ADAPTIVE EXPERIMENTAL CONDITION SELECTION IN EVENT-RELATED FMRI

C. Bakhous[1,3]*, F. Forbes*[1]*, T. Vincent*[1,2]*, L. Chaari*[1]*, M. Dojat*[3]*, P. Ciuciu*[2]

[1] INRIA, MISTIS, Grenoble University, LJK, Grenoble, France
[2] CEA/DSV/I[2]BM/Neurospin, LNAO, Gif-Sur-Yvette, France
[3] INSERM, U836, Grenoble University, GIN, Grenoble, France

ABSTRACT

Standard Bayesian analysis of event-related functional Magnetic Resonance Imaging (fMRI) data usually assumes that all delivered stimuli possibly generate a BOLD response everywhere in the brain although activation is likely to be induced by only some of them in specific brain areas. Criteria are not always available to select the relevant conditions or stimulus types (e.g. visual, auditory, etc.) prior to estimation and the unnecessary inclusion of the corresponding events may degrade the results. To face this issue, we propose within a Joint Detection Estimation (JDE) framework, a procedure that automatically selects the conditions according to the brain activity they elicit. It follows an improved activation detection that we illustrate on real data.

***Index Terms*—** Model specification, Stimulus type selection, Joint detection-estimation, Bayesian hierarchical modelling, Functional magnetic resonance imaging

1. INTRODUCTION

Functional MRI is based on the analysis of Blood Oxygen Level Dependent (BOLD) signals that reflect neuronal mass activity in the brain [1]. Within-subject analysis relies on both (i) a detection step to localize which brain regions are activated by a given stimulus type, and on (ii) an estimation step to recover the underlying BOLD signal dynamics through the estimation of the Hemodynamic Response Function (HRF). In [2], a Bayesian detection-estimation approach (JDE) which jointly addresses (i)-(ii) in a brain region (parcel)-based manner, has been proposed for event-related experimental designs. An important challenge in designing and analyzing such event-related fMRI experiments is how to optimize the accuracy with which the activation probabilities can be evaluated and the event-related hemodynamic response to different stimuli estimated. Most approaches have related this issue to the optimization of the inter-stimulus-interval (ISI) as this is one of the only relevant factors genuinely under the experimenter's control [3]. Once the data has been acquired, a crucial issue concerns the design of the activation model: most often, this question is addressed in the General Linear Model (GLM) context by comparing different model structures using Fisher tests between reduced and full models. In the JDE framework, this model definition becomes parcel-specific so as to account for spatial hemodynamic variability. Then, to optimize the activation probabilities, it makes sense to consider the most sparse model by assessing the relevance to include or not each of the experimental conditions in a specific parcel. The activation of interest is likely to be induced by only a subset of these conditions that may depend on the functional segregation of local brain areas and the cognitive task. The relevant stimulus types may fluctuate across regions making this optimization procedure a complex combinatorial task.

In this paper, we address this issue as a model specification problem within the JDE framework. In contrast to model selection approaches (such as in [4]) that require to compare the performance of several models and select the most appropriate one, we propose a single parsimonious procedure that includes the automatic selection of the experimental conditions that best explain brain activity. This is done by introducing for each stimulus type an additional binary variable as a measure of its relevance (in terms of evoked activity). In a regression context, the idea of adding such indicator variables is usually referred to as variable selection (see *e.g.* [5]) and has been used in [6, 7] to assess evoked brain activity. In the JDE framework, this activity detection task is already handled within the model in a more general way. Activated and unactivated voxels are modelled using a two-class Gaussian mixture instead of a Bernoulli-Gaussian prior [6, 7]. Our use of binary variables is then rather oriented toward the selection of stimulus types which has to be done across the whole set of voxelwise regressions.

In Section 2, we detail how the JDE framework allows a straightforward characterization of relevant stimulus types, and therefore how we optimize the degree of model sparsity by selecting the appropriate variables. Section 3 is devoted to the stochastic inference of our adaptive model. In Section 4, the proposed approach is validated on a real fMRI dataset acquired during fast event-related design: we show a better determination of activated brain regions from noisy fMRI time series. A conclusion is drawn in Section 5.

