MODELLING THE NEUROVASCULAR HABITUATION EFFECT ON fMRI TIME SERIES

Philippe Ciuciu[1,2]*, Stéphane Sockeel*[1,2]*, Thomas Vincent*[1,2] *and Jérôme Idier*[3]

[1] NeuroSpin/CEA, F-91191 Gif-sur-Yvette, France
[2] IFR 49, Institut d'Imagerie Neurofonctionnelle, Paris, France
[3] IRCCyN/CNRS, 1 rue de la Noë 44300 Nantes, France
[1] firstname.lastname@cea.fr, [3] Jerome.Idier@irccyn.ec-nantes.fr

ABSTRACT

In this paper, a novel non-stationary model of functional Magnetic Resonance Imaging (fMRI) time series is proposed. It allows us to account for some putative habituation effect arising in event-related fMRI paradigms that involves the so-called repetition-suppression phenomenon [1] and induces decreasing magnitude responses over successive trials. Akin to [2], this model is defined over functionnally homogeneous regions-of-interest (ROIs) and embedded in a joint detection-estimation approach of brain activity. Importantly, its non-stationarity character is embodied in the trial-varying nature of the BOLD response magnitude. Habituation and activation maps are then estimated within the Bayesian framework in a fully unsupervised MCMC procedure. On artificial fMRI datasets, we show that habituation effects can be accurately recovered in activating voxels.

***Index Terms*—** Bayes procedures, Biomedical signal detection, functional MRI, non-stationary model, repetition suppression effect.

1. INTRODUCTION

Since the first report of the Blood Oxygenated Level Dependent (BOLD) effect in human [3], fMRI has represented a powerful tool to non-invasively study the relation between cognitive task and the hemodynamic (BOLD) response. Within-subject analysis in fMRI consists first in *localizing* brain activation in response to a given stimulus type or experimental tasks and second in *estimating* the underlying brain dynamics. In [2,4], a Joint Detection-Estimation (JDE) approach has been proposed to address both issues simultaneously in a region-based analysis, i.e. on a set of prespecified regions of interest (ROI). In a few words, the most salient features of the JDE framework is that this methodology allows us to loosen the classical assumptions usually employed in fMRI data analysis based on Generalized Linear Models (GLM) [5]:*(i)* the HRF shape is not assumed to be fixed but estimated instead [2]; *(ii)*: the model of the BOLD signal varies in space [2]; *(iii)*: spatial Gaussian filtering of the fMRI dataset is no longer necessary [4].

A first characterization of brain dynamics is usually given by an estimate of the Hemodynamic Response Function (HRF), which corresponds to the impulse response of the neurovascular coupling. This characterization assumes that the BOLD signal modelling is linear and time-invariant. Recent works [6, 7] have also demonstrated that transient neurodynamics can be recovered from fMRI time series at the expense of a more complex modelling of the BOLD signal. In the same vein, we extend here the most recent JDE contribution [4] in order to account for the trial-by-trial variability. The proposed region-based BOLD signal model is no longer stationary in time. It can actually account for neurovascular habituation effects either due to the paradigm (repetition suppression effect in case of short inter-stimulus intervals (ISIs)) or to the subject (strategy, tiredness, attentional effects). Our model generalizes the classical convolution BOLD model since it provides the neuroscientist with a trial-varying magnitude of the BOLD response (see Section 2). Nonetheless, instead of considering independent BOLD response magnitudes across trials, our modelling of nonstationarity is parsimonious in order to avoid overfitting: all these parameters are linked to a *mean habituation speed* through a parametric model. Combined to this model, a Bernoulli-uniform prior is introduced in the Bayesian framework within the JDE methodology and associated to mean habituation speeds (MHS) for activating voxels only. We then derive the posterior distribution of interest, from which realizations are drawn using a Gibbs sampler. The posterior mean (PM) estimates of the HRF, the MHS, the BOLD magnitudes, and the corresponding voxel states are computed from the generated samples. In Section 3, our algorithm is validated on realistic artificial fMRI time series.

