Robust voxel-wise Joint Detection Estimation of Brain Activity in fMRI

Lotfi Chaari, Florence Forbes, Thomas Vincent, Philippe Ciuciu

ROBUST VOXEL-WISE JOINT DETECTION ESTIMATION OF BRAIN ACTIVITY IN fMRI

Lotfi Chaari[1,2], *Florence Forbes*[1], *Thomas Vincent*[1], *Philippe Ciuciu*[2]

[1] INRIA Rhône-Alpes, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, France (firstname.lastname@inria.fr)
[2] CEA/DSV/I^{2}BM/Neurospin, CEA, Saclay, Bbt. 145, Point Courrier 156, 91191 Gif-sur-Yvette cedex, France (firstname.lastname@cea.fr)

ABSTRACT

We address the issue of jointly detecting brain activity and estimating brain hemodynamics from functional MRI data. To this end, we adopt the so-called Joint-Detection-Estimation (JDE) framework introduced in [1] and augmented in [2]. An inherent difficulty is to find the right spatial scale at which brain hemodynamics estimation makes sense. The voxel level is clearly not appropriate as estimating a full hemodynamic response function (HRF) from a single voxel time course may suffer from a poor signal-to-noise-ratio and lead to potentially misleading results (non-physiological HRF shapes). More robust estimation can be obtained by considering groups of voxels (*i.e.* parcels) with some functional homogeneity properties. Current JDE approaches are therefore based on an initial parcellation but with no guarantee of its optimality or goodness. In this work, we propose a joint parcellation-detection-estimation (JPDE) procedure that incorporates an additional parcel estimation step solving this way both the parcellation choice and robust HRF estimation issues. As in [3], inference is carried out in a Bayesian setting using variational approximation techniques for computational efficiency.

Index Terms— Variational EM, MRF, Biomedical signal detection, Magnetic resonance imaging.

1. INTRODUCTION

Functional Magnetic Resonance Imaging (fMRI) is a powerful tool to non-invasively study the relation between cognitive task and cerebral activity through the analysis of the hemodynamic BOLD signal [4]. Within-subject analysis in event-related fMRI first relies on (i) a detection step to localize which parts of the brain are activated by a given stimulus type, and second on (ii) an estimation step to recover the temporal dynamics of the brain response. Most approaches to detect neural activity rely on a single *a priori* model for the temporal dynamics of activated voxels also known as the hemodynamic response function (HRF) [5]. A canonical HRF is usually assumed for the whole brain although there has been evidence that this response can vary in space or region, across subjects and groups [6]. In addition, a robust and accurate estimation of the HRF is possible only in regions that elicit an evoked response to an experimental stimulus [7]. Both issues of properly detecting evoked activity and estimating the HRF then play a central role in fMRI data analysis. They are usually dealt with independently with no possible feedback although both issues are strongly connected one to another. To introduce more flexibility regarding the assumptions on the HRF model, a novel approach referred to as the Joint Detection Estimation (JDE) framework has been introduced in [1] and extended in [2] to account for spatial correlation between neighboring voxels in the brain volume (regular lattice in 3D). In this latter approach, the HRF can be estimated while simultaneously detecting activity, in a region-based analysis, that is on a set of pre-specified regions also named *parcels* partitioning the whole brain data set. This approach is mainly based on: (i) the non-parametric modelling of the HRF at a regional spatial scale (parcel-level) that provides a fair compromise between homogeneity of the BOLD signal (reproducibility of the HRF) and the increase of the signal-to-noise ratio by spatially aggregating signals; (ii) prior information about the temporal smoothness of the HRF to be estimated; and (iii) the modelling of spatial correlation between neighboring voxels within each parcel using condition-specific hidden Markov fields. In [1,2], posterior inference is carried out in a Bayesian setting using Monte Carlo Markov Chain (MCMC) methods. In order to overcome the high computational cost of this MCMC-based approach, an alternative with similar performance has been proposed in [3] based on Variational Expectation Maximization (VEM) algorithm.

However, one current limitation of all these JDE approaches lies in the prior decomposition of the brain into functionally homogeneous regions (parcels) required so as to estimate parcel-based HRFs. Those parcels should be small enough to guarantee the invariance of the HRF within each parcel but large enough to contain reliable information for the estimation. This spatial scale trade-off is currently decided on an appropriate external a priori parcellation. Ideally, the parcellation should instead be reconsidered while estimating the HRF directly from the BOLD signal. We propose then to further constrain the HRF estimation by incorporating for parcel estimation an additional layer in the model and to carry out inference in a variational setting. In this novel Joint-Parcellation-Detection-Estimation (JPDE) model, a parcel is a set of connected voxels with different individual HRFs but that can be seen as local perturbations of the same parcel-specific HRF pattern. The hemodynamics estimation reduces then to the identification of a limited number (say K) of such HRF patterns. We relax the assumption of a unique HRF model by parcel by considering voxel-varying HRFs but to reduce the number of parameters to be estimated, these HRFs are modelled as realizations of a K-component mixture model. As a by product, a parcellation estimation can be derived by considering that each set of connected voxels in the same HRF group defines a parcel. The group variables will be governed by a hidden Markov Model to enforce spatial correlation, *i.e.* favor group assignments to vary smoothly. Finally, the overall scheme will iteratively and alternately identify parcels and HRF groups. The proposed approach makes the JDE framework fully automatic and more flexible. It is based on a VEM algorithm to derive estimates of the HRF, the Neural Response Levels (NRLs), the corresponding labels (activated/non-activated voxels) and the HRF group labels.