2. PARSIMONIOUS JOINT DETECTION ESTIMATION

A vector is by convention a column vector. The transpose is denoted by $^{\mathrm{t}}$. Unless stated otherwise, subscripts j, m and i are respectively indexes over voxels, stimulus types and mixture components (activation classes). The Gaussian distribution with mean μ and variance $\mathbf{\Sigma}$ is denoted using $\mathcal{N}(\mu, \mathbf{\Sigma})$.

2.1. Missing and observed variables

The parcel-based model of the BOLD signal described in [2] can be recast in a missing data framework. For a given brain parcel γ, the observed data is denoted by $y = \{y_j, j \in \gamma\}$ where y_j is a N-dimensional vector representing the fMRI time course measured at voxel $j \in \gamma$. Additional non observed variables are introduced: 1) The Neural Response Levels (NRLs) $a = \{a^m, m = 1 : M\}$ with $a^m = \left\{a_j^m, j \in \gamma\right\}$ where M is the number of experimental conditions (or stimulus types) and $a_j = \left\{a_j^m, m = 1 : M\right\}$; 2) The HRF function denoted by $h = [h_0, h_{\Delta t},, \ldots, h_{D\Delta t}]^{\mathrm{t}}$ is a $(D+1)$-real valued vector with Δt the sampling period of the HRF; 3) The activation class assignments $q = \{q^m, m = 1 : M\}$ where $q^m = \left\{q_j^m, j \in \gamma\right\}$ represent the *activation classes* with $q_j^m = i$ meaning that voxel j lies in activation class i for the mth experi-

mental condition. Typically the number of classes is 2 for activated ($i = 1$) and unactivated ($i = 0$) voxels, while the case of deactivations has been addressed in [8].

In addition, in our parsimonious context, our goal is to account for the fact that only a subset of the M stimulus types are necessary to explain the evoked BOLD signal in a given parcel. A stimulus type will be identified as irrelevant for the data under consideration if the number of activated voxels for this stimulus type is too small. In this case, we consider that such evoked activity is artifactual and decide that the stimulus type should be discarded from the model definition. To encode such information, we then add another set of M missing binary variables $w = \{w^m, m = 1 : M\}$ where $w^m = 1$ means that the mth stimulus type is relevant while $w^m = 0$ means that it can be discarded. The observed and missing variables are then linked through the following generative model implying additional parameters to be estimated or fixed as specified below.

$$\forall j \in \gamma, \quad y_j = \sum_{m=1}^{M} w^m a_j^m \boldsymbol{X}^m h + \varepsilon_j, \tag{1}$$

where $\boldsymbol{X}^m$ denotes the $N \times (D+1)$ binary matrix that codes the arrival times of the events of type m which are approximated to fit a Δt-sampled grid, ε_j's stand for the noise (σ_j^2) and physiological artifacts (as accounted for by $\boldsymbol{P}$ a low frequency orthogonal $N \times L$ matrix) and are independent and normally distributed, $\varepsilon_j \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_j^{-1})$, with $\boldsymbol{Q}_j = \frac{1}{\sigma_j^2}\left(\boldsymbol{I}_N - \boldsymbol{P}\boldsymbol{P}^t\right)$ ($\boldsymbol{I}_N$ is the $N \times N$ identity matrix). More details can be found in [9].

2.2. Hierarchical model of the complete data distribution

With standard additional assumptions, not detailed here and denoting by θ the whole set of unknown parameters, the joint distribution $p(y, w, a, h, q, \theta)$ can be decomposed as follows:

$$p(y \,|\, w, a, h, \theta)\; p(a \,|\, w, q, \theta)\; p(w \,|\, q, \theta)\; p(h \,|\, \theta)\; p(q|\theta)\; p(\theta).$$

To fully define the model, we now specify each term in turn.

The $p(y \,|\, w, a, h, \theta)$ term. From (1), it comes that:

$$p(y \,|\, w, a, h, \theta) = \prod_{j \in \gamma} p(y_j \,|\, w, a_j, h, \theta)$$

with $(y_j \,|\, w, a_j, h, \theta) \sim \mathcal{N}\big(\sum_{m=1}^{M} w^m a_j^m \boldsymbol{X}^m h, \boldsymbol{Q}_j^{-1}\big)$.

