# MODEL SELECTION FOR HEMODYNAMIC BRAIN PARCELLATION IN FMRI

*Mohanad Albughdadi*[1], *Lotfi Chaari*[1], *Florence Forbes*[2], *Jean-Yves Tourneret*[1] *and Philippe Ciuciu*[3]

[1] University of Toulouse, IRIT
INP-ENSEEIHT, France
firstname.lastname@enseeiht.fr

[2] INRIA, MISTIS, Grenoble
University, LJK, Grenoble, France
florence.forbes@inria.fr

[3] CEA/NeuroSpin and INRIA Saclay,
Parietal, France
philippe.ciuciu@cea.fr

## ABSTRACT

Brain parcellation into a number of hemodynamically homogeneous regions (parcels) is a challenging issue in fMRI analyses. This task has been recently integrated in the joint detection estimation [1] resulting in the so-called joint parcellation detection estimation (JPDE) model [2]. JPDE automatically estimates the parcels from the fMRI data but requires the desired number of parcels to be fixed. This is potentially critical in that the chosen number of parcels may influence detection-estimation performance. In this paper, we propose a model selection procedure to automatically set the number of parcels from the data. The selection procedure relies on the calculation of the free energy corresponding to each concurrent model, within the variational expectation maximization framework. Experiments on synthetic and real fMRI data demonstrate the ability of the proposed procedure to select the optimal number of parcels.

***Index Terms***— fMRI, JDE, JPDE, Parcellation, VEM

## 1. INTRODUCTION

Functional MRI (fMRI) is an imaging technique that indirectly measures neural activity through the blood-oxygen-level-dependent (BOLD) signal [3], which captures the variation in blood oxygenation arising from an external stimulation. This variation also allows the estimation of the underlying dynamics, namely the characterization of the so-called hemodynamic response function (HRF). The hemodynamic characteristics are likely to spatially vary, but can be considered constant up to a certain spatial extent. Hence, it makes sense to estimate a single HRF shape for any given area of the brain. To this end, parcel-based approaches that segment fMRI data into functionally homogeneous regions and perform parcel-wise fMRI data analysis provide an appealing framework [4].

In [1, 5, 6], a joint detection-estimation (JDE) approach has been proposed for simultaneously localizing evoked brain activity and estimating HRF shapes at a parcel-level. This parcel scale allows one to make a spatial compromise between hemodynamics reproducibility and signal aggregation, the latter operation enhancing the inherent low signal-to-noise ratio (SNR) of fMRI data. This aggregation has to be well spatially controlled because of the known fluctuation of hemodynamics across brain regions [7, 8]. Hence, the JDE approach operates on a prior partitioning of the brain into functionally homogeneous parcels in which the hemodynamics is assumed constant. A robust parcellation is needed to ensure a good JDE performance. A few attempts have been proposed to cope with this issue [2, 4, 8–11] among which some do not account for hemodynamics variability [4, 9, 10]. In this work, we focus on the JDE extension proposed in [2]. This extension performs online parcellation during the detection-estimation step, resulting in the so-called joint parcellation detection estimation (JPDE) framework. This was accomplished by proposing the concept of hemodynamic territory, which is a set of parcels that share a common HRF pattern. HRFs are voxel-wise and are defined as local stochastic perturbations of the same HRF pattern, which is used to define the hemodynamic territory. However, this approach still requires to fix the number of parcels and lacks the possibility to determine this number automatically from the data. Indeed, as reported in [2], prior parcellations may suffer frorm over granularity, which means that some of the brain regions are assumed to be in different parcels while they share the same HRF properties. The JPDE model allows to reduce this over-parcellation effect. However, the user still has to manually specify the number of parcels to be considered.

In this paper, we propose a variational scoring for JPDE model selection. It relies on calculating the free energy of the model with different numbers of parcels. The model with the highest free energy corresponds to the best fit. The selection criterion is integrated in the variational expectation maximization algorithm already developed in [2]. The algorithm is tested on both synthetic and real data and the results show its ability to select the best model that fits the data automatically. The rest of the paper is organized as follows. Section 2 describes the JPDE framework. The model selection procedure is then described in Section 3 and validated in Section 4 both on synthetic and real data. Finally, some conclusions and perspectives are drawn in Section 5.