The rest of this paper is organized as follows. In Section 2, the JDE framework is presented and extended to JPDE with a whole-brain model. The new model inference is detailed in Section 3. Finally, results on realistic artificial fMRI data are reported in Section 4, and some conclusions are drawn in Section 5.

2. A JOINT DETECTION-ESTIMATION MODEL

Matrices (resp. vectors) are denoted with bold upper (resp. lower) case letters such as $\boldsymbol{P}$ (resp. $\boldsymbol{h}$). A vector is by convention a column vector. The transpose is denoted by $^{\mathrm{t}}$. Unless stated otherwise, subscripts j, m, k, i and n are respectively indexes over voxels, stimulus

types, HRF groups, activation classes and time point. The Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ is denoted using $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

2.1. Observation BOLD model

Let us partition the set of voxels ($\boldsymbol{\mathcal{P}}$) in the brain into K groups, each of them having homogeneous functional properties, *i.e* a unique HRF model: $\boldsymbol{\mathcal{P}} = (\mathcal{P}_k)_{k=1:K}$. At a voxel j, the fMRI time series $\boldsymbol{y}_j$ is measured at times $(n\,TR)_{n=1:N}$, N being the number of scans and TR the time of repetition. The number of different stimulus types or experimental conditions is M. For a given voxel, a BOLD signal model (the same for all voxels) is used in order to link the observed data $\boldsymbol{Y} = \{\boldsymbol{y}_j \in \mathbb{R}^N, j \in \boldsymbol{\mathcal{P}}\}$ to the voxel-dependent HRFs $\boldsymbol{h}_j \in \mathbb{R}^{D+1}$ and to the response magnitudes $\boldsymbol{A} = \{\boldsymbol{a}^m, m = 1 : M\}$ with $\boldsymbol{a}^m = \{a_j^m, j \in \boldsymbol{\mathcal{P}}\}$ and a_j^m being the magnitude at voxel j for condition m. More specifically, the observation model at each voxel $j \in \boldsymbol{\mathcal{P}}$ is expressed as follows:

$$\boldsymbol{y}_j = \boldsymbol{S}_j\boldsymbol{h}_j + \boldsymbol{P}\ell_j + \boldsymbol{\varepsilon}_j, \quad \text{with} \quad \boldsymbol{S}_j = \sum_{m=1}^{M} a_j^m \boldsymbol{X}_m \qquad (1)$$

where $\boldsymbol{S}_j\boldsymbol{h}_j$ is the summation of the stimulus-induced components of the BOLD signal. The binary matrix $\boldsymbol{X}_m = \{x_m^{n-d\Delta t}, n = 1 : N, d = 0 : D\}$ is of size $N \times (D+1)$ and provides information on the stimuli occurrences for the m-th experimental condition, $\Delta t < TR$ being the sampling period of the unknown HRFs $\boldsymbol{h}_j = (h_{d\Delta t})_{d=0:D}$. We denote by $\boldsymbol{H} = \{\boldsymbol{h}_j, j \in \boldsymbol{\mathcal{P}}\}$ the set of all HRFs. The scalar a_j^m's are weights that model the transition between stimulations and the neurovascular response. This response is a consequence of the neuronal excitation which is commonly assumed to occur with stimulations. It follows that the a_j^m's are generally referred to as Neural Response Levels (NRL). The rest of the signal is made of matrix $\boldsymbol{P}$, which corresponds to physiological artifacts accounted for via a low frequency orthonormal function basis of size $N \times O$. With each voxel j is associated a vector of low frequency drifts $\ell_j \in \mathbb{R}^O$ which has to be estimated. These vectors may be grouped into the same matrix $\boldsymbol{L} = \{\ell_j, j \in \boldsymbol{\mathcal{P}}\}$. Regarding the observation noise, the $\boldsymbol{\varepsilon}_j$'s are assumed to be independent with $\boldsymbol{\varepsilon}_j \sim \mathcal{N}(0, \boldsymbol{\Gamma}_j^{-1})$ at voxel j (see Section 2.2.1 for more details). The set of all unknown precision matrices is denoted by $\boldsymbol{\Gamma} = \{\boldsymbol{\Gamma}_j, j \in \boldsymbol{\mathcal{P}}\}$.

Detection is handled through the introduction of activation class assignments $\boldsymbol{Q} = \{\boldsymbol{q}^m, m = 1 : M\}$ where $\boldsymbol{q}^m = \{q_j^m, j \in \boldsymbol{\mathcal{P}}\}$ and q_j^m represents the *activation class* at voxel j for experimental condition m. Without loss of generality, the number of classes considered here is $I = 2$ for activated ($i = 2$) and non-activated ($i = 1$) voxels. Finally, joint parcellation is performed by introducing another set of hidden variables $\boldsymbol{Z} = \{z_j, j \in \boldsymbol{\mathcal{P}}\}$ where $z_j \in \{1 : K\}$ denotes the group or HRF class at voxel j ($z_j = k$ means that voxel j belongs to the k-th group).