The $p(a \,|\, w, q, \theta)$ term. As usually assumed [2], different types of stimuli induce statistically independent NRLs. The assignment variables q_j^m are then introduced to segregate activated from unactivated voxels. Among voxels, the NRLs are assumed to be independent conditionally on the q_j^m's so that putting together all stimulus types we get: $p(a \,|\, w, q, \theta) = \prod_{m=1}^{M} \prod_{j \in \gamma} p(a_j^m \,|\, w^m, q_j^m, \theta)$, where we further assume that $\big(a_j^m | w^m = 1, q_j^m = i, \theta\big) \sim \mathcal{N}(\mu_i^m, v_i^m)$ for $i \in \{0, 1\}$ and $\big(a_j^m | w^m = 0, q_j^m = i, \theta\big) \sim \mathcal{N}(\mu_0^m, v_0^m)$. The Gaussian parameters $\{\mu_1^m, v_1^m, v_0^m, m = 1 : M\}$ need to be estimated but we set $\mu_0^m = 0$ for all m. The idea is that for a relevant stimulus type ($w^m = 1$), the distribution of a_j^m depends on the activation state q_j^m in voxel j while for an irrelevant stimulus type ($w^m = 0$), q_j^m has no influence on a_j^m, which is distributed around 0 to account for the absence of response to stimulus type m.

The $p(w \,|\, q, \theta)$ term. The binary variables w are independent across stimulus types, $p(w \,|\, q, \theta) = \prod_{m=1}^{M} p(w^m \,|\, q^m, \theta)$ and follow Bernoulli distributions whose probabilities of success are given via a logit link to the number of activated voxels as given by q^m, $P(w^m = 1 \,|\, q^m, \theta) = \mathcal{F}(\sum_{j \in \gamma} q_j^m)$,where $\mathcal{F}$ is the sigmoid function $\mathcal{F}(x; \tau_1, \tau_2) = (1 + \exp(-\tau_1(x - \tau_2)))^{-1}$ with τ_1 controlling the slope of the sigmoid and τ_2 the inflection point that can be seen as a relevance threshold above which the stimulus type will be considered as relevant with a high probability.

The $p(h|\theta)$ term. Akin to [10, 2], we introduce constraints in the HRF prior that favor smooth variations in h (see [10] for details).

The $p(q \,|\, \theta)$ term. As in [2], we assume prior independence between stimulus types regarding the activation class assignments. It follows that $p(q \,|\, \theta) = \prod_{m=1}^{M} p(q^m \,|\, \beta^m)$ where we assume in addition that $p(q^m \,|\, \beta^m)$ is a 2-class Potts model with interaction parameter β^m (see [2] for details).

The $p(\theta)$ term. $\theta = \{\sigma_j^2, \mu_1^m, v_1^m, v_0^m, \tau_1, \tau_2, \beta^m, j \in \gamma, m = 1 : M\}$. Following [2], these parameters will be inferred upon the posterior distribution except for the additional parameters τ_1 and τ_2 involved in the $\mathcal{F}$ logit link which are fixed as indicated in Section 4.

3. MARKOV CHAIN MONTE CARLO ESTIMATION

Our Bayesian model is too complex to be amenable to analytical calculations. Hence, we resort to Gibbs sampling to sample from the posterior distribution $p(w, a, h, q, \theta \,|\, y)$. Compared to [2], this implies adding a sampling block for the binary variables w and discussing their impact on other blocks. Here, of particular interest are the estimates of the posterior probabilities $p(w^m = 1 \,|\, y)$ for the stimulus types and $p(q_j^m = 1 \,|\, y)$ for the voxels class assignments. Both are obtained by averaging after some burn-in period a series of Monte Carlo iterates from the joint posterior distribution. We can then derive respectively a *relevance profile* (a M-dimensional vector of 0 and 1 depending on whether a stimulus type is relevant or not) and an activation map for each stimulus type by thresholding the corresponding probability estimate. Following [6], the threshold is in practice set to 0.872 which approximately corresponds to a p-value of 0.05.

