

EVALUATING MODELS OF DYNAMIC FUNCTIONAL CONNECTIVITY USING PREDICTIVE CLASSIFICATION ACCURACY

Søren Føns Vind Nielsen¹, Yuri Levin-Schwartz², Diego Vidaurre³,
Tulay Adali², Vince D. Calhoun^{5,6}, Kristoffer H. Madsen^{1,4}, Lars Kai Hansen¹ and Morten Mørup¹

¹ Department of Applied Mathematics and Computer Science, Technical University of Denmark

² Department of CSEE, University of Maryland, Baltimore County, USA

³ OHBA, Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, UK

⁴ Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Denmark

⁵ The Mind Research Network, Albuquerque, USA

⁶ Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, USA

ABSTRACT

Dynamic functional connectivity has become a prominent approach for tracking the changes of macroscale statistical dependencies between regions in the brain. Effective parametrization of these statistical dependencies, referred to as brain states, is however still an open problem. We investigate different emission models in the hidden Markov model framework, each representing certain assumptions about dynamic changes in the brain. We evaluate each model by how well they can discriminate between schizophrenic patients and healthy controls based on a group independent component analysis of resting-state functional magnetic resonance imaging data. We find that simple emission models without full covariance matrices can achieve similar classification results as the models with more parameters. This raises questions about the predictability of dynamic functional connectivity in comparison to simpler dynamic features when used as biomarkers. However, we must stress that there is a distinction between characterization and classification, which has to be investigated further.

Index Terms— Dynamic functional connectivity, Hidden Markov models, Classification, Schizophrenia

1. INTRODUCTION

In the study of how the brain integrates information, communication between disjoint regions is often described using *functional connectivity* (FC). Over the last two decades, FC analysis has relied on a stationary assumption, i.e. that the statistical dependencies between regions do not change over

time. This assumption has been shown to disregard a potential wealth of information in the changes in between-region connectivity, especially in resting-state functional magnetic resonance imaging (rs-fMRI) where this analysis approach has been coined dynamic functional connectivity (dFC) [1, 2].

The most widely used approach in the dFC literature is the sliding-window correlation (SWC) [3], in which the (regularized) correlation matrix was estimated in windows slid one time-step at a time on group independent component analysis (gICA) time-courses from rs-fMRI from healthy subjects. After applying a k-means clustering to the estimated windowed correlation matrices they found that the seven clusters extracted varied especially in their connectivity within the default mode network.

However, SWC has been criticized because the choice of window-length has a large influence on the results thus questioning the reliability of the extracted dynamics [4, 5, 6]. Furthermore, the lack of consensus on what drives the underlying neurological changes questions what is the appropriate model for dFC. As an alternative to windowing setting the window length to 1 and imposing smoothness in the state transitions leads to a hidden Markov model (HMM), which has been used for modeling dFC in several recent publications [7, 8, 9, 10, 11, 12].

Recently, dFC approaches have been applied in the context of schizophrenic patients and shown promise in characterizing the differences between patient and healthy controls. In fMRI-studies, the focus has been on understanding heredity of the disease [13], the disease influence on working memory [14] as well as hallucinations [15] and resting-state dFC differences between medicated patient populations and controls [16, 17, 18, 19].

In this paper, we investigate how different HMM modeling assumptions on the dynamics in resting state fMRI translate into classification accuracy using a cohort of schizophrenic patients (SZ) and healthy controls (HC). We accomplish

Corresponding author: sfvn@dtu.dk. This work was supported by the Lundbeckfonden grant no. R105-9813, the Novo Nordisk Foundation Interdisciplinary Synergy Program 2014 (BASICS) grant no. NNF14OC0011413 and by the NIH grant R01-EB-020407. We thank Qunfang Long for assistance with the group ICA.

this using the Bayesian hidden Markov model framework [7, 11, 20] with different emission models and investigate their ability to discriminate between SZ and HC. The different emission models, that each encode different assumptions on dynamics, will be compared using classification accuracy on held-out data. The purpose of this paper, then, is to use classification performance as a tool to evaluate the utility of different modeling assumptions about dynamic functional connectivity. Thus we pose the following research questions to be answered: 1) How do different assumptions on dynamic functional connectivity models influence classification performance? 2) To what extent does modeling dynamic (as opposed to static) functional connectivity influence classification performance?