2.2. Hierarchical model of the complete data distribution

With standard additional assumptions [1–3], and omitting the dependence on the parameters to be specified later, the distribution of both the observed and hidden variables can be decomposed as $p(\boldsymbol{Y}, \boldsymbol{A}, \boldsymbol{H}, \boldsymbol{Z}, \boldsymbol{Q}) = p(\boldsymbol{Y} \mid \boldsymbol{A}, \boldsymbol{H})\, p(\boldsymbol{A} \mid \boldsymbol{Q})\, p(\boldsymbol{H} \mid \boldsymbol{Z})\, p(\boldsymbol{Q})\, p(\boldsymbol{Z})$.

2.2.1. Likelihood

In [8], an autoregressive (AR) noise model has been adopted to account for serial correlations in fMRI time series. It has also been shown in [8] that a spatially-varying AR noise model helped to control false positive rate. In the same context, we will therefore assume $\boldsymbol{\varepsilon}_j \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Gamma}_j^{-1})$ with $\boldsymbol{\Gamma}_j = \sigma_j^{-2}\boldsymbol{\Lambda}_j$ where $\boldsymbol{\Lambda}_j$ is a tridiagonal symmetric matrix which depends on the AR(1) parameter ρ_j [1]: $(\boldsymbol{\Lambda}_j)_{1,1} = (\boldsymbol{\Lambda}_j)_{N,N} = 1$, $(\boldsymbol{\Lambda}_j)_{n,n} = 1+\rho_j^2$ for $n = 2 : N-1$ and $(\boldsymbol{\Lambda}_j)_{n+1,n} = (\boldsymbol{\Lambda}_j)_{n,n+1} = -\rho_j$ for $n = 1 : N-1$. The likelihood can therefore be written as:

$$p(\boldsymbol{Y} \mid \boldsymbol{A}, \boldsymbol{H}; \boldsymbol{L}, \boldsymbol{\theta}_0) \propto \prod_{j \in \boldsymbol{\mathcal{P}}} |\boldsymbol{\Gamma}_j|^{1/2} \exp\big(-\frac{1}{2}\overline{\boldsymbol{y}}_j^{\mathrm{t}}\boldsymbol{\Gamma}_j\overline{\boldsymbol{y}}_j\big), \qquad (2)$$

where $|\boldsymbol{\Gamma}_j| = \sigma_j^{-2N}\,|\boldsymbol{\Lambda}_j|$ and $|\boldsymbol{\Lambda}_j| = 1-\rho_j^2$, $\boldsymbol{\theta}_0 = \{\rho_j, \sigma_j^2, j \in \boldsymbol{\mathcal{P}}\}$ and $\overline{\boldsymbol{y}}_j = \boldsymbol{y}_j - \boldsymbol{P}\ell_j - \boldsymbol{S}_j\boldsymbol{h}_j$.

2.2.2. Model priors

Neuronal response levels. Akin to [1–3], the NRLs are assumed to be statistically independent across conditions: $p(\boldsymbol{A}; \boldsymbol{\theta}_{\boldsymbol{a}}) = \prod_m p(\boldsymbol{a}^m; \boldsymbol{\theta}_m)$ where $\boldsymbol{\theta}_{\boldsymbol{a}} = \{\boldsymbol{\theta}_m, m = 1 : M\}$ and $\boldsymbol{\theta}_m$ gathers the parameters for the m-th condition. A mixture model is then adopted by using the allocation variables q_j^m to segregate activated voxels ($q_j^m = 2$) from non-activated ones ($q_j^m = 1$). For the m-th condition, and conditionally to the assignment variables $\boldsymbol{q}^m$, the NRLs are assumed to be independent: $p(\boldsymbol{a}^m \mid \boldsymbol{q}^m; \boldsymbol{\theta}_m) = \prod_{j \in \boldsymbol{\mathcal{P}}} p(a_j^m \mid q_j^m; \boldsymbol{\theta}_m)$ with $p(a_j^m \mid q_j^m = i; \boldsymbol{\theta}_m) \sim \mathcal{N}(\mu_{mi}, v_{mi})$ and $\boldsymbol{\theta}_m = \{\mu_{mi}, v_{mi}, i = 1, 2\}$. We also write $\boldsymbol{\mu} = \{\boldsymbol{\mu}_m, m = 1 : M\}$ with $\boldsymbol{\mu}_m = \{\mu_{m1}, \mu_{m2}\}$ and $\boldsymbol{v} = \{\boldsymbol{v}_m, m = 1 : M\}$ with $\boldsymbol{v}_m = \{v_{m1}, v_{m2}\}$. For non-activated voxels we set for all m, $\mu_{m1} = 0$. The other parameters have to be estimated.

Activation classes. As in [2, 3], we assume prior independence between the M experimental conditions regarding the activation class assignments. It follows that $p(\boldsymbol{Q}) = \prod_{m=1}^{M} p(\boldsymbol{q}^m; \beta_m)$ where we assumed in addition that $p(\boldsymbol{q}^m; \beta_m)$ is a spatial Markov prior, namely an Ising model with interaction parameter β_m [2]:

$$p(\boldsymbol{q}^m; \beta_m) \propto \exp\big(\beta_m U(\boldsymbol{q}^m)\big), \qquad (3)$$

where $U(\boldsymbol{q}^m) = \sum_{j \sim j'} \delta(q_j^m, q_{j'}^m)$ and $\forall (a, b) \in \mathbb{R}^2$, $\delta(a, b) = 1$ if $a = b$ and 0 otherwise. The notation $j \sim j'$ means that the summation is over all neighboring voxels. The neighboring system covers a 3D scheme through the brain. The unknown parameters are denoted by $\boldsymbol{\beta} = \{\beta_m, m = 1 : M\}$. In what follows, we will consider a 6-connexity 3D neighboring system.