To focus on the relevance variables, within the Gibbs sampler, the w^m iterates are sampled from $p(w^m \,|\, y, w^{\setminus m}, a, q^m, h, \theta)$ (where $w^{\setminus m} = \{w^{m'}, m' \neq m\}$) which is proportional to

$$p(w^m \,|\, q^m, \theta) \prod_{j \in \gamma} p(y_j \,|\, w, a_j, h, \theta) \prod_{j \in \gamma} p(a_j^m \,|\, w^m, q_j^m, \theta)\ .$$

To compute the probability of w^m to be sampled at 0, we set $w^m = 0$ in the above expression that then simplifies: $p(w^m = 0 \,|\, q^m, \theta) = 1 - \mathcal{F}(\sum_{j \in \gamma} q_j^m)$, $p(y_j \,|\, w, a_j, h, \theta)$ does not depend on a_j^m anymore and $p(a_j^m \,|\, w^m, q_j^m, \theta) = \mathcal{N}(a_j^m; 0, v_0^m)$ does not depend on q_j^m. It follows that the probability to generate $w^m = 0$ can be impacted in turn by the number of activated voxels for type m *i.e.* $\sum_{j \in \gamma} q_j^m$ via the first term above, by the NRLs for type m *i.e.* $\{a_j^m, j \in \gamma\}$ via the third term and by $\{y_j, a_j^{m'}, m' \neq m, j \in \gamma, h\}$ via the second term. Each of these groups of variables act independently on different parts of the sampled conditional probability. More specifically, the first term increases when the number of activated voxels decreases but the other terms are not affected by q^m. Similarly, if $a_j^m \to 0$, the third term increases but the others are not impacted. Then, the second term increases when the y_j's are well explained by the model without type m or equivalently when the current noise model parameters can accommodate the absence of stimulus type m.

The generated values of w^m have in turn an effect on the other simulated variables and in particular on the activation class assignments q^m which are more likely to favor inactivity when $w^m = 0$ (this is illustrated in Figs. 1-2). If $w^m = 1$, w^m has no direct effect on the class assignments as both inactivity and activity can arise depending on the observed data.

4. RESULTS ON REAL FMRI DATASETS

We considered real unsmoothed fMRI data recorded during an experiment designed to map auditory, visual and motor brain functions as well as higher cognitive tasks such as number processing and language comprehension. It consists of a single session of $N = 128$ scans lasting $2.4s$ each, yielding 3-D volumes composed of $79\times95\times46$ voxels. The paradigm is a fast event-related design comprising eighty auditory, visual and motor stimuli of ten types (auditory and visual sentences, auditory and visual calculations, left/right auditory and visual clicks, horizontal and vertical checkerboards). The average ISI is of 3.76s with a standard deviation of 1.99s.

To better assess the impact of the relevance variables w, it is necessary to focus on a brain region γ where some of the types above are likely to be irrelevant. To this end, we focused on the left visual area (390 voxels) where it is well known that visual stimuli are more likely to induce activity than auditory ones. We compared the proposed model referred to as the *parsimonious model* to the *complete model* where all the 10 stimulus types are included without introducing w and allowing selection. Posterior mean estimates have been computed over 150000 realizations of the Gibbs sampler after a burn-in period of 50000 iterations. For our first model assessment, we did not investigate the efficiency of the MCMC sampling scheme. Regarding the model specification, we set $\tau_1 = 1$ and $\tau_2 = 273$ which is equivalent to 70% of the studied region. The value of τ_2 impacts the prior probability of a stimulus type to be relevant. For illustrative purposes and to emphasize differences between the two models, τ_2 is set to a relatively high value which *a priori* slightly favors the apparition of irrelevant stimulus types. The exact influence of τ_2 depends on the underlying Potts models for the activation class assignments q and would required more investigations to be specified in more details. Note however, that for smaller τ_2 values, the relevant stimulus types are still detected but possibly with additional ones.

The *parsimonious model* reports as irrelevant the auditory calculation, left/right auditory clic and auditory sentence types with estimated posterior probabilities $p(w^m = 0 \,|\, y)$ of 0.78, 0.94, 0.88 and 0.92. The other 6 visual stimulus types are estimated as relevant with probability 1. This observation is consistent with our prior knowledge and suggests the ability of the *parsimonious model* to correctly select the relevant stimulus types and discard the others. The corresponding activation class assignments (q_j^m's) are then more likely estimated to 0 for irrelevant types (see Fig. 1(b) and Fig. 2(a) where no voxels are above the 0.872 threshold). In contrast, the *complete model* tends to find meaningless activations for all conditions as illustrated in Fig. 1(a) and Fig. 2(a) where activations are found for a number of voxels (301 are above the 0.872 threshold) in the left visual region for auditory calculation, which should not induce evoked activity. Fig. 2 also shows the histograms of the estimated $p(q_j^m = 1 \,|\, y)$ for a relevant stimulus type (vertical checkboard). In this case, the activation maps and the numbers of voxels over the 0.872 threshold are similar for the *parsimonious* (261 voxels) and *complete* (267 voxels) models.

