A METROPOLIS-WITHIN-GIBBS SAMPLER TO INFER TASK-BASED FUNCTIONAL BRAIN CONNECTIVITY

M Faizan Ahmad, James Murphy, Simon Godsill†  Deniz Vatansever, Emmanuel A Stamatakis*

† Department of Engineering, University of Cambridge, UK
* Department of Clinical Neuroscience, Addenbrookes Hospital, Cambridge, UK

ABSTRACT

Examining the dynamic aspects of functional networks in the brain is imperative in order to obtain a thorough description and to gain a better insight into its several features. Present methods of analysing brain data in task-conditions mainly include concatenation followed by temporal correlation. We employ Markov Chain Monte Carlo methods, namely Metropolis within Gibbs sampling, on a stochastic model to infer dynamic functional connectivity in such conditions. By using a Bayesian probabilistic framework, distributional estimates of the linkage strengths are obtained as opposed to point estimates, and the uncertainty of the existence of such links is accounted for. The methodology is applied to fMRI data from a finger opposition paradigm with task and fixation conditions, investigating the dynamics of the well characterised somato-motor network while using the visual network as a control case.

Index Terms— functional connectivity, fMRI, Gibbs sampling, Metropolis-Hastings, ROI

1. INTRODUCTION

Neuroimaging research has highlighted the significance of quantifying and tracking brain dynamics in order to obtain a description of the mechanisms behind its diverse set of operations. In this regard, the brain has been functionally categorised into distinct large-scale networks, typically through the application of functional connectivity to BOLD data collected during fMRI scanning, by identifying the statistical interdependencies of signals from remote brain regions [1][2][3].

In task-conditions, it is believed that the functional networks in the brain corresponding to the nature of the task will behave differently when the task is being executed in comparison to when it is not. To calculate functional connectivity, BOLD time-series data in such conditions is usually concatenated followed by the application of Independent Component Analysis (ICA) [4][5][6] or pairwise correlation [7][8][9]. Though simple to use and easily scalable to large sized networks, these methods are highly sensitive to noise and outliers, and do not incorporate uncertainty into their computations. Their results are simple point estimates which are very subjective to the experimental characteristics of the data used. We introduce a more sophisticated approach of inferring functional connectivity by employing Markov Chain Monte Carlo (MCMC) methods, namely Metropolis within Gibbs sampling, on a probabilistic model which is currently used to describe causal relations among neuronal states. This scheme sets the problem in distributional terms and accounts for experimental error as well as the uncertainty on prior assumptions.

We apply our methodology to fMRI data from a finger opposition paradigm with task and fixation conditions investigating dynamics of the somato-motor network, while using the visual network as a control case. Our results are able to capture the dynamic differences in the interactions of the motor network regions in comparison with the fixation condition, which are not present to the same extent in the visual network, hence verifying the efficacy of our method.

2. METHODS

2.1. Model

The BOLD signal \(x_{i,t}\) from a brain location \(i\) is modelled such that the change in the signal value at a given time \(t\) is the result of a summation of forces, which are dependent on BOLD signals from other nodes, linked through interaction parameters \(\phi_{i,j}\). This is described as:

\[
\Delta x_{i,t} = \left[\sum_{j=1}^{N} \Sigma_{j \neq i} \phi_{i,j} x_{j,t} - \gamma_i x_{i,t}\right] \Delta t + \sigma_i^2 W_{i,t},
\]

where \(x_{i,t}\) is magnitude of the BOLD signal at node \(i\) at time \(t\), \(\phi_{i,j}\) is the interaction parameter between nodes \(i\) and \(j\), \(S_{i,j} \in \{0, 1\}\) is an indicator variable to specify the existence of a link between nodes \(i\) and \(j\), \(W_{i,t}\) represents Brownian motion with unit variance, \(\sigma_i^2\) is the variance of this noise for node \(i\), \(\gamma_i\) is an optional mean-reverting or self-inhibition term to prevent the signal from taking very large values, and \(N\) is the total number of nodes in the network.