HRF groups. In order to promote parcellation regularity, we use here a spatial Markov prior, namely an Ising model with interaction parameter β_z:

$$p(\boldsymbol{Z}; \beta_z) \propto \exp\big(\beta_z U(\boldsymbol{Z})\big), \qquad (4)$$

where $U(\boldsymbol{Z}) = \sum_{j \sim j'} \delta(z_j, z_{j'})$. We use here the same 6-connexity 3D neighboring system as before.

HRF. In contrast to [1–3] where a unique HRF is used for a whole parcel and where a smoothness constraint is imposed for the HRF by controlling its second order derivative, we define here $p(\boldsymbol{H}|\boldsymbol{Z}) = \prod_{j \in \boldsymbol{\mathcal{P}}} p(\boldsymbol{h}_j \mid z_j)$ where $p(\boldsymbol{h}_j \mid z_j = k) \sim \mathcal{N}(\bar{\boldsymbol{h}}_k, \bar{\boldsymbol{\Sigma}}_k)$. The distribution on $\boldsymbol{h}_j$ is expressed conditionally to the HRF group variable z_j. Regularity across neighbouring voxels is then favored via the Markov prior on $\boldsymbol{Z}$.

For the complete model, the whole set of parameters is denoted by $\boldsymbol{\Theta} = \big\{\boldsymbol{\Gamma}, \boldsymbol{L}, \boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{\beta}, \beta_z, (\bar{\boldsymbol{h}}_k, \bar{\boldsymbol{\Sigma}}_k)_{1 \le k \le K}\big\}$ and belong to a set $\underline{\boldsymbol{\Theta}}$.

3. ESTIMATION BY VARIATIONAL EM

We propose to use an Expectation-Maximization (EM) framework to deal with the missing data namely, $\boldsymbol{A} \in \mathcal{A}$, $\boldsymbol{H} \in \mathcal{H}$, $\boldsymbol{Q} \in \mathcal{Q}$, $\boldsymbol{Z} \in$

$\mathcal{Z}$. Let $\mathcal{D}$ be the set of all probability distributions on $\mathcal{A}\times\mathcal{H}\times\mathcal{Q}\times\mathcal{Z}$. EM can be viewed as an alternating maximization procedure of a function $\mathcal{F}$ on $\mathcal{D}$, $\mathcal{F}(p,\boldsymbol{\Theta}) = \mathrm{E}_p\big[\log p(\boldsymbol{Y},\boldsymbol{A},\boldsymbol{H},\boldsymbol{Q},\boldsymbol{Z}\,|\,\boldsymbol{\Theta})\big] + \mathcal{G}(p)$ where $\mathrm{E}_p\big[.\big]$ denotes the expectation with respect to p and $\mathcal{G}(p) = -\mathrm{E}_p\big[\log p(\boldsymbol{A},\boldsymbol{H},\boldsymbol{Q},\boldsymbol{Z})\big]$ is the entropy of p. At iteration (r), denoting the current parameter values by $\boldsymbol{\Theta}^{(r-1)}$, the alternating procedure proceeds as follows:

$$\textbf{E-step: } p^{(r)}_{A,H,Q,Z} = \arg\max_{p\in\mathcal{D}}\ \mathcal{F}(p,\boldsymbol{\Theta}^{(r-1)}) \tag{5}$$

$$\textbf{M-step: } \boldsymbol{\Theta}^{(r)} = \arg\max_{\Theta\in\underline{\boldsymbol{\Theta}}}\ \mathcal{F}(p^{(r)}_{A,H,Q,Z},\boldsymbol{\Theta}) \tag{6}$$

However, the optimization step in Eq. (5) leads to $p^{(r)}_{A,H,Q,Z} = p(\boldsymbol{A},\boldsymbol{H},\boldsymbol{Q},\boldsymbol{Z}\,|\,\boldsymbol{Y},\boldsymbol{\Theta}^{(r-1)})$, which is intractable for our model. Hence, we resort to a variational EM variant in which the intractable posterior is approximated as a product of four pdfs on $\mathcal{A}$, $\mathcal{H}$, $\mathcal{Q}$ and $\mathcal{Z}$ respectively. The intractable E-step is instead solved over $\tilde{\mathcal{D}}$, a restricted class of probability distributions chosen as the set of distributions that factorize as $\widetilde{p}_{A,H,Q,Z} = \widetilde{p}_A\widetilde{p}_H\widetilde{p}_Q\widetilde{p}_Z$ where $\widetilde{p}_A$, $\widetilde{p}_H$, $\widetilde{p}_Q$ and $\widetilde{p}_Z$ are probability distributions on $\mathcal{A}$, $\mathcal{H}$, $\mathcal{Q}$ and $\mathcal{Z}$ respectively. It follows then that our E-step becomes an approximate E-step, which can be further decomposed into four stages that consist of updating the four pdfs, $\widetilde{p}_H$, $\widetilde{p}_A$, $\widetilde{p}_Q$ and $\widetilde{p}_Z$, in turn using four equivalent expressions of $\mathcal{F}$ when p factorizes as in $\tilde{\mathcal{D}}$. At iteration (r), with current estimates denoted by $q_A^{(r-1)}, q_Q^{(r-1)}, q_Z^{(r-1)}$ and $\boldsymbol{\Theta}^{(r-1)}$, the updating rules become (using the Kullback-Leibler divergence properties):