This works has been partially supported by ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02. The work of Philippe Ciuciu was partially supported by the CIMI (Centre International de Mathmatiques et d'Informatique) Excellence program during winter and spring 2013.

## 2. JOINT PARCELLATION DETECTION ESTIMATION MODEL

### 2.1. The JPDE model

The adopted JPDE model is the one proposed in [2], which can be fragmented as follows. Let $\boldsymbol{\mathcal{P}}$ be the set of voxels ($J$ voxels) and $y_j$ the fMRI time series for the voxel $\#j$ at times $\{t_n,\ n = 1,\dots,N\}$, where $t_n = nTR$, $N$ is the number of scans and $TR$ is the repetition time. The whole data set is denoted by $\boldsymbol{Y} = \{\boldsymbol{y}_j, j \in \boldsymbol{\mathcal{P}}\}$ and we consider $M$ different stimulus types (experimental conditions). The HRFs are assumed voxel-dependent and belong to a set denoted by $\boldsymbol{H} = \{\boldsymbol{h}_j, j \in \boldsymbol{\mathcal{P}}\}$ with $\boldsymbol{h}_j \in \mathbb{R}^{D+1}$. Each $\boldsymbol{h}_j$ is associated with one among $K$ considered HRF groups (parcels). These groups are encoded using a set of hidden labels $\boldsymbol{Z} = \{z_j, j \in \boldsymbol{\mathcal{P}}\}$ (where $z_j \in \{1,\dots,K\}$) which follows a K-class Markov field (Potts) model with interaction parameter $\beta_z$. This parameter accounts for spatial information and is assigned an exponential prior with hyperparameter $\lambda_z$. We assume that in the group $\#k$, the HRF $\boldsymbol{h}_j$ is a stochastic perturbation of an HRF pattern $\bar{\boldsymbol{h}}_k$. More precisely, we assign to $\boldsymbol{h}_j$ a Gaussian prior distribution $\mathcal{N}(\bar{\boldsymbol{h}}_k, \nu_k \mathrm{Id})$. To favor smooth HRF patterns, the temporal HRF mean $\bar{\boldsymbol{h}}_k$ is also assumed to be distributed according to a common centered isotropic Gaussian distribution defined by a scalar variance parameter $\sigma_h^2$. We then consider the following observation model

$$\forall j \in \boldsymbol{\mathcal{P}}, \quad \boldsymbol{y}_j = \sum_{m=1}^{M} a_j^m \boldsymbol{X}_m \boldsymbol{h}_j + \boldsymbol{P}\ell_j + \boldsymbol{\varepsilon}_j, \qquad (1)$$

