Functional brain connectivity evaluated by an effective and more sufficient estimator based on extreme events

Djalel-Eddine Meskaldji[1,2], Stephan Morgenthaler[3], Dimitri Van De Ville[1,2]

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

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

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

djalel.meskaldji@epfl.ch, stephan.morgenthaler@epfl.ch
dimitri.vandeville@epfl.ch

Résumé − Une méthodologie commune pour mesurer la connectivité fonctionnelle du cerveau (CF) est d'estimer la corrélation entre les paires des signaux IRMf. Cet estimateur de la CF ne reflète pas toutes les l'informations que nous voulons extraire concernant l'activité spontanée des différentes régions du cerveau, dont on sait qu'elle est non-stationnaire. Nous proposons d'estimer la CF, de manière robuste, en modélisant les activations et désactivations des signaux IRMf comme des événements extrêmes et en mesurant les co-activations entre ces événements à l'aide d'estimateurs non-paramétriques. Le nouvel estimateur de la CF fait la distinction entre conformité et discordance d'activations pour chaque paire de regions, ce qui en fait un bon candidat pour représenter la non-stationnarité de la CF car il contient des informations clés supplémentaires qui reflètent de manière concise les caractéristiques dynamiques des fluctuations spontanées de l'activité cérébrale. La méthode proposée est prometteuse car elle augmente la sensibilité de détecter des effets entre différents groupes.

Abstract − A common methodology to measure brain functional connectivity (FC) is to estimate pairwise correlations between fMRI time courses. Such FC estimator does not fully reflect the information we want to extract about the spontaneous activity in the different regions of the brain, because activity has been proved non-stationary. Furthermore, the estimated FC could be highly influenced by the co-deactivation parts of the fMRI signals, which might induce connectivity misspecification and misleading interpretations. We propose to estimate FC, in a robust way, by modeling activation and deactivation parts of fMRI time courses as extreme events and by measuring the co-activations between these events using non-parametric estimators. The new FC estimator distinguishes between accordance and discordance of pairwise activations and deactivations, which makes the new estimator a good candidate to represent non-stationarity as it contains additional key information that concisely reflects dynamical features of spontaneous fluctuations of the brain activity. The proposed method is promising because it brings more sensitivity to detect group effects.

1 Introduction

Functional connectivity (FC) is aimed to describe spontaneous fluctuations of brain activity. FC between a pair of regions is defined as the statistical dependence between their time courses. One important modality of FC is functional MRI, which is based on the Bold signal. Before measuring FC, fMRI signals are pre-processed to remove data acquisition artifacts and other non-desirable confounds. Then, given an anatomical segmentation of the brain cortex, voxel-wise signals are averaged across each region to end up with a single time course per region of interest (ROI). Conventionally, FC is estimated by Pearson correlations between pairs of fMRI time courses of all brain regions [1, 2]. Spearman correlation is a non parametric alternative estimator of FC, which can lead to different values compared to Pearson correlation. Whole brain connectivity is represented by the so-called FC matrix, also referred to as the functional connectome. The matrix that we obtain is usually full and the functional connectome represents a complete graph.

Some important questions arise in this context. Which measure of FC is better in describing spontaneous activity and dynamics? Do the correlation, the covariance or even the partial correlation/covariance reflect well the spontaneous fluctuations or co-activations of brain regions? Whether we normalize fMRI signals or not to keep sig-

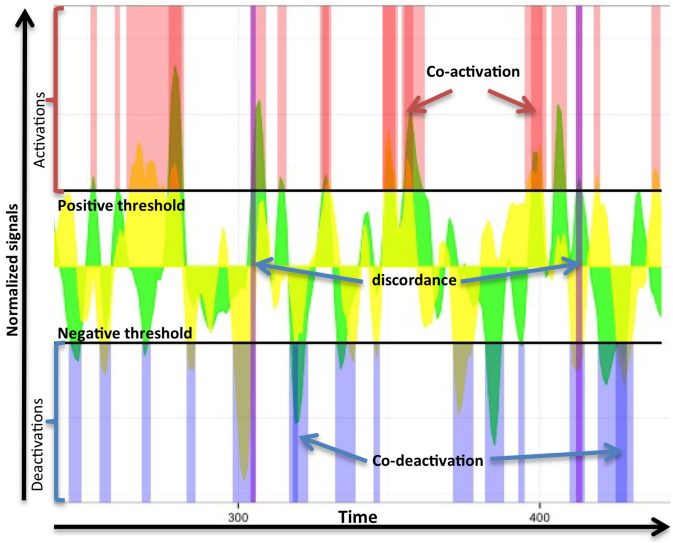

FIG. 1: Illustration of the different cases in the construction of the new FC estimator.

