NEW MEASURES OF BRAIN FUNCTIONAL CONNECTIVITY BY TEMPORAL ANALYSIS OF EXTREME EVENTS

Djalel-Eddine Meskaldji[⋆†] *Stephan Morgenthaler*[‡] *Dimitri Van De Ville*[⋆†]

[⋆] Medical Image Processing Lab (MIPLAB), Institute of Bioengineering, EPFL, CH-1015 Lausanne, Switzerland

[†] Department of Radiology and Medical Informatics, University of Geneva, CH-1211 Geneva, Switzerland

[‡] Applied statistics, Institute of mathematics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

ABSTRACT

Understanding brain structure and function can benefit from studying functional connectivity. A common methodology to measure functional connectivity between two brain regions is to estimate the correlation between their corresponding average time courses. Usually, these correlations are computed either via the Pearson estimator or the non-parametric Spearman estimator. However, these two measures do not fully reflect the information we want to extract about the spontaneous activity in the different areas of the brain. In this paper, we propose to estimate functional connectivity between two regions by modeling the activation parts of the time course as the extreme events and by measuring the co-activation between these events. We show that our new measure of functional connectivity contains key information about the co-activations, which is lost when using common functional connectivity measures; i.e., Pearson or Spearman correlation.

***Index Terms*—** Functional connectivity, extreme events, sufficiency, neuroimaging, fMRI.

1. INTRODUCTION

The study of dynamics of MRI signals has become essential to advance our understanding of brain function. Functional connectivity reflects the spontaneous fluctuations of brain activity by measuring correlation between fMRI time courses [1]. Whole brain connectivity is represented by the so-called functional connectivity matrix, also termed the functional connectome. After preprocessing of the fMRI data aiming to remove data acquisition artifacts and other non-desirable confounds, the connectivity is conventionally estimated by Pearson correlations between pairs of fMRI time courses of all brain regions [2, 3]. The matrix that we obtain is usually full and the functional connectome represents a complete graph. A more sparse representation of the functional connectome is obtained using regularized estimators. Often, these regularizations are applied to the coefficients of the inverse covariance matrix called the precision matrix, which is directly linked to the partial correlations between time courses [4]. Other recent studies of brain connectivity considered the fMRI activation signal as a phase transition time process by considering its extreme values [5].

Here, we also consider the activation signals and define (positive or negative) extreme values on the basis of a fixed threshold. We then present a new estimator of the functional connectivity: for each pair of regions, we measure two values: (1) the accordance, which measures the co-activation and the co-disactivation of a pair of tome courses, and (2) the discordance, a measure of activation-disactivation of a pair of time courses. We show that the new estimator reflects the dynamical features of spontaneous fluctuations of the brain activity better than the common estimators such as correlation. The proposed method is promising for the emerging interest in non-stationary behavior of fMRI signals.

2. METHODS

Functional connectivity (FC) is a measure of relationship between functional data. FC summarizes this relationship for the whole time interval by only one value (univariate or multivariate) for each pair of brain regions. Let $\mathbf{X} = \mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_T}$ be a multivariate stochastic process with $\mathbf{x} = x^{(1)}, \ldots, x^{(N)} \in \mathbb{R}^N$, observed in time points indexed by $\mathcal{T} = \{1, \ldots, T\}$. Let Θ be the FC that we would like to estimate from the data. Let η be an estimator (a function of the observed data). We note the estimated functional connectivity by $\eta(\mathbf{X}) = \hat{\Theta}_{\mathbf{X}}$. The most-used FC estimator is the sample Pearson correlation matrix $\eta(\mathbf{X}) = \hat{R}_{\mathbf{X}} = \left[\mathrm{diag}(\hat{S}_{\mathbf{X}})\right]^{-1} \hat{S}_{\mathbf{X}} \left[\mathrm{diag}(\hat{S}_{\mathbf{X}})\right]^{-1}$, where $\hat{S}_{\mathbf{X}} = \frac{1}{T}\mathbf{X}'\mathbf{X}$ is the covariance matrix estimate of the stochastic process $\mathbf{X}$. It is well known that this estimator is a good estimator for the true correlation matrix. However, is the correlation matrix itself a good candidate to represent FC? In other words, does the correlation matrix contain all the desired information about the FC. In order to clarify this question we introduce the notion of sufficiency, a well known

This work was supported in part by the Swiss National Science Foundation (grant number PP00P2-146318), and in part by the Center for Biomedical Imaging (CIBM) of the Geneva-Lausanne Universities and the EPFL, as well as the Leenaards and Louis-Jeantet foundations.

concept in estimation theory.