2. METHODS

We use the variational Bayes hidden Markov model (VB-HMM) from [7, 20] (and the accompanying MATLAB implementation¹). The VB-HMM with K states has the generative model for the observations $\mathbf{x}_t \in \mathbb{R}^P$ for $t = 1 \dots T$,

$$\boldsymbol{\pi}_0 \sim \text{Dir}(\boldsymbol{\kappa}_0) \quad (1)$$

$$\boldsymbol{\pi}^{(k)} \sim \text{Dir}(\boldsymbol{\kappa}^{(k)}), \quad (2)$$

$$z_t | z_{t-1} \sim \text{Multinomial}(\boldsymbol{\pi}^{(z_{t-1})}), \quad (3)$$

$$\boldsymbol{\Sigma}^{(k)-1} \sim \mathcal{W}(\boldsymbol{\Sigma}_0, \nu_0), \quad (4)$$

$$\boldsymbol{\mu}^{(k)} \sim \mathcal{N}(\boldsymbol{\mu}_0, \lambda^{-1} \boldsymbol{\Sigma}^{(k)}), \quad (5)$$

$$\mathbf{x}_t \sim \mathcal{N}(\boldsymbol{\mu}^{(z_t)}, \boldsymbol{\Sigma}^{(z_t)}), \quad (6)$$

in which $\boldsymbol{\pi}_0$ is the initial state distribution vector of length K , $\text{Dir}(\cdot)$ is the Dirichlet distribution, $\boldsymbol{\kappa}_0$ is the prior vector for the initial distribution, $\boldsymbol{\pi}^{(k)}$ is a row of the transition matrix, $\boldsymbol{\kappa}^{(k)}$ is the associated prior to that row, z_t is the integer valued state taking possible values from $1 \dots K$ at time point t , $\boldsymbol{\Sigma}^{(k)-1}$ is the precision matrix from the k 'th state assumed to be Wishart distributed (\mathcal{W}) with priors $\boldsymbol{\Sigma}_0$ and ν_0 whereas $\boldsymbol{\mu}_0$ is the prior on the mean of each state with associated scaling parameter λ .

To create a classifier from an HMM-model we use a density-based approach. For that we need the predictive likelihood on held out subjects. We use a VB-approximation [11, 21] to the predictive likelihood by calculating the free-energy on the test set keeping the transition matrix and state specific parameters fixed from training, and neglecting the terms in the free-energy that have not changed from training. This corresponds to for given training data \mathbf{X} and test data \mathbf{X}^* to the following bound multiplying by $\frac{Q_{X^*}(\mathbf{z}^*)}{Q_{X^*}(\mathbf{z}^*)} = 1$ and

using Jensen's inequality,

$$\begin{aligned} \ln p(\mathbf{X}^* | \mathbf{X}) &\approx \ln \int \int \int \int [p(\mathbf{X}^* \mathbf{z}^* | \boldsymbol{\pi}_0, \boldsymbol{\pi}, \boldsymbol{\theta}_{obs}) \\ &Q_X(\boldsymbol{\pi}_0) Q_X(\boldsymbol{\pi}) Q_X(\boldsymbol{\theta}_{obs})] d\boldsymbol{\pi}_0 d\boldsymbol{\pi} d\boldsymbol{\theta}_{obs} d\mathbf{z}^* \\ &\geq \langle \ln p(\mathbf{X}^*, \mathbf{z}^* | \boldsymbol{\pi}_0, \boldsymbol{\pi}, \boldsymbol{\theta}_{obs}) \rangle_{Q_X(\boldsymbol{\pi}_0) Q_X(\boldsymbol{\pi}) Q_X(\boldsymbol{\theta}_{obs}) Q_{X^*}(\mathbf{z}^*)} \\ &\quad - \langle \ln Q_{X^*}(\mathbf{z}^*) \rangle_{Q_{X^*}(\mathbf{z}^*)}, \end{aligned} \quad (7)$$

in which $\boldsymbol{\theta}_{obs}$ is all the parameters in the emission model, $Q_X(\cdot)$ is the fitted variational distribution to the training set and $Q_{X^*}(\cdot)$ is the corresponding distribution for the test set, whereas \mathbf{z}^* is the state sequence of the test set.