where $\boldsymbol{P}\ell_j$ corresponds to low frequency drifts (more details are given in [1, 6, 12]). For the $m^{th}$ experimental condition, $\boldsymbol{X}_m = \{x_m^{n-d\Delta t}, n = 1,\dots,N, d = 0,\dots,D\}$ is a binary matrix that provides information on the stimulus occurrences, $\Delta t < TR$ being the sampling period of the unknown HRFs. The neural response levels (NRL) are denoted by $\boldsymbol{A} = \{\boldsymbol{a}^m, m = 1,\dots,M\}$ with $\boldsymbol{a}^m = \{a_j^m, j \in \boldsymbol{\mathcal{P}}\}$. The $\boldsymbol{a}^m$'s follow spatial Gaussian mixtures defined by a set of parameters $\boldsymbol{\theta_a}$ and governed by binary Markov fields. More specifically, each NRL is assigned to one of the activation classes encoded by the variables $\boldsymbol{Q} = \{\boldsymbol{q}^m, m = 1,\dots,M\}$ where $\boldsymbol{q}^m = \{q_j^m, j \in \boldsymbol{\mathcal{P}}\}$ is a binary Markov field with interaction parameter $\beta_m$ distributed according to an exponential distribution with parameter $\lambda_m$. Two classes are considered, $q_j^m = 1$ (resp. $q_j^m = 2$) if voxel $j$ is not activated (resp. activated) by condition $m$. Akin to [1, 2, 6, 12], an additive Gaussian noise $\boldsymbol{\varepsilon}_j$ is considered with covariance matrix $\boldsymbol{\Gamma}_j^{-1}$. The JPDE model is summarized in the directed acyclic graph displayed in Fig. 1. We will denote by $\boldsymbol{\Phi} = \{\boldsymbol{L}, \boldsymbol{\Gamma}, \boldsymbol{\theta_a}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \lambda_z, \sigma_h^2\}$ the set of hyperparameters of the proposed model, where $\boldsymbol{L}$ (resp. $\boldsymbol{\Gamma}, \boldsymbol{\nu}$) is a shortcut for $\boldsymbol{L} = \{\ell_j, j \in \mathcal{P}\}$ (resp. $\boldsymbol{\Gamma} = \{\boldsymbol{\Gamma}_j, j \in \boldsymbol{\mathcal{P}}\}, \boldsymbol{\nu} = \{\nu_k, k = 1,\dots,K\}$). The other random parameters for which we consider priors are denoted by $\boldsymbol{\Psi} = \{\bar{\boldsymbol{h}}, \boldsymbol{\beta}, \beta_z\}$ with $\bar{\boldsymbol{h}} = \{\bar{\boldsymbol{h}}_k, k = 1,\dots,K\}$ and $\boldsymbol{\beta} = \{\beta_m, m = 1,\dots,M\}$. The set of all model parameters is finally denoted as $\boldsymbol{\Theta} = \{\boldsymbol{\Psi}, \boldsymbol{\Phi}\}$.

Note here that the proposed model depends on $K$, the number of parcels, and hence the number of Gaussian distributions involved in the prior of $\boldsymbol{h}_j$. This parameter has to be known to apply the estimation procedure developed in [2] (see dashed square in Fig. 1). However, several models with different values of $K$ may be in competition to better fit the fMRI data. The main contribution of this paper is to investigate a method allowing the best value of $K$ to be selected. This method is detailed in Section 3.

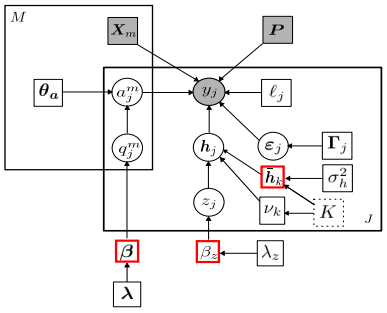


**Fig. 1**. Graphical model describing dependencies between variables involved in the JPDE model. Circles and squares indicate random variables and model parameters, respectively. Red squares indicate the random model parameters. The dashed square indicates the fixed parameter $K$.

### 2.2. Variational expectation maximization estimation

Using the same assumptions as in [2], joint distribution of the proposed model is

$$p(\boldsymbol{Y}, \boldsymbol{A}, \boldsymbol{Q}, \boldsymbol{H}, \boldsymbol{Z}, \boldsymbol{\Psi}; \boldsymbol{\Phi}) = p(\boldsymbol{Y} \mid \boldsymbol{A}, \boldsymbol{H}; \boldsymbol{\Phi})\ p(\boldsymbol{A} \mid \boldsymbol{Q}; \boldsymbol{\Phi}) \times p(\boldsymbol{Q}|\boldsymbol{\beta})\ p(\boldsymbol{H} \mid \boldsymbol{Z}, \bar{\boldsymbol{h}}; \boldsymbol{\Phi})\ p(\boldsymbol{Z}|\beta_z)\ p(\bar{\boldsymbol{h}}; \boldsymbol{\Phi}) p(\boldsymbol{\beta}; \boldsymbol{\Phi}) p(\beta_z; \boldsymbol{\Phi}), \qquad (2)$$

