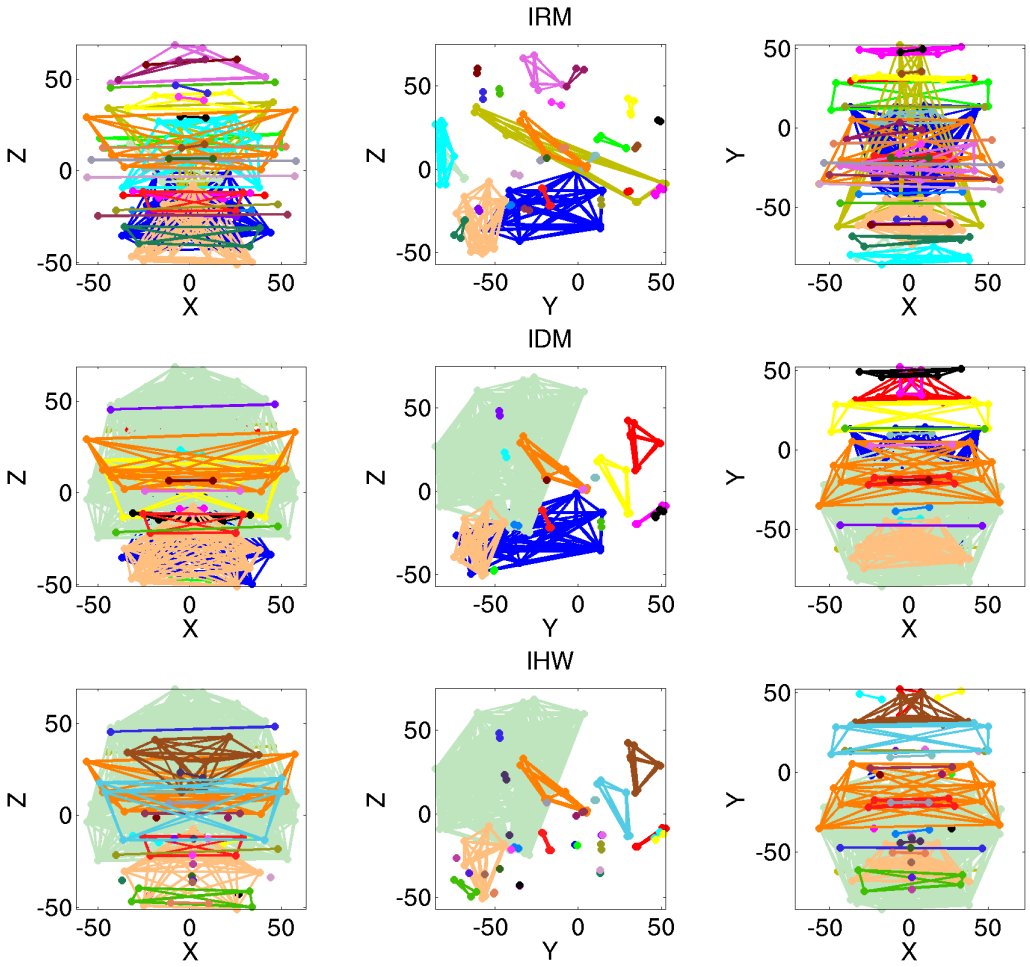


**Fig. 5**. Resting state fMRI data. Community structure for each of the three models. Top row IRM model, middle row IDM model, bottom row IHW model. For each model the consensus of 400 runs are plotted at three different planes in the 3 dimensional stereotaxic MNI space (X: left to right; Y: posterior to anterior; Z: ventral to dorsal).

ducibility and predictability compared with the other models. The IDM and IHW models yields comparable predictability but the IHW shows better reproducibility. Fig. 5 shows the layout of the identified communities in the stereotaxic MNI space in coronal, sagittal, and axial planes respectively. Lines connecting nodes in a community are shown with the same color. For all three models the communities are highly symmetric across hemispheres. The two simpler models IDM and IHW produce a huge cluster in the occipital and parietal lobes, while especially the IHW produces a large number of small communities. The huge component is segregated in several components with the IRM model, made possible by the more expressive parametrization.

## 4. DISCUSSION AND CONCLUSION

We used the NPAIRS framework for quantification of the predictability and reproducibility of three models for structural inference in complex brain networks. For synthetic data generated from each of the models we show that the IRM model identifies a general relational structure when such exist in network data. The two simpler models fail in both predicting unseen data and reproduce across datasets. When data is generated using the simpler models the IRM still produce comparable predictability and reproducibility results and thus does not seem to produce severe overfits.

Based on resting state fMRI data from 30 subjects extracted in 116 brain regions covering the whole brain we show that the IRM model is superior in terms of both predictability and reproducibility compared with the less expressive models. Visual inspection of the inferred communities indeed indicates that the richer representation of the IRM produces a more detailed yet reproducible connectivity pattern in the resting human brain.

