OUTLIER DETECTION FOR ROBUST REGION-BASED ESTIMATION OF THE HEMODYNAMIC RESPONSE FUNCTION IN EVENT-RELATED fMRI

Philippe Ciuciu[1,3]*, Jérôme Idier*[2]*, Alexis Roche*[1,3] *and Christophe Pallier*[1,3]

[1]SHFJ/CEA and INSERM U562, 4 Place du Général Leclerc, 91406 Orsay, France
[2]IRCCyN (CNRS), 1 rue de la Noë, BP 92101 44321 Nantes cedex 3, France
[3] IFR 49 , Institut d'Imagerie Neurofonctionnelle, Paris, France
[1] lastname@shfj.cea.fr, [2] Jerome.Idier@irccyn.ec-nantes.fr

ABSTRACT

In functional Magnetic Resonance Imaging (fMRI), the Hemodynamic Response Function (HRF) represents the impulse response of the *neurovascular* system. Its identification is essential for a deeper understanding of the dynamics of cerebral activity. In [1, 2], we developed a voxelwise approach *i.e.,* based on a single time-course. In this paper, we propose an extension to cope with region-based HRF estimation. We introduce a *spatial homogeneous* model that assumes the same HRF shape for a majority of voxels within a given region-of-interest (ROI). A *Least Trimmed Squares* estimator is employed to select those voxels. A Bayesian HRF estimation is then performed with the corresponding time courses. Our approach is tested on real fMRI data to illustrate the gain in robustness achieved with the region-based estimate.

1. INTRODUCTION

Functional MRI (fMRI) is a non-invasive technique allowing for the evolution of brain processes to be dynamically followed in various cognitive or sensori-motor tasks. In the most common fMRI technique, based on the so-called Blood Oxygenated Level Dependent (BOLD) contrast [3], the measure is only indirectly related to neuronal activity through a process that has not been completely understood so far [4]. For this reason, a convenient way to analyze BOLD fMRI data consists of modeling the whole brain as a stationary, linear "black box" system characterized by its impulse response, also called the Hemodynamic Response Function (HRF).

Estimation of the HRF has received considerable attention for the last decade [5] since it can give a deep insight into the underlying dynamics of brain activation. Recently, non-parametric, Bayesian approaches have been developed to define an accurate estimation of a task-dependent response for event related designs [1, 2]. Such methods were based on a single time course regardless of the origin of the signal (a voxel or a region-of-interest (ROI)). To define a region-based HRF estimate, the usual technique consists of summarizing the brain activity over a given ROI by computing its mean fMRI signal. However, some voxels may have a different hemodynamic response within a given ROI. In particular, closeness to large veinules or partial volume effect may have strong influence on this response. Therefore, averaging data from different voxels may lead to a suboptimal estimate. In this paper, we propose an extension to address HRF estimation in a given ROI accounting for the spatial dimension of the data. We introduce a region-based modeling for the HRF: within a given ROI, we assume that a majority of voxels respond with the same HRF shape. To identify this subset, we resort to *Least Trimmed Squares* (LTS) estimation [6]. We therefore reject the time courses of voxels which represent *outliers* in the ROI. We then compute an *unsupervised* Bayesian HRF estimate, based upon the remaining time courses. Following [1], we resort to the Expectation Maximization (EM) algorithm for hyperparameter estimation. Finally, we demonstrate the robustness of our approach on real fMRI data acquired in a speech perception experiment.

2. SPATIAL MODELING IN fMRI

2.1. Standard voxel-based formulation

Let an fMRI experiment be composed of K sessions, each session involving M different stimulus types. Define $\boldsymbol{y}_j^k = (y_{j,t_n}^k)_{1\leq n\leq N_k}^{\mathrm{t}}$ as the BOLD fMRI time course of voxel (*i.e.,* volume element) V_j at (non-necessarily uniformly sampled) times t_n for session k and $\boldsymbol{x}_k^m = (x_{k,t}^m)_{t\geq t_0}$ the corresponding binary time series, composed of the mth stimulus onsets (*i.e.,* arrival times). Following [1, 2], a convolution model is assumed between the stimuli and the data:

$$y_{j,t_n}^k = \sum_{m=1}^{M}\sum_{p=0}^{P} h_{j,p\Delta t}^m\, x_{k,t_n-p\Delta t}^m + \sum_{q=1}^{Q_k} d_{t_n,q}\, l_{q,j}^k + b_{j,t_n}^k\,.$$