$$\textbf{E-H: } \widetilde{p}_H^{(r)}(\boldsymbol{h}) \propto \exp\Big(\mathrm{E}_{\widetilde{p}_A^{(r-1)}\widetilde{p}_Z^{(r-1)}}\big[\log p(\boldsymbol{h}\,|\,\boldsymbol{Y},\boldsymbol{A},\boldsymbol{Z};\boldsymbol{\Theta}^{(r-1)}\big]\Big) \tag{7}$$

$$\textbf{E-A: } \widetilde{p}_A^{(r)}(\boldsymbol{A}) \propto \exp\Big(\mathrm{E}_{\widetilde{p}_H^{(r)}\widetilde{p}_Q^{(r-1)}}\big[\log p(\boldsymbol{A}\,|\,\boldsymbol{Y},\boldsymbol{H},\boldsymbol{Q};\boldsymbol{\Theta}^{(r-1)})\big]\Big) \tag{8}$$

$$\textbf{E-Q: } \widetilde{p}_Q^{(r)}(\boldsymbol{Q}) \propto \exp\Big(\mathrm{E}_{\widetilde{p}_A^{(r)}}\big[\log p(\boldsymbol{Q}\,|\,\boldsymbol{Y},\boldsymbol{A};\boldsymbol{\Theta}^{(r-1)})\big]\Big) \tag{9}$$

$$\textbf{E-Z: } \widetilde{p}_Z^{(r)}(\boldsymbol{Z}) \propto \exp\Big(\mathrm{E}_{\widetilde{p}_H^{(r)}}\big[\log p(\boldsymbol{Z}\,|\,\boldsymbol{Y},\boldsymbol{H};\boldsymbol{\Theta}^{(r-1)}\big]\Big). \tag{10}$$

The **M-step** writes (since $\boldsymbol{\Theta}$ and $\mathcal{G}(p^{(r)}_{A,H,Q,Z})$ are independent):

$$\textbf{M: } \boldsymbol{\Theta}^{(r)} = \arg\max_{\boldsymbol{\Theta}}\ \mathrm{E}_{\widetilde{p}_A^{(r)}\widetilde{p}_H^{(r)}\widetilde{p}_Q^{(r)}\widetilde{p}_Z^{(r)}}\big[\log p(\boldsymbol{Y},\boldsymbol{A},\boldsymbol{H},\boldsymbol{Q},\boldsymbol{Z};\boldsymbol{\Theta})\big]. \tag{11}$$

For the **E-H** and **E-A** steps it follows from standard algebra that $q_H^{(r)}$ and $q_A^{(r)}$ are both Gaussian distributions: $\widetilde{p}_H^{(r)} = \prod_{j\in\boldsymbol{\mathcal{P}}}\widetilde{p}_{H_j}^{(r)}$ and $\widetilde{p}_A^{(r)} = \prod_{j\in\boldsymbol{\mathcal{P}}}\widetilde{p}_{A_j}^{(r)}$, where $\widetilde{p}_{H_j}^{(r)} \sim \mathcal{N}(\boldsymbol{m}_{H_j}^{(r)},\boldsymbol{\Sigma}_{H_j}^{(r)})$ and $q_{A_j}^{(r)} \sim \mathcal{N}(\boldsymbol{m}_{A_j}^{(r)},\boldsymbol{\Sigma}_{A_j}^{(r)})$. More specifically, assuming current values for the $\boldsymbol{m}_{A_j}^{(r-1)}$, $\boldsymbol{\Sigma}_{A_j}^{(r-1)}$ and $\widetilde{p}_{Q_j^m}^{(r-1)}$, the r^{th} iteration starts with:

• **E-H step:** Compute

$$\boldsymbol{\Sigma}_{H_j}^{(r)-1} = \boldsymbol{V}_1 + \boldsymbol{V}_2 \tag{12}$$

$$\boldsymbol{m}_{H_j}^{(r)} = \boldsymbol{\Sigma}_{H_j}^{(r)}(\boldsymbol{m}_1 + \boldsymbol{m}_2), \tag{13}$$