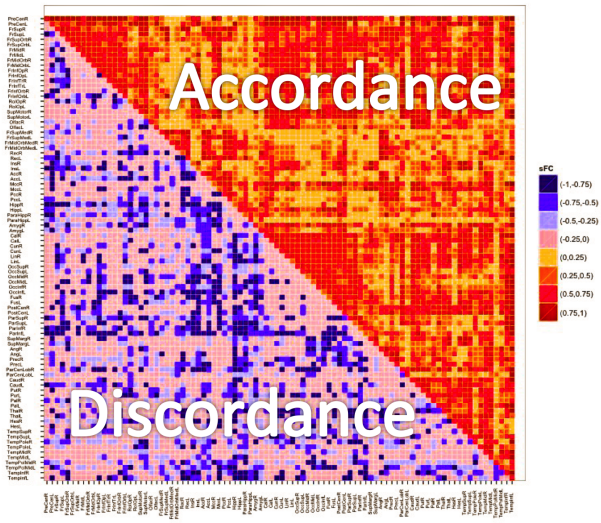

FIG. 2: An example of an estimated FC matrix. The upper triangular part represents the values of accordance, while the lower triangular part represents the values of discordance. The diagonal contains the proportion of activation of each ROI.

nal amplitude information? We are not going to address all these questions in this paper. However, we try to improve one of the aspects.

The main purpose of FC estimation is to detect activation in brain ROIs and, second, to measure how much pairs of ROIs are activated at the same time in resting state or during a task. Some recent studies tried to cope with the first problem by modeling the fMRI signals as a point process by considering its extreme values [3]. In this paper, we also consider extreme values of the signals and we propose to estimate FC by measuring how much two regions are co-activated (de-activated) and how much they are not. For each pair of regions, we measure two values: (1) the accordance, which measures the co-activation (and the co-deactivation) of a pair of time courses, and (2) the discordance, a measure of activation-deactivation of a pair of time courses. We obtain a bi-variate estimator of FC that is robust, well defined and that better reflects features of non-stationarity of spontaneous fluctuations of the brain activity. We show how the new estimator is different from the correlation based estimators in terms of sensitivity of detecting global and local differences between groups.

2 Methods

2.1 FC estimation

Let $\mathbf{Y} = \mathbf{y_1}, \mathbf{y_2}, ..., \mathbf{y_T}$ be a multivariate stochastic process that represents the fMRI signals, with $\mathbf{y} = y^{(1)}, \ldots, y^{(N)} \in \mathbb{R}^N$, observed in time points indexed by $\mathcal{T} = \{1, \ldots, T\}$, where N is the number of ROIs and T is the total acquisition time. According to our conceptualization of FC, a good estimator has to consistently estimate the statistical dependence and should reflect dynamics and non-stationarity features of FC. Then a good estimator, has to be robust, consistent and exhaustive. For the robustness, we propose to threshold the fMRI signals to keep only significant activations and deactivations of the fMRI signals. Concretely, we consider only extreme events of the observed time courses, assuming that these extreme events represent significant activations or deactivations of the corresponding brain regions. Practically, after normalizing each time course by subtracting the mean and dividing by the standard deviation of the time course, the normalized time courses $\mathbf{x}^{(i)}, i = 1, \ldots, N$, are compared to a positive 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 $T_i^- = \{t \in \{1, \ldots, T\} : x_t^{(i)} < \Phi^{-1}(1-q)\}$, for positive and negative extreme events, respectively, where Φ is the CDF of the Gaussian distribution. Doing so, we aim to eliminate spurious fluctuations of fMRI time courses, which are considered as noisy observations of brain activity.

Now we turn to statistical estimation of FC time courses. Our estimator should be exhaustive (sufficient), that is, it contains all information we want to extract from our signals to represent FC. It is important when studying FC between two brain regions to know how much each region is activated during the time course and how much these two regions are activated or deactivated at the same time, and how much they are in a opposite situation.