This model is akin that used widely in neuroscience for dynamic causal modelling DCM in which the underlying neuronal states and physiological parameters are inferred via the observed haemodynamic response [10][11][12][13]; however, here it is used to model interaction among BOLD signals directly. Inclusion of sparsity through the indicator variables in our analysis is useful as we expect our particular application to brain networks to naturally have a sparse representation, because not every pair of nodes is connected through a significant link. In addition to providing a more representative description through a lesser number of important parameters, it yields more efficient results and requires lesser storage. The concept of sparsity has been previously employed in neuroimaging to analyse fMRI and EEG data using mainly regularisation schemes [14][15][16][17][18][19]; however, it has not been used in a full Bayesian network setting to infer functional connectivity.

The system over all nodes in Equation 1 can be described as a linear stochastic differential equation of the form:

\[
\Delta X_t = AX_t + BW_t,
\]

where \(X_t \in R^{N \times 1}\) is \([x_{i,1}, x_{i,2}, \cdots, x_{i,N}]^T\) and is the state of all nodes at time \(t\), and \(W_t\) represents noise which could be due to car-

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) grant number EP/K020153/1 and Yousef Jameel Scholarship Programme.
diadic motion, respiratory volume, blood flow etc.

The matrix \( A \in R^{N \times N} \) is defined as:

\[
A = \Delta t \begin{bmatrix}
-\gamma_1 & S_{1,2} \phi_{1,2} & S_{1,3} \phi_{1,3} & \cdots & \cdots \\
S_{1,2} \phi_{2,1} & -\gamma_2 & S_{2,3} \phi_{2,3} & \cdots & \cdots \\
S_{1,3} \phi_{3,1} & S_{2,3} \phi_{3,2} & -\gamma_3 & \cdots & \cdots \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
\end{bmatrix}
\]  

(3)

and \( B \in R^{N \times N} \) is an identity matrix.

If \( \Phi \in R^{M \times 1} \) is defined as \( [\phi_{1,2}, \phi_{1,3}, \cdots, \phi_{N-1,N}]^T \) where \( \phi_{i,j} = \phi_{j,i} \) and \( M = N \times (N - 1)/2 \), and \( S \in R^{M \times 1} \) is defined as \( [S_{1,2}, S_{1,3}, \cdots, S_{N-1,N}]^T \), then for a given set of linkage parameters values and sparsity variables, the transition density of \( X_t \) is given by the following Normal distribution:

\[
p(X_t|X_{t-1}, S, \Phi) = \mathcal{N}(X_t|F(S, \Phi)X_{t-1}, Q),
\]

(4)

where \( F \) and \( Q \) are transition and covariance matrices respectively, and are given as follows:

\[
F(S, \Phi) = A(S, \Phi) + I,
\]

(5)

\[
Q = BB^T,
\]

(6)

where \( I \) is an identity matrix and \( P \in R^{N \times N} \) is \( \sigma^2 I \).

The observation model is specified as:

\[
z_{i,t} = x_{i,t} + v_{i,t},
\]

(7)

where \( v_{i,t} \) is random additive noise having a Gaussian distribution with a mean of zero and variance \( \sigma^2_{\Phi} \) and represents measurement noise. Then the observation probability density for the joint state is given by:

\[
p(Z_t|X_t) = \mathcal{N}(Z_t|X_t, \sigma^2_{\Phi} I),
\]

(8)

where \( Z_t \in R^{N \times 1} \) is \( [z_{1,i}, z_{2,i}, \cdots, z_{N,i}]^T \) and describes the observations at time \( t \).

By specifying the model in terms of state and observation equations, the problem can be cast in probabilistic terms. In order to estimate the values of the underlying linkages \( \phi_{i,j} \) from observations and thereby infer the networks, the aim is to calculate the posterior probability distribution \( p(S, \Phi|Z_{1:t}) \) where \( Z_{1:t} \) is the observation set up to time \( t \). Altered conditions will render different networks and hence different values for the parameters. As explained in the next section, this inference is carried out for two different states, namely activation and fixation.