where $\boldsymbol{V}_1 = \sum_{m,m'}\boldsymbol{\Sigma}_{A_j}^{(r-1)}(m,m')\boldsymbol{X}_m^{\mathrm{t}}\boldsymbol{\Gamma}_j^{(r-1)}\boldsymbol{X}_{m'} + \widetilde{\boldsymbol{S}}_j^{\mathrm{t}}\boldsymbol{\Gamma}_j^{(r-1)}\widetilde{\boldsymbol{S}}_j$, $\boldsymbol{V}_2 = \sum_{k=1}^{K}\widetilde{p}_{Z_j}^{(r-1)}(k)\bar{\boldsymbol{\Sigma}}_k^{-1}$, $\boldsymbol{m}_1 = \widetilde{\boldsymbol{S}}_j^{\mathrm{t}}\boldsymbol{\Gamma}_j^{(r-1)}(\boldsymbol{y}_j - \boldsymbol{P}\ell_j^{(r-1)})$ and $\boldsymbol{m}_2 = \sum_{k=1}^{K}\bar{\boldsymbol{\Sigma}}_k^{-1}\widetilde{p}_{Z_j}^{(r-1)}(k)\bar{\boldsymbol{h}}_k$. Hereabove, $\widetilde{\boldsymbol{S}}_j = \sum_{m=1}^{M} m_{A_j^m}^{(r-1)}\boldsymbol{X}_m$ and $m_{A_j^m}^{(r-1)}$, $\boldsymbol{\Sigma}_{A_j}^{(r-1)}(m,m')$ denote respectively the m and (m,m') entries of the mean vector $\boldsymbol{m}_{A_j}^{(r-1)}$ and covariance matrix $\boldsymbol{\Sigma}_{A_j}^{(r-1)}$ of the current $\widetilde{p}_{A_j}^{(r-1)}$.

• **E-A step:** Compute $\boldsymbol{\Sigma}_{A_j}^{(r)} = (\sum_{i=1}^{I}\boldsymbol{\Delta}_{ij}^{(r)} + \tilde{\boldsymbol{H}}_j^{(r)})^{-1}$ and $\boldsymbol{m}_{A_j}^{(r)} = \boldsymbol{\Sigma}_{A_j}^{(r)}(\sum_{i=1}^{I}\boldsymbol{\Delta}_{ij}^{(r)}\boldsymbol{\mu}_i^{(r)} + \widetilde{\boldsymbol{X}}_j^{(r)\mathrm{t}}\boldsymbol{m}_{H_j}^{(r)})$
with $\boldsymbol{\mu}_i^{(r)} = \Big[\mu_{1i}^{(r)}\dots\mu_{Mi}^{(r)}\Big]^{\mathrm{t}}$
and $\boldsymbol{\Delta}_{ij}^{(r)} = \mathrm{diag}_M\Big[\widetilde{p}_{Q_j^1}^{(r-1)}(i)/v_{1i}^{(r)},\dots,\widetilde{p}_{Q_j^M}^{(r-1)}(i)/v_{Mi}^{(r)}\Big]$ where $\widetilde{\boldsymbol{X}}_j^{(r)} = \big[g_1^{\mathrm{t}}\,|\,\dots\,|\,g_M^{\mathrm{t}}\big]^{\mathrm{t}}$ with $g_m = \boldsymbol{\Gamma}_j^{(r)}(y_j - \boldsymbol{P}\ell_j^{(r)})^{\mathrm{t}}\boldsymbol{X}_m$ and $\tilde{\boldsymbol{H}}_j^{(r)}$ is a $M\times M$ matrix whose (m,m') entry is

$$\begin{aligned}\tilde{\boldsymbol{H}}_j^{(r)}(m,m') = {} & \mathit{trace}\,(\boldsymbol{\Sigma}_{H_j}^{(r)}\boldsymbol{X}_m{}^{\mathrm{t}}\boldsymbol{\Gamma}_j^{(r)}\boldsymbol{X}_{m'}) \\ & + \boldsymbol{m}_{H_j}^{(r)\mathrm{t}}\boldsymbol{X}_m^{\mathrm{t}}\boldsymbol{\Gamma}_j^{(r)}\boldsymbol{X}_{m'}\boldsymbol{m}_{H_j}^{(r)}\end{aligned}$$

• **E-Q step:** It comes

$$\widetilde{p}_Q^{(r)}(\boldsymbol{Q}) = \prod_{m=1}^{M}\widetilde{p}_{Q^m}^{(r)}(\boldsymbol{q}^m) \tag{14}$$

$$\text{with}\quad \widetilde{p}_{Q_m}^{(r)}(\boldsymbol{q}^m) = p_m(\boldsymbol{q}^m\,|\,\boldsymbol{a}^m = \boldsymbol{m}_A^{(r)}(m);\boldsymbol{v}_m^{(r)},\beta_m^{(r)})$$

where p_m is a Potts model with interaction parameter $\beta_m^{(r)}$ and external field $\boldsymbol{\alpha}_m^{(r)} = \{\boldsymbol{\alpha}_{mj}^{(r)}, j\in\boldsymbol{\mathcal{P}}\}$
with $\boldsymbol{\alpha}_{mj}^{(r)} = \frac{-1}{2}\boldsymbol{\Sigma}_{A_j}^{(r)}(m,m)\ \Big[1/v_{m1}^{(r)}, 1/v_{m2}^{(r)}\Big]^{\mathrm{t}}$ i.e.

$$\begin{aligned}p_m(\boldsymbol{q}^m;\boldsymbol{v}_m^{(r)},\beta_m^{(r)}) \propto \exp\{\sum_{j\in\boldsymbol{\mathcal{P}}}\Big(&\boldsymbol{\alpha}_{mj}^{(r)}(q_j^m) \\ &+ \frac{\beta_m^{(r)}}{2}\sum_{j\sim i}\delta(q_i^m,q_j^m)\Big)\}\,.\end{aligned}$$