Vector $\boldsymbol{h}_j^m = \left(h_{j,p\Delta t}^m\right)_{p=0,..,P}^{\mathrm{t}}$ represents the mth unknown HRF to be estimated in voxel V_j, sampled every Δt. All HRFs have been assumed to be space-varying but constant across sessions so far. Denote $\boldsymbol{x}_{k,t_n}^m = \left(x_{k,t_n-p\Delta t}^m\right)_{0\leq p\leq P}^{\mathrm{t}}$ such that matrix $\boldsymbol{X}_k^m = \left[\boldsymbol{x}_{k,t_1}^m,\ldots,\boldsymbol{x}_{k,t_{N_k}}^m\right]^{\mathrm{t}}$ is the $N_k\times(P+1)$ design matrix, consisting of the lagged stimulus covariates. In matrix $\boldsymbol{D}_k = \left[\boldsymbol{d}_1^k,\ldots,\boldsymbol{d}_{Q_k}^k\right]$, of size $N_k\times Q_k$, are the values at times t_n of an orthonormal basis (*i.e.,* $\boldsymbol{D}_k^{\mathrm{t}}\boldsymbol{D}_k = \boldsymbol{I}_{N_k}$) of Q_k functions $\boldsymbol{d}_q^k = (\boldsymbol{d}_{q,t_n}^k)^{\mathrm{t}}$ that take a potential session-dependent drift and any other nuisance effect into account. Vector $\boldsymbol{l}_j^k = (l_{q,j}^k)_{1\leq q\leq Q_k}^{\mathrm{t}}$ contains the corresponding unknown coefficients. Vector $\boldsymbol{b}_j^k = (b_{j,t_n}^k)^{\mathrm{t}}$ accounts for uncertainty in voxel V_j for session k and is supposed to be $\mathcal{N}(\boldsymbol{0},\sigma_j^2\boldsymbol{I}_{N_k})$-distributed and independent of the HRFs. Variance σ_j^2 is unknown and assumed to be constant across sessions [1]. Denote $\mathbb{N}_\star = \{1,\ldots,\star\}$, in matrix form, this relation reads

$$\begin{aligned}\boldsymbol{y}_j^k &= \sum_{m=1}^{M} \boldsymbol{X}_k^m\boldsymbol{h}_j^m + \boldsymbol{D}_k\boldsymbol{l}_j^k + \boldsymbol{b}_j^k, \quad \forall(k,j)\in\mathbb{N}_K\times\mathbb{N}_J \\ &= \boldsymbol{X}_k\boldsymbol{h}_j + \boldsymbol{D}_k\boldsymbol{l}_j^k + \boldsymbol{b}_j^k \end{aligned} \tag{1}$$

with $\boldsymbol{X}_k = [\boldsymbol{X}_k^1|\ldots|\boldsymbol{X}_k^M]$ and $\boldsymbol{h}_j = \left[(\boldsymbol{h}_j^1)^{\mathrm{t}},\ldots,(\boldsymbol{h}_j^M)^{\mathrm{t}}\right]^{\mathrm{t}}$. As hypothetized by (1), voxel-based approaches [1, 2] have addressed HRF identification using several time courses acquired in the same voxel V_j during successive runs. However, when a functionnally homogeneous ROI is analyzed, it has been observed that the HRF shape is similar for voxels close from each other [7]. We therefore introduce a generalization of (1) that takes the spatial dimension of fMRI data into account.

2.2. Region-based models

Let $\mathcal{R} = \{V_1,\ldots,V_J\}$ be a J-dimensional homogeneous ROI, which may correspond to a functional activation cluster detected using SPM99[1]. A straightforward extension of model (1) that accounts for voxel-dependent noise level but assumes a constant HRF shape over the region is given by:

$$\mathbb{y}_j = \mathbb{X}\boldsymbol{h} + \mathbb{D}\mathbb{l}_j + \mathbb{b}_j \tag{2}$$
$$\mathbb{X} = \left[\boldsymbol{X}_1^{\mathrm{t}}|\ldots|\boldsymbol{X}_K^{\mathrm{t}}\right]^{\mathrm{t}},\ \mathbb{D} = \mathrm{diag}\left[\boldsymbol{D}_1,\ldots,\boldsymbol{D}_K\right]$$
$$\mathbb{y}_j = \left((\boldsymbol{y}_j^k)_{k\in\mathbb{N}_K}^{\mathrm{t}}\right)^{\mathrm{t}}, \mathbb{l}_j = \left((\boldsymbol{l}_j^k)_{k\in\mathbb{N}_K}^{\mathrm{t}}\right)^{\mathrm{t}}, \mathbb{b}_j = \left((\boldsymbol{b}_j^k)_{k\in\mathbb{N}_K}^{\mathrm{t}}\right)^{\mathrm{t}}.$$

Model (2) implicitely assumes that the HRF time course $\boldsymbol{h}$ is stable enough across sessions and that vector $\mathbb{b}_j$ is composed of independent and identically distributed (i.i.d.) realizations $\boldsymbol{b}_j^k$ of the same noise process. If the goal was to investigate the variability of the HRF across sessions (due for instance to learning effects), we could consider a session-dependent regional model.

In functional neuroimaging, ROIs are often defined from an anatomical reference in order to identify its functional role. Some voxels of such an ROI may have a space-varying functional activity. To select the subset of voxels within a given ROI that are the most functionnally homogeneous, we identify the time courses that match model (2) at best using an LTS algorithm [6].