where we have distinguished parameters from hyperparameters by using $\boldsymbol{\Phi}$ whenever there was a dependence on some of the hyperparameters. For inference, we use as in [2], a variational expectation maximization (VEM) [13, 14] procedure in which we compute approximate posterior distributions for the missing variables namely $\{\boldsymbol{A}, \boldsymbol{H}, \boldsymbol{Q}, \boldsymbol{Z}\}$. Formally, we could also compute posterior distributions for the parameters $\boldsymbol{\Psi}$ but we choose to restrict for them to point estimates, which amounts to seeking for posteriors that are Dirac ($\delta$) functions. It follows that the VEM procedure is based on the alternating maximization of the following free energy term

$$\mathcal{F}(\tilde{p}, \boldsymbol{\Theta}) = \mathrm{E}_{\tilde{p}}\big[\log p(\boldsymbol{Y}, \boldsymbol{A}, \boldsymbol{Q}, \boldsymbol{H}, \boldsymbol{Z}, \boldsymbol{\Psi}; \boldsymbol{\Phi})\big] + \mathcal{G}(\tilde{p}), \quad (3)$$

where we optimize alternatively with respect to a distribution $\tilde{p}$ (VE-step) and the parameter vector $\Phi$ (VM step). Note that $\tilde{p}$ has the following product form $\tilde{p}(\boldsymbol{A}, \boldsymbol{H}, \boldsymbol{Q}, \boldsymbol{Z}, \boldsymbol{\Psi}) =$

$\tilde{p}_A(\boldsymbol{A})\ \tilde{p}_H(\boldsymbol{H})\ \tilde{p}_Q(\boldsymbol{Q})\ \tilde{p}_Z(\boldsymbol{Z})\ \delta_{\tilde{\boldsymbol{\Psi}}}(\boldsymbol{\Psi})$ and over $\boldsymbol{\Phi}$ (VM-step). $\mathrm{E}_{\tilde{p}}\big[.\big]$ denotes the expectation with respect to $\tilde{p}$ and $\mathcal{G}(\tilde{p}) = -\mathrm{E}_{\tilde{p}}\big[\log \tilde{p}(\boldsymbol{A}, \boldsymbol{H}, \boldsymbol{Q}, \boldsymbol{Z}, \boldsymbol{\Psi})\big]$ is the entropy of $\tilde{p}$. For the product form above, it simplifies to $\mathcal{G}(\tilde{p}) = \mathcal{G}(\tilde{p}_A) + \mathcal{G}(\tilde{p}_H) + \mathcal{G}(\tilde{p}_Q) + \mathcal{G}(\tilde{p}_Z)$. Starting from initial distributions $\tilde{p}_A^{(0)}, \tilde{p}_H^{(0)}, \tilde{p}_Q^{(0)}, \tilde{p}_Z^{(0)}$ and initial values $\boldsymbol{\Psi}^{(0)}, \boldsymbol{\Phi}^{(0)}$, we then alternate the two optimization E and M steps to obtain the final posteriors denoted by $\tilde{p}_A^{(\infty)}, \tilde{p}_H^{(\infty)}, \tilde{p}_Q^{(\infty)}, \tilde{p}_Z^{(\infty)}$ and point estimates $\boldsymbol{\Psi}^{(\infty)}, \boldsymbol{\Phi}^{(\infty)}$, where $\boldsymbol{\Psi}^{(\infty)}$ takes into account the prior distributions for the parameters in $\boldsymbol{\Psi}$. Point estimates for the missing variables can then be also straightforwardly derived. In the next section, we explain that the free energy can be in addition used as a model selection criterion to assess the *best* model that fits the data.

