

DIRECTED INTERACTIVITY OF LARGE-SCALE BRAIN NETWORKS: INTRODUCING A NEW METHOD FOR ESTIMATING RESTING-STATE EFFECTIVE CONNECTIVITY MRI*

Nan Xu, R. Nathan Spreng, and Peter C. Doerschuk

School of Electrical and Computer Engineering (NX and PCD)
Department of Human Development (RNS)
Cornell University
Ithaca, NY 14853

ABSTRACT

Resting-state functional MRI (rs fMRI) is widely used to non-invasively study human brain networks. Network functional connectivity is estimated by calculating the standard correlation between blood-oxygen-level dependent (BOLD) signals in specific regions of interests (ROIs). However, standard correlation fails to characterize the causality and the direction of information flow between regions, which are important factors in characterizing a network. Here, we use causal linear time-invariant models, with the impulse response duration estimated by Information Criteria, to describe the effective connectivity between ROIs. To do so, we replace the standard correlation between BOLD signals with a correlation between a BOLD signal and a prediction via the model of that BOLD signal. Prediction correlation is then used in a network analysis similar to that used with standard correlation. Our results include the causality information, the direction of information flow, and the possibility of delays in information flow. This approach replicates the local and distributed network architecture of the human brain previously observed with standard correlations, as well as providing novel insight into the directed interactivity of regions comprising these networks.

Index Terms— functional connectivity, functional connected network, resting-state fMRI, neuroimaging

1. INTRODUCTION

The human brain is a complex network comprising many interacting regions. Substantial progress has been made in delineating large-scale functional brain networks using rs fMRI data [1–4]. One fruitful approach has been to examine the correlations among pairs of BOLD timeseries in *a priori* determined ROIs and use the correlations to compute interacting networks [5, 6]. However, the direction of information flow (unidirectional or bidirectional between a pair of ROIs) and the causality of information flow are not completely contained in the standard correlation matrices even though di-

rection and causality are believed to be important aspects of functional brain networks.

In order to capture at least some information related to the direction of information flow and the causality of information flow, the replacement of the correlation between two BOLD signals by a correlation between the first BOLD signal and a prediction of the first signal from the second signal is proposed. Unlike the standard correlation, this prediction correlation is asymmetrical in the two signals and so includes some information about directionality. Furthermore, because the prediction can depend on delayed values of the BOLD signal, information from all lags is optimally (by a prediction criteria) combined into a single number rather than computing standard correlations at each of the different lags with no obvious way to combine the information. If only the present value of the BOLD signal on which the prediction is based is used, the prediction correlation used in this paper is identical to the standard correlation. With this approach, we successfully replicate prior characterizations [3] of the modular organization of the human brain. Additionally, we determine patterns of directed network interactivity. Details of the method are described in three stages in Section 2.2.

2. METHODS

2.1. Participants, data preprocessing and terminology

A cohort containing $N_s = 132$ subjects aged 25-30 years (mean age = 26.8 ± 1.65 years, 73 women) was extracted from the 1000 Functional Connectomes Project (http://www.nitrc.org/projects/fcon_1000/) [7].

The MRI scans were performed with a range of TR values in the interval 1–3 s (mean \pm standard deviation of 2.3 ± 0.4 s). Measured in terms of the number of TR periods, the total scan duration N_x was in the interval $\{119, \dots, 295\}$ (mean \pm standard deviation of 167.5 ± 41.7). The whole brain data was preprocessed [8], linearly detrended and band-pass filtered (retaining signal between 0.001 and 0.1 Hz), and motion scrubbed [9] with the threshold set to 0.2. The preprocessed rs fMRI BOLD signals were extracted from $N_{ROI} = 264$ spherical ROIs each with a 10mm diameter.

*SUPPORTED BY NSF 1217867.