2.1. Sufficiency

Suppose that we collected functional data $\mathbf{X} = \mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_T}$ in order to estimate the parameter Θ. Let $f_\Theta(\mathbf{X})$ be the probability density function (PDF) for $\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_T}$. Let $\eta = \eta(\mathbf{X})$ be an estimator based on $\mathbf{X}$. Let $g_\Theta(\eta)$ be the PDF for $\eta(\mathbf{X})$. If the conditional PDF

$$h_\Theta(\mathbf{X}) = \frac{f_\Theta(\mathbf{X})}{g_\Theta[\eta(\mathbf{X})]}$$

is independent of Θ, then $\eta(\mathbf{X})$ is a sufficient statistic for Θ. In other words, $h_\Theta(\mathbf{X}) = h(\mathbf{X})$, and Θ does not appear in $h(\mathbf{X})$. Intuitively, this means that $\eta(\mathbf{X})$ contains all the information contained in $\mathbf{X}$ to estimate Θ, that is, knowing $\eta(\mathbf{X})$ (i.e., conditioning $f_\Theta(\mathbf{x})$ on $\eta(\mathbf{X})$) is sufficient for estimating the true unknown parameter Θ.

Often, a sufficient statistic for Θ is a summary statistic of $\mathbf{X} = \mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_T}$. If such a summary statistic is sufficient for Θ, then knowing this one statistic is just as useful as knowing all the T observations of the process for estimating Θ. The correlation matrix estimator $\hat{R}(\mathbf{X})$ is a sufficient estimator for the true correlation matrix. However, the best estimator η is an estimator that extracts all the information about the FC from the the available data $\mathbf{X}$. Furthermore, an estimator η_1 is better then a second estimator η_2 if it contains more information about the true unknown parameter Θ. We say in this case that η_1 is more sufficient than η_2.

Input The normalized observed multivariate process $\mathbf{X} = \mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_T}$, where $\mathbf{x_t} = x_t^{(1)}, \ldots, x_t^{(N)} \in \mathbb{R}^N$ and $t = 1, \ldots, T$.
A quantile threshold q.
Output An estimation of the functional connectivity, $\hat{\Theta}$.
Initialization $\hat{\Theta} = \mathbf{0} \in \mathbb{R}^{N \times N}$.
for $i \in \{1, ..., N\}$ **do**
  $T_i^+ = \{t \in \{1, \ldots, T\} : x_t^i > \Phi^{-1}(q)\}$.
  $T_i^- = \{t \in \{1, \ldots, T\} : x_t^i < -\Phi^{-1}(q)\}$.
  $\hat{\Theta}_{i,i} = \frac{1}{T}\left|T_i^+\right|$.
end
for $i \in \{1, ..., N-1\}$ **do**
  for $j \in \{(i+1), ..., N\}$ **do**
    $\cup T_{i,j}^+ = T_i^+ \cup T_j^+$;
    $\cap T_{i,j}^+ = T_i^+ \cap T_j^+$;
    $\cup T_{i,j}^- = T_i^- \cup T_j^-$;
    $\cap T_{i,j}^- = T_i^- \cap T_j^-$;
    $\cap T_{i,j}^\pm = \{T_i^+ \cap T_j^-\} \cup \{T_i^- \cap T_j^+\}$;
    $\hat{\Theta}_{i,j} = \left|\{\cap T_{i,j}^+\} \cup \{\cap T_{i,j}^-\}\right| / \left|\{\cup T_{i,j}^+\} \cup \{\cup T_{i,j}^-\}\right|$;
    $\hat{\Theta}_{j,i} = \left|\cap T_{i,j}^\pm\right| / \left|\{\cup T_{i,j}^+\} \cup \{\cup T_{i,j}^-\}\right|$.
  end
end

Algorithm 1: Estimation of the proposed FC.