## 3. JPDE MODEL SELECTION

As stated in the introduction, selecting the number of parcels is a challenging issue for fMRI data analysis. In the above mentioned JPDE framework, it reduces to a model selection issue in which we seek for the value of $K$ that provides the best fit to the data under consideration. When comparing models with different parameter sets (typically the number of JPDE parameters increases with $K$), the likelihood cannot be directly used as a model score because it does not account for the model complexity. Existing well known alternatives based on penalized likelihoods include the Bayesian information criterion (BIC) and the minimum description length (MDL) criterion [15]. We propose another automatic approach that allows retrieving the number of useful parcels based on the calculated free energy ($\mathcal{F}$). Indeed, it has been stated in [16, 17] that the variational objective function $\mathcal{F}$ penalizes model complexity and thus can be used for model selection. It is the idea investigated in this paper which is directly related to the proposed VEM framework. Note that the value of the free energy converges to the popular BIC and MDL criteria when the sample size increases, illustrating the interest of this measure for model selection.

Let us consider $\Omega$ different competing models. For each model, the number of considered parcels (HRF groups) will be denoted by $K^\omega$ ($\omega \in \{1, \ldots, \Omega\}$). For each model $\omega$, a VEM procedure is run as explained in the previous section leading to some estimated posteriors $\tilde{p}_A^{(\infty)}, \tilde{p}_H^{(\infty)}, \tilde{p}_Q^{(\infty)}, \tilde{p}_Z^{(\infty)}$ and point estimates $\boldsymbol{\Psi}^{(\infty)}, \boldsymbol{\Phi}^{(\infty)}$. To refer to model $\omega$, the previous quantities will be indexed by the superscript $\omega$, typically writing $p_A^\omega$ instead of $\tilde{p}_A^{(\infty)}$. The free energy in (3) can then be rewritten as

$$\mathcal{F}(p^\omega, \boldsymbol{\Psi}^\omega, \boldsymbol{\Phi}^\omega) = \mathrm{E}_{p^\omega}\big[\log p(\boldsymbol{Y}, \boldsymbol{A}, \boldsymbol{Q}, \boldsymbol{H}, \boldsymbol{Z}, \boldsymbol{\Psi}^\omega; \boldsymbol{\Phi}^\omega)\big] + \mathcal{G}(p^\omega) \quad (4)$$

where $p^\omega$ is $p^\omega(\boldsymbol{A}, \boldsymbol{Q}, \boldsymbol{H}, \boldsymbol{Z}) = p_A^\omega(\boldsymbol{A}) p_Q^\omega(\boldsymbol{Q})\ p_H^\omega(\boldsymbol{H})\ p_Z^\omega(\boldsymbol{Z})$, or equivalently as

$$\begin{aligned}\mathcal{F}(p^\omega, \boldsymbol{\Psi}^\omega, \boldsymbol{\Phi}^\omega) &= \mathrm{E}_{p_A^\omega p_H^\omega}\big[\log p(\boldsymbol{Y} \mid \boldsymbol{A}, \boldsymbol{H}; \boldsymbol{\Phi}^\omega)\big] \\ &+ \mathrm{E}_{p_A^\omega p_Q^\omega}\big[\log p(\boldsymbol{A} \mid \boldsymbol{Q}; \boldsymbol{\Phi}^\omega)\big] + \mathrm{E}_{p_Q^\omega}\big[\log p(\boldsymbol{Q}|\boldsymbol{\beta}^\omega)\big] \\ &+ \mathrm{E}_{p_H^\omega p_Z^\omega}\big[\log p(\boldsymbol{H} \mid \boldsymbol{Z}, \bar{\boldsymbol{h}}^\omega; \boldsymbol{\Phi}^\omega)\big] + \mathrm{E}_{p_Z^\omega}\big[\log p(\boldsymbol{Z}|\beta_z^\omega)\big] \\ &+ \log p(\bar{\boldsymbol{h}}^\omega; \boldsymbol{\Phi}^\omega) + \log p(\boldsymbol{\beta}^\omega; \boldsymbol{\Phi}^\omega) + \log p(\beta_z^\omega; \boldsymbol{\Phi}^\omega) \\ &+ \mathcal{G}(p_A^\omega) + \mathcal{G}(p_Q^\omega) + \mathcal{G}(p_H^\omega) + \mathcal{G}(p_Z^\omega). \end{aligned} \quad (5)$$

Each term in the sum above can be easily computed using the VEM results due to the tractable (*e.g.* Gaussian) form of the posteriors. The reader is invited to consult [2] for details on these posteriors. After running JPDE for the $\Omega$ models, the model giving the highest free energy is selected as the most relevant one with respect to the parcellation task.

