Hemodynamic estimation based on Consensus Clustering

Solveig Badillo, Gaël Varoquaux, Philippe Ciuciu

Hemodynamic estimation based on Consensus Clustering

Solveig Badillo*, Gael Varoquaux* and Philippe Ciuciu*

* *Parietal team, INRIA Saclay Ile-de-France, NeuroSpin, Bât 145, 91191 Gif-sur-Yvette cedex, France*
Email: firstname.lastname@inria.fr

Abstract—**Modern cognitive experiments in functional Magnetic Resonance Imaging (fMRI) often aim at understanding the temporal dynamics of the brain response in regions activated by a given stimulus. The study of the variability of the hemodynamic response function (HRF) and its characteristics can provide some answers. In this context, we aim at improving the accuracy of the HRF estimation. To do so, we relied on a Joint-Detection-Estimation (JDE) framework that enables robust detection of brain activity as well as HRF estimation, in a Bayesian setting [2]. So far, the hemodynamic results provided by the JDE formalism have depended on a prior parcellation of the data performed before JDE inference. In this study, we propose a new approach to relax this prior knowledge: using consensus clustering techniques based on random parcellations of the data, we combine hemodynamics results provided by different parcellations, so as to robustify the HRF estimation.**

Keywords-**fMRI; Consensus Clustering; Random parcellation; Hemodynamic estimation; Bayesian inference**

I. INTRODUCTION

In fMRI studies, two main concerns arise at the subject-level analysis: i.) a precise localization of evoked brain activity elicited by sensorimotor or cognitive tasks, and ii.) a robust estimation of the underlying hemodynamic response associated with these activations. Since these two steps are inherently linked, the Joint Detection-Estimation (JDE) approach has been proposed to address both issues in a coordinated formalism [1]. This formalism performs a multivariate inference for both detection and estimation through a regional bilinear generative model of the BOLD response, which embeds spatio-temporal regularization within the Bayesian framework. The efficiency and usefulness of this approach has been validated at the group level in [2]. In particular, the estimation of the HRF is more adequately performed using the JDE formalism. As the HRF is potentially linked to the magnitude, latency and duration of neural activity, it is of primary interest to accurately estimate its shape to make inference about neural processes.

However, so far the JDE formalism has relied on a prior parcellation of fMRI data into functionally homogeneous parcels. The functional mask of each subject's brain is *a priori* divided in K functionally homogeneous *parcels* using a parcellation technique proposed in [3]. Usually, K is chosen so that the parcel size is equal to $2.3cm^3$, as done in [3]. The hemodynamics results thus depend on this prior parcellation to a large extent, and several differences in

hemodynamic estimation were observed depending on the prior parcellation [4]. Moreover, a potential bias can be introduced using the classical GLM that does not enable important fluctuations of the HRF shape throughout the brain. Previous works have extended the JDE formalism to a joint parcellation-detection-estimation (JPDE) approach [5] which targets to optimally infer the parcellation from the JDE results iteratively by imposing some spatial constraints on the definition of hemodynamic territoires. However, this formalism is really computationally demanding and not fully validated on real data yet. Therefore, we propose here another approach to circumvent this problem by resorting to random parcellation techniques of decimated fMRI data.

The rest of the paper is organized as follows. Section II summarizes the standard JDE framework. The consensus clustering technique is presented in Section III and then applied to artificial and real data sets in Section IV and V. Finally, conclusions are drawn in Section VI.

II. THE JOINT-DETECTION ESTIMATION FRAMEWORK

The JDE model relies on a prior parcellation of fMRI data into $(\mathcal{P}_\gamma)_{\gamma=1:\Gamma}$ parcels. In voxel $j \in \mathcal{P}_\gamma$, the fMRI time series $\boldsymbol{y}_j$ is measured at times $(t_n)_{n=1:N}$ where $t_n = n\text{TR}$, N being the number of scans and TR the time of repetition. In [1], the BOLD signal is modelled in a given parcel $\mathcal{P}_\gamma$ as follows:

$$\forall j \in \mathcal{P}_\gamma, \quad \boldsymbol{y}_j = \sum_{m=1}^{M} a_j^m \boldsymbol{X}^m \boldsymbol{h}_\gamma + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j. \quad (1)$$