2.2. Estimation of the functional connectivity

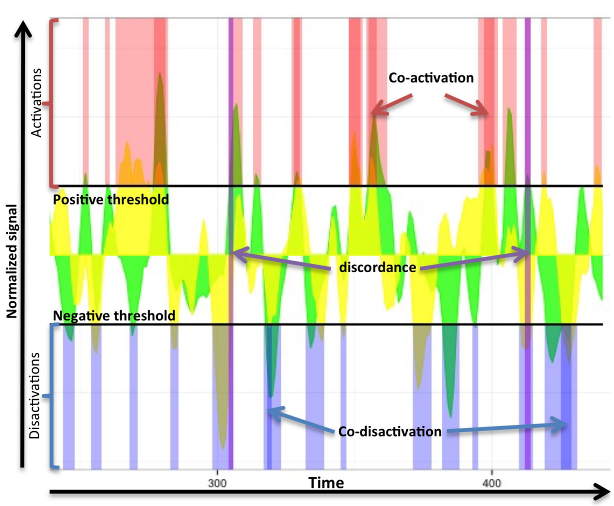
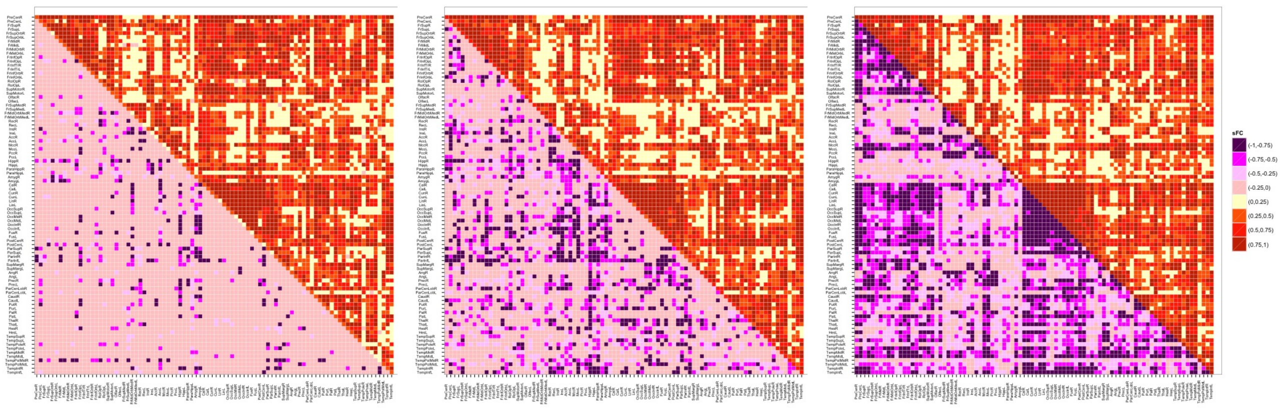


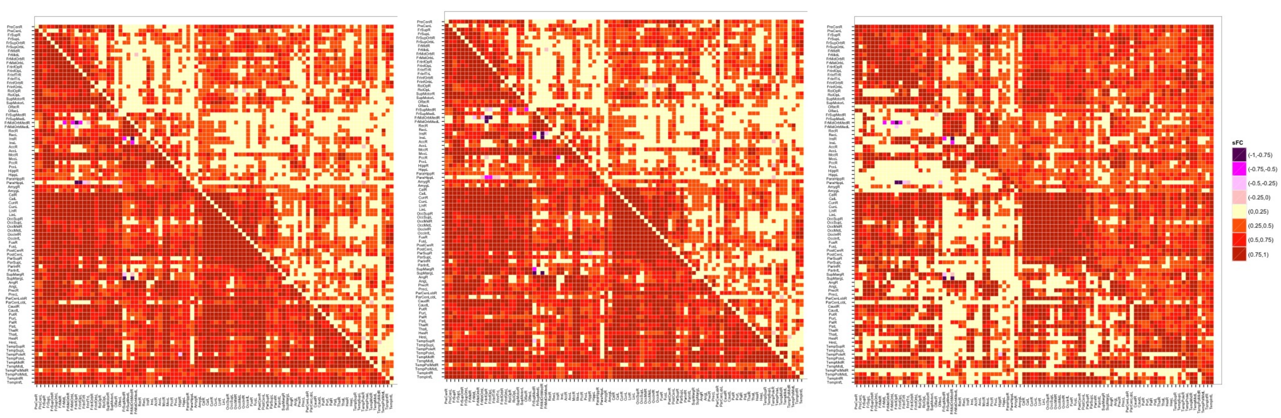
Fig. 1. Illustration of the different cases in the construction of the new FC estimator. The green and yellow curves represent a pair of normalized fMRI time courses.

