Testing group differences in state transition structure of dynamic functional connectivity models

Author1*, Author2*,†,
Author3* and Author4*, ‡
* Address 1
† Address 2
‡ Address 3
Corresponding email: oneauthor@email.com

***Abstract*—Understanding the origins of intrinsic time-varying functional connectivity remains a challenge in the neuroimaging community. However, some associations between dynamic functional connectivity (dFC) and behavioral traits have been observed along with gender differences. We propose a permutation testing framework to investigate dynamic differences between groups of subjects. In particular, we investigate differences in fractional occupancy, state persistency and the full transition probability matrix. We demonstrate our framework on resting state functional magnetic resonance imaging data from 820 healthy young adults from the Human Connectome Project considering two prominent dFC models, namely sliding-window k-means and the Gaussian hidden Markov model. The variables showing consistent significant dynamic differences were limited to gender and the degree of motion in the scanner. We observe for the data considered that a large sample size (here 500 subjects) is needed to to draw reliable conclusions about the significance of those variables. Our results point to dynamic features providing limited information with regard to behavioral traits despite a relatively large sample size.**

I. INTRODUCTION

Neuroimaging has over the last decade moved from localizing brain function, mainly using statistical parametric mapping (SPM), and into characterizing *functional connectivity* (FC), i.e. the statistical dependencies between segregated brain regions. Especially in resting-state (rs) functional magnetic resonance imaging (fMRI) a lot of research papers have investigated how we can explain FC-differences in healthy populations [1], [2] and how we can use FC as biomarkers for neuropsychiatric diseases [3].

More recently, the implicit/explicit assumption of temporal stationary FC in rs-fMRI has been questioned and investigated [4], [5], [6], [7], which has fueled the modeling of so-called dynamic FC (dFC) *states*. dFC states are a discrete set of FC patterns that reoccur in time, both within subjects [8], [9] and across a population [5], [10]. The two most prominent methods for modeling dFC states is the sliding-window k-means (SWKM) [5] and the Gaussian hidden Markov model (HMM) [11], [10].

Recently, the temporal characteristics of dFC and their relation to cognitive measures has been investigated. Ma et al. [12] investigated the use of SWKM on rs-fMRI data from a cohort of patients diagnosed with schizophrenia and healthy controls. They found qualitative differences between the two groups in the so-called *transition matrix* estimated post-hoc from dFC states. The transition matrix quantifies for all time steps the probability of switching between any two states. Ma et al. [12] did not apply statistical testing of the transition differences.

Vidaurre et al. [10] trained a Gaussian hidden Markov model (HMM) on rs-fMRI from healthy young adults in the Human Connectome Project (HCP). The HMM is a probabilistic generative model of the data that assumes a latent discrete state space which is 1st order Markovian. Here the transition matrix is estimated directly from the HMM fitting procedure. They found that the states extracted had a hierarchical structure in terms of time each subject spent in each state, denoted fractional occupancy (FO). Two meta-states from the top of the hierarchy were then extracted and the difference in FO between the two meta-states, called the meta-state profile, associated with behavioral data significantly better than random (obtained through permutation testing). Furthermore, a comparison of the transition matrix in subgroups defined by their meta-state FO was carried out and showed qualitative differences. Vidaurre et al. [10] did however not carry out a quantitative analysis of the relationship between transition features and behavioral data.

In this paper, we propose a permutation framework for testing for group differences in the transition dynamics of dFC. We apply the framework to the rs-fMRI data from the Human Connectome Project [13] considering both SWKM and HMM. We do this by first training a dFC model (HMM and SWKM) on the entire population (820 subjects), and subsequently we estimate the transition matrix in two subgroups based on behavioural data (gender, motion, personality traits, etc). To characterise the difference between the two transition matrices we use the total variation (TV) distance of probability measures. We further contrast the performance

---

Thanks a lot. Funding funding funding Funding funding fundingFunding funding fundingFunding funding fundingFunding funding funding. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University

to simple properties of the dFC models given by fractional occupancy (FO) and a measure of self-transitions denoted global state persistency (GSP). In particular we investigate; 1) What behavioral variables significantly influence transition dynamics and how influenced is transition dynamics of head motion? 2) What aspects of the transition dynamics are important for characterizing these differences? 3) In this population, how many subjects are necessary to reliably detect group differences?