We set for each vector $\mathbf{x}$, the thresholded vector $\mathbf{x^u}$ such

that $x_t^u = 0$ if $x_t < u$ and $x_t^u = 1$ otherwise. Similarly, $\mathbf{x^l}$ derived from the vector $\mathbf{x}$ as $x_t^l = 0$ if $x_t > l$ and $x_t^l = -1$ otherwise. In the following algorithm we use this notation $< \mathbf{x}, \mathbf{y} >$ for the inner product of $\mathbf{x}$ and $\mathbf{y}$.
The ratio of the union of the significant positive extreme

Input The normalized observed multivariate process $\mathbf{X}^T = \mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N}$, where $\mathbf{x_i} = x_1^{(i)}, \ldots, x_T^{(i)} \in \mathbb{R}^T$ and $t = 1, \ldots, T$.
A quantile threshold q.
Output An estimation of FC.
for $i \in \{1, ..., N\}$ **do**
  Define $u = \Phi^{-1}(q)$ and $l = \Phi^{-1}(-q)$.
  $\pi_i = (\mathbf{x^u}_i * \mathbf{x^u}_i)/(\mathbf{x^u}_i * \mathbf{x^u}_i + \mathbf{x^l}_i * \mathbf{x^l}_i)$.
end
for $i \in \{1, ..., N-1\}$ **do**
  for $j \in \{(i+1), ..., N\}$ **do**
    $E_{i,j} = \sqrt{(\mathbf{x^u}_i * \mathbf{x^u}_i) + (\mathbf{x^l}_i * \mathbf{x^l}_i)}\sqrt{(\mathbf{x^u}_j * \mathbf{x^u}_j) + (\mathbf{x^l}_j * \mathbf{x^l}_j)}$
    $a_{i,j} = (\mathbf{x^u}_i * \mathbf{x^u}_j + \mathbf{x^l}_i * \mathbf{x^l}_j)/E_{i,j}$
    $d_{i,j} = (\mathbf{x^u}_i * \mathbf{x^l}_j + \mathbf{x^l}_i * \mathbf{x^u}_j)/E_{i,j}$
  end
end

Algorithm 1: Estimation of the proposed FC measures.

sub-intervals over the whole significant time interval length measures the proportion of significant activation of the corresponding brain region π_i. If the FC is represented by $N \times N$ matrix, this value is stored in the diagonal element i in the estimated FC matrix (see Figure 2). 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-deactivation interval times and we normalize by the size of the union of significant activation and deactivation interval times of the two time courses. The obtained value indicates the *accordance* of co-activation and co-deactivation of the corresponding pair of brain regions, and is stored in the upper-triangular part of the FC matrix. Similarly, we obtain the measure of *discordance* between two time courses by considering the size of positive-negative and negative-positive extreme interval times, also normalized by the size of the union of activation and deactivation interval times of the two time courses. This measure is stored in the lower-triangular part of the FC matrix. The FC estimator is summarized in Algorithm 1. Figure 1 illustrates some of the concepts introduced in this section.
We can easily show that $0 \leq \pi_i \leq 1$, $0 \leq a_{i,j} \leq 1$, $-1 \leq d_{i,j} \leq 0$ and $-1 \leq a_{i,j} + d_{i,j} \leq 1$, for all i and j.
We can also prove that for a given time course $\mathbf{x}$, the following holds: $a(\mathbf{x}, \mathbf{x}) = 1$, $a(\mathbf{x}, -\mathbf{x}) = 0$ and $d(\mathbf{x}, \mathbf{x}) = 0$, $d(\mathbf{x}, -\mathbf{x}) = -1$.

2.2 Statistical comparaisons

The new FC estimator could be used in different brain connectivity applications. Here we present a group study and we compare the sensitivity of our estimator against Pearson correlation estimator. The study consists in comparing brain FC between 13 healthy controls (HC) and 15 patients with multiple sclerosis (MS) [4]. We consider two levels of comparison: the local level and the global level [5]. Local comparison gives more refined results, but suffers from lack of power because of multiplicity correction. We study local differences by considering each single connection between a pair of ROIs. In the Pearson FC case, the local connectivity measure is the Pearson correlation, that is, $\rho_{i,j}^{(k)}$ for a given subject k and for each pair of ROIs i and j. We also used different local measures derived from our new FC estimator. For each subject k, we consider these two local measures: (1) the accordance $a_{i,j}^{(k)}$ (2) the bivariate measure $(a_{i,j}^{(k)}, d_{i,j}^{(k)})^T$. As a global measure, we take the empirical mean of FC. For the Pearson correlation case, the mean is $\bar{C}^{(k)} = 2/(N(N-1)) \sum_{i \in \{1,...,N-1\}} \sum_{j \in \{i,...,N\}} \rho_{i,j}^{(k)}$. For the proposed FC estimator, the summary statistic is either a bivariate corresponding to the empirical mean of accordance $\bar{a}^{(k)}$ and the empirical mean of discordance $\bar{d}^{(k)}$, or a trivariate measure by adding the empirical mean of significant activation time $\bar{\pi}_i^{(k)}$.
In all univariate cases, we used the Student t-test to derive significance p-values. While in the multivariate cases, we used the t^2 Hotelling test statistic [6]. Finally, in the local level, p-values are corrected for multiplicity using either the Bonferroni procedure which controls the family wise error rate (FWER), or the Benjamini and Hochberg procedure to control the false discovery rate (FDR) [7].