FMRI time courses are considered as noisy observations of brain activity. In order to eliminate spurious fluctuations, we consider only extreme events of the observed time courses. We suppose that these extreme events represent significant activations or disactivations of the corresponding brain regions. Practically, after normalizing each time course, i.e., subtracting the mean and dividing by the standard deviation, the normalized time courses $\mathbf{x}^{(i)}, i = 1, \ldots, N$, are compared to a positive threshold and a negative threshold based on a predefined quantile q. More specifically, for each time course $\mathbf{x}^{(i)}$, we identify the sub-intervals corresponding to extreme events by $T_i^+ = \{t \in \{1, \ldots, T\} : x_t^i > \Phi^{-1}(q)\}$ and $T_i^- = \{t \in \{1, \ldots, T\} : x_t^i < -\Phi^{-1}(q)\}$, for positive and negative extreme events, respectively, where Φ is the CDF of the Gaussian distribution. Other distribution could be used depending on the assumptions. The ratio of the union of the significant positive extreme sub-intervals over the whole time interval length measures the proportion of significant activation of the corresponding brain region. This value is stored as the diagonal element i in the estimated FC matrix. Then, for each pair of time courses, $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$, we determine the size of the union of co-activation and co-disactivation interval times and we normalize by the size of the union of significant activation and disactivation interval times of the two time courses. The obtained value measures the *accordance* of co-activation and co-disactivation of the corresponding pair of brain regions, and is stored in the upper-triangular part of the



(a) FC with $q = 0.95$. (b) FC with $q = 0.90$. (c) FC with $q = 0.5$.

Fig. 2. Three FC matrices of one subject, estimated by Algorithm 1. The three matrices correspond to quantile threshold $q = 0.9$, $q = 0.95$, and $q = 0.50$, respectively. The upper-triangular part indicates the accordance in co-activation and co-disactivation, while the lower-triangular part indicates the discordance. The color map is based on the 0.25, 0.5, 0.75 and 1 quantiles of positive values (accordance) and negative values (discordance), respectively. The diagonal of the FC matrix indicates the percentage of the activation parts of each time course.



(a) FC with $q = 0.95$ vs. Pearson correlation (b) FC with $q = 0.95$ vs. Spearman correlation (c) FC with quantile $q = 0.50$ vs. Pearson correlation

Fig. 3. Comparison between FC derived with our algorithm (upper-triangular) and FC derived either by (a and c) Pearson correlation or by (b) Spearman correlation (lower-triangular). The FC estimated by our algorithm is summarized in the upper triangular part by adding the discordance (negative) values to the accordance (positive) values.

FC matrix. Similarly, we obtain the measure of *discordance* between two time courses be considering the size of positive-negative and negative-positive extreme interval times, also normalized by the size of the union of activation and disactivation interval times of the two time courses. This measure is stored in the lower-triangular part of the FC matrix. Note that all values of the estimated FC are normalized by construction between -1 and $+1$. The FC estimator is summarized in Algorithm 1. Figure 1 illustrates some of the concepts introduced in this section.

3. RESULTS AND DISCUSSION

We applied the new FC estimation algorithm to resting state (RS) fMRI data of healthy subjects from a previous study [6], and we compared the obtained FC matrices to those estimated by pair-wise Pearson and Spearman correlations. For each subject, time courses were obtained from 90 brain regions [7] by regional averaging. Figure 2 shows two FC matrices for the same subject, estimated with Algorithm 1, using three different values of the quantile threshold, i.e., $q = 0.95$,

$q = 0.90$ and $q = 0.5$. These matrices are of course non-symmetric. The upper-triangular part represents the accordance value, while the lower triangular part represents the discordance value. The diagonal indicates the proportion of activation for each time course. Both pieces of information are relevant for understanding brain functional connectivity. The estimation of FC is highly influenced by the threshold quantile. The higher is the threshold, the higher is the sparsity of the matrix obtained. High thresholds give more robust estimation in the sense that only highly significant activations and disactivations are considered. Two extremal values of quantile threshold are possible. The value $q = 1$ in which case, the estimated FC is identically the zero matrix $\mathbf{0}$. On the other hand, the value of $q = 0.5$, which corresponds to set both the positive threshold and the negative threshold to 0, leads to FC matrices in which the parts that contribute to the negative correlations are separated from the parts that contribute to the positive correlations. In this case, if we symmetrize the FC matrices by summing the upper-triangular part and the lower-triangular part we obtain FC matrices that resemble more to those obtained by Pearson or Spearman correlations (see Figure 3).