The unknown HRF shape $\boldsymbol{h}_\gamma$ is constant within each parcel $\mathcal{P}_\gamma$. The Neural response levels a_j^m, which model the magnitude of activation, may vary in space and across stimulus types or experimental conditions. Here, each Neural response level is assumed to be in one of the two groups specified by activation class assignment variables $\boldsymbol{Q} = \{\boldsymbol{q}^m, m = 1 : M\}$ where $\boldsymbol{q}^m = \{q_j^m, j \in \mathcal{P}\}$, q_j^m representing the *activation class* at voxel j for condition m ($q_j^m = 2$ for activated voxels and $q_j^m = 1$ for non-activated voxels). $\boldsymbol{X}^m$ denotes the $N \times (D+1)$ binary matrix $\boldsymbol{X}_m = \{x_m^{n-d\Delta t}, n = 1 \ldots N, d = 0 \ldots D\}$ that provides information on the stimulus occurrences for the mth experimental condition. $\Delta t < TR$ is the sampling period of the unknown HRF in the parcel γ: $\boldsymbol{h}_\gamma \in \mathbb{R}^{D+1}$. Finally, matrix $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_O]$ comprises the values at times t_n of an orthonormal basis consisting of functions that

take a potential drift and any other nuisance effect into account. Vector $\boldsymbol{\ell}_j = (l_{o,j})^{\text{t}}_{1\leqslant o\leqslant O}$ contains the corresponding unknown regression coefficients in V_j. Vector $\boldsymbol{b}_j = (b_{j,t_n})^{\text{t}}$ defines the noise term in voxel j and is supposed to be white, normally-distributed with variance $v_{\boldsymbol{b}_j}$ and independent of the HRFs. The parameters are estimated in a Bayesian framework as described in [1], using a stochastic inference scheme.

III. JDE RESULTS BASED ON CONSENSUS CLUSTERING
A. The different steps involved in the procedure

The consensus clustering techniques consist in running multiple times a clustering algorithm on different perturbations of the original data and combining the resulting clusters so as to assess their stability [6], [7]. Moreover, to avoid any arbitrary choice of the number of clusters, the procedure can be repeated for different values of this number. In our study, the perturbations were generated by randomly undersampling the data along the temporal axis. This undersampling is done without replacement to avoid a too significant change of the data structure. Following [7], 75% of the temporal points were kept in each voxel.

Since the fMRI data are usually very noisy, we have performed the clustering after extracting temporal features of interest. Moreover, as the objective is to recover the hemodynamic territories at best, the clustering algorithm was based on voxel-wise HRFs inferred by the Regularized Finite Impulse Response approach (RFIR). Akin to the JDE model, the latter does not assume any functional form for the HRF and amounts to estimating a large number of parameters to identify its prototypical properties. The different steps of our procedure are summarized in the Algorithm 1. In short, the hemodynamic results from different parcellations are weighted by their posterior probability and by information from consensus clustering.

B. The Regularized FIR modeling

For the sake of completeness, let us briefly summarize how the voxelwise RFIR approach proceeds. The generative BOLD signal model reads as follows:

$$\forall j \in [\![1; J]\!], \qquad \boldsymbol{y}_j = \sum_{m=1}^{M} \boldsymbol{X}^m \boldsymbol{h}_j^m + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j \quad (2)$$

Vector $\boldsymbol{h}_j^m = \left(h_{j,d\Delta t}^m\right)^{\text{t}}_{d=0,..,D}$ represents the unknown HRF time course in voxel j which is associated with the mth experimental condition and sampled every Δt. The other terms are the same as in Eq. (1), except the Neural Response Levels, here are embodied in the HRF. The HRF estimation is *regularized* with a penalization that prevents from sharp fluctuations: $\boldsymbol{h}_\gamma \sim \mathcal{N}(\boldsymbol{0}, v_{\boldsymbol{h}}\boldsymbol{R})$ with $\boldsymbol{R} = (\boldsymbol{D}_2^{\text{t}}\boldsymbol{D}_2)^{-1}$.