The expression in (14) is intractable but a number of approximation techniques are available. In particular, we can use a mean-field like algorithm (fixing the neighbours to their mean value) as described in [9] in which $\widetilde{q}_{Q^m}(\boldsymbol{q}^m)$ can be approximated by $\widetilde{q}_{Q^m}(\boldsymbol{q}^m) = \prod_{j\in\boldsymbol{\mathcal{P}}}\widetilde{q}_{Q_j^m}(\boldsymbol{q}_j^m)$ with, if $q_j^m = i$, $\widetilde{q}_{Q_j^m}(i) \propto \mathcal{N}(\boldsymbol{m}_{A_j}(m);\mu_{mi},v_{mi})p_m(Q_j^m = i\,|\,\tilde{q}_{\sim j}^m;\beta_m,\boldsymbol{v}_m)$, where $\tilde{\boldsymbol{q}}^m$ is a particular configuration of $\boldsymbol{q}^m$ updated at each iteration according to a specific scheme, $\sim j$ denotes neighbouring voxels to j, and $p_m(q_j^m\,|\,\tilde{q}_{\sim j}^m;\beta_m,\boldsymbol{v}_m) \propto \exp\{\boldsymbol{\alpha}_{mj}(q_j^m) + \beta_m\sum_{l\sim j}\delta(\tilde{q}_j^m,q_l^m)\}$. See [9] for details.

• **E-Z step:** This step is similar to the **E-Q** one. Calculations yield to similar form of $\widetilde{p}_Z(\boldsymbol{z}) = \prod_{j\in\boldsymbol{\mathcal{P}}}\widetilde{p}_{Z_j}(z_j)$ where $\widetilde{p}_{Z_j}(k) \propto \mathcal{N}(\boldsymbol{m}_{H_j};\bar{h}_k^{(r-1)},\bar{\boldsymbol{\Sigma}}_k^{(r-1)})p_m(Z_j = k\,|\,\tilde{z}_{\sim j};\beta_z)$ where $\tilde{\boldsymbol{z}}$ is a particular configuration of $\boldsymbol{z}$ updated at each iteration according to a specific scheme and $p_m(z_j\,|\,\tilde{z}_{\sim j};\beta_z) \propto \exp\{-\frac{1}{2}\mathrm{trace}(\boldsymbol{\Sigma}_{H_j}^{(r)}\bar{\boldsymbol{\Sigma}}_{z_j}^{(r-1)-1}) + \beta_z\sum_{j\sim l}\delta(z_j,\tilde{z}_l)\}$.

• **M step:** The maximization step can also be divided into four sub-steps involving separately $(\boldsymbol{\mu},\boldsymbol{\sigma})$, $\boldsymbol{\beta}$, $(\ell,\boldsymbol{\Gamma})$ and $(\bar{\boldsymbol{h}}_k,\bar{\boldsymbol{\Sigma}}_k)_{1\le k\le K}$. For the $(\boldsymbol{\mu},\boldsymbol{\sigma})$ and $(\bar{\boldsymbol{h}}_k,\bar{\boldsymbol{\Sigma}}_k)_{1\le k\le K}$ sub-steps, closed forms can be analytically derived for the updates. However, numerical resolution is required for the other sub-steps. For more details, the interested reader can refer to [3].

4. ILLUSTRATIONS

In this section, the JPDE is validated and compared to the parcel-based JDE approach. Experiments have been conducted on artificial fMRI signal generated according to the observation model in Eq. (1). We simulated a random mixed sequence of indexes coding for $M = 2$ different stimuli. These two sets of trials (30 trials per stimulus) were then multiplied by stimulus-dependent and space-varying NRLs, which were generated according to the prior distribution in Section 2.2.2. To this end, we generated 2D slices composed of 20 x 20 binary labels $\boldsymbol{Q}^m$ (activated and non-activated voxels) for each stimulus type m. Then, we simulated normally-distributed NRLs: $a_j^1 \,|\, q_j^1 = 1 \sim \mathcal{N}(0, 0.5)$, $a_j^1 \,|\, q_j^1 = 2 \sim \mathcal{N}(2, 0.5)$, $a_j^2 \,|\, q_j^2 = 1 \sim \mathcal{N}(0, 0.5)$, and $a_j^1 \,|\, q_j^2 = 2 \sim \mathcal{N}(2.8, 0.5)$. HRFs have been also simulated for each voxel conditionally to a parcellation mask and according to the prior distribution in Section 2.2.2. This parcellation mask ($\boldsymbol{Z}$) is composed of two parcels having the same size. Within each parcel, all voxels share the same HRF prior parameters $\bar{\boldsymbol{h}}_k$ and $\bar{\boldsymbol{\Sigma}}_k$. The first experiment conducted here aims at demonstrating the robustness of the proposed approach especially in estimating the HRFs in comparison with parcel-based JDE. In this simulation, the same HRF prior parameters have been used for both parcels (1 actual HRF group). For estimation we assume two HRF groups (one for each parcel) for JPDE. For JDE, the two parcels are considered as one parcel (with a single HRF model). Regarding estimated NRLs, results (not displayed here) show that JDE is more robust in terms of Mean Square Error (MSE), which was expected due to the model mis-specification in the JPDE case ($MSE_{\text{JDE}} = 0.020$ and $MSE_{\text{JPDE}} = 0.025$). However, estimated HRFs show that the new approach allows retrieving more accurate HRFs due to the new mixture prior. Fig. 1[left] shows reference and estimated HRFs using both methods (red and yellow curves correspond to the HRF expectation across each parcel). From a qualitative viewpoint, it is worth noticing that the proposed method better estimates the HRF tail compared to JDE. In terms of MSE, obtained values confirm the superiority of our approach from a quantitative viewpoint ($MSE_{\text{JDE}} = 1.7010^{-5}$ and $MSE_{\text{JPDE}} = 3.0910^{-7}$).