Finally, we could say that the new estimator of FC is more suitable since it brings extra relevant information about functional connectivity. Especially, the new estimator can differentiate between two situations where time courses could give weak correlations. The first case is when the correlation is weak along the time courses, while the second case is when strong positive correlations are annihilated by strong negative correlations. The multivariate property of the new estimator might bring more statistical power in group difference studies, and the extra information that it contains reflects concisely a dynamical feature of FC, which is an emerging topic in the field usually assessed using sliding-window correlation technique.

4. CONCLUSION

We proposed a new estimator of FC derived from fMRI time courses. The new estimator is more closely related to co-activation of brain regions by estimating accordance and discordance of co-activations and co-disactivations separately. This information is lost in common estimators of FC, which makes our estimator more sufficient in the estimation-theory sense. We also presented a simple algorithm to construct the new estimates of FC. We expect that expressing FC using our new estimator affords more accurate interpretations of the brain function, and helps to better disentangle between brain states and fMRI modalities, or even between different groups of subjects. The brain networks representing connectivity matrices as estimated by our method could be compared using adaptive multimodal statistical methods, such as the adaptive two step strategy [8, 9, 10, 11].

5. REFERENCES

[1] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *The Journal of Neuroscience*, vol. 26, no. 1, pp. 63–72, 2006.

[2] S. M. Smith, D. Vidaurre, C. F. Beckmann, M. F. Glasser, M. Jenkinson, K. L. Miller, T. E. Nichols, E. C. Robinson, G. Salimi-Khorshidi, M. W. Woolrich, D. M. Barch, K. Uğurbil, and D. C. V. Essen, "Functional connectomics from resting-state fmri," *Trends in Cognitive Sciences*, vol. 17, no. 12, pp. 666–682, 2013.

[3] J. Richiardi, H. Eryilmaz, S. Schwartz, P. Vuilleumier, and D. Van De Ville, "Decoding brain states from fMRI connectivity graphs," *NeuroImage*, vol. 56, no. 2, pp. 616–626, 2011.

[4] G. Varoquaux, A. Gramfort, J. baptiste Poline, and B. Thirion, "Brain covariance selection: better individual functional connectivity models using population prior," in *Advances in Neural Information Processing Systems 23*, J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, Eds., pp. 2334–2342. Curran Associates, Inc., 2010.

[5] E. Tagliazucchi, P. Balenzuela, D. Fraiman, and D. R. Chialvo, "Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis," *Frontiers in physiology*, vol. 3, pp. 15, 2012.

[6] H. Eryilmaz, D. Van De Ville, S. Schwartz, and P. Vuilleumier, "Impact of transient emotions on functional connectivity during subsequent resting state: A wavelet correlation approach," *NeuroImage*, vol. 54, no. 3, pp. 2481–2491, February 2011.

[7] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in {SPM} using a macroscopic anatomical parcellation of the {MNI} {MRI} single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273 – 289, 2002.

[8] D.-E. Meskaldji, L. Vasung, D. Romascano, J.-P. Thiran, P. Hagmann, S. Morgenthaler, and D. V. D. Ville, "Improved statistical evaluation of group differences in connectomes by screening–filtering strategy with application to study maturation of brain connections between childhood and adolescence," *NeuroImage*, vol. 108, no. 0, pp. 251 – 264, 2015.

[9] D.-E. Meskaldji and D. Van De Ville, "Multimodal graph theoretical analysis of functional brain connectivity using adaptive two-step strategy," *Proceedings of the Eleventh IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 919–922, 2014.

[10] D.-E. Meskaldji, E. Fischi-Gomez, A. Griffa, P. Hagmann, S. Morgenthaler, and J.-P. Thiran, "Comparing connectomes across subjects and populations at different scales," *NeuroImage*, vol. 80, no. 0, pp. 416 – 425, 2013, Mapping the Connectome.

[11] D.-E. Meskaldji, M.-C. Ottet, L. Cammoun, P. Hagmann, R. Meuli, S. Eliez, J.-P. Thiran, and S. Morgenthaler, "Adaptive Strategy for the Statistical Analysis of Connectomes," *PLoS ONE*, vol. 6, no. 8, pp. e23009, 08 2011.