IV. RESULTS ON SIMULATION

We validate the proposed methodology on artificial datain order to assess the gain in robustness we achieved using

---

Algorithm 1 JDE-based on Consensus Clustering procedure

1: **Require: Input:** datasets $\boldsymbol{Y} \in \mathbb{R}^{n\times p}$, list of number of clusters to test $\{K = K_{min}, ..., K = K_{max}\}$, number of resampling S, undersampling fraction $\pi \in [0, 1]$. J voxels are contained in the mask.
2: **for** $K = K_{min}$ to K_{max} **do**
3:     **for** $s = 1$ to S **do**
4:         Down-sample the data along temporal axis: $\boldsymbol{Y}^{K,s} \leftarrow \boldsymbol{Y}$ where $\boldsymbol{Y}^{K,s} \in \mathbb{R}^{n\times\pi p}$.
5:         Extract the HRF in each voxel with RFIR model: $\boldsymbol{h}_j^{K,s^{\text{RFIR}}}$.
6:         Cluster the features: cluster $\boldsymbol{h}_j^{K,s^{\text{RFIR}}}$ into K clusters.
7:         Compute connectivity matrix $M^{K,s}$: $M^{K,s}(i;j) = 1$ if voxels i and j belong to the same cluster, 0 otherwise.
8:     **end for**
9:     Perform consensus clustering: compute consensus matrix $\mathcal{M}^K$ from $\boldsymbol{M} = \{M^{K,1}, ..., M^{K,S}\}$: $\mathcal{M}^K(i;j) = \sum_s M^{K,s}(i;j)$, indicating the number of times i and j are assigned to the same cluster.
10:     Cluster consensus matrix $\mathcal{M}^K$ to obtain parcellation $\mathcal{P}_{cons}^K$.
11:     For each parcellation $\mathcal{P}_{cons}^K$ and in each cluster $\mathcal{C}_k^K$ ($\mathcal{C}_k^K \in \mathcal{P}_{cons}^K$, $k \in [\![K_{min}, K_{max}]\!]$), compute the average index between all pairs of voxels: $cc^K = \{cc^K(k=0), ..., cc^K(k=K)\}$ [6].
12:     Fit the JDE model on the parcellation P_{cons}^K: one HRF estimate is computed in each cluster $\mathcal{C}_k^K$. For all the parcellation P_{cons}^K, the hemodynamics results are $\{\boldsymbol{h}_{k=0}^K, ..., \boldsymbol{h}_{k=K}^K\}$.
13:     Compute posterior probability associated with each HRF estimate: pp_k^K.
14:     Compute weight $\omega_k^K = cc_k^K \times pp_k^K$ associated with each HRF h_k^K in cluster $\mathcal{C}_k^K$.
15:     Recompute weights and HRFs in each voxel: $\forall j \in \mathcal{C}_k^K, \omega_j^K = \omega_k^K$ and $\boldsymbol{h}_j^K = h_k^K$.
16: **end for**
17: Compute final result: $\forall j \in [\![1, J]\!], \boldsymbol{h}_j = \sum_{K=K_{min}}^{K=K_{max}} .\boldsymbol{h}_j^K * \omega_j^K$
18: **return** $\{\boldsymbol{h}_j\}_{j=1}^{j=J}$

---

the CC-JDE approach instead of running the JDE model on GLM-based parcellations. A 4D-BOLD artificial signal of $N = 135$ time points was simulated according to the observation model in Eq. (1). 8 hemodynamic territories were simulated, with a different HRF shape simulated for each one. $\boldsymbol{P}$ was a polynomial basis of order 4 and $\boldsymbol{X}^{m,s}$ encodes a fast event-related paradigm comprising ten conditions (ISI=3.5 s.). Neural response levels were drawn according to: $(a_j^m|q_j^m = 2) \sim \mathcal{N}(2.5, 0.3)$ (activated voxels) and $(a_j^{m,s}|q_j^m = 1) \sim \mathcal{N}(0, 0.3)$ (unactivated voxels). Drift coefficients were drawn as $\boldsymbol{\ell}_j \sim \mathcal{N}(0, 3.2\boldsymbol{I}_4)$. Noise realizations follow $\boldsymbol{b}_j \sim \mathcal{N}(0, 1.1\boldsymbol{I}_N)$, so as to simulate more realistic data with a low SNR. Finally, activation states were set by a hand-drawn map, as shown in Fig. 1(a)-(c).