2. THE JOINT DETECTION-ESTIMATION FRAMEWORK

2.1. Classical BOLD signal modelling

First, the regional forward bilinear model of the BOLD signal introduced in [2] is summarized. It allows us to specify voxel-dependent and stimulus-specific magnitudes of the BOLD response, also called *Neural Response Levels* (NRLs). This time-invariant model characterises each and every functionally homogeneous brain region $\mathcal{R} = (V_j)_{j=1:J}$ by a single HRF shape $\boldsymbol{h}$ and a NRL a_j^m for each voxel V_j and stimulus type m. This means that $\boldsymbol{h}$ is assumed constant within $\mathcal{R}$, while its magnitude a_j^m can vary in space and across experimental conditions as suggested by:

$$\boldsymbol{y}_j = \sum_{m=1}^{M} \underbrace{a_j^m \boldsymbol{X}^m \boldsymbol{h}}_{\boldsymbol{s}_j^m} + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j, \tag{1}$$

where $\boldsymbol{y}_j = (y_{j,n})_{n=1:N}$ denotes the BOLD fMRI time course measured in voxel V_j at times $n = 1{:}N$ (N is the number of scans). M is the total number of experimental conditions in the experiment. Vector $\boldsymbol{h} = (h_{d\Delta t})_{d=0:D}$ represents the unknown HRF shape in $\mathcal{R}$ ($D+1$ is the number of HRF coefficients). Δt is the sampling period of the HRF. Matrix $\boldsymbol{X}^m = (x^m_{n-d\Delta t})_{n=1:N,d=0:D}$ is a $N \times (D+1)$ binary matrix coding for the occurrences of the mth stimulus type. Hence, the activation time course associated to the mth stimulus type in voxel V_j is given by $\boldsymbol{s}_j^m$. Columns of the $N \times Q$ matrix $\boldsymbol{P}$ are orthogonal vectors that span a subspace of low

The autors thank the ANR for funding the OPTIMED Project.

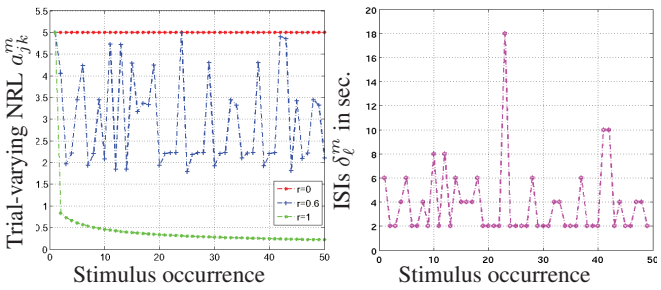

Fig. 1. (g) Illustration of the habituation effect on the trial-specific BOLD magnitudes a_{jk}^m in three different voxels corresponding to a zero (red curve), a medium (blue curve) and a large (green curve) normalized habituation speed r_j^m.

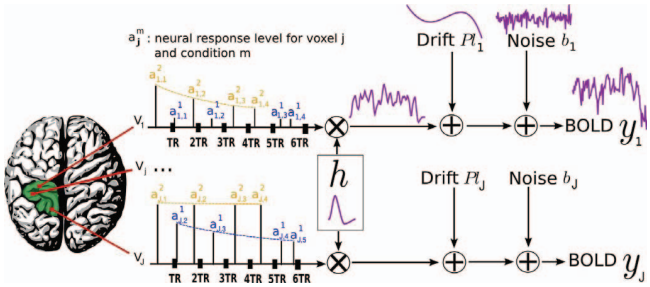

Fig. 2. Non-stationary ROI-based model accounting for voxel-based and stimulus-specific habituation effects.

frequency signals. To each voxel is attached an unknown weighting vector $\boldsymbol{\ell}_j$ to estimate the trend in V_j. $\boldsymbol{b}_j \in \mathbb{R}^N$ is the noise component in voxel V_j and is assumed first-order autoregressive: $\boldsymbol{b}_j \sim \mathcal{N}(\mathbf{0}, \sigma_j^2 \boldsymbol{\Lambda}_j^{-1})$ where $\boldsymbol{\Lambda}_j$ is tridiagonal (see [2] for details) and depends on the AR parameter ρ_j ($|\rho_j| < 1$).