To further demonstrate the gain induced by selecting stimulus types, we focus on the improvement provided by the *parsimonious model* on relevant stimulus types compared to the *complete model* whose estimates can be degraded by overfitting i.e., by modelling the numerous irrelevant stimulus types. Fig. 3 shows normalized contrasts maps of the 6 visual conditions versus the 4 auditory conditions. The *parsimonious model* leads to more sensitive results as activations in the left visual region are better highlighted with the *parsimonious model* than with the *complete model*. A larger number of voxels shows close to maximum contrasts with a lower dispersion in the parsimonious case as also emphasized in Fig. 4 with the histograms of the normalized contrasts for values greater than 2.

Considering the estimation of the Gaussian prior mixture model, as shown in Fig. 5(a), for the auditory stimulus types (suspected as irrelevant for the brain region under consideration), the *complete model* leads to very overlapped estimated components with a higher variance for the unactivated class. In addition, for relevant visual stimulus types (*e.g.* Fig. 5(c)), the *complete model* shows more uncertainty with higher variances. In contrast, the *parsimonious model* yields consistent and better separated estimates of the Gaussian components for all stimulus types (Fig. 5(b)-(d)). Note that for visual stimuli, the two models give similar mean values.

Finally, Fig. 6 shows the estimated HRF shapes for both models in the visual area under study. The proximity of the main HRF features (peak value, time-to-peak and time-to-undershoot) suggests that the modifications introduced in our *parsimonious model* act more on the spatial features of the activation maps than on the recovery of the HRF shape.

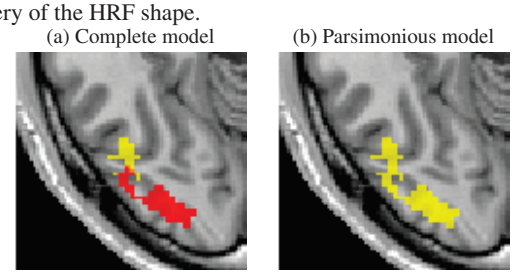


Fig. 1. Maps of activation class assignments for irrelevant auditory calculation stimuli within the considered left occipital region of interest superimposed to the anatomical image: the *complete model* (a) misleadingly shows a lot of activated voxels (in red) while the *parsimonious model* (b) shows unactivated ones (in yellow). **Neurological convention: left is left**.

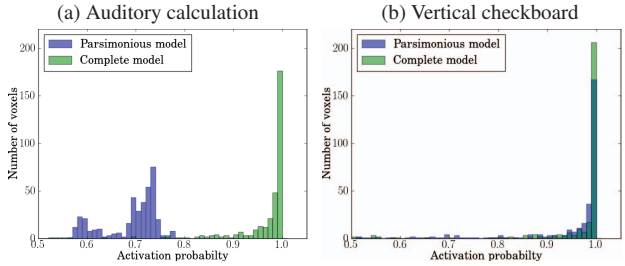


Fig. 2. Histograms of the estimated posterior probabilities of activation $p(q_j^m = 1 \,|\, y)$ for an irrelevant (a) and a relevant (b) stimulus type.

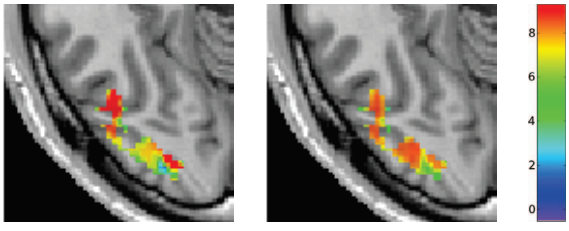


Fig. 3. Normalized contrast maps of the 6 visual conditions vs the 4 auditory ones in the left occipital region superimposed on the anatomical image: (a) *complete model* and (b) *parsimonious model*. More voxels show close to maximum contrasts in the parsimonious case. **Neurological convention: left is left**.