II. METHODS

A. Dynamic functional connectivity state transitions

Dynamic functional connectivity (dFC) models in general describe the changes in temporal correlation between two brain regions (i and j), c_{ij}. Thus at each time point, $t = 1 \dots T$, we have a snapshot of the FC between all pairs of regions, $\mathbf{C}^{(t)}$. A dFC *state* model further assumes that the $\mathbf{C}^{(t)}$'s can be clustered into K states, which yields a *state sequence*, $\mathbf{z}$, i.e. the assignment of each time point to one of the K states, $z_t \in \{1, 2, \dots, K\}$.

Assume that we have obtained a state sequence from a dFC state model (e.g. HMM or SWKM), then the K-by-K transition matrix, $\mathbf{P}$, can be written as,

$$\mathbf{P}_{k,k'} = \frac{\sum_{t=2}^{T} \delta(z_t = k', z_{t-1} = k)}{\sum_{t=2}^{T} \delta(z_{t-1} = k)}. \qquad (1)$$

The element $\mathbf{P}_{k,k'}$ thus quantifies the probability of transitioning from state k to state k'. Furthermore, each row is a probability distribution meaning that it sums to one, $\sum_{k'} \mathbf{P}_{k,k'} = 1$.

We also quantify the overall persistency of all the states, which we will denote the *global state persistency* (GSP), by taking the mean of the diagonal of the transition matrix, i.e.

$$\text{GSP} = \frac{\sum_j \mathbf{P}_{j,j}}{K}. \qquad (2)$$

Finally, we consider fractional occupancy FO_k which is a commonly used statistic to characterize clustering solutions [10], which can be calculated as,

$$\text{FO}_k = \frac{\sum_t \delta(z_t = k)}{T}, \qquad (3)$$

i.e. this quantifies how much of the total time is spent in the state k. Notice that this also sums to 1 over states and thereby forms a probability distribution, however it disregards the temporal structure of the state sequence.

B. Permutation testing using group information

To assess statistical differences between the dFC transition features of two groups we use approximate nonparametric permutation testing [14]. We investigate dFC transitions at a population level where we have data from S subjects, where each subject's state sequence can be denoted $\mathbf{z}^{(s)}$ for $s = 1 \dots S$. This state sequence is obtained by a population-level analysis, i.e. all subjects (regardless of grouping) have been concatenated into one long sequence. Given the grouping information, $\mathbf{g} \in [1, 2]^S$, we want to post-hoc estimate the difference in transition patterns between the groups. Each group's transition matrix is estimated on the collection of state sequences, i.e. group 1 has the transition matrix $\mathbf{P}^{(1)}$ estimated from $\mathbf{Z}^{(1)} = \{\mathbf{z}^{(s)} \; \forall \; s \; : \; \mathbf{g}_s = 1\}$. Another approach would be to train the dFC model with a transition matrix for each group, however, this approach is computationally expensive as we would need to retrain the model for each permutation.

As a distance measure between the transition probability matrices we use the total variation measure (TV) summed over the rows of $\mathbf{P}$. The TV between two probability distributions corresponds to the largest difference in probability which the two distributions assigns to the same event [15]. Another way to measure "closeness" is the Kullback-Leibler (KL) divergence, which is related to TV through Pinsker's inequality. However, this has the disadvantage that it is not symmetric and degenerates when an element has zero probability mass (due to a logarithm). We investigated using a regularized version of KL instead of TV with no difference in the conclusions of this paper. The TV distance can be written as,

$$\text{TV}(\mathbf{P}^{(1)}, \mathbf{P}^{(2)}) = \sum_{k=1}^{K} \frac{1}{2} \sum_{j=1}^{K} \left| \mathbf{P}^{(1)}_{k,j} - \mathbf{P}^{(2)}_{k,j} \right|. \qquad (4)$$

This same measure can be applied to the fractional occupancies (FO) for each group. For the GSP measure we take the absolute value of the difference between the two groups' GSP. For the permutation testing we permute the group labels and reestimate the transition matrices (and FO) for the permuted groups and calculate the distance between them. We thereby obtain a null-distribution of the considered measure between the groups by repeating the procedure for a large number of permutations as defined by the smallest p-value obtainable [16]. We used 10^5 permutations for our main analysis, which lets us obtain a minimum p-value of 10^{-5}.