2.2. Determination of the network

2.2.1. Stage 1: Causal linear model of information flow

In the first stage, all ordered pairs of ROIs in each subject are considered with the exception of pairs that are closer than a distance r_0 ($r_0 = 20\text{mm}$ in our calculations). Let x_i denote the BOLD signal at the i^{th} ROI (abbr. the i^{th} signal), and x_j denote the BOLD signal at the j^{th} ROI (abbr. the j^{th} signal). For the ordered pair of signals (x_i, x_j) , x_j is predicted from x_i by computing a linear time-invariant causal dynamic model with x_i as the input and the prediction of x_j as the output. The prediction is denoted by $\hat{x}_{j|i}$. Because both the ordered pair (x_i, x_j) and the ordered pair (x_j, x_i) are considered, there are two linear systems connecting ROIs i and j . Properties of this dynamic model include: (a) the system can be described by an impulse response, denoted $\mathbf{h}_{j|i}$, since it is linear and time-invariant; (b) the first potentially non-zero term in the impulse response occurs at time index 0, because the model is causal; (c) we assume that $\mathbf{h}_{j|i}$ has finite duration $N_{h_{j|i}}$. Assume that x_i has the finite length N_x . The basic approach to estimate the coefficients of $\mathbf{h}_{j|i}$ is to minimize the least squares cost $\mathcal{J}(\mathbf{h}_{j|i}) = \sum_{n=0}^{N_x-1} (x_j[n] - \hat{x}_{j|i}[n])^2$. However, it is also necessary to determine the duration $N_{h_{j|i}}$ of the impulse response. Therefore, we restate the least squares problem as a maximum likelihood estimation problem with a known signal variance and we replace the known signal variance by its sample variance. In the maximum likelihood framework, we can estimate $N_{h_{j|i}}$ by balancing fitting of the current data (i.e., minimizing \mathcal{J}), which is best done by large values of $N_{h_{j|i}}$, and predicting when presented with new data, which is best done by smaller values of $N_{h_{j|i}}$. In particular, we quantify the balancing by applying the Bayesian information criteria (BIC) [10]. BIC realizes this balancing goal by minimizing the sum of two terms, one term that characterizes the prediction error of the dynamic system and a second term that depends on the duration $N_{h_{j|i}}$ and N_x :

$$\text{BIC} = N_x \log\left(\frac{2\pi}{N_x - N_{h_{j|i}}}\mathcal{J}\right) + N_x - N_{h_{j|i}} + N_{h_{j|i}} \log(N_x). \quad (1)$$

Joint minimization of Eqs. 1 with respect to both $\mathbf{h}_{j|i}$, which occurs only in the \mathcal{J} term, and $N_{h_{j|i}}$ is our goal in Stage 1. As is often done, the integer minimization over $N_{h_{j|i}}$ is computed by testing each value in a predetermined range of values. Here, we restricted the temporal window for directional influence between ROIs to less than 30s and so consider $N_{h_{j|i}}$ no larger than $\lfloor 30/\text{TR} \rfloor$. Note that TR varies from subject to subject. Then, for each value of $N_{h_{j|i}}$, the minimization with respect to $\mathbf{h}_{j|i}$ involves only minimizing \mathcal{J} . For reasons described in the final paragraph of Section 2.2.2, we consider constrained estimation, which allows only nonnegative $\mathbf{h}_{j|i}$ values. For N_{ROI} ROIs, the result of this calculation is a set of N_{ROI}^2 impulse responses, where pairs of ROIs closer than r_0 have impulse responses that are identically zero. Since the

dynamical system describing how x_i influences x_j is separate from the dynamical system describing how x_j influences x_i , the approach described here can lead to a directed rather than undirected graph of interactions between ROIs.

2.2.2. Stage 2: Prediction correlation (p-correlation)

In the second stage, all ordered pairs of ROIs are again considered with the exception of pairs that are closer than a distance r_0 . For the (x_i, x_j) ordered pair, recall that $\hat{x}_{j|i}$ is the prediction of x_j given x_i which is the output of the dynamical system determined in Stage 1. The correlation used in this paper, denoted by $\rho_{j|i}$, is the correlation between x_j and $\hat{x}_{j|i}$. We use “correlation” and $\rho_{j,i}$ for the standard approach (i.e., the standard correlation between x_j and x_i) and “p-correlation” and $\rho_{j|i}$ for the approach proposed in this paper. This is an asymmetric definition since, in general, $\rho_{j|i} \neq \rho_{i|j}$. Furthermore, this definition includes lags of the x_i signal since the dynamical system output at time n , $\hat{x}_{j|i}[n]$, depends on the input at its current and previous times, i.e., $x_i[n]$, $x_i[n-1]$, \dots , $x_i[n - N_{h_{j|i}} + 1]$. If $N_{h_{j|i}} = 1$ (i.e., no lags) then $\rho_{j|i}$ is the correlation between x_j and x_i so that $\rho_{j|i} = \rho_{j,i}$ and the approach of this paper exactly reduces to the standard approach.