2.2. A parametric approach to habituation modelling

The stationary model (1) assumes that each trial k of a given experimental condition m evokes a BOLD response constant in shape and in magnitude. Recently, it was suggested that this might not always be the case [8]. Here, we propose an extension of (1) that is able to account for a task-related *trial-by-trial* variability of the BOLD response while assuming a constant HRF shape. In contrast to [9], our modelling of this variability source relies on a single normalized mean habituation parameter $r_{jm} \in [0, 1]$ for voxel V_j and condition m that introduces a *parametric* relationship between the trial-dependent NRLs a_{jk}^m. Our parsimonious model aims at mimicking the repetition suppression effect, which means that decreasing NRLs over trials can be obtained in case of short ISIs. However, more flexible patterns can be observed in other paradigms (see Fig. 1 for details). A key feature is to progressively forget the past events. Hence, our parametric habituation model depends on the paradigm as follows:

$$\forall k \geq 2,\ \boldsymbol{a}_{jk}^m = \boldsymbol{\gamma}_{jk}^m \boldsymbol{a}_{j1}^m \text{ with } \boldsymbol{\gamma}_{jk}^m = \left(1 + \sum_{l=1}^{k-1} \boldsymbol{a}_{j\ell}^m \boldsymbol{r}_{jm}^{\delta_\ell^m}\right)^{-1} \quad (2)$$

where $\delta_\ell^m = \tau_{\ell+1}^m - \tau_\ell^m$ and $(\tau_\ell^m)_\ell$ define the successive ISIs and onsets of the mth stimulus. The activation signal $\boldsymbol{s}_j^m$ in Eq. (1) becomes:

$$\boldsymbol{s}_j^m = \sum_{k=1}^{K_m} a_{jk}^m \boldsymbol{X}_k^m \boldsymbol{h} = a_{j1}^m \widetilde{\boldsymbol{X}}_j^m \boldsymbol{h} \text{ with } \widetilde{\boldsymbol{X}}_j^m = \sum_{k=1}^{K_m} \gamma_{jk}^m \boldsymbol{X}_k^m, \quad (3)$$

where $\gamma_{j1}^m = 1$ and $\boldsymbol{X}_k^m$ is the kth trial-specific submatrix of $\boldsymbol{X}^m$. The ensuing region-based model of fMRI time series is depicted in Fig. 2. It clearly indicates that the habituation speed may vary in space and across stimulus types. As expected, when $r_{jm} \to 0$, the proposed extension becomes stationary since $a_{jk}^m \to a_{j1}^m, \forall k \geqslant 2$ (cf red line in Fig. 1(a)). When $r_{jm} \in (0, 1)$, the sequence $(a_{jk}^m)_k$ is non-monotonous and the between-trial variability is strongly influenced by the ISI values (see blue line in Fig. 1(a) and Fig. 1(b)) . Finally, when $r_{jm} \to 1$, the sequence $(a_{jk}^m)_k \to 0$ whatever the ISIs, as shown by the green curve in Fig. 1.

2.3. Likelihood definition

Although the noise structure is correlated in space [10], we neglect such spatial dependency and consider the fMRI time series $\mathbb{y} = (\boldsymbol{y}_j)_{j=1:J}$ independent in space but not identically distributed (iid):

$$p(\mathbb{y} \,|\, \boldsymbol{h}, \mathbb{a}_1, \mathbb{r}, \mathbb{l}, \boldsymbol{\theta}_0) \propto \prod_{j=1}^{J} \det \boldsymbol{\Lambda}_j{}^{1/2} \sigma_j^{-N} \exp\Big(-\frac{\tilde{\boldsymbol{y}}_j^{\mathrm{t}} \boldsymbol{\Lambda}_j \tilde{\boldsymbol{y}}_j}{2\sigma_j^2}\Big) \quad (4)$$

where $\mathbb{l} = (\boldsymbol{\ell}_j)_{j=1:J}$, $\boldsymbol{\theta}_0 = (\rho_j, \sigma_j^2)_{j=1}^J$ and $\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j - \sum_m \boldsymbol{s}_j^m - \boldsymbol{P}\boldsymbol{\ell}_j$. Here, $\mathbb{a}_1$ and $\mathbb{r}$ make reference to the NRLs associated to the first trial and to the MHS, respectively.

2.4. Priors

The Bayesian approaches developed in [2, 4] introduce proper priors on the unknown parameters $(\boldsymbol{h}, \mathbb{a})$ in order to recover a robust estimate of brain activity (localization and activation profile). Akin to [2], the prior density for the hemodynamic filter $\boldsymbol{h}$ remains Gaussian, $\boldsymbol{h} \sim \mathcal{N}(\mathbf{0}, \sigma_{\boldsymbol{h}}^2 \boldsymbol{R})$ with $\boldsymbol{R} = (\boldsymbol{D}_2^{\mathrm{t}} \boldsymbol{D}_2)^{-1}$ and $\boldsymbol{D}_2$ the second-order finite difference matrix penalizing therefore abrupt changes. Moreover, the extreme time points of the HRF can be constrained to zero if necessary [2].