III. RESULTS

We investigate the above permutation testing framework on resting state (rs) functional magnetic resonance imaging (fMRI) data from the Human Connectome Project (HCP) 820 subject release [13]. The data has been parcellated into 50 components using a group independent component analysis (ICA) publicly available through the HCP website[1]. Data were temporally concatenated and standardised such that each IC time-course within a subject had zero mean and unit variance. Afterwards, we ran the variational Bayes hidden Markov model (HMM) using the `HMM-MAR` MATLAB-toolbox[2] with $K = 12$ states and the stochastic inference engine [11]. All states had individual mean and full covariance in order to be comparable to the analysis carried out by [10]. We in addition ran the sliding-window k-means (SWKM) with the same number of states as the HMM ($K = 12$) using a window of length $W = [60, 100, 150]$ convolved with a Gaussian ($\sigma = 3TR$) [5] sliding the window one TR at a

---

[1] https://db.humanconnectome.org/

[2] https://github.com/OHBA-analysis/HMM-MAR

time. We did not use shorter window lengths because this necessitates regularization of the correlation matrix, such as the sparse-inverse regularization approach from [5], which was too computationally demanding. For the k-means inference we used the `litekmeans` implementation [17]. We investigated grouping the subjects into two groups according to 10 behavioral variables; gender, the five factor traits [18], two self reported measures of stress, fluid intelligence and a measure of head motion estimated from the realignment procedure (session average). All of the continuous variables were thresholded to match the proportions in the gender variable. We discarded four subjects that had missing values among the behavioral data we chose to investigate.

The results of the permutation testing for the fractional occupancy (FO), global state persistency (GSP) and transition probability matrix (TPM) can be seen in Figure 1. In general gender and motion (unsurprisingly) yield significant differences in almost all of the measures and models with few interesting exceptions. Looking at the results for the HMM we note that in the FO the largest difference is observed for the motion variable whereas in the transition matrix this is true for the gender variable. And looking at the GSP it is only the gender variable that overall shows significant differences.

For the SWKM (three rightmost columns of the figure) we observe very similar results for the short window lengths ($W = [60, 100]$), however when we increase the window length to $W = 150$ the GSP no longer shows significant results for any of the behavioral variables. Furthermore, FO and TPM differences for Gender and Motion move closer to the tail of the null-distribution as compared to the shorter window lengths.

IV. DISCUSSION

The neural origins of dFC in resting state fMRI is still not very well understood. To what extent it is best explained by cognitive differences in the subjects, ongoing cognitive processing, anatomical differences or noise confounds remains an open question. In this paper we have presented a framework for investigating dFC transition differences in groups of subjects using permutation testing. We applied this to healthy adults' resting state fMRI data from the Human Connectome Project.

Overall, we found no statistical evidence to support dFC differences in groups defined by higher-order cognitive and psychological traits (such as the five factor model) in neither of the dFC models considered (HMM and SWKM). However, we acknowledge that the thresholding we have applied to the continuous variables reduces the resolution, such that detection of transition differences is no longer possible. Gender showed significant differences in (almost) all of our analyses; however, this was expected since sex differences in anatomy are quite large, which could lead to systematic differences in the BOLD signal [19]. Recently, a machine learning model based on neuroanatomical features was trained on 967 subjects was trained to predict their gender, and achieved a 86% (cross-validated) classification accuracy [20]. Our analysis revealed that grouping subjects by how much they moved inside the scanner also gave significant differences in dFC features. This is also fairly unsurprising as motion has been put forward as a strong bias in discovering behavior and static FC relationships [21]. Furthermore, in the domain of dFC head motion has been attributed the strongest source of dFC variance by Laumann et al [22]. We investigated three dFC temporal features derived from the state sequence. FO and TPM differences were significant for gender and motion, whereas for GSP gender was the only significant variable across both HMM and SWKM. This indicates that head motion influences transitions to new states and overall time spent in particular states more compared to state persistency.

Our empirical investigation into the power of the permutation testing framework shows that we need quite a lot of subjects (> 500) to get reliable significant differences in the transition probability matrices (cf. bottom of Figure1). However, the absolute differences between the elements of the TPMs between males and females were very low (on the order of 10^{-3}). This shows that the effect is very small but reliable enough to be detectable in the large sample size. Future work will include using prediction of continuous behavioral variables on held-out subjects to investigate and disentangle FC and dFC features.