The JDE model was applied to the synthetic data either using the consensus clustering procedure or relying on parcellations obtained by clustering of GLM results (in this second case, one JDE model is being run for each number of clusters chosen for the parcellation: $K \in [\![5; 11]\!]$)

The GLM-based parcellations shown Fig. 1(c)-(e) were compared to the ground truth hemodynamic territories shown in Fig. 1(a). We used the Normalized Information Criterion (NMI) to measure the agreement of two parcellations, ignoring permutations ($NMI = 1$ if identical parcellations). We also clustered the hemodynamic results
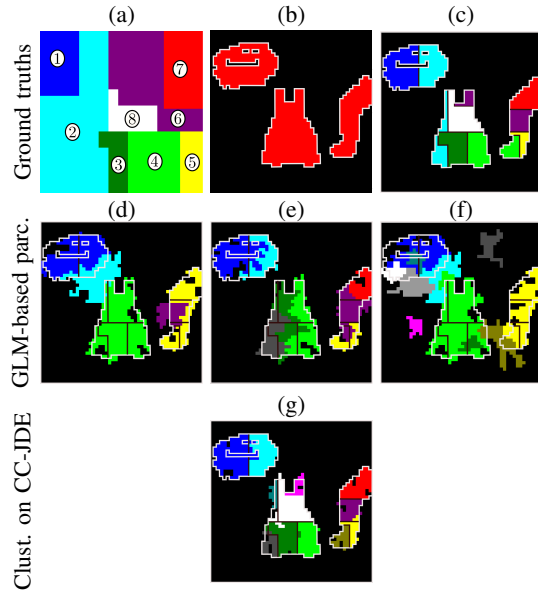
Figure 1. **Top row:** Ground truth with simulated hemodynamic territories (a), simulated activation labels (b) and superimposition of activation patterns with hemodynamic territories. **Middle row:** Parcellations obtained by clustering the results obtained from the GLM fit of the data, for $K = 6$ (a), $K = 8$ (b), $K = 10$ (c). **Bottom row:** Clustering of the hemodynamic results obtained by JDE modeling. The CC-JDE clustering technique described in Algo. 1 was applied to the synthetic data.

obtained using the CC-JDE procedure, and compare this clustering, denoted CC-JDE, with the ground truth.
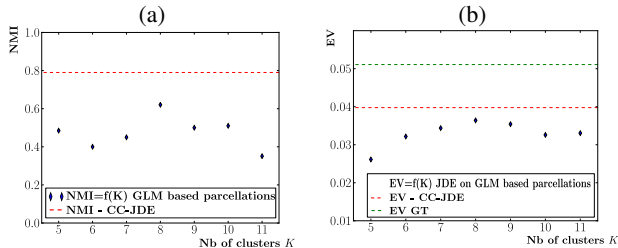

Figure 2. **(a):** Comparisons of the GLM-based parcellations, depending on the total number of clusters K, with the Ground Truth ($NMI = f(K)$ in blue points) and evaluation of the clustering of CC-JDE results (red dashed line). **(b):** Percentages of variance explained by the JDE model, using the GLM-based parcellations (blue points) or after the CC-JDE procedure (red dashed line). The ground truth (GT) is shown in green dashed line.