For a given training and test split we end up with two models each only trained to their respective group, \mathcal{M}_{SZ} and \mathcal{M}_{HC} . Now we can evaluate for a new data set, X^* , what model/group was most likely to generate the data by Bayes rule, $p(\mathcal{M}_{SZ} | X^*) = \frac{p(X^* | \mathcal{M}_{SZ}) p(\mathcal{M}_{SZ})}{\sum_{c \in \{HC, SZ\}} p(X^* | \mathcal{M}_c) p(\mathcal{M}_c)}$, in

which $p(X^* | \mathcal{M}_{SZ})$ is the predictive likelihood on test set X^* by model \mathcal{M}_{SZ} and $p(\mathcal{M}_{SZ})$ is our prior of observing that model. We set this to the empirical proportions in the training data, i.e. $p(\mathcal{M}_{SZ}) = \frac{\#SZ}{\#SZ + \#HC}$.

It is unclear what characterizes differences between SZ and HC. We therefore consider the six different emission parameterizations given in Table 1 each based on different characterizations of dFC. The differences could be driven by changes in interaction between ICA components accounted for by having the full covariance $\boldsymbol{\Sigma}^{(k)}$ ("Mean+Cov" and "Zero-Mean"), or potentially only by within component differences not taking interactions into account ("Diag-Cov" and "Diag-Cov Zero-Mean"), or solely changes in mean activity with stationary (co-)variance ("Stationary Cov" and "Stationary Diag-Cov"). By varying the model order we further quantify if differences are best characterized by static differences between groups ($K = 1$) or relies on the dynamic characterizations ($K > 1$). We thus use the classification accuracy to quantify which parameterization best discriminates between SZ and HC.

3. RESULTS

In the following we will present the results from a synthetic study and a resting-state fMRI data set containing schizophrenic patients and healthy controls. In all of the analyses we set the priors in the HMM models to their defaults as explained in [7].

Equivalence of Different Emission Models: There are many different ways of parameterizing the underlying brain dynamics. In the six emission models we have chosen there are some equivalences in the representations which we have to take into consideration when interpreting the results. To illustrate this we have generated two data sets (mimicking two groups for classification) from two "Stationary Diag-Cov" models in Figure 1 (left panel). Both models have two states

¹Code was downloaded from the following Github repository: <https://github.com/OHBA-analysis/HMM-MAR> in July 2016

Name [Free Parameters] (Description)	Parameterization
<i>Mean+Cov</i> [$K(p + p(p + 1)/2)$] (Bias and component interaction)	$\Sigma^{(k)-1} \sim \mathcal{W}(\Sigma_0, \nu_0), \quad k = 1 \dots K$ $\mu^{(k)} \sim \mathcal{N}(\mu_0, \lambda^{-1} \Sigma^{(k)}) \quad k = 1 \dots K$
<i>Zero-Mean</i> [$Kp(p + 1)/2$] (No bias but only component interaction)	$\Sigma^{(k)-1} \sim \mathcal{W}(\Sigma_0, \nu_0), \quad k = 1 \dots K$ $\mu^{(k)} = \mathbf{0} \quad k = 1 \dots K$
<i>Diag-Cov</i> [$2Kp$] (Bias and within component modulation)	$\sigma_i^{(k)-1} \sim \mathcal{G}(a_0, b_0), \quad i = 1..p, \quad k = 1 \dots K$ $\Sigma^{(k)-1} = \text{diag} \left(\left[\sigma_1^{(k)-1}, \sigma_2^{(k)-1}, \dots, \sigma_p^{(k)-1} \right] \right), \quad k = 1 \dots K$ $\mu^{(k)} \sim \mathcal{N}(\mu_0, \lambda^{-1} \Sigma^{(k)}) \quad k = 1 \dots K$
<i>Diag-Cov Zero-Mean</i> [Kp] (No bias but only within component modulation)	$\sigma_i^{(k)} \sim \mathcal{G}(a_0, b_0), \quad i = 1..p, \quad k = 1 \dots K$ $\Sigma^{(k)-1} = \text{diag} \left(\left[\sigma_1^{(k)-1}, \sigma_2^{(k)-1}, \dots, \sigma_p^{(k)-1} \right] \right), \quad k = 1 \dots K$ $\mu^{(k)} = \mathbf{0} \quad k = 1 \dots K$
<i>Stationary Cov</i> [$p(p + 1)/2 + Kp$] (Bias with stationary component interaction)	$\Sigma^{-1} \sim \mathcal{W}(\Sigma_0, \nu_0)$ $\mu^{(k)} \sim \mathcal{N}(\mu_0, \lambda^{-1} \Sigma) \quad k = 1 \dots K$
<i>Stationary Diag-Cov</i> [$p + Kp$] (Bias with stationary within component modulation)	$\sigma_i \sim \mathcal{G}(a_0, b_0), \quad i = 1..p$ $\Sigma^{-1} = \text{diag}([\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_p^{-1}])$ $\mu^{(k)} \sim \mathcal{N}(\mu_0, \lambda^{-1} \Sigma) \quad k = 1 \dots K$