The implications of requiring $h_{j|i}[n] \geq 0$ are considered in this paragraph. Let $R_{j|i}$ be the covariance of x_j and $\hat{x}_{j|i}$, which is related to the covariance of $x_j[n]$ and $x_i[n-m]$ (i.e., the m -lagged covariance of the two signals, denoted by $R_{j,i}[m]$) by $R_{j|i} = \sum_{m=0}^{N_{h_{j|i}}-1} R_{j,i}[m]h_{j|i}[m]$. The covariance $R_{j|i}$ is also the numerator of $\rho_{j|i}$. Therefore, if all the lagged covariances are positive and we require the estimated values of $h_{j|i}[m]$ to be positive then we are assured of getting a nonnegative value for $R_{j|i}$ and for the p-correlation $\rho_{j|i}$. In many neuroscience problems, only positive correlations are relevant, thus the nonnegative “constrained” estimation is appropriate for such applications. Even with $h_{j|i}[n] \geq 0$, it may be that the p-correlation is not positive because one or more of the m -lagged covariance values are negative. Therefore, we define the “constrained” p-correlation by $\rho_{j|i}^+ = \rho_{j|i}$ if $\rho_{j|i} \geq 0$, and $\rho_{j|i}^+ = 0$ otherwise.

2.2.3. Stage 3: Networks from p-correlation

The third stage applies the widely-used [3, 9, 11] Infomap graph analytical algorithm [12] to compute the graph among the ROIs as a set of interacting networks. This graph is based on a $N_{\text{ROI}} \times N_{\text{ROI}}$ matrix of connection weights. Each network shares more information within the network than it shares with other networks.

The signal-to-noise ratio in the resting-state fMRI data is low. Therefore the $\rho_{j|i}^+$ values, which are computed in Stage 2 for each subject, are averaged over all subjects to give averaged values denoted by $\bar{\rho}_{j|i}$. Then a thresholding operation is applied, specifically, the set of $\bar{\rho}_{j|i}$ values over all pairs of ROIs are ordered and the top s percent of values are unaltered

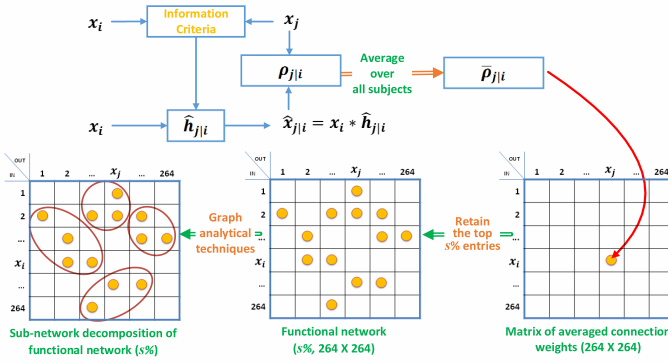


Fig. 1: Flow chart describing the computational method.

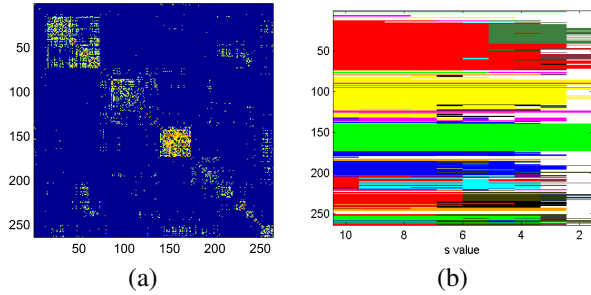


Fig. 2: Matrix of connection weights at $s = 5$ (Panel (a)) and the stability of networks across various thresholding criteria (s) (Panel (b)).

and the lower $100 - s$ percent of values are replaced by zeros. This result, denoted by $c_{j|i}$, is the matrix of connection weights that is the input to Infomap for the computation of a weighted graph. The entire computational methodology of the above three stages is summarized in Fig. 1.

3. RESULTS

The methods described in this paper are implemented in Matlab [13] software which is available upon request.