3. OUTLIER DETECTION IN A GIVEN ROI

Details about the LTS algorithm are provided for model (2).

3.1. Principle

Let r be the number of outliers, which is assumed to be known *a priori* ($0 \leq r < J/2$) (see section 5 for a discussion about this point). Denote $\mathcal{I} \subset \mathcal{R}$ a subset of voxels of size $I = J - r$. Our goal is to identify the subset that best fits (2), say $\widehat{\mathcal{I}}$, by solving:

$$(\widehat{\boldsymbol{h}}^{\mathrm{LTS}}, \widehat{\mathbf{l}}, \widehat{\mathcal{I}}) = \arg\min_{\boldsymbol{h},\mathbf{l},\mathcal{I}} \left[\sum_{j\in\mathcal{I}} \|\mathbb{y}_j - \mathbb{X}\boldsymbol{h} - \mathbb{D}\mathbb{l}_j\|^2\right]$$

with $\mathbf{l} = \left((\mathbb{l}_j^{\mathrm{t}})_{j\in\mathbb{N}_I}\right)^{\mathrm{t}}$. Since this criterion is quadratic with respect to (w.r.t.) $(\boldsymbol{h},\mathbf{l})$, it can be shown that $\widehat{\mathcal{I}}$ is given by

$$\widehat{\mathcal{I}} = \arg\min_{\mathcal{I}} \left[L''(\mathcal{I}) = \sum_{j\in\mathcal{I}} \|\widetilde{\mathbb{y}}_j\|^2 - \left\|\sum_{j\in\mathcal{I}} \mathbb{z}_j\right\|^2\right] \tag{3}$$

where $\widetilde{\mathbb{y}}_j = \mathbb{P}\mathbb{y}_j$, $\mathbb{P} = \mathrm{diag}\,[\boldsymbol{P}_1,\ldots,\boldsymbol{P}_K]$, $\boldsymbol{P}_k = \boldsymbol{I}_{N_k} - \boldsymbol{D}_k\boldsymbol{D}_k^{\mathrm{t}} = \boldsymbol{P}_k^{\mathrm{t}}$, $\mathbb{z}_j = \mathbb{Q}\widetilde{\mathbb{y}}_j$ and $\mathbb{Q} = (\mathbb{X}^{\mathrm{t}}\mathbb{P}\mathbb{X})^{-1}\mathbb{X}^{\mathrm{t}}\mathbb{P}/\sqrt{I}$. Computing $\widehat{\mathcal{I}}$ is an acknowledged NP-hard problem. It requires time-consuming algorithms, especially for large datasets as in fMRI. Here, we resort to a fast but suboptimal numerical scheme to get a local minimizer.

[1] http://www.fil.ion.ucl.ac.uk

3.2. Single Most Likely Replacement (SMLR) type algorithm

SMLR algorithms originate from sparse spike trains deconvolution in geophysics [8]. An *inlier* subset $\widehat{\mathcal{I}}$ that locally minimizes (3) is identified as follows:

- Set up $\mathcal{I}$ and compute $\mathbb{s}_{\mathcal{I}} = \sum_{j\in\mathcal{I}} \mathbb{z}_j$
- Define $\mathcal{I}_{mp} = \mathcal{I}\backslash\{m\} \cup \{p\}$ for the $m \times p$ possible permutations between $m \in \mathcal{I}$ and $p \in \mathcal{R}\backslash\mathcal{I}$
- Compute the local minimizer $\widehat{\mathcal{I}} = \mathcal{I}_{\widehat{mp}}$ such that $L''(\mathcal{I}_{\widehat{mp}}) = \min_{(m,p)} \left[\delta L''_{mp} = L''(\mathcal{I}_{mp}) - L''(\mathcal{I})\right]$ with $\delta L''_{mp} = \|\widetilde{\mathbb{y}}_p\|^2 - \|\widetilde{\mathbb{y}}_m\|^2 - 2(\mathbb{z}_p - \mathbb{z}_m)^{\mathrm{t}}\mathbb{s}_{\mathcal{I}} - \|\mathbb{z}_p - \mathbb{z}_m\|^2$
- Update $\mathbb{s}_{\mathcal{I}} = \mathbb{s}_{\mathcal{I}_{\widehat{mp}}}$ and $\mathcal{I}$; iterate until $\delta L''_{\widehat{mp}} < 0$.

4. REGION-BASED HRF ESTIMATION

Once $\widehat{\mathcal{I}}$ is known, a region-based HRF estimate may be derived. The likelihood of model (2) reads:

$$p(\mathbf{y}\,|\,\boldsymbol{h},\mathbf{l}\,;\,\boldsymbol{\sigma}) \propto \prod_{j=1}^{I} \sigma_j^{-N} \exp\left(-\|\mathbb{y}_j - \mathbb{X}\boldsymbol{h} - \mathbb{D}\boldsymbol{l}_j\|^2/2\sigma_j^2\right).$$

with $\mathbf{y} = \left(\left(\mathbb{y}_j^{\mathrm{t}}\right)_{j\in\mathbb{N}_I}\right)^{\mathrm{t}}$, $N = \sum_k N_k$. In the following, we show how prior physiological information can be modelled in the Bayesian framework to get a more accurate HRF estimate.