### 2.2. Algorithm

When studies are conducted in the field of Neuroscience to investigate neurological responses to various kinds of stimuli, the subjects are typically instructed to engage in a task intermittently with periods of rest in between. For the case where networks are deemed to change abruptly, such as in the case of motor task based experimental data examined later, the most suitable approach is to divide the data set according to the associated experimental conditions and to treat each concatenated subset as a static network. For instance, the model switches between two modes \( G \) and \( H \) depending on the external conditions. This reduces the problem to inferring constant parameters in each of these subsets with known change-points using offline methods.

\[
Model \Rightarrow \begin{cases}
G \leftarrow \text{conditions } g \\
H \leftarrow \text{conditions } h
\end{cases}
\]

A Markov Chain Monte Carlo (MCMC) algorithm [20][21] is used to infer the different networks via sampling from the joint distribution \( p(S, \Phi|Z_{1:t}) \) without the need to sample the BOLD signals \( X_{1:t} \) as these are marginalised out. As direct sampling from the joint distribution is difficult, Gibbs sampling is used to obtain samples approximating the posterior distribution of these linkage parameters. This samples from the full conditional distribution of each variable in turn, conditional on the current values of the other variables. For variables for which this full conditional distribution is difficult to sample, Metropolis-within-Gibbs sampling can be used in which a new sample is proposed from an easy to sample distribution and accepted with a probability given by the Metropolis-Hastings algorithm.

In our case, we would like to sample each \( \phi_{i,j} \) from its full conditional distribution \( p(\phi_{i,j}|S, \Phi_{all/i,j}, Z_{1:t}) \).

\[
p(\phi_{i,j}|S, \Phi_{all/i,j}, Z_{1:t}) = p(Z_{1:t}|\phi_{i,j}, \Phi_{all/i,j}, S)p(\phi_{i,j}|\Phi_{all/i,j}, S)
\]

(9)

As it is difficult to sample directly from this distribution, Metropolis within Gibbs sampling is employed in order to make inference.

If the corresponding indicator variable \( S_{i,j} \) is currently active, i.e. has a value of 1, then \( \phi_{i,j} \) is given from a random walk proposal.

\[
q(\phi^*_i, j|\phi^{-1}_{i, j}, \Phi_{all/i,j}) = \mathcal{N}(\phi^*_i, j|\phi^{-1}_{i, j}, \sigma^2_{\phi})
\]

(10)

where \( \phi^*_i, j \) is the proposed value, \( \phi^{-1}_{i, j} \) is the previous sample, and \( \sigma^2_{\phi} \) is the sampling variance.

The next step is to compute the likelihood of data given the new sample and compare it with the current value, in order to calculate the Metropolis-Hastings ratio:

\[
K = \frac{p(Z_{1:t}|S, \Phi_{all/i,j}, \phi^*_i, j)}{p(Z_{1:t}|S, \Phi_{all/i,j}, \phi^{-1}_{i, j})}
\]

(11)

As the proposal density is Normal and thus symmetric, this reduces to:

\[
K = \frac{p(Z_{1:t}|S, \Phi_{all/i,j}, \phi^*_i, j)}{p(Z_{1:t}|S, \Phi_{all/i,j}, \phi^{-1}_{i, j})}
\]

(12)

The observation likelihood \( p(Z_{1:t}|S, \Phi) \) can be found using quantities calculated from the Kalman filter such that:

\[
p(Z_{1:t}|S, \Phi) = p(Z_1)\Pi_{h=2}^{h=1}N(Z_h|\mu_{h|1:h-1}, \Sigma_{h|1:h-1})
\]

(13)

where \( \mu_{h|1:h-1} \) and \( \Sigma_{h|1:h-1} \) is the predictive mean and covariance of the observation respectively, found through Kalman filtering recursive equations. This distribution is an integration over the hidden state \( X_t \) since:

\[
p(Z_{1:t}|S, \Phi) = p(Z_1|S, \Phi)\Pi_{h=2}^{h=1}p(Z_h|Z_{h|1:h-1}, S, \Phi)
\]

(14)

and

\[
p(Z_t|Z_{t-1}, S, \Phi) = \int p(Z_t|X_t)p(X_t|Z_{1:t-1}, S, \Phi)dX_t,
\]