## 4. EXPERIMENTAL VALIDATION

In order to validate the proposed model selection procedure, different experiments have been conducted both on real and synthetic data.

### 4.1. Synthetic data

Three experiments (denoted as Exp.1, Exp.2 and Exp.3) have been conducted here to validate the proposed model selection procedure. For each experiment, a different parcellation mask is considered with two, three and four parcels (see Fig. 3[left]). Based on these original masks, a BOLD signal has been generated according to the observation model in (1) for each experiment (using the PYHRF software[1]). Two experimental conditions ($M = 2$) have been considered with 30 trials for each of them. The reference activation labels are illustrated in Fig. 2. These maps of size $20 \times 20$ have been chosen so that each parcel covers a part of the simulated activity. As regards NRLs, they have been drawn according to their prior distribution. For the HRFs, different groups have been considered, each with $K^\omega = \omega + 1$ parcels where $\omega \in \Omega$. Data resulting from the three experiments have been processed by the proposed JPDE framework while performing model selection based on the proposed free energy scoring. For each experiment, a set of $\Omega = 1, \ldots, 4$ models has been investigated. After evaluating the free energies for all models, it turns out that the retained models are $\omega = 1$, $\omega = 2$ and $\omega = 3$, respectively. The corresponding values of the free energy are provided in Tab. 1, where the highest value for each experiment appears in bold font. The estimated and initial parcellation masks are displayed in Figs. 3[middle] and 3[right] , respectively. This figure shows accurate estimated masks from a visual point of view.

Quantitatively speaking, MSE (mean square error) values for NRLs and activation maps are reported in Tab. 2. The same table also shows the parcellation error percentage for the

[1] www.pyhrf.org

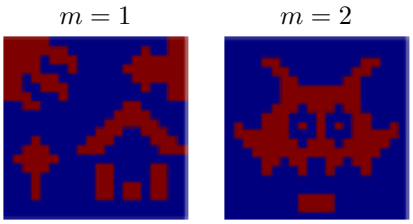


**Fig. 2**. Reference activation labels for the two experimental conditions.

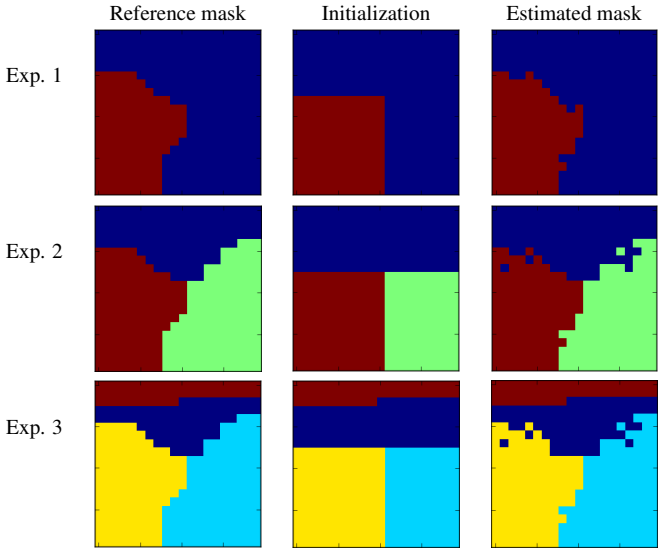


**Fig. 3**. Reference, estimated and initial parcellation masks for the three experiments.

selected model in each experiment. The results are promising and illustrate the good detection accuracy and parcellation associated with the proposed JPDE for the selected model. The values reported in Tab. 1 also show the slight error increase with respect to the number of parcels, which is actually expected. Note that these results are reproducible across runs.