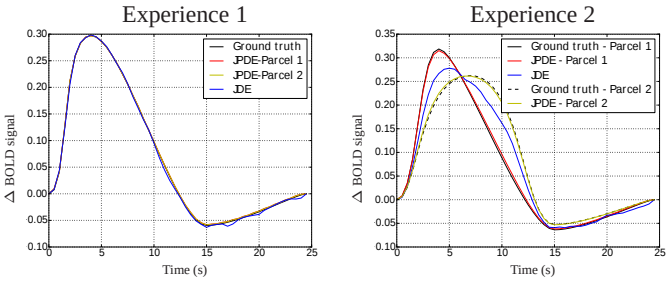

Fig. 1. Ground truth and estimated HRFs using JDE and JPDE.

The second experiment has been conducted to validate the robustness of the proposed approach when the HRF properties of the considered parcels are really different (two actual HRF groups). For estimation, $K = 2$ is still assumed for JPDE while JDE cannot account for more than one HRF group. Fig. 1[right] illustrates reference and estimated HRFs with both methods (red and yellow curves correspond to the HRF expectation across each parcel). It is clear that the proposed approach allows retrieving accurate HRF estimates for each parcel, while JDE estimates a HRF which lies between the two groups (Parcel 1: $MSE_{\text{JPDE}} = 5.3410^{-6}$, $MSE_{\text{JDE}} = 1.1010^{-4}$; Parcel 2: $MSE_{\text{JPDE}} = 1.4410^{-6}$, $MSE_{\text{JDE}} = 1.3610^{-4}$). This results confirms the superiority of the proposed approach in allowing variability of HRFs across voxels compare to the single HRF model assumed in JDE.

As regards activation detection, Fig. 2 plots the ROC curves for both algorithms. These activation classification results show that JPDE allows well detecting activations, and even slightly outperforms JDE.

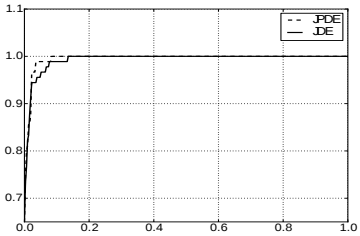

Fig. 2. ROC curves for activation classification using JDE and JPDE.

5. CONCLUSION

We proposed an extension (JPDE) to the joint detection-estimation framework that does not depend on the choice of an initial brain parcellation. Preliminary experiments showed that this approach achieved similar and even better results than the standard JDE and future work includes the application of our method to real 3D datasets on the whole brain. In addition, this new framework raises the question of model selection, in particular the issue of choosing the right number of HRF groups at best *i.e.* in a sparse manner so as to capture the spatial variability in hemodynamic territories while enabling the reproducibility of parcel identification across fMRI datasets. This question should be the most critical to validate our approach but also the most interesting to neuroscientists in case of success. For this specific point, we shall investigate variational approximations of standard information criteria such as the Bayesian Information Criterion.

6. REFERENCES

[1] S. Makni, J. Idier, T. Vincent, B. Thirion, G. Dehaene-Lambertz, and P. Ciuciu, "A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI," *Neuroimage*, vol. 41, no. 3, pp. 941–969, 2008.

[2] 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, pp. 1059–1074, 2010.

[3] L. Chaari, F. Forbes, T. Vincent, M. Dojat, and P. Ciuciu, "Variational solution to the joint detection estimation of brain activity in fMRI," in *MICCAI*, 2011, vol. 6892, pp. 260–268.

[4] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Nat. Acad. Sci.*, vol. 87, pp. 9868–72, 1990.

[5] K. J. Friston, A. P. Holmes, J.-B. Poline, P. J. Grasby, S. C. R. Williams, R. S. J. Frackowiak, and R. Turner, "Analysis of fMRI time-series revisited," *Neuroimage*, vol. 2, no. 1, pp. 45–53, 1995.

[6] D. A. Handwerker, J. M. Ollinger, and D. Mark, "Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses," vol. 21, pp. 1639–1651, 2004.

[7] J. Kershaw, B. A. Ardekani, and I. Kanno, "Application of Bayesian inference to fMRI data analysis," *IEEE Trans. on Med. Imag.*, vol. 18, no. 12, pp. 1138–1152, Dec. 1999.

[8] S. Makni, P. Ciuciu, J. Idier, and J.-B. Poline, "Joint detection-estimation of brain activity in fmri using an autoregressive noise model," in *IEEE Int. Symp. on Biomed. Imag. (ISBI)*, Arlington, VA, 6-9 April 2006, pp. 1048 – 1051.

[9] G. Celeux, F. Forbes, and N. Peyrard, "EM procedures using mean field-like approximations for Markov model-based image segmentation," *Patt. Rec.*, vol. 36, pp. 131–144, 2003.