Table 1: Overview of the six different HMM emission model parameterizations tested. The model is written for the emission space \mathbb{R}^p , i.e. we observe time series from $i = 1 \dots p$ regions or independent components, and we model that with K states. The diag-operator used above takes a p -dimensional vector as input and produces a $p \times p$ matrix with the input vector in the diagonal and zeros elsewhere. Furthermore, $\mathcal{G}(a, b)$ denotes the gamma distribution.

($K = 2$) but the states differ in their mean values over the two groups making the classification task possible. However, in the bottom of Figure 1 we show the static covariance matrix for each group, i.e., equivalent to fitting the "Zero-Mean" model with one state. We notice that the classification task is still feasible since the two covariance matrices are very different even though we have a model mismatch in terms of which model generated the data.

To investigate this more systematically, we generated a synthetic dataset containing two groups with 100 subjects in each; the individual subjects data were generated with $T = 150$ timepoints (matching the data used in the subsequent analysis) in $p = 3$ dimensions. We used the state-means from the synthetic data illustrated in Figure 1 (left panel), and otherwise identical parameters across the two groups (i.e. π , π_0 , and diagonal covariance). The data were demeaned and set to unit variance as done in the GIFT-toolbox (cf. section below). The classification accuracy obtained from 10-fold stratified cross-validation can be seen in Figure 1 (right panel). We see that all the emission-models can achieve perfect classification accuracy, except the "Diag-Cov Zero-Mean" model that is unable to account for the dynamic difference in mean activation present across the two groups.

Schizophrenia Classification We ran our analysis on a cohort consisting of 192 subjects' resting-state fMRI data (COBRE) [22]. Of those, 101 subjects were diagnosed as schizophrenic or schizoaffective (SZ) and 91 subjects were healthy controls (HC). We ran a gICA using the GIFT tool-

box [23] with the ERBM algorithm [24] and 85 components. We restarted the algorithm 25 times and chose the best run using the minimum spanning tree (MST) criterion [25]. Afterwards we calculated the fractional amplitude of low-frequency fluctuation (fALFF) [26] of each component and removed all components with a fALFF lower than 3, yielding 48 components. Finally, we visually inspected the spatial maps and removed four additional components that had spatial overlap known noise sources (e.g. ventricles), such that we ended up with 44 ICs. Note that GIFT by default standardizes the time-series to have zero mean and unit variance which will become important when we compare the different model parameterizations. We estimate the accuracy of the classifiers by stratified 10-fold cross-validation. Each HMM-model was initialized 5 times, and the model with the best free-energy was chosen for the subsequent classification step. In Figure 1 we report the mean accuracy over folds, and the standard error, i.e. the standard deviation on the mean. We also report the performance of the baseline classifier, that assigns all data points in the test set to the largest class from the training set.

From the performance curves of the different HMM models we observe that all models except the "Diag-Cov Zero-Mean" have similar classification accuracies (the errorbars overlap). There seems to be a low influence on how many states we choose; there are intervals, i.e. from 4-6 states, where the more complex models "Mean+Cov" and "Zero-Mean" pull ahead in average classification accuracy, however