3 Results and discussion

Tab. 1: Number of significantly different connections (after Bonferroni or FDR correction) when using $(a_{i,j}; d_{i,j})^T$ as local connectivity measure. We also report the minimum (non-corrected) p-value. Different quantile thresholds are used. P-values are Bonferroni or FDR corrected.

Threshold	0.75	0.80	0.9	0.95
Bonferroni	2	2	0	0
FDR	22	6	0	0
min p-value	$3 \cdot 10^{-7}$	$3 \cdot 10^{-6}$	$2 \cdot 10^{-3}$	$4 \cdot 10^{-3}$

Using Pearson correlation as a measure of local connectivity ends up with no significant results after both Bonferroni and FDR correction. The minimum p-value in this case is $2.83 \cdot 10^{-5}$. The situation is the same

TAB. 2: Number of significantly different connections (after Bonferroni or FDR correction) when using the accordance $a_{i,j}$ as local connectivity measure. We also report the minimum (non-corrected) p-value. Different quantile thresholds are used.

Threshold	0.75	0.80	0.9	0.95
Bonferroni	1	2	0	0
FDR	48	34	0	0
min p-value	$3 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$1 \cdot 10^{-4}$	$2 \cdot 10^{-4}$

TAB. 3: P-values of global comparison using different summary statistics with different quantile threshold values.

Threshold	0.75	0.8	0.9	0.95
T^+	0.02	0.0059	0.0032	0.055
$\bar{a}$	0.0015	0.0025	0.047	0.44
$\bar{d}$	0.077	0.011	0.0084	0.0088
$\bar{a}/\bar{d}$	0.001	0.0058	0.041	0.048
$(\bar{a}; \bar{d})^T$	0.0070	0.02	0.017	0.0025
$(\bar{a}; \bar{d}, \bar{\pi})^T$	0.0074	0.00052	0.0012	0.0032

when using either the accordance $a_{i,j}$ or the bi-variate measure $(a_{i,j}; d_{i,j})^T$, with quantile threshold $q = 0.9$ or 0.95. However, when the quantile threshold is 0.75 or 0.80, sensitivity increases, especially when using only the accordance as a local measure (Table 1 and Table 2). At the global level, the comparison sensitivity is also influenced by the quantile threshold q. The most significant p-values correspond to quantile threshold $q = 0.75$ or 0.80. The bi-variate and the tri-variate summary statistics lead to high significance (Table 3) compared to the global correlation mean that gives a significance p-value of 0.0143. Small values of the quantile threshold give less sparse FC. Even with the smallest possible value, which is 0.5, the new measure still have a sensitivity advantage. Note that in this case, that is $q = 0.5$, if we sum the accordance value with the discordance value, we end up with a FC estimation that is close to the one obtained with Pearson correlation.

These results show that the new FC estimator is more sensitive than common FC estimators. This sensitivity is increased because the new estimator contains extra relevant information about FC that is lost with common estimators such as Pearson correlation or Spearman correaltion. Thus, the multivariate property together with the robustness of the new estimator might be the key ingredients that bring more statistical power to detect group differences. Then, we can say that it is more suitable in brain connectivity studies.

4 Conclusion

We proposed a new measure of FC derived from fMRI time courses, and presented a simple algorithm to construct an estimate of FC. The new estimator measures accordances and discordances of co-activations and co-deactivations of fMRI signals, separately, which makes our estimator exhaustive compared to common FC estimators. This might contribute to more complete interpretations in brain function studies, and to better differentiate between brain states. One could extend the new estimator to partial accordance and partial discordance as an analogy with correlation and partial correlation. We also presented a comparison study to detect group differences and we showed how the detection sensitivity is increased when using our new estimator. The sensitivity could be further improved by using adaptive unimodal and multimodal statistical methods, such as those proposed in [6, 8, 9].

References

[1] 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.

[2] 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.

[3] 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, p. 15, 2012.

[4] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, M. Schluep, P. Vuilleumier, and D. Van De Ville, "Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, pp. 937–950, 2013.

[5] 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.

[6] 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.

[7] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society. Series B. Methodological*, vol. 57, no. 1, pp. 289–300, 1995.

[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, 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, p. e23009, 08 2011.