REFERENCES

[1] S. M. Smith et al., "A positive-negative mode of population covariation links brain connectivity, demographics and behavior," *Nat. Neurosci.*, vol. 18, no. 11, pp. 1565–1567, Nov. 2015.

[2] S. Smith, "Linking cognition to brain connectivity," *Nat. Neurosci.*, vol. 19, no. 1, pp. 7–9, Jan. 2016.

[3] T. Yamada et al., "Resting-State functional Connectivity-Based biomarkers and functional MRI-Based neurofeedback for psychiatric disorders: A challenge for developing theranostic biomarkers," *Int. J. Neuropsychopharmacol.*, vol. 20, no. 10, pp. 769–781, Oct. 2017.

[4] C. Chang and G. H. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI," *Neuroimage*, vol. 50, no. 1, pp. 81–98, Mar. 2010.

[5] 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.

[6] R. M. Hutchison et al., "Dynamic functional connectivity: promise, issues, and interpretations," *Neuroimage*, vol. 80, pp. 360–378, Oct. 2013.

[7] V. D. Calhoun et al., "The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery," *Neuron*, vol. 84, no. 2, pp. 262–274, Oct. 2014.

[8] J. Gonzalez-Castillo et al., "Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 28, pp. 8762–8767, July 2015.

[9] S. F. V. Nielsen et al., "Predictive assessment of models for dynamic functional connectivity," *Neuroimage*, vol. 171, pp. 116–134, Dec. 2017.

[10] D. Vidaurre et al., "Brain network dynamics are hierarchically organized in time," *Proc. Natl. Acad. Sci. U. S. A.*, Oct. 2017.

[11] D. Vidaurre et al., "Discovering dynamic brain networks from big data in rest and task," *Neuroimage*, June 2017.

[12] 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.

[13] S. M. Smith et al., "Resting-state fMRI in the human connectome project," *Neuroimage*, vol. 80, pp. 144–168, Oct. 2013.

[14] T. E. Nichols and A. P. Holmes, "Nonparametric permutation tests for functional neuroimaging: a primer with examples," *Hum. Brain Mapp.*, vol. 15, no. 1, pp. 1–25, Jan. 2002.

[15] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times: Second Edition*, American Mathematical Soc., Oct. 2017.

[16] A. M. Winkler et al., "Permutation inference for the general linear model," *Neuroimage*, vol. 92, pp. 381–397, May 2014.