(15)

where \( p(X_t|Z_{1:t-1}, S, \Phi) \) is the predictive distribution of the next
state. In the linear Gaussian case, this integral can be computed in closed form to give the one-step observation likelihood, which is Gaussian with mean and variance given by that in Equation 13. The sample in Equation 10 is then accepted with a probability of

$$p(x_{t}\mid y_{1:T}, \theta_{alt}, \mu_{alt}, \Sigma_{alt/1:t}) = \frac{p(y_{t}\mid x_{t}, \theta_{alt}, \mu_{alt}, \Sigma_{alt/1:t})p(x_{t}\mid \theta_{alt}, \mu_{alt}, \Sigma_{alt/1:t})}{p(y_{t}\mid x_{t}, \theta_{alt}, \mu_{alt}, \Sigma_{alt/1:t})}$$

(17)

The prior \(p(S_{i,j} \mid \phi_{alt}, \mu_{alt}, \Sigma_{alt/1:t})\) can, in the simplest case, be given by an independent Bernoulli distribution for each case of \(S_{i,j}\) where a small value of \(v\) favours sparse network structures while a value of 0.5 represents the completely uninformined case.

$$p(S_{i,j} = s \mid \phi_{alt}, \mu_{alt}, \Sigma_{alt/1:t}) = \begin{cases} v & s = 1 \\ 1 - v & s = 0 \end{cases}$$

(18)

### 3. EXPERIMENTAL DATA

The algorithm was first tested on synthetic samples in order to verify its robustness and efficiency, and confirm its ability to estimate meaningful results when applied to experimental fMRI data. It is complex to benchmark the performance of simulated data results with those obtained from current methods employing subtractive analysis, as the realms of models and underlying assumptions are different and any direct comparison of numerical errors obtained would not be a fair and accurate approach. However, the similarity of results acquired when applied to real data in this section yield an insight into the employability of our proposed methodology.

Twenty-two healthy subjects with an age range of 19 to 57 and a mean of 35.0 participated in a self-paced, right-handed finger opposition task-based, boxcar design experiment with five alternating cycles of task and fixation blocks. The task comprised of touching one’s fingers sequentially from index to little finger with the right thumb, and continuing to do so till the end of the task period. A visual move command was an indication to carry out the task while rest was to stop it in order to enter the fixation state. Each cycle lasted for 30 seconds making the entire period of data acquisition for one participant to be five minutes.

fMRI data was obtained using a Siemens Trio 3T scanner with whole-brain echo planar imaging (TR = 2000 ms; TE = 30 ms; flip angle = 78\(^\circ\); FOV read = 192 mm; voxel size = 3.0 x 3.0 x 3.0 mm; volumes = 160; slices per volume = 32). Preprocessing of imaging data involved standard slice-time and motion corrections, normalization to the Montreal Neurological Institute (MNI) space and an a priori grey matter template, smoothing with an 8 mm FWHM Gaussian kernel, and low-pass filtering (0.009 - 0.08 Hz).

The experiment was conducted at the Wolfson Brain Imaging Centre, Cambridge, UK and was approved by the local ethics committee with all participants having given informed consent in writing.

BOLD time-series data acquired for five regions of interest (ROIs) corresponding to motor function and five regions that relate to visual activity is analysed. The names and MNI co-ordinates of these ROIs are listed in Table 1. Given the nature of the task involving motor skills and dexterity, it is believed that the network formed by the motor ROIs would undergo a more significant change in the two modes of activation and fixation, than that formed by the visual nodes.

Figure 1 shows the results of applying our algorithm to fMRI measurements taken from one subject undergoing this experiment. For the entire cohort, the average change in all links found over the entire sample for the two kind of networks is displayed in Table 2. To validate this computation, a similar analysis is conducted on the same data using temporal correlation, and the results are included. As expected, the absolute change in the motor network is greater on average than that in the visual network. It can be noted that the mean proportional change in the interaction strengths matches very well to the one found by correlation analysis. This offers as a confirmation that our novel approach of finding functional connectivity provides sensible results and could thus be used for further analysis.