Here, the main originality lies in the definition of the joint prior involving the initial NRLs $\mathbb{a}_1$ and the habituation parameters $\mathbb{r}$. First, we still assume that different stimulus type induce statistically independent NRLs and NHS: $p(\mathbb{a}_1, \mathbb{r} \,|\, \boldsymbol{\theta}_{\boldsymbol{a}}) = \prod_m p(\boldsymbol{a}_1^m, \boldsymbol{r}^m \,|\, \boldsymbol{\theta}^m)$ with $(\mathbb{a}_1, \mathbb{r}) = (\boldsymbol{a}_1^m, \boldsymbol{r}^m)_{m=1}^M$, $(\boldsymbol{a}_1^m, \boldsymbol{r}^m) = (a_{j1}^m, r_j^m)_{j=1}^J$ and $\boldsymbol{\theta}_{\boldsymbol{a}} = (\boldsymbol{\theta}^m)_{m=1}^M$. Vector $\boldsymbol{\theta}^m$ denotes the set of unknown hyperparameters related to the mth stimulus type. Then, following [4], the present paper is based upon spatial mixture models (SMM). This means that if q_j^m is the allocation variable that states whether voxel V_j is activated ($q_j^m = 1$) or not ($q_j^m = 0$) in response to stimulus m, the probability of activation $\Pr(q_j^m = 1)$ depends on the states of q_p^m with $p \sim j$ (the set of neighbors of V_j). Here, spatial correlation is incorporated in the probabilities of activation through a *hidden* Ising random field on $\boldsymbol{q}^m$:

$$\Pr(\boldsymbol{q}^m \,|\, \beta^m) = \frac{\exp\big(\beta^m U(\boldsymbol{q}^m)\big)}{Z(\beta^m)},\ U(\boldsymbol{q}^m) = \sum_{j \sim p} \omega_{jp} I(q_j^m = q_p^m)$$

with $I(A) = 1$ if A is true and $I(A) = 0$ otherwise. The constants ω_{jp} weight the interaction between voxels (V_j, V_p) according to the neighborhood system, which is typically defined in 3D from the region under study $\mathcal{R}$. Here, we only consider the 6-connexity. The

partition function $Z(\cdot)$ is not estimated here since parameter β^m, which controls the amount of spatial regularization, is set by hand: large values of β^m associate higher probabilities to configurations containing clusters of like-valued neighboring binary variables $\boldsymbol{q}^m$. An unsupervised extension is however available [11].

NRLs and NHSs still remain independent conditionally upon $\boldsymbol{q}^m$. This means that $p(\boldsymbol{a}_1^m, \boldsymbol{r}^m \,|\, \boldsymbol{q}^m, \boldsymbol{\theta}^m) = \prod_j p(a_{j1}^m, r_j^m \,|\, q_j^m, \boldsymbol{\theta}^m)$ $\forall m$. Moreover, these parameters are supposed independent of each other: $p(a_{j1}^m, r_j^m \,|\, q_j^m, \boldsymbol{\theta}^m) = p(a_{j1}^m \,|\, q_j^m, \boldsymbol{\theta}^m)\, p(r_j^m \,|\, q_j^m)$. Akin to [4], the NRLs $\boldsymbol{a}_1^m$ follow a two-class Gaussian mixture: $a_{j1}^m \,|\, q_j^m = i \sim f_{i,m} \equiv \mathcal{N}(\mu_{i,m}, v_{i,m})$, with $i = 0, 1$. The mean of the NRLs in non-activating voxels is fixed ($\mu_{0,m} = 0$), while the other parameters $\boldsymbol{\theta}^m = [v_{0,m}, \mu_{1,m}, v_{1,m}]$ remain free and need to be estimated. To get rid of identifiability problems, the MHSs $\boldsymbol{r}^m$ are constrained to zero in non-activating voxels: $\forall m, r_j^m = 0 \,|\, q_j^m = 0$. It seems a priori meaningless to fit habituation effect on noise-only time series. In contrast, for activating voxels, the MHSs $\boldsymbol{r}^m$ are assumed identically and uniformly distributed: $(r_j^m \,|\, q_j^m = 1) \sim \mathcal{U}([0,1])$. The compound prior mixture therefore reads for every stimulus type:

$$p(\boldsymbol{a}_1^m, \boldsymbol{r}^m \,|\, \boldsymbol{\theta}^m) = \sum_{\boldsymbol{q}^m} \Big(\prod_{j=1}^{J} f_{i,m}(a_{j1}^m)\, p_i(r_j^m) \Big) \Pr(\boldsymbol{q}^m \,|\, \beta^m).$$