4.1. Prior information

The HRF. Following [1, 2], we assume that: *(i)* each HRF $\boldsymbol{h}^m$ starts and ends at 0, so that only $P-1$ parameters are unknown. *(ii)* A smoothness prior $p(\boldsymbol{h}^m\,;\,\tau_m) = \mathcal{N}(0, \tau_m\boldsymbol{R})$, is chosen to enforce the low time-varying behavior of the hemodynamic response to the mth stimulus type. Covariance matrix $\boldsymbol{R}$ is a measure of second-order smoothness: $\boldsymbol{R} = (\boldsymbol{D}_2^{\mathrm{t}}\boldsymbol{D}_2)^{-1}$, with $\boldsymbol{D}_2$ the truncated second-order finite difference matrix to account for $h_0^m = h_{P\Delta t}^m = 0$. *(iii)* No prior dependence is assumed between HRFs, so that $p(\boldsymbol{h},\boldsymbol{\tau}) = \prod_m p(\boldsymbol{h}^m\,;\,\tau_m)$. Prior variances $\boldsymbol{\tau} = (\tau_m)^{\mathrm{t}}$ are unknwon and estimated from fMRI data.

The nuisance variables. We assume that $\mathbf{l}$ is an i.i.d. random process, independent of $\boldsymbol{h}$, with common probabilit density function (pdf) a Gaussian law $p(\mathbf{l}\,;\,\boldsymbol{\mu}) = \prod_{j,k} p(\boldsymbol{l}_j^k;\mu_k)$ such that $p(\boldsymbol{l}_j^k;\mu_k) \sim \mathcal{N}(\mathbf{0}, \mu_k\boldsymbol{I}_{Q_k})$, with unknown parameters $\mu = (\mu_k)^{\mathrm{t}}$.

4.2. MAP HRF estimator

Let $\boldsymbol{\theta} = \{\boldsymbol{\sigma},\boldsymbol{\tau},\boldsymbol{\mu}\}$, we obtain the joint posterior pdf of $(\boldsymbol{h},\mathbf{l})$ using Bayes rule:

$$p(\boldsymbol{h},\mathbf{l}\,|\,\mathbf{y}\,;\,\boldsymbol{\theta}) \propto p(\mathbf{y}\,|\,\boldsymbol{h}\,;\,\mathbf{l},\boldsymbol{\sigma})p(\boldsymbol{h}\,;\,\boldsymbol{\tau})\,p(\mathbf{l}\,;\,\boldsymbol{\mu}).$$

Since $p(\boldsymbol{h},\mathbf{l}\,|\,\mathbf{y}\,;\,\boldsymbol{\theta}) \sim \mathcal{N}\big(\widehat{\boldsymbol{h}},\widehat{\mathbf{l}}\,;\,\boldsymbol{\Sigma}\big)$, we retain the maximum of this Gaussian pdf as the regionwise HRF estimate:

$$\boldsymbol{\Sigma} \triangleq \mathrm{Cov}\,(\boldsymbol{h},\mathbf{l}\,|\,\mathbf{y}) = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{h}|\mathbf{y}} & (\boldsymbol{R}_{\boldsymbol{h}\mathbb{l}_j|\mathbf{y}})_j \\ (\boldsymbol{R}_{\boldsymbol{h}\mathbb{l}_j|\mathbf{y}})_j & (\boldsymbol{R}_{\mathbb{l}_j\mathbb{l}_{j'}|\mathbf{y}})_{jj'} \end{bmatrix} \tag{4}$$

$$\alpha_{jk} = (1 + \sigma_j^2/\mu_k)^{-1}, \quad \boldsymbol{\Omega}_j^k = \boldsymbol{I}_{N_k} - \alpha_{jk}\boldsymbol{D}_k\boldsymbol{D}_k^{\mathrm{t}},$$
$$\boldsymbol{R}_{\boldsymbol{h}|\mathbf{y}}^{-1} = \sum_{j,k} \boldsymbol{X}_k^{\mathrm{t}}\boldsymbol{\Omega}_j^k\boldsymbol{X}_k/\boldsymbol{\sigma}_j^2 + \boldsymbol{R}_{\boldsymbol{h}}^{-1} \tag{5}$$

$$\boldsymbol{R}_{\boldsymbol{h}l_j^k|\mathrm{y}} = -\alpha_{jk}\boldsymbol{R}_{\boldsymbol{h}|\mathrm{y}}\boldsymbol{X}_k^{\mathrm{t}}\boldsymbol{D}_k, \quad (6)$$