**Table 1**. Estimated free energy for the three experiments and the four different models. Bold values indicate the highest free energy.

| | Exp.1 | Exp.2 | Exp.3 |
|---|---|---|---|
| $\omega = 1$ (2 parcels) | **78614.04** | 59459.32 | 6185.50 |
| $\omega = 2$ (3 parcels) | 77126.72 | **86584.94** | 83424.95 |
| $\omega = 3$ (4 parcels) | 78365.17 | 86492.43 | **89146.02** |
| $\omega = 4$ (5 parcels) | 75651.11 | 85180.44 | 87305.96 |

### 4.2. Real data

A gradient-echo EPI (echo planar imaging) sequence (Echo Time=30ms / Repetition Time=2.4s / slice thickness = 3mm / Field Of View=192mm$^2$) was used to acquire the real fMRI data at 3T during a localizer experiment [18]. This paradigm involved sixty auditory, visual and motor stimuli, defined in ten experimental conditions ($M = 10$). During the used paradigm, 128 scans have been acquired at a $2 \times 2 \times 3$mm$^3$ 3D spatial resolution. Akin to the experiment conducted in [2], we focus on the auditory condition that generates activations in the temporal lobe. Specifically, we focus on the two regions of interest (ROI) presented in Fig. 4 denoted as ROI 1 (left ROI) and ROI 2 (right ROI). $\Omega = 5$ different models have been tested with $K^1 = 1$, $K^2 = 2$, $K^3 = 3$, $K^4 = 5$ and $K^5 = 8$. For each model, the initial parcellation has been obtained by merging neighboring parcels obtained by applying the method of [4]. After running all the models separately and calculating the final free energies, it turns out that $\omega = 2$ (2 parcels) is the best model that fits the fMRI data. The initialization and estimated masks with the selected model are illustrated in Figs. 4[a] and 4[b]. This figure shows the two estimated parcels. Fig. 4[c] also shows the estimated NRLs with the same model. These results are consistent with those obtained in [2] where three parcels have been identified, two of them being very similar. In our results, these two parcels have been successfully merged in the two-parcel model.

**Table 2**. MSE values of NRLs and activation maps, in addition to the parcellation error percentage. The reported values correspond to the estimated model for each experiment.

| | | Exp.1 | Exp.2 | Exp.3 |
|---|---|---|---|---|
| NRLs | $m = 1$ | 0.016 | 0.017 | 0.017 |
| | $m = 2$ | 0.012 | 0.012 | 0.012 |
| Labels | $m = 1$ | 0.0034 | 0.011 | 0.011 |
| | $m = 2$ | 0.0026 | 0.0026 | 0.0027 |
| Parcellation | | 1.5% | 2.75% | 3.25% |

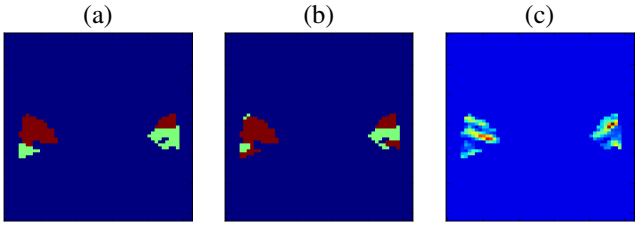


**Fig. 4**. (a): initial parcellation mask; (b) estimated parcellation mask with the two-parcel model ($\omega = 2$); (c) estimated NRLs with the retained model ($\omega = 2$). ROI 1: left side; ROI 2: right side.