To complete the Bayesian model, priors are required for all the remaining parameters. Without any prior knowledge, the noise and drift parameters $(\rho_j, \sigma_j^2, \boldsymbol{\ell}_j)_j$ are assumed independent in space and: $\boldsymbol{\ell}_j \sim \mathcal{N}(\boldsymbol{0}, \sigma_\ell^2 \boldsymbol{I}_Q)$, $p(\rho_j, \sigma_j^2) = \sigma_j^{-1} \mathbb{1}_{(-1,1)}(\rho_j)$, to ensure stability of the AR(1) noise process. Non-informative priors are retained for the drift and HRF shape variances: $p(\sigma_{\boldsymbol{h}}^2, \sigma_\ell^2) = (\sigma_{\boldsymbol{h}} \sigma_\ell)^{-1}$. In contrast, conjugate priors are considered for $\boldsymbol{\theta}^m$.

2.5. Bayesian inference

Considering the constructed model and assuming no further prior dependence between parameters, Bayes' rule gives us:

$$\begin{aligned} p(\boldsymbol{h}, \mathbb{a}, \mathbb{l}, \boldsymbol{\Theta} \,|\, \mathbb{y}) &\propto p(\mathbb{y} \,|\, \boldsymbol{h}, \mathbb{a}_1, \mathbb{r}, \mathbb{l}, \boldsymbol{\theta}_0)\, p(\mathbb{a}_1, \mathbb{r} \,|\, \boldsymbol{\theta}_{\mathbb{a}})\; p(\boldsymbol{h} \,|\, \sigma_{\boldsymbol{h}}^2) \times \\ &\quad \times p(\mathbb{l} \,|\, \sigma_{\boldsymbol{\ell}}^2)\, p(\boldsymbol{\Theta}) \qquad (5) \\ &\propto \sigma_{\boldsymbol{h}}^{-D} \sigma_\ell^{-JQ} \prod_{j=1}^{J} \Big(\frac{(1-\rho_j^2)^{1/2}}{\sigma_j^{N+1}}\, \mathbb{1}_{(-1,1)}(\rho_j) \Big) \\ &\quad \exp\Big(-\sum_{j=1}^{J} \Big(\frac{1}{2\sigma_j^2} \tilde{\boldsymbol{y}}_j^{\mathrm{t}} \boldsymbol{\Lambda}_j \tilde{\boldsymbol{y}}_j + \frac{1}{2\sigma_\ell^2} \|\boldsymbol{\ell}_j\|^2 \Big) \Big) \\ &\quad \exp\big(-\frac{\boldsymbol{h}^{\mathrm{t}} \boldsymbol{R}^{-1} \boldsymbol{h}}{2\sigma_{\boldsymbol{h}}^2} \big) \prod_{m=1}^{M} p(\boldsymbol{a}_1^m, \boldsymbol{r}^m \,|\, \boldsymbol{\theta}^m)\, p(\boldsymbol{\theta}^m). \end{aligned}$$

To get samples of the posterior pdf, we use a Gibbs sampler whose target distribution is (5). Since the full conditional posterior $\pi_1^\star(r_j^m) = p(r_j^m \,|\, \boldsymbol{y}_j, q_j^m = 1, \mathit{rest})$ reads:

$$\pi_1^\star(r_j^m) \propto \exp\big(-\|\boldsymbol{y}_j - C_{j\setminus m} - \boldsymbol{s}_j^m\|_{\boldsymbol{\Lambda}_j}^2 / 2\sigma_j^2 \big) \mathcal{U}_{(0,1)}(r_j^m),$$