$$\boldsymbol{R}_{l_j^k l_j^k|\mathrm{y}} = \alpha_{jk}\sigma_j^2\boldsymbol{I}_{Q_k} + \alpha_{jk}^2\boldsymbol{D}_k^{\mathrm{t}}\boldsymbol{X}_k\boldsymbol{R}_{\boldsymbol{h}|\mathrm{y}}\boldsymbol{X}_k^{\mathrm{t}}\boldsymbol{D}_k, \quad (7)$$

$$\boldsymbol{R}_{l_j^k l_{j'}^{k'}|\mathrm{y}} = \alpha_{jk}\alpha_{j'k'}\boldsymbol{D}_k^{\mathrm{t}}\boldsymbol{X}_k\boldsymbol{R}_{\boldsymbol{h}|\mathrm{y}}\boldsymbol{X}_{k'}^{\mathrm{t}}\boldsymbol{D}_{k'}, \ (j',k') \neq (j,k) \quad (8)$$

$$\widehat{\boldsymbol{h}} = \boldsymbol{R}_{\boldsymbol{h}|\mathrm{y}}\sum_{j,k}\boldsymbol{X}_k^{\mathrm{t}}\boldsymbol{\Omega}_j^k\boldsymbol{y}_j^k/\sigma_j^2, \quad (9)$$

$$\widehat{\boldsymbol{l}}_j^k = \alpha_{jk}\boldsymbol{D}_k^{\mathrm{t}}(\boldsymbol{y}_j^k - \boldsymbol{X}_k\widehat{\boldsymbol{h}}) \quad (10)$$

where $\boldsymbol{R}_{\boldsymbol{h}} = \mathrm{diag}\,[\tau_1\boldsymbol{R},\ldots,\tau_M\boldsymbol{R}]$ and $(j,j',k,k') \in \mathbb{N}_I^2 \times \mathbb{N}_K^2$.

4.3. Hyperparameter estimation

ML estimator. Since $\widehat{\boldsymbol{h}}$ depends on $\boldsymbol{\theta}$, we rely on Maximum Likelihood (ML) to estimate $\boldsymbol{\theta}$, *i.e.,* we compute $\widehat{\boldsymbol{\theta}}^{\mathrm{ML}}$ by maximization of the likelihood $p(\mathbf{y}\,;\,\boldsymbol{\theta})$:

$$\widehat{\boldsymbol{\theta}}^{\mathrm{ML}} = \arg\min_{\boldsymbol{\theta}}\big[-\ln p(\mathbf{y}\,;\,\boldsymbol{\theta}) = \ln|\boldsymbol{\Xi}| + \mathbf{y}^{\mathrm{t}}\boldsymbol{\Xi}^{-1}\mathbf{y}\big], \quad (11)$$

$$\boldsymbol{\Xi} = (\boldsymbol{I}_N - \boldsymbol{\Upsilon}\boldsymbol{\Sigma}\boldsymbol{\Upsilon}^{\mathrm{t}}\boldsymbol{R}_{\mathrm{b}}^{-1})^{-1}\boldsymbol{R}_{\mathrm{b}},$$

$$\boldsymbol{\Upsilon} = [\mathbf{1}_I \otimes \mathbb{X}|\boldsymbol{I}_I \otimes \mathbb{D}], \quad \boldsymbol{R}_{\mathrm{b}} = \mathrm{diag}\big[\sigma_1^2,\ldots,\sigma_I^2\big] \otimes \boldsymbol{I}_N.$$

where $\otimes$ stands for the Kronecker product. In practice, we only locally maximize the likelihood using an EM algorithm [9].

EM algorithm. The EM algorithm is a general locally convergent method that ensures the increasing of the likelihood function $p(\mathbf{y}\,;\,\boldsymbol{\theta})$ of a parameter vector $\boldsymbol{\theta}$ given the data $\mathbf{y}$ at each iteration. Starting from $\boldsymbol{\theta}^0$, we generate a series of estimates $\boldsymbol{\theta}^k$ by alternating the following two steps:

$$\text{Evaluate (E-step):} \quad \mathcal{Q}(\boldsymbol{\theta},\boldsymbol{\theta}^k) = \mathrm{E}_{\boldsymbol{h},\mathbf{l}}\big[\ln p(\mathbf{y},\boldsymbol{h},\mathbf{l}\,;\,\boldsymbol{\theta})|\mathbf{y}\,;\,\boldsymbol{\theta}^k\big]$$

$$\text{Maximize (M-step):} \quad \boldsymbol{\theta}^{k+1} = \arg\max_{\boldsymbol{\theta}}\mathcal{Q}(\boldsymbol{\theta},\boldsymbol{\theta}^k)$$