3.1. P-correlation results

In Stage 1, using BIC we estimate the impulse response $\hat{h}_{j|i}$ and its duration $\hat{N}_{h_{j|i}}$ for each ordered pair (x_j, x_i) of ROI signals. In order to compute the sample mean of the typical duration, the durations are ranked and the mean of the lowest 80% of the ranked durations is $2.33s$ demonstrating the importance of non-zero lags which are missing in the standard correlation approach (e.g., Ref. [3]). In Stage 2, using the impulse responses computed in Stage 1, the predictions $\hat{x}_{j|i}$ are computed. It was verified that p-correlation with impulse response of duration 1 equaled standard correlation as described in the second-to-last paragraph of Section 2.2.2. In Stage 3, thresholding the averaged “constrained” p-correlations by the threshold values $s \in \{2, 3, \dots, 10\}$ percent [3] results in the connection weights $c_{j|i}$ which are shown in Fig. 2(a) for $s = 5$. The connection weights have a great deal of structure, e.g.,

s	10	9	8	7	6	5	4	3	2
corr	18	19	20	22	24	27	31	43	67
p-corr	17	19	20	22	24	28	32	41	66

Table 1: The number of networks at different thresholds s (percent) for correlation and p-correlation.

corr \ p-corr	$x_i \rightarrow x_j$	$x_i \leftarrow x_j$	$x_i \leftrightarrow x_j$	$x_i \not\leftrightarrow x_j$
	$x_i \leftrightarrow x_j$	27	20	1686
$x_i \not\leftrightarrow x_j$	24	36	3	33174

Table 2: Comparison of the standard correlation and the p-correlation in terms of the connections computed by Infomap using p-correlation at $s = 5$. The entries are the number of unidirectional connections from x_i to x_j ($x_i \rightarrow x_j$) or from x_j to x_i ($x_i \leftarrow x_j$), the number of bidirectional connections between x_i and x_j ($x_i \leftrightarrow x_j$), and the number of no connections between x_i and x_j ($x_i \not\leftrightarrow x_j$). Every pair (x_i, x_j) where $i \in \{1, \dots, 264\}$ and $j \in \{i + 1, \dots, 264\}$ is considered.

the blocks visible along the diagonal, which Infomap can exploit to give a graph of networks.

3.2. Networks: stability, identity and interactions

We use the graph analysis tool Infomap-0.11.5 [14, 15] to compute a directed weighted graph of directed weighted networks based on our asymmetrical matrix of real-valued connection weights. A primary parameter in the calculation of the connection weights in Section 2.2 is the threshold s . As a function of the value of s , Infomap creates a variable number of networks and the number is tabulated in Table 1 as s varies between 2 and 10. In summary, the number of networks increases as the value of s decreases.

Following Ref. [3, Fig. 1], the network stability over a range of threshold $s \in \{2, \dots, 10\}$ using p-correlations is shown in Fig. 2(b), in which different networks are represented by different colors. Similar to Ref. [3], we note that the assignment of ROIs to networks remains relatively constant over all values of the threshold s , illustrated by the constant horizontal bands in different colors. Also, networks are hierarchically refined as s rises. Table 2 contains a comparison of the standard correlation (e.g., Ref. [3]) and the p-correlation in terms of the connections computed by Infomap. At $s = 5$, about 2.7% of the connections that are unavoidably bidirectional when using correlation have become unidirectional when using p-correlation. The percentage of unidirectional connections, while small, is little altered by changing the value of s .

In addition, we note that the maps of the networks are in a good agreement with the brain functional systems [3]. The software tool BrainViewer [16] can be used to visualize the modular architecture of the local and distributed networks. In particular, Fig. 3 illustrates the directed connectivity of the somatosensory motor network, the incoming information to this network (Fig. 3(a)), the outward flow of information from this