where $\boldsymbol{s}_j^m$ depends on r_j^m and $C_{j\setminus m} = \boldsymbol{P}\boldsymbol{\ell}_j - \sum_{n \neq m} \boldsymbol{s}_j^n$ is independent of r_j^m, it cannot be sampled directly. Hence, we resort to a *Metropolis-Hastings* step with an uncentered Laplacian density, truncated over [0,1], as proposal: $f(r|r_0) = Z_{\beta,r_0}^{-1} e^{-\beta|r-r_0|} \mathbb{1}_{[0,1]}(r)$, where $r_0 = r_j^{m,(t-1)}$. At iteration t, the MH acceptance ratio is given by $\alpha(r_0 \to r) = \min\big[1, \frac{\pi_1^\star(r)}{\pi_1^\star(r_0)} \frac{Z_{\beta,r_0}}{Z_{\beta,r}}\big]$ which requires to generate the trial-varying NRLs (2) at points r and r_0. Finally, posterior mean (PM) estimates are computed from the samples as follows: $\widehat{x}^{\mathrm{PM}} = (T-I)^{-1} \sum_{t=I+1}^{T} x^{(k)}$, $\forall\, x \in \{\boldsymbol{h}, \mathbb{a}, \mathbb{q}, \boldsymbol{\Theta}\}$ where I stands for the length of the burn-in period. Note also that detection is performed according to the Maximum *A Posteriori* criterion: $(\widehat{q}_j^m)^{\mathrm{MAP}} = \arg\max_i \Pr(q_j^m = i \,|\, \boldsymbol{y}_j)$. The MHS estimate is derived as follows. In every non-activating voxel ($(\widehat{q}_j^m)^{\mathrm{MAP}} = 0$), we impose $\widehat{r}_j^m = 0$. For activating voxels ($(\widehat{q}_j^m)^{\mathrm{MAP}} = 1$), the MHS estimate is computed as the average of samples $r_j^{m,(k)}$ over iterations k that satisfy $q_j^{(k),m} = 1$. The goal is to avoid mixing effects with zero-valued MHS samples in case of label switching.

3. SIMULATION RESULTS

3.1. Synthetic fMRI time series

Artificial fMRI time series were generated according to the additive principle defined in model (1). A random mixed sequence coding two stimus types ($M = 2$) comprising 30 trials each ($K_m = 30, m = 1:2$) was simulated to define the onsets and thus matrices $\boldsymbol{X}^m$. Then, the voxel-based BOLD signals $\boldsymbol{s}_j^m$ were computed from Eqs. (2)-(3) on a 2D 5x5 grid. The true parameters $(\mathbb{a}_1, \mathbb{r}, \mathbb{q})$ involved in the compound prior mixtures are shown in Fig. 3. The labels $\mathbb{q}$ were set manually and not simulated according to the Ising prior. Activating and non-activating voxels appear in red and black, respectively in Fig. 3. Also, the same statistical parameters $\boldsymbol{\theta}^m$ were used to generate the NRLs $\mathbb{a}_1$ $(\mu_{1,m}, v_{1,m}, v_{0,m}) = (5.5, 0.3, 0.4)$. Hence, the same *contrast-to-noise ratio* (CNR) holds for both stimulus types. Following Fig. 2, the trial varying NRLs a_{jk}^m were then convolved with a HRF $\boldsymbol{h}$, whose exact shape appears in Fig. 5(a). Space-varying AR(1) Gaussian noise $\boldsymbol{b}_j$ and low-frequency drift $\boldsymbol{P}\boldsymbol{\ell}_j$[1] were then superimposed to $\sum_m \boldsymbol{s}_j^m$ in every voxel of the grid. The amount of noise was varied in space according to a *signal-to-noise ratio* ranging from -10 to 12 dB. This simulation is thus representative of real fMRI datasets.

3.2. Results

MAP and PM estimates were computed from the outputs of the Gibbs sampler with $I = 500$ and $T = 2000$ iterations and obtained with $\beta^m = \beta = 0.3, \forall m$ (reasonable trade-off between sen-

[1] $\boldsymbol{P}$ is a cosine transform basis and $\boldsymbol{\ell}_j$ is a zero-mean Gaussian vector.