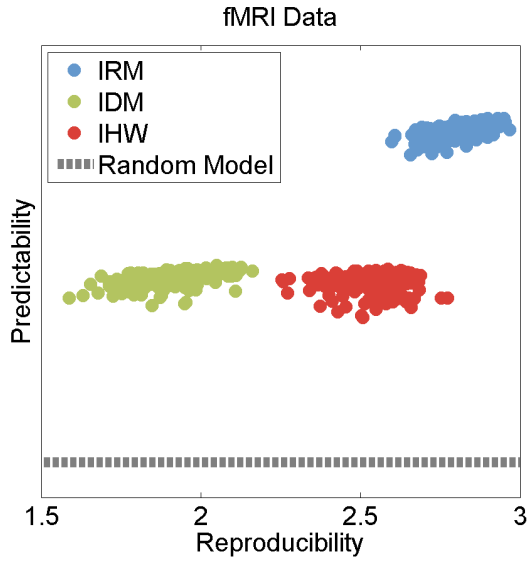


**Fig. 3**. Reproducibility-predictability plot of the fMRI dataset. Dotted line indicates log likelihood for a random model in which all elements in the $\rho$ matrix are identical.

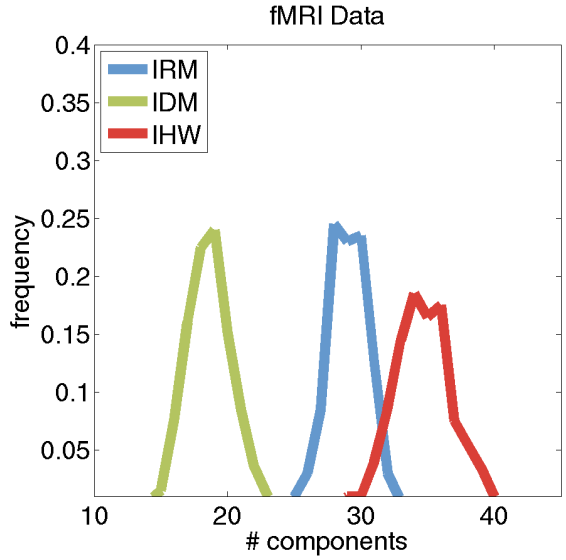


**Fig. 4**. Number of components estimated using the IRM, IDM, and IHW models for the 400 half-splits of the data.

## Acknowledgement

This work is funded by the Lundbeck Foundation (grant-nr R48 A4846).

### 5. REFERENCES

[1] S.M. Blinkov and I.I. Glezer, *The human brain in figures and tables: a quantitative handbook*, Basic Books New York:, 1968.

[2] A.M. Galaburda, G.D. Rosen, and G.F. Sherman, "Individual variability in cortical organization: its relationship to brain laterality and implications to function," *Neuropsychologia*, vol. 28, no. 6, pp. 529–546, 1990.

[3] K.J. Friston, A.P. Holmes, K.J. Worsley, J.P. Poline, C.D. Frith, and R.S.J. Frackowiak, "Statistical parametric maps in functional imaging: a general linear approach," *Human brain mapping*, vol. 2, no. 4, pp. 189–210, 1994.

[4] Olaf Sporns, "The human connectome: a complex network.," *Annals of the New York Academy of Sciences*, vol. 1224, no. 1, pp. 109–25, Apr. 2011.

[5] Edward T Bullmore and Danielle S Bassett, "Brain graphs: graphical models of the human brain connectome.," *Annual review of clinical psychology*, vol. 7, pp. 113–40, Apr. 2011.

[6] M Mørup, K.H. Madsen, A.M. Dogonowski, H. Siebner, and L.K. Hansen, "Infinite Relational Modeling of Functional Connectivity in Resting State fMRI," *Neural Information Processing Systems 23*, 2010.

[7] M. Mørup and M. N. Schmidt, "Bayesian community detection," *Neural Computation*, 2012.

[8] S.C. Strother, J. Anderson, L.K. Hansen, U. Kjems, R. Kustra, J. Sidtis, S. Frutiger, S. Muley, S. LaConte, and D. Rottenberg, "The quantitative evaluation of functional neuroimaging experiments: The npairs data analysis framework," *NeuroImage*, vol. 15, no. 4, pp. 747–771, 2002.

[9] D.J. Jacobsen, L.K. Hansen, and K.H. Madsen, "Bayesian model comparison in nonlinear bold fmri hemodynamics," *Neural computation*, vol. 20, no. 3, pp. 738–755, 2008.

[10] Charles 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*. 2006, vol. 21, pp. 381–388, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

[11] Zhao Xu, Volker Tresp, Kai Yu, and H.P. Kriegel, "Infinite hidden relational models," *In Proceedings of the 22nd International Conference on Uncertainity in Artificial Intelligence*, 2006.

[12] Jake M Hofman and Chris H Wiggins, "A Bayesian approach to network modularity, 2008," *Phys. Rev. Lett*, vol. 100, no. 25, pp. 1–4, 2008.

[13] Sonia Jain and Radford M Neal, "A Split-Merge Markov chain Monte Carlo Procedure for the Dirichlet Process Mixture Model," *Journal of Computational and Graphical Statistics*, vol. 13, no. 1, pp. 158–182, 2004.

[14] N Tzourio-Mazoyer, B Landeau, and D Papathanassiou, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuroimage*, vol. 15, pp. 273–289, 2002.