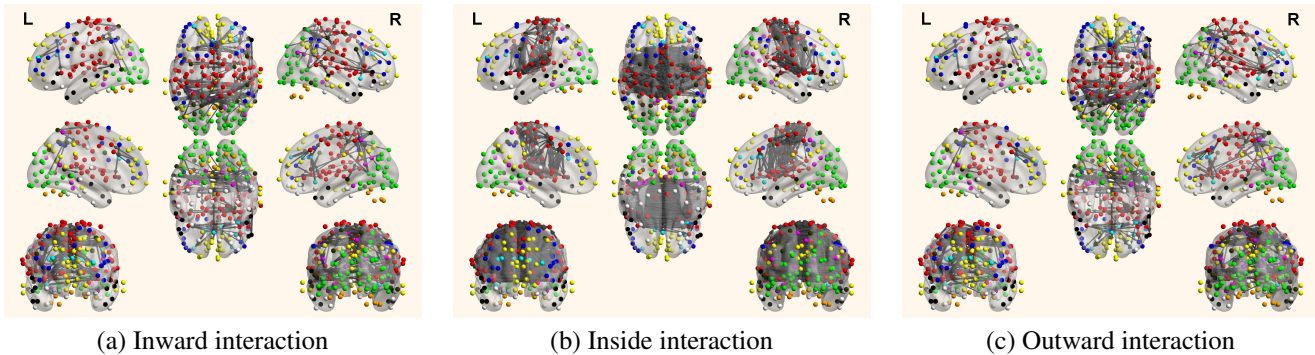


Fig. 3: Interaction of somatosensory-motor network with brain cortex at $s = 5$ for p-correlation computed by BIC. The brain hemisphere is visualized from eight orthogonal directions.

network (Fig. 3(c)), and the dominant structure of connections within this network (Fig. 3(b)). In order to reduce the number of visualizations, only the results for the threshold value $s = 5$ is shown.

4. DISCUSSION

This paper describes a methodology for analyzing resting-state fMRI data in terms of prespecified Regions of Interest (ROIs) using a generalization of well-established correlation ideas. The generalization, denoted by “p-correlation” (“p” for “prediction”), is to compute the correlation between the j^{th} signal and an optimal linear time-invariant causal estimate of the j^{th} signal based on the i^{th} signal. Features of p-correlation include the ability to indicate the directionality of the interaction between two ROIs and the ability to include the possibility that the interaction would depend on the past values of the i^{th} signal. This is a generalization of standard correlation ideas because, if the estimate of the j^{th} signal based on the i^{th} signal is restricted to use only the current value of the i^{th} signal, then the p-correlation is the same as the standard correlation.

Many approaches have been introduced to assess effective connectivity of resting-state MRI data. Ref. [4] evaluated the validity of many approaches [4, Fig. 4] using simulated BOLD signals and a variety of measures of performance. The algorithms tend to have different levels of performance for different measures of performance, e.g., detection of a connection versus determination of the direction of a connection. The p-correlation approach introduced in this paper depends on causal dynamical models and so we focus on this particular aspect of previous works. Several versions of Granger causality analysis, based on multivariate vector autoregressive modeling, were tested in Ref. [4] (see also Ref. [17, p. 20]) and performed poorly. However, the key in Granger causality analysis is statistical tests to determine whether the regressor is significantly different from zero in a statistical sense and such tests are not used in our approach, which is a generalization of correlation analysis. Dynamic Causal Modeling

(DCM) has been used with some success to assess causal dynamics in fMRI data by relying on sophisticated models of neural dynamics. Recent developments in DCM [18] have made it unnecessary to specify the driving inputs, making the DCM approach more appropriate for the analysis of resting-state fMRI data. However, models of large-scale network interactivity observed by resting-state fMRI using 264 ROIs far exceed computational limits for a DCM analysis. In contrast, the p-correlation approach described in this paper scales similarly to a correlation approach for which hundreds of ROIs are not a challenge.

Examining spontaneous low-frequency BOLD signal fluctuations across the human brain using fMRI reveals dissociable functional-anatomic networks [1]. Recently, significant advances in identifying the brain’s intrinsic functional architecture have been made (e.g., [2, 3]). Differentiation of intrinsic networks has revealed specialized information processing modules, but directed patterns of connectivity are largely unknown. As increasing numbers of dissociable and functionally specialized intrinsic networks are identified, characterizing the flow of information between regions is increasingly important. Using the p-correlation approach, we have successfully replicated the modular architecture of the local and distributed networks previously reported using standard correlation [2, 3]. Extending these findings, for the first time we map the directed connectivity within and among these networks, revealing the regional flow of information into, within, and from them. The detailed connectivity profile of the somatosensory-motor network (Fig. 3) is largely convergent with known connectivity profiles of these regions revealed through tracer studies (e.g., [19]). While structural and functional coherence cannot be assumed, methodological convergence within the somatosensory-motor cortex offers preliminary support for this approach. Efforts to map directed network interactivity have faced significant methodological challenges. Here we introduce a novel approach, the p-correlation, to estimate effective connectivity within well-characterized large-scale functional brain networks.