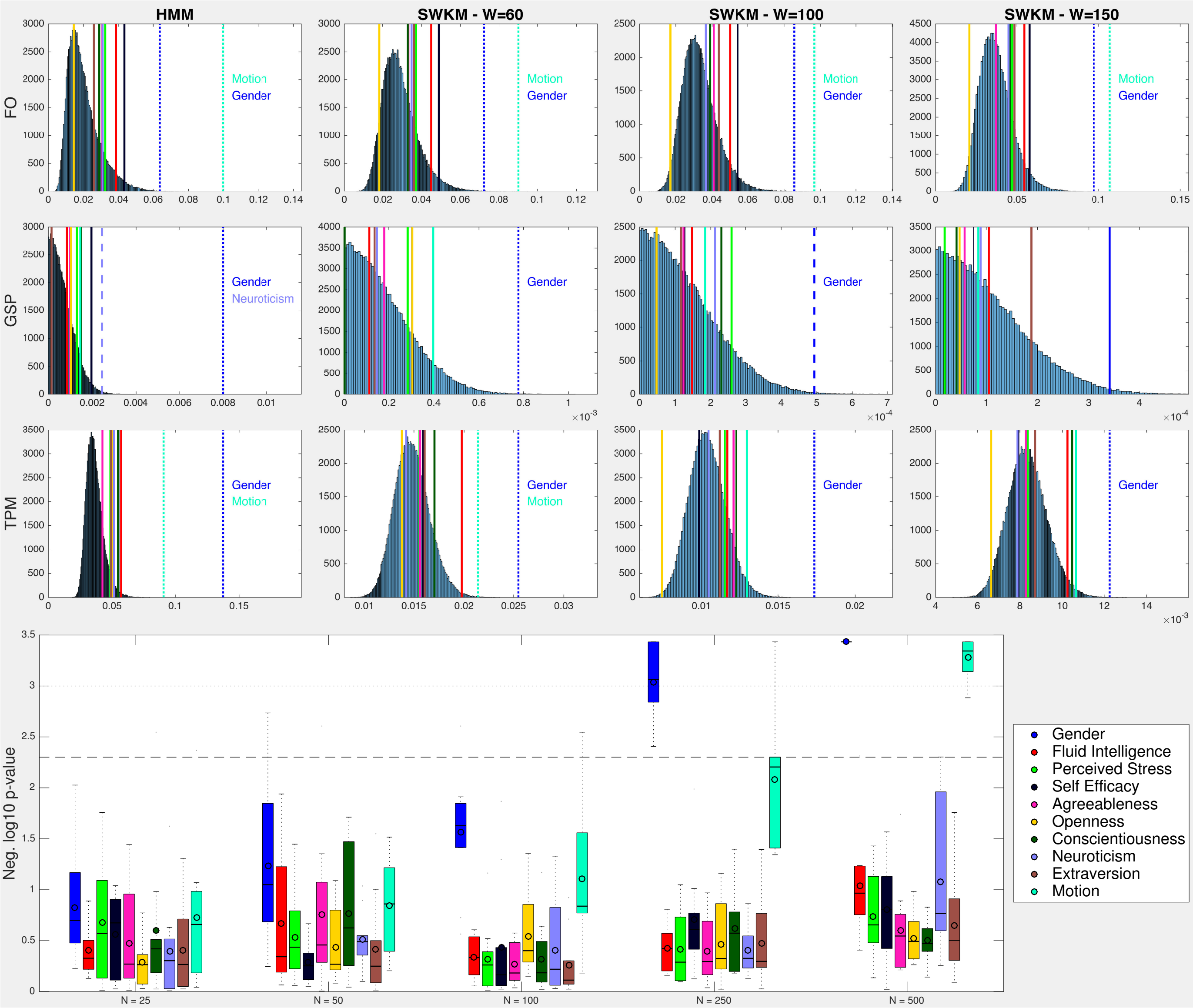

Fig. 1: *Top figure:* Permutation results for the SWKM and HMM using 10^5 permutations in the null-distribution. The SWKM and HMM were each retrained 10 times and we chose the best model according to the minimum cost function value. Behavioral variables were split to match the proportions in the *Gender* variable, such that hypothesis testing could be carried out using the same null-distribution for all behaviorals. For each behavioral it has been indicated by the linestyle if significant differences were detected (Bonferroni corrected with $\alpha = 0.05$ for dashed lines and $\alpha = 0.01$ for dotted lines). The significant ($\alpha = 0.05$ level) variable names are indicated in each subplot.
Bottom figure: We analyse the influence of the number of subjects ($N = [25, 50, 100, 250, 500]$) on the permutation framework. Using the best HMM solution (out of 10 restarts), trained on the entire HCP820 data, we calculated TV on the TPMs for different groups (described above) and we report the estimated p-value using 10^4 permutations. The boxplot above is the negative log p-value over 10 random subsets (H0: grouping yields same TV). The dashed black line indicates significance level $\alpha = 0.05$ and the dotted line significance level $\alpha = 0.01$ (both Bonferroni-corrected).

[17] D. Cai et al., "Document clustering using locality preserving indexing," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 12, pp. 1624–1637, Dec. 2005.

[18] J. M. Digman, "Personality structure: Emergence of the Five-Factor model," *Annu. Rev. Psychol.*, vol. 41, no. 1, pp. 417–440, Jan. 1990.

[19] B. B. Biswal et al., "Toward discovery science of human brain function," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 10, pp. 4734–4739, Mar. 2010.

[20] F. Sepehrband et al., "Neuroanatomical morphometric characterization of sex differences in youth using statistical learning," *Neuroimage*, vol. 172, pp. 217–227, Feb. 2018.

[21] J. S. Siegel et al., "Data quality influences observed links between functional connectivity and behavior," *Cereb. Cortex*, vol. 27, no. 9, pp. 4492–4502, Sept. 2017.

[22] T. O. Laumann et al., "On the stability of BOLD fMRI correlations," *Cereb. Cortex*, Sept. 2016.