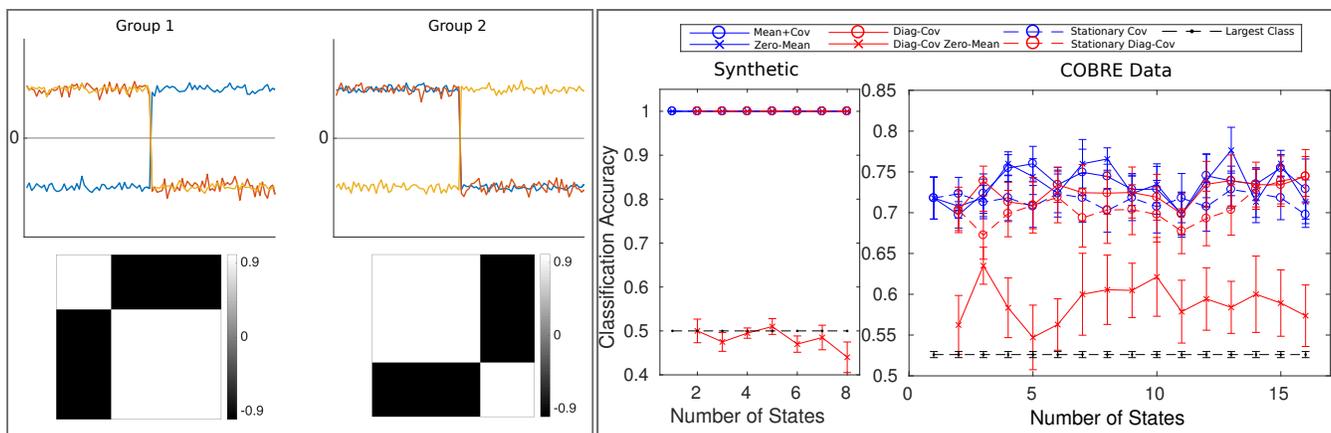


Fig. 1: *Left panel:* Synthetic toy-data showing the equivalence of different emission models. We generated two datasets in three dimensions each from a “Shared Diag-Cov” emission model (see top). Each model contains two states with a change point in the middle of the sequence. Below, we visualize the empirical correlation matrix for each data set. *Right panel:* Classification accuracy as a function of the number of states used in the different HMM-models, where the errorbars indicate the standard error over folds. Accuracy was estimated based on stratified 10-fold cross-validation. Note that the diagonal covariance models (“Diag-Cov”, “Diag-Cov Zero-Mean” and “Shared Diag-Cov”) start in $K = 2$ due to time-series standardization employed by the GIFT-toolbox, which makes the one state model unable to discriminate.

the errorbars still overlap with some of the simpler models.

4. DISCUSSION

In this work, we evaluated different assumptions on dFC within the HMM framework, by their ability to discriminate between schizophrenic patients (SZ) and health controls (HC) based on a short resting state fMRI scan.

Answer to research question 1: The performance gap between the full-parameterized model (with both mean and covariance for each state) and the more constrained models was fairly low. Only the “Diag-Cov Zero-Mean” model assuming that only the variance of the components varies over time, gave a noticeable drop in classification accuracy when compared to the other models. As simple models accounting for dynamic changes in the mean performed on par with models accounting for interactions between components this could indicate that the ICA we have employed as a “preprocessing” step has sufficiently demixed the problem. Thus the discriminative signal is mainly characterized by within component differences, and not in their coupling. From Figure 1 it seems that different model parameterizations can carry the same discriminative information. For example, if a certain state is characterized by one region having above mean activation and another region having below mean activation, this can be modeled in several ways in the HMM as illustrated in the synthetic data. The most natural way would be to do this using an emission model with a mean, however, a zero-mean model with full covariance could also model this by a large negative covariance between the two regions in question. Since we do not see a large discriminative effect in the

more complex emission models compared to their constrained counterparts this could make the case that the differences between SZ and HC is adequately captured by non-stationary mean IC activation.

Answer to research question 2: We saw that the performance of the different models was not highly influenced by the number of states chosen in the model. This could again be an effect of the representation that we have chosen, i.e. the ICA. If all the “dynamics” are captured by the ICs and we are in a sufficiently demixed space then there is no need to subsequently fit a temporally dynamic model like the HMM beyond how these ICs are stationary coupled (i.e., the “Zero-Mean” emission model for $K = 1$).

We stress that there can be a distinction between the model that is best for *classification* and the model that best *characterizes* the data. The conclusion we make about the dynamic models here are based on their ability to discriminate between two populations, and even though we conclude that a simple emission model can bring us a long way, this does not mean that full-covariance models should be ruled out. However, care should be taken when estimating many parameters (such as full covariance matrices in the complex emission models) when data is limited. The difference between characterization and classification has to be investigated further, along with the relationship between different subspace representations, such as PCA, ICA as well as atlas parcellations into functional units, and how these representations influence the estimated dynamic functional connectivity. We argue that the presently considered predictive classification accuracy is an important complementary tool to tools quantifying models ability to characterize data [11].