## 5. CONCLUSION

In this paper, a model selection procedure was proposed to assess the problem of selecting the number of parcels in the

**Table 3**. Evaluated free energy values for the 5 models tested on real data.

| | $\omega = 1$ (1 par.) | $\omega = 2$ (2 par.) | $\omega = 3$ (3 par.) | $\omega = 4$ (5 par.) | $\omega = 5$ (8 par.) |
|---|---|---|---|---|---|
| ROI 1 | -307292 | **-306941** | -309810 | -309580 | -310089 |
| ROI 2 | -226216 | **-224608** | -226127 | -227406 | -226571 |

JPDE framework. The proposed procedure relied on calculating the free energy of different concurrent models. The model with the highest free energy yields best fit. The proposed procedure was included in the VEM framework already adopted in [2]. Validation on synthetic and real data confirmed the ability of the proposed model selection procedure to select the model that best fits the data. Future work will focus on applying the JPDE framework with the proposed model selection procedure to both region-based and whole brain analyses.

## 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," *Neuroim.*, vol. 41, no. 3, pp. 941–969, 2008.

[2] L. Chaari, F. Forbes, T. Vincent, and P. Ciuciu, "Hemodynamic-informed parcellation of fMRI data in a variational joint detection estimation framework," in *Medical Image Computing and Computer-Assisted Intervention*, N. Ayache et al., Ed. 2012, vol. 7512, pp. 180–188, Springer.

[3] S. Ogawa, T. Lee, A. Kay, and D. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proc. Natl. Acad. Sci.*, vol. 87, no. 24, pp. 9868–9872, 1990.

[4] B. Thirion, G. Flandin, P. Pinel, A. Roche, P. Ciuciu, and J.-B. Poline, "Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fMRI datasets," *Hum. Brain Mapp.*

[5] 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, pp. 1059–1074, Apr. 2010.

[6] 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. on Med. Imag.*, vol. 32, no. 5, pp. 821–837, May 2013.

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

[8] S. Badillo, G. Varoquaux, and Ciuciu, "Hemodynamic estimation based on consensus clustering," in *Proc. IEEE Int. Symp. on Biomed. Imag. (ISBI)*, San Francisco, USA, Apr. 2013, pp. 1512–1515.

[9] G. Flandin, F. Kherif, X. Pennec, D. Rivire, N. Ayache, and J.-B. Poline, "A new representation of fMRI data using anatomo-functional constraints," in *Proc. I nt. Conf. on Funct. Mapp. of the Hum. Brain*.

[10] T. Vincent, P. Ciuciu, and B. Thirion, "Sensitivity analysis of parcellation in the joint detection-estimation of brain activity in fMRI," in *Proc. IEEE Int. Symp. on Biomed. Imag. (ISBI)*, Paris, France, May 2008, pp. 568–571.

[11] A.-L. Fouque, P. Ciuciu, and L. Risser, "Multivariate spatial Gaussian mixture modeling for statistical clustering of hemodynamic parameters in functional MRI," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 445–448.

[12] T. Vincent, L. Risser, and P. Ciuciu, "Spatially adaptive mixture modeling for analysis of fMRI time series," *IEEE Trans. Med. Imag.*, vol. 29, pp. 1059–1074, 2010.

[13] R. M. Neal and G. E. Hinton, "A view of the EM algorithm that justifies incremental, sparse and other variants," in *Lear. in Graph. Mod.*, Jordan, Ed., pp. 355–368. 1998.

[14] 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.

[15] D. Maxwell Chickering and D. Heckerman, "Efficient approximations for the marginal likelihood of bayesian networks with hidden variables," *Mach. Lear.*, vol. 29, pp. 181–212, 1996.

[16] H. Attias, "A variational bayesian framework for graphical models," in *Proc. Advances in Neural Info. Process. Syst.* 2000, pp. 209–215, MIT Press.

[17] M. J. Beal, *Variational Algorithms for Approximate Bayesian Inference*, Ph.D. thesis, Gatsby Computational Neuroscience Unit, University College London, May 2003.

[18] P. Pinel, B. Thirion, S. Mériaux, A. Jobert, J. Serres, D. Le Bihan, J.-B. Poline, and S. Dehaene, "Fast reproducible identification and large-scale databasing of individual functional cognitive networks," *BMC Neurosci.*, vol. 8, no. 1, pp. 91, 2007.