As shown in Fig. 2(a), the NMI criterion of GLM-based parcellations varies around 0.5 depending on the total number of clusters we chose. Thus, the GLM-based parcellations seem to fail to detect the intrinsic structure of the response dynamics. In contrast, the NMI value associated with our CC-JDE clustering is very high i.e., close to 0.8. This corresponds to the visual observation of Fig. 1 (middle and top rows): the hemodynamic territories are far better recovered using the CC-JDE clustering. Moreover, the estimation relying on the GLM-based parcellations is less accurate than that performed by the CC-JDE procedure. Indeed, as can be seen on Fig. 2(b), the percentage of variance explained by the model (EV) is larger using this procedure ($\%EV \sim 0.4$ closer to the ground truth

$\%EV \sim 0.5$). For each model, the given percentages of explained variance were averaged over all voxels.

V. RESULTS ON REAL DATA

Real fMRI data were recorded on an adult. The experiment was designed to map auditive and visual brain functions and consisted of a session of $N = 128$ scans lasting $TR = 2.4s$, each yielding a 3-D volume composed of $64 \times 64 \times 32$ voxels. The paradigm was fast event-related comprising eighty auditory and visual stimuli. A region of interest (ROI) in the left temporal lobe was defined along the Superior Temporal Sulcus (STS). For the clustering steps based upon RFIR results, only HRF corresponding to the auditory conditions were used.

As done on the synthetic data, the JDE model was fitted either using GLM-based parcellations or the CC-JDE methodology. The GLM-based activation maps were clusterized using a total number of clusters K varying between 20 and 40. Thus, 21 GLM-based parcellations were obtained, before being used as the prior parcellation in JDE analysis.
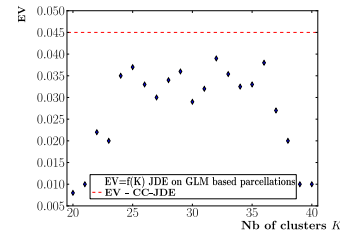

Figure 3. Percentages of variance explained by the JDE model, using the GLM-based parcellations (blue points) or after the CC-JDE procedure (red dashed line). $\%EV$ are averaged over all voxels.

We first compare the accuracy of the estimation using the percentages of explained variance. As shown in Fig. 3, the mean percentage of explained variance obtained after using GLM-based parcellation (in blue) is always lower than that obtained after the CC-JDE procedure ($\%EV = 4.5\%$, in red), irrespective of K. This result shows that the CC-JDE procedure outperforms the previous way of clustering the brain before running JDE analysis.

To evaluate the potential neuroscientific impact, we observed the values of time-to-peak (TTP) of HRF estimates in the region defined along the STS. Fig. 4 shows the values of TTP obtained in this region, either using the GLM-based parcellation comprising 30 voxels (see Fig. 4(a)) or using the CC-JDE procedure (see Fig. 4(b)). On Fig. 4(b), we retrieved a gradient of activations along the superior temporal regions which was previously described in adults [3] and children [1]. The fastest responses are observed near Heschl's gyrus and the middle STS (time-to-peak of the response around 6s.) and slow down along the caudal-rostral axes. In the posterior and anterior part of the ROI, the TTP are much slower since within the range from 8 to 11s. Importantly, the TTPs obtained using the CC-JDE procedure vary more smoothly along the STS as compared to those

retrieved by GLM-based parcellations. Also, the latter did not enable to clearly identify the TTP gradient along the STS. Hence, we observed that the prior parcellation of the data impacts the hemodynamic results.
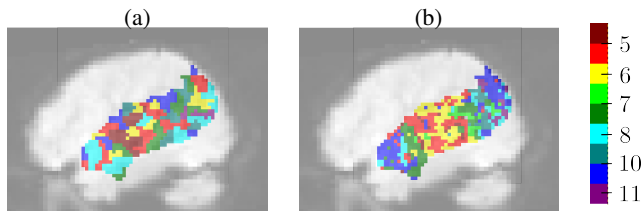

Figure 4. Time-To-Peak values (in s.) estimated along the STS from real data by JDE analysis performed either on a GLM based parcellation comprising $K = 30$ clusters (a) or on the output of the CC-JDE procedure (b).