5. REFERENCES

- [1] R. M. Hutchison et al., “Dynamic functional connectivity: promise, issues, and interpretations,” *Neuroimage*, vol. 80, pp. 360–378, Oct. 2013.
- [2] V. D. Calhoun and T. Adali, “Time-Varying brain connectivity in fMRI data: Whole-brain data-driven approaches for capturing and characterizing dynamic states,” *IEEE Signal Process. Mag.*, vol. 33, no. 3, pp. 52–66, May 2016.
- [3] E. A. Allen et al., “Tracking whole-brain connectivity dynamics in the resting state,” *Cereb. Cortex*, vol. 24, no. 3, pp. 663–676, Mar. 2014.
- [4] T. O. Laumann et al., “On the stability of BOLD fMRI correlations,” *Cereb. Cortex*, Sept. 2016.
- [5] S. Shakil et al., “Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states,” *Neuroimage*, vol. 133, pp. 111–128, Mar. 2016.
- [6] R. Hindriks et al., “Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI?,” *Neuroimage*, vol. 127, pp. 242–256, Feb. 2016.
- [7] D. Vidaurre et al., “Spectrally resolved fast transient brain states in electrophysiological data,” *Neuroimage*, vol. 126, pp. 81–95, Feb. 2016.
- [8] A. P. Baker et al., “Fast transient networks in spontaneous human brain activity,” *Elife*, 2014.
- [9] S. Ryali et al., “Temporal dynamics and developmental maturation of salience, default and Central-Executive network interactions revealed by variational bayes hidden markov modeling,” *PLoS Comput. Biol.*, vol. 12, no. 12, pp. e1005138, Dec. 2016.
- [10] S. F. V. Nielsen et al., “Nonparametric modeling of dynamic functional connectivity in fMRI data,” in *NIPS 2015 Workshop on Machine Learning and Interpretation in Neuroimaging*, I. Rish et al., Eds. Jan. 2016, arxiv.org.
- [11] S. F. V. Nielsen et al., “Predictive assessment of models for dynamic functional connectivity,” *Neuroimage*, vol. 171, pp. 116–134, Dec. 2017.
- [12] J. Ou et al., “Characterizing and differentiating brain state dynamics via hidden markov models,” *Brain Topogr.*, vol. 28, no. 5, pp. 666–679, Sept. 2015.
- [13] J. Su et al., “Hereditary characteristics of schizophrenia shown by dynamic functional connectivity analysis of resting-state functional MRI scans of unaffected siblings,” *Neuroreport*, vol. 27, no. 11, pp. 843–848, Aug. 2016.
- [14] C. M. Cassidy et al., “Dynamic connectivity between brain networks supports working memory: Relationships to dopamine release and schizophrenia,” *J. Neurosci.*, vol. 36, no. 15, pp. 4377–4388, Apr. 2016.
- [15] S. Lefebvre et al., “Network dynamics during the different stages of hallucinations in schizophrenia,” *Hum. Brain Mapp.*, vol. 37, no. 7, pp. 2571–2586, July 2016.
- [16] H. Shen et al., “Internetwork dynamic connectivity effectively differentiates schizophrenic patients from healthy controls,” *Neuroreport*, vol. 25, no. 17, pp. 1344–1349, Dec. 2014.
- [17] X. Wang et al., “Aberrant intra-salience network dynamic functional connectivity impairs large-scale network interactions in schizophrenia,” *Neuropsychologia*, vol. 93, no. Pt A, pp. 262–270, Dec. 2016.
- [18] Y. Du et al., “Dynamic functional connectivity impairments in early schizophrenia and clinical high-risk for psychosis,” *Neuroimage*, Oct. 2017.
- [19] E. Damaraju et al., “Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia,” *Neuroimage Clin*, vol. 5, pp. 298–308, July 2014.
- [20] D. Vidaurre et al., “Discovering dynamic brain networks from big data in rest and task,” *Neuroimage*, June 2017.
- [21] M. J. Beal, *Variational Algorithms for Approximate Bayesian Inference*, Ph.D. thesis, 2003.
- [22] S. Ma et al., “Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis,” *Neuroimage*, vol. 90, pp. 196–206, Apr. 2014.
- [23] V. D. Calhoun et al., “A method for making group inferences from functional MRI data using independent component analysis,” *Hum. Brain Mapp.*, vol. 14, no. 3, pp. 140–151, Nov. 2001.
- [24] X. L. Li and T. Adali, “Blind spatiotemporal separation of second and/or higher-order correlated sources by entropy rate minimization,” in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. Mar. 2010, pp. 1934–1937, ieeexplore.ieee.org.
- [25] W. Du et al., “The role of diversity in complex ICA algorithms for fMRI analysis,” *J. Neurosci. Methods*, vol. 264, pp. 129–135, May 2016.
- [26] Q.-H. Zou et al., “An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF,” *J. Neurosci. Methods*, vol. 172, no. 1, pp. 137–141, July 2008.