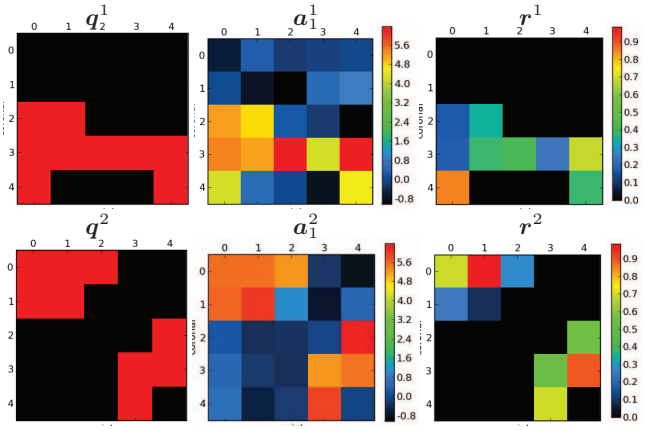

Fig. 3. True parameters (left: labels $\mathbb{q}$, middle: NRLs $\mathbb{a}_1$ and right: MHS $\mathbb{r}$) for simulating fMRI time series according to Eqs. (2)-(3).

sibility and specificity). As shown in Fig. 4, only two false negatives (FNs) appear in $\widehat{\mathbb{q}}^{\text{MAP}}$ in positions $\{(4,0);(0,2)\}$ for $m=1,2$, respectively. No false positive has emerged for these simulation settings while this could occur for lower CNRs. The reason for which two FNs were estimated lies in the values of habituation parameter, where strong MHS values remarkably occur. Regarding the NRLs, the true parameters are accurately estimated by $\widehat{\mathbb{a}}_1{}^{\text{PM}}$ except for the two abovementioned voxels. However, it is worth noting that there is a slight shift in magnitude between $\mathbb{a}$ and its PM estimate ($\widehat{\mathbb{a}}_1{}^{\text{PM}}$ being larger), which is compensated by an inverse trend occuring on the HRF (the peak of $\widehat{\boldsymbol{h}}^{\text{PM}}$ being smaller). This is a direct consequence of the scale ambiguity problem arising in model (1), which is bilinear with respect to $\mathbb{a}_1$ and $\boldsymbol{h}$. Hence, the pair of solutions $(\boldsymbol{h}^\star, \mathbb{a}_1^\star)$ is known up to a multiplicative constant. Finally, note that the habituation parameters $\widehat{\mathbb{r}}^{\text{PM}}$ are well recovered in activating voxels which have been properly detected, while zero-valued habituation speeds are estimated in non-activating voxels. Fig. 6 illustrates on three different voxels the profiles of the trial-varying NRLs $\widehat{a}_j k^m$ for different MHS values. Low habituation speed reported in Fig. 6(a) does not strongly affect the stationarity assumption. In contrast, medium and large MHs values emphasize the decrease of the trial-varying NRLs as pointed out in Fig. 6(b)-(c).

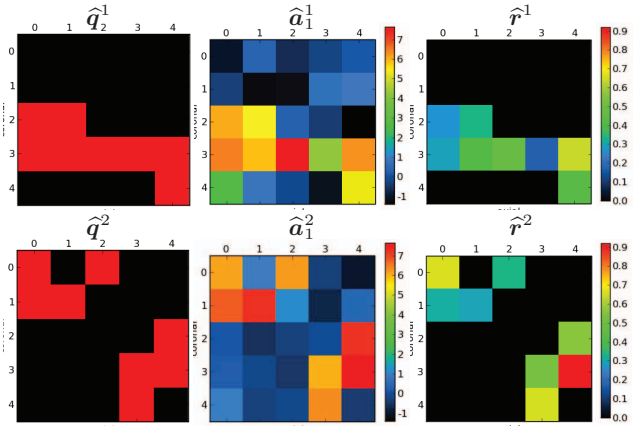

Fig. 4. PM parameter estimates (left: labels $\widehat{\mathbb{q}}^{\text{MAP}}$, middle: NRLs $\widehat{\mathbb{a}}_1{}^{\text{PM}}$ and right: MHS $\widehat{\mathbb{r}}^{\text{PM}}$).

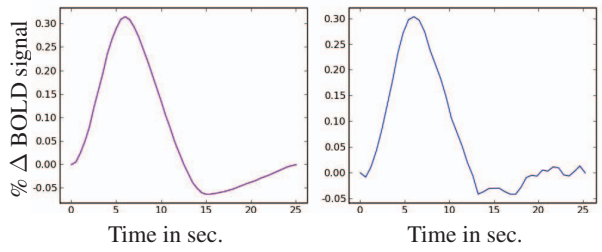

Fig. 5. Left: True HRF shape. Right: HRF estimate $\widehat{\boldsymbol{h}}^{\text{PM}}$.