From $p(\mathbf{y},\boldsymbol{h},\mathbf{l}\,;\,\boldsymbol{\theta}) = p(\mathbf{y}\,|\,\boldsymbol{h},\mathbf{l}\,;\,\boldsymbol{\sigma})\,p(\boldsymbol{h}\,;\,\boldsymbol{\tau})\,p(\mathbf{l}\,;\,\boldsymbol{\mu})$, we deduce $\mathcal{Q}(\boldsymbol{\theta},\boldsymbol{\theta}') = \mathcal{Q}_{\mathrm{y}}(\boldsymbol{\sigma},\boldsymbol{\theta}') + \mathcal{Q}_{\boldsymbol{h}}(\boldsymbol{\tau},\boldsymbol{\theta}') + \mathcal{Q}_{\mathrm{l}}(\boldsymbol{\mu},\boldsymbol{\theta}')$ with:

$$\mathcal{Q}_{\mathrm{y}}(\boldsymbol{\sigma},\boldsymbol{\theta}') = -\frac{N}{2}\sum_j\ln\sigma_j^2 - \frac{1}{2\sigma_j^2}\mathrm{E}\big[\|\mathrm{y}_j - \mathbb{X}\boldsymbol{h} - \mathbb{D}\mathbb{l}_j\|^2\,|\mathbf{y};\boldsymbol{\theta}'\big]$$

$$\mathcal{Q}_{\boldsymbol{h}}(\boldsymbol{\tau},\boldsymbol{\theta}') = \frac{1-P}{2}\sum_m\ln\tau_m - \frac{M}{2}\ln|\boldsymbol{R}| - \frac{1}{2}\mathrm{E}\big[\boldsymbol{h}^{\mathrm{t}}\boldsymbol{R}_{\boldsymbol{h}}^{-1}\boldsymbol{h}|\mathbf{y};\boldsymbol{\theta}'\big]$$

$$\mathcal{Q}_{\mathrm{l}}(\boldsymbol{\mu},\boldsymbol{\theta}') = \sum_k\Big(-\frac{IQ_k}{2}\ln\mu_k - \frac{1}{2\mu_k}\sum_j\mathrm{E}\Big[(\boldsymbol{l}_j^k)^{\mathrm{t}}\boldsymbol{l}_j^k\,|\,\mathbf{y}\,;\,\boldsymbol{\theta}'\Big]\Big).$$

The M-step decouples and thus amounts to searching for $\widehat{\boldsymbol{\theta}}$ that cancels the partial derivative $\partial\mathcal{Q}(\boldsymbol{\theta},\boldsymbol{\theta}')/\partial\boldsymbol{\theta}$. The optimal parameters are given by

$$\widehat{\sigma}_j^2 = \big(\|\mathrm{y}_j - \mathbb{X}\widehat{\boldsymbol{h}} - \mathbb{D}\widehat{\mathbb{l}}_j\|^2 + \mathrm{tr}\big(\mathbb{C}^{\mathrm{t}}\boldsymbol{\Sigma}_j\mathbb{C}\big)\big)/N, \ j \in \mathbb{N}_I \quad (12)$$

$$\mathbb{C} = [\mathbb{X}\,|\,\mathbb{D}] \quad \text{and} \quad \boldsymbol{\Sigma}_j = \begin{bmatrix}\boldsymbol{R}_{\boldsymbol{h}|\mathrm{y}} & \boldsymbol{R}_{\boldsymbol{h}\mathbb{l}_j|\mathrm{y}}\\ \boldsymbol{R}_{\boldsymbol{h}\mathbb{l}_j|\mathrm{y}}^{\mathrm{t}} & \boldsymbol{R}_{\mathbb{l}_j\mathbb{l}_j|\mathrm{y}}\end{bmatrix},$$

$$\widehat{\tau}_m = \mathrm{tr}\big(\big(\widehat{\boldsymbol{h}}_m\widehat{\boldsymbol{h}}_m^{\mathrm{t}} + \boldsymbol{R}_{\boldsymbol{h}^m|\mathrm{y}}\big)\boldsymbol{R}_{\boldsymbol{h}}^{-1}\big)/(P-1), \ m \in \mathbb{N}_M \quad (13)$$

$$\widehat{\mu}_k = \sum_{j=1}^{I}\big(\|\widehat{\boldsymbol{l}}_j^k\|^2 + \mathrm{tr}\big(\boldsymbol{R}_{l_j^k l_j^k|\mathrm{y}}\big)\big)/IQ_k, \ k \in \mathbb{N}_K. \quad (14)$$

The EM algorithm consists in iteratively computing $\boldsymbol{\Sigma}$, $\widehat{\boldsymbol{h}}$ and $\widehat{\mathbf{l}}$ using (4)-(10), and then updating sucessively $\widehat{\boldsymbol{\sigma}}^2$, $\widehat{\boldsymbol{\tau}}$ and $\widehat{\boldsymbol{\mu}}$ with (12)-(14). The unsupervised HRF estimate is finally obtained when local convergence is achieved.