It has to be noted that the temporal undersampling performed on the data may result in non-smooth HRF estimates. Indeed, the following situation might occur: if all time points removed from the time series correspond exactly to the 5th second after the presentation of the trials of a given condition, the corresponding HRF will be impacted by the presence of a discontinuity around this time point. However, three ingredients may compensate for this potential problem. First, in the current experiment, 128 scans are recorded during a fast event-related paradigm, each stimulus type being randomly repeated 10 times. This situation is thus unlikely. Second, as many parcellations based upon different undersamplings are performed, the poor estimation performed for one particular undersampling could be compensated by the other estimates. Finally, the clustering procedure is usually based upon HRF estimates computed over several conditions, which limits the impact of the above mentioned situation.

VI. Conclusion

In this paper, a consensus clustering-based procedure was introduced to robustify the hemodynamic estimation performed in the context of the JDE formalism. Instead of relying on a prior parcellation of the data, different clusterings are performed after perturbating the data and extracting the corresponding hemodynamic features. Results from artificial and real data showed that the new procedure is better adapted to recover the hemodynamics feature of the BOLD signal. To improve the hemodynamics estimation, we plan to further extend our procedure using Weighted Ensemble Clustering techniques [8]. These methods jointly use different representations for temporal data clustering algorithms. The resulting clusters are weighted and combined to form a final clustering. Moreover, we also plan to compare our approach with the other alternative to the prior parcellation, the Joint parcellation-detection-estimation framework [5].

Besides, to further justify the practical value of the proposed method, further validation on real fMRI dataset are currently performed. Thus, we plan to benefit from the advantages of the JDE framework in cognitive studies focusing on cerebral specialization such as linguistic processing in infants. In the latter context, the CC-JDE model allows an extensive and more robust analysis of the hemodynamic variability along the STS.

Finally, the JDE framework is computationally demanding when infered using a stochastic inference scheme: this makes the computational cost of the CC-JDE procedure very high. This can be partially compensated by alternative variational inference schemes [9] and also by the fact that the CC-JDE approach is especially devoted to study specific areas of the brain and not the whole brain in an exploratory manner: this limits the required time to carry out the algorithm.

References

[1] T. Vincent, L. Risser, and P. Ciuciu, "Spatially adaptive mixture modeling for analysis of within-subject fMRI time series," *IEEE Trans. Med. Imag.*, vol. 29, no. 4, Apr. 2010.

[2] S. Badillo, T. Vincent, and P. Ciuciu, "Impact of the joint detection-estimation approach on random effects group studies in fMRI," in *8th Proc. IEEE ISBI*, Chicago, IL, Apr. 2011.

[3] B. Thyreau, B. Thirion, G. Flandin, and J.-B. Poline, "Anatomo-functional description of the brain: a probabilistic approach," in *Proc. 31th Proc. IEEE ICASSP*, Toulouse, France, May 2006.

[4] T. Vincent, P. Ciuciu, and B. Thirion, "Sensitivity analysis of parcellation in the joint detection-estimation of brain activity in fMRI," in *5th Proc. IEEE ISBI*, Paris, France, May 2008.

[5] L. Chaari, F. Forbes, T. Vincent, and P. Ciuciu, "Adaptive hemodynamic-informed parcellation of fMRI data in a variational joint detection estimation framework," in *15th Proc. MICCAI*, Nice, France, Oct. 2012.

[6] S. Monti, P. Tamayo, J. Mesirov, and T. Golub, "Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data," *Machine learning*, 2003.

[7] G. Varoquaux, A. Gramfort, and B. Thirion, "Small-sample brain mapping : sparse recovery on spatially correlated designs with randomization and clustering," in *Proc. 29th ICML*, Edinburgh, Scotland, June 2012.

[8] C. Domeniconi and MunaG. Al-Razgan, "Weighted cluster ensembles," *ACM Transactions on Knowledge Discovery from Data*, vol. 2, no. 4, 2009.

[9] L. Chaari, T. Vincent, F. Forbes, M. Dojat, and P. Ciuciu, "Fast joint detection-estimation of evoked brain activity in event-related fMRI using a variational approach," *IEEE Trans. Med. Imag.*, vol. 32, no. 5, May 2013.