4. CONCLUSION

The proposed contribution has achieved the goal of handling trial-varying magnitudes of the BOLD response while accounting for habitutation effects in a parsimonious manner at the expense of a higher computational cost. Validation has to be conducted on real fMRI datasets, for instance acquired during a comprehension study involving sentence repetition. The goal is to emphasize the habituation gradient existing along the left superior temporal sulcus in right-handed subjects.

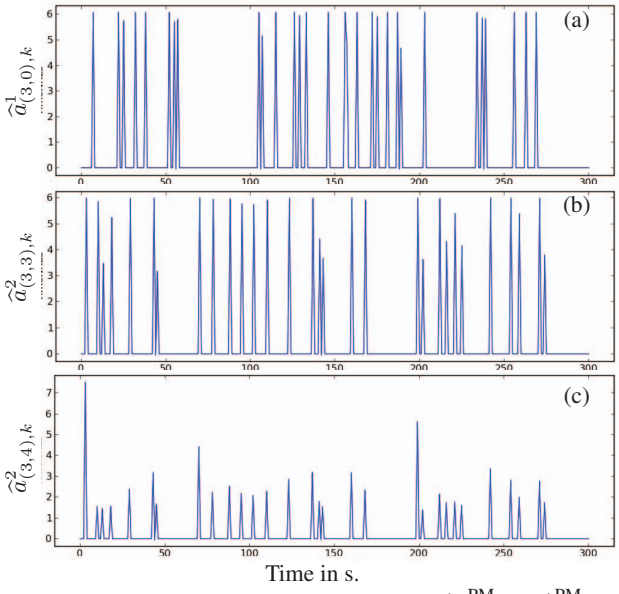

Fig. 6. Trial-varying NRLs reconstructed from $\widehat{\mathbb{a}}_1{}^{\text{PM}}$ and $\widehat{\mathbb{r}}^{\text{PM}}$: **(a)**: $m=1$, $V_j=(3,0)$ and $\widehat{r}^1_{(3,0)}=0.29$. **(b)**: $m=2$, $V_j=(3,3)$ and $\widehat{r}^2_{(3,3)}=0.52$. **(c)**: $m=2$, $V_j=(3,4)$ and $\widehat{r}^2_{(3,4)}=0.92$.

5. REFERENCES

[1] L. Naccache and S. Dehaene, "Unconscious semantic priming extends to novel unseen stimuli," *Cognition*, vol. 80, pp. 215-229, 2001.

[2] S. Makni, J. Idier, T. Vincent, B. Thirion, Ghislaine 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.

[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] T. Vincent, P. Ciuciu, and J. Idier, "Spatial mixture modelling for the joint detection-estimation of brain activity in fMRI," in *32th Proc. IEEE ICASSP*, Hawaii, 2007, vol. I, pp. 325-328.

[5] J. Ashburner, K.J. Friston, and W. Penny, Eds., *Human Brain Function, 2nd Edition*, Academic press, 2004.

[6] W. Penny, Z. Ghahramani, and K. Friston, "Bilinear dynamical systems," *Philos Trans R Soc Lond B Biol Sci*, vol. 360, no. 1457, pp. 983–993, May 2005.

[7] S. Makni, Ch. Beckmann, S. Smith, and M. Woolrich, "Bayesian deconvolution fMRI data using bilinear dynamical systems," *Neuroimage*, vol. 42, no. 4, pp. 1381–1396, 2008.

[8] J. Duann, T. Jung, W. Kuo, TC Yeh, S. Makeig, J. Hsieh and T. Sejnowski, "Single-trial variability in event-related BOLD signals," *Neuroimage*, vol. 15, no. 4, pp. 823–35, Apr. 2002.

[9] S. Donnet, M. Lavielle, and J.-B. Poline, "Are fMRI event-related response constant in time? A model selection answer," *Neuroimage*, pp. 1169–1176, Apr. 2006.

[10] M. Woolrich, M. Jenkinson, J. M. Brady, and S. Smith, "Constrained linear basis set for HRF modelling using variational Bayes," *Neuroimage*, vol. 21, no. 4, pp. 1748–1761, 2004.

[11] T. Vincent, L. Risser, J. Idier, and P. Ciuciu, "Spatially adaptive mixture modelling for analysis of fMRI time series," submitted to IEEE Trans. Med. Imag., Gif-sur-Yvette, France, Sep. 2008.