5. RESULTS ON REAL fMRI DATA

5.1. Description of the paradigm

The method was assessed on real data acquired in a speech discrimination experiment. The experiment consisted of $K = 6$ sessions comprising $N_k = 100$ trials lasting 3.3 seconds each. In each trial, the participant heard two pseudo-words over headphones. His task was to indicate whether he had perceived or not a difference between the two stimuli. There were $M = 3$ types of trials: 'Phonological' ($\boldsymbol{h}^1$), 'Acoustic' ($\boldsymbol{h}^2$), 'Control' ($\boldsymbol{h}^3$). In trials belonging to the 'Control' condition, the two auditory stimuli in the pair were exactly the same. In the 'Phonological' condition, the stimuli differed along a contrast used to distinguish words in the language of the participant (it was *linguistically relevant*: path *vs* bath in English). In the 'Acoustic' condition, the stimuli also differed but the contrast between the stimuli was not linguistically relevant (beat *vs* beet in English).

The stimuli pairs were presented during the silent gaps lasting 2 seconds between two succesive acquisitions (the TR was 3.3 s and the time of acquisition of one volume was 1.3 s). The onsets of events were aligned with the start of the second stimulus in a pair (*i.e.,* at 1.65 s), which felt in between two successive acquisitions. $\Delta t = \mathrm{TR}/2$ is therefore an appropriate choice for the sampling period of the HRF.

5.2. Discussion

ROI definition. The ROI $\mathcal{R}$ was defined from a t-map[2] computed from the (Phono.-Cont.) contrast exclusively masked by the (Phono.-Acous.) t-map. This region consisted of $J = \mathrm{Card}\,[\mathcal{R}] = 15$ voxels that performed a differential treatment when the stimuli differed, regardless of the type of difference. Fig. 1 shows that $\mathcal{R}$ is located in the planum temporale of the left hemisphere.

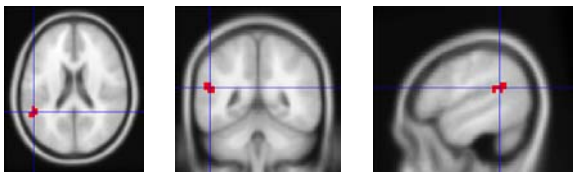

Fig. 1. $\mathcal{R}$ definition. The cross indicates the most significant voxel $V_1 = (-52,-48,20)$ mm in Talairach coordinate system.

Outlier detection. Region $\mathcal{R}$ was functionnally homogeneous since all voxels except V_4 elicited very similar responses for $\widehat{\boldsymbol{h}}_j^1$, $\widehat{\boldsymbol{h}}_j^2$ and $\widehat{\boldsymbol{h}}_j^3$, respectively ($j \in \mathcal{R}$). Four out of the $J =$ voxel-based HRF estimates are presented in Fig. 2. They were obtained from all sessions as in [1]. Not surprisingly, for $j \in \mathcal{R}$, $\widehat{\boldsymbol{h}}_j^1$ and $\widehat{\boldsymbol{h}}_j^2$ are very close to each other, whereas $\widehat{\boldsymbol{h}}_j^3$ slightly differs by its lower magnitude and its shorter peak duration. The error bars also depicted on Fig. 2 correspond to the square root of the diagonal of matrices $\boldsymbol{R}_{\boldsymbol{h}_j^m}$, $m \in \mathbb{N}_3$. As suggested by Fig. 2, the closeness of the HRF shapes justifies the interest of a region-based modeling.

The multisession region-based HRF estimates are plotted in Fig. 3. The results depicted in Fig. 3(a) were obtained without discarding voxels. The outlier detection was then performed by

[2] thresholded at $p = .05$ corrected for multiple comparisons.

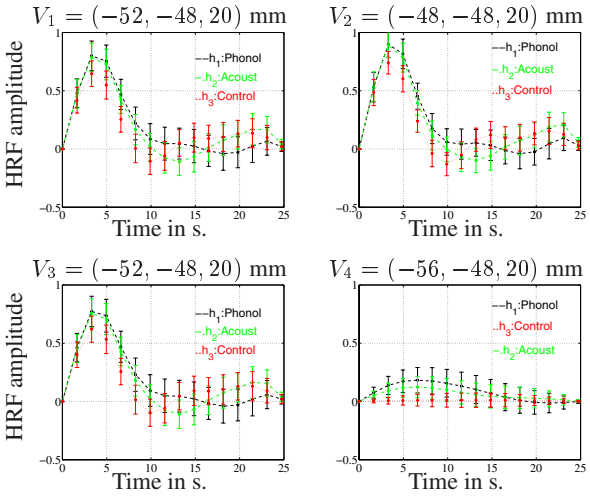

Fig. 2. Voxel-based HRF estimates in four voxels of $\mathcal{R}$ ($j \in \mathbb{N}_4$). $\widehat{\boldsymbol{h}}_j^1$, $\widehat{\boldsymbol{h}}_j^2$ and $\widehat{\boldsymbol{h}}_j^3$ are plotted in (- -), (-.) and ($\cdot$) lines, respectively.