In order to complete this study, a paired t-test is applied to our results. This test is used to determine whether the two sets are significantly different from each other, the null hypothesis being that the difference has a distribution with zero mean and unknown variance. Paired t-test is applied to the motor and visual networks at a 5% significance level as shown in Table 3. Unlike for the visual case, the \(p\)-value, which is the probability of observing the given result if

<table>
<thead>
<tr>
<th>ROI</th>
<th>Name</th>
<th>Abbreviation</th>
<th>Co-ordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor</td>
<td>Supplementary Motor Area</td>
<td>SMA</td>
<td>[-4 -2 54]</td>
</tr>
<tr>
<td>Motor</td>
<td>Left Precentral Gyrus</td>
<td>PRECGL</td>
<td>[-36 -22 64]</td>
</tr>
<tr>
<td>Motor</td>
<td>Right Precentral Gyrus</td>
<td>PRECGR</td>
<td>[60 8 28]</td>
</tr>
<tr>
<td>Motor</td>
<td>Left Postcentral Gyrus</td>
<td>POCGL</td>
<td>[-40 -26 52]</td>
</tr>
<tr>
<td>Motor</td>
<td>Right Postcentral Gyrus</td>
<td>POCHR</td>
<td>[56 -16 38]</td>
</tr>
<tr>
<td>Visual</td>
<td>Left Lingual Gyrus</td>
<td>LINGL</td>
<td>[-15 -72 -8]</td>
</tr>
<tr>
<td>Visual</td>
<td>Right Lingual Gyrus</td>
<td>LINGR</td>
<td>[18 -47 -10]</td>
</tr>
<tr>
<td>Visual</td>
<td>Left Calcarine Gyrus</td>
<td>LICAL</td>
<td>[-18 -68.5]</td>
</tr>
<tr>
<td>Visual</td>
<td>Right Calcarine Gyrus</td>
<td>LICAL</td>
<td>[8 -72 -8]</td>
</tr>
<tr>
<td>Visual</td>
<td>Right Fusiform Gyrus</td>
<td>FFG</td>
<td>[27 -59 -9]</td>
</tr>
</tbody>
</table>

Table 1: Details of Regions of Interest

### Table 2: Average change in links obtained from Metropolis within Gibbs and correlation analysis

<table>
<thead>
<tr>
<th>Mean absolute change</th>
<th>Correlation analysis</th>
<th>Metropolis with Gibbs sampler</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor network</td>
<td>0.0376</td>
<td>0.0707</td>
</tr>
<tr>
<td>Visual network</td>
<td>0.0245</td>
<td>0.0497</td>
</tr>
<tr>
<td>Visual/Motor</td>
<td>0.70</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Fig. 1: Sample results obtained from experimental data of one subject.

<table>
<thead>
<tr>
<th></th>
<th>p-value</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor</td>
<td>0.0033</td>
<td>-0.0345</td>
<td>0.0099</td>
<td>0.0500</td>
</tr>
<tr>
<td>Visual</td>
<td>0.2613</td>
<td>-0.0530</td>
<td>-0.0121</td>
<td>0.0461</td>
</tr>
</tbody>
</table>

Table 3: Results from paired t-tests

the null hypothesis is true, calculated for the motor network is much lower than 0.05 indicating that it undergoes a statistically significant change between the activated and fixated periods. These findings further substantiate the validity of our work.

4. CONCLUSION

We have presented a method of inferring functional connectivity using a Bayesian framework in a full network form. A case with a clear hypothesis has been examined, in which the motor network was believed to change more significantly than the visual network. In order to investigate the implications of our work in the context of providing network-level diagnostic tools, data from patients with traumatic brain injury (TBI) is being analysed for disruptive effects, which could serve as a guide for more comprehensive therapeutic treatments. Having validated the proposed methodology and demonstrated its potential use, we can then delve into the unknown realms of resting-state data and attempt to infer parameters that are smoothly varying with time. This more complicated case would require the use of Sequential Monte Carlo methods [22] and could yield an even more comprehensive understanding of brain functions.

We aim to further improve our algorithms for obtaining more accurate inference as well as extending to networks of larger sizes. By providing network level descriptors of brain functions, our work has the potential of offering novel disease markers for diagnostic as well as therapeutic purposes which are not only limited to TBI [23][24][25], but also include neurological disorders such as major depression [26][27], schizophrenia [28][29][30] and Alzheimer’s disease [31][32][33].

5. REFERENCES