5. REFERENCES

- [1] Bharat Biswal, F Zerrin Yetkin, Victor M Haughton, and James S Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar mri," *Magnetic resonance in medicine*, vol. 34, no. 4, pp. 537–541, 1995.
- [2] BT Thomas Yeo, Fenna M Krienen, Jorge Sepulcre, Mert R Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L Roffman, Jordan W Smoller, Lilla Zöllei, Jonathan R Polimeni, et al., "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *Journal of neurophysiology*, vol. 106, no. 3, pp. 1125–1165, 2011.
- [3] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al., "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.
- [4] Stephen M Smith, Karla L Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F Beckmann, Thomas E Nichols, Joseph D Ramsey, and Mark W Woolrich, "Network modelling methods for fMRI," *Neuroimage*, vol. 54, no. 2, pp. 875–891, 2011.
- [5] Gagan S. Wig, Bradley L. Schlaggar, and Steven E. Petersen, "Concepts and principles in the analysis of brain networks," *Annals of the New York Academy of Sciences*, vol. 1224, no. 1, pp. 126–146, 2011.
- [6] Olaf Sporns, "The human connectome: a complex network," *Annals of the New York Academy of Sciences*, vol. 1224, no. 1, pp. 109–125, 2011.
- [7] Bharat B Biswal, Maarten Mennes, Xi-Nian Zuo, Suril Gohel, Clare Kelly, Steve M Smith, Christian F Beckmann, Jonathan S Adelstein, Randy L Buckner, Stan Colcombe, et al., "Toward discovery science of human brain function," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4734–4739, 2010.
- [8] Jeffrey S Anderson, T Jason Druzgal, Melissa Lopez-Larson, Eun-Kee Jeong, Krishnaji Desai, and Deborah Yurgelun-Todd, "Network anticorrelations, global regression, and phase-shifted soft tissue correction," *Human brain mapping*, vol. 32, no. 6, pp. 919–934, 2011.
- [9] Jonathan D. Power, Kelly A. Barnes, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen, "Spurious but systematic correlations in functional connectivity mri networks arise from subject motion.," *NeuroImage*, vol. 59, no. 3, pp. 2142–2154, 2012.
- [10] Gideon Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.
- [11] Juha M Lahnakoski, Enrico Glerean, Juha Salmi, Iiro P Jääskeläinen, Mikko Sams, Riitta Hari, and Lauri Nummenmaa, "Naturalistic fmri mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception," *Frontiers in human neuroscience*, vol. 6, 2012.
- [12] Andrea Lancichinetti and Santo Fortunato, "Community detection algorithms: A comparative analysis," *Phys. Rev. E*, vol. 80, pp. 056117, Nov 2009.
- [13] Mathworks, "URL," <http://www.mathworks.com/>.
- [14] Martin Rosvall and Carl T Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, 2008.
- [15] Martin Rosvall, Daniel Axelsson, and Carl T Bergstrom, "The map equation," *The European Physical Journal Special Topics*, vol. 178, no. 1, pp. 13–23, 2009.
- [16] Mingrui Xia, Jinhui Wang, and Yong He, "BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics," *PLoS ONE*, vol. 8, no. 7, pp. e68910+, July 2013.
- [17] Karl J Friston, "Functional and effective connectivity: A review," *Brain Connectivity*, vol. 1, no. 1, pp. 13–36, 2011.
- [18] J. Daunizeau, K.J. Friston, and S.J. Kiebel, "Variational bayesian identification and prediction of stochastic nonlinear dynamic causal models," *Physica D: Nonlinear Phenomena*, vol. 238, no. 21, pp. 2089 – 2118, 2009.
- [19] LD Selemon and PS Goldman-Rakic, "Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network subserving spatially guided behavior," *The Journal of Neuroscience*, vol. 8, no. 11, pp. 4049–4068, 1988.