setting $r = 1$ since V_4 responded differently. In a couple of iterations, the minimum of $L''(\mathcal{I})$ was achieved by $\widehat{\mathcal{I}} = \mathcal{R} \backslash V_4$ from which we got the region-based HRF estimate appearing in Fig. 3(b). Compared to Fig. 3(a), no significant improvement is obtained using $\widehat{\mathcal{I}}$ instead of $\mathcal{R}$. This outlier detection procedure should therefore be tested on less homogeneous ROI (*e.g.,* anatomical ROI) to assess whether this step becomes necessary when the outlier number grows. Since r was supposed to be known, we tested different values. We observed smoother HRF shapes in case of over-estimation (*e.g.,* $r = 7$ in the presence of one outlier), as when the signal-to-noise ratio was decreased.

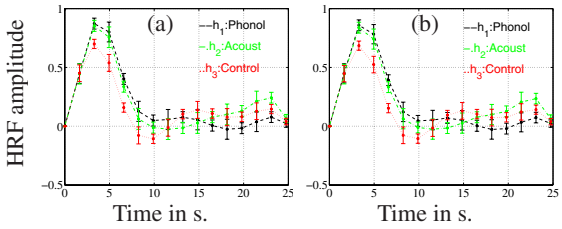

Fig. 3. Region-based HRF estimates $\widehat{\boldsymbol{h}}$. (a): no outlier detection. (b) $r = 10\%$ *i.e.,* one voxel (V_4) is identified as an outlier.

Analysis of the between-session variability. As outlined in Subsection 2.2, our region-based HRF estimation method makes the study of the between-session variability feasible. Fig. 4 (top row) shows session-dependent HRF estimates computed from a single-time course (the mean signal over $\mathcal{R}$). For comparison, the proposed region-based HRF estimates are reported in Fig. 4 (bottom row). These results highlight the dramatic improvement achieved using spatial information, especially for session 1.

6. CONCLUSION

We have proposed a region-based modeling and an unsupervised estimation method of the HRF that accounts for spatial information. For a potential misspecification of this region, we have introduced a preliminary outlier detection step. Future works should investigate whether this step is a keypoint or not. The automatic tuning of the outlier number should also be addressed. Finally, for a given ROI the identification of session-dependent HRFs make the investigatation of putative learning effects possible.

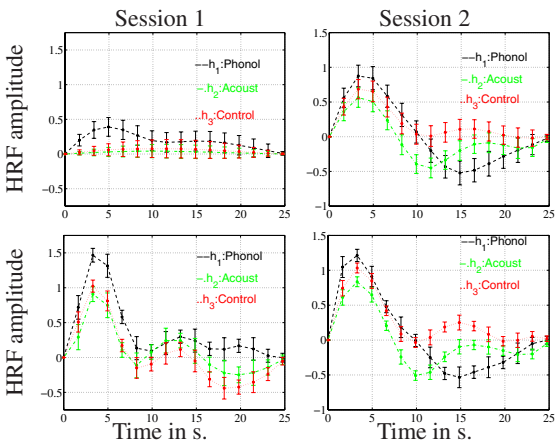

Fig. 4. Session-dependent HRF estimates. Top row: identification from the mean signal; bottom row: region-based estimates.

7. REFERENCES

[1] P. Ciuciu, J.-B. Poline, G. Marrelec, J. Idier, Ch. Pallier, and H. Benali, "Unsupervised robust non-parametric estimation of the hemodynamic response function for any fMRI experiment," *IEEE TMI*, vol. 22, no. 10, pp. 1235–1251, Oct. 2003.

[2] G. Marrelec, P. Ciuciu, M. Pélégrini-Issac, and H. Benali, "Estimation of the hemodynamic response function in event-related functional MRI: Directed acyclic graphs for a general Bayesian inference framework," in *Proc. 18th IPMI*, Ambleside, UK, 2003, LNCS-2732, pp. 635–646, Springer Verlag.

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

[4] A. Aubert and R. Costalat, "A model of the coupling between brain electrical activity, metabolism, and hemodynamics: application to the interpretation of functional neuroimaging," *Neuroimage*, vol. 17, pp. 1162–1181, 2002.

[5] G. Glover, "Deconvolution of impulse response in event-related BOLD fMRI,"*Neuroimage*,vol.9,pp.416–429, 1999.

[6] P. Rousseeuw and A. Leroy, *Robust regression and outlier detection*, Wiley, 1987.

[7] C. Gössl, L. Fahrmeir, and D. Auer, "Bayesian modeling of the hemodynamic response function in BOLD fMRI," *Neuroimage*, vol. 14, pp. 140–148, 2001.

[8] J. Kormylo and J. Mendel, "Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE IT*, vol. 28, pp. 482–488, 1982.

[9] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc. B*, vol. 39, pp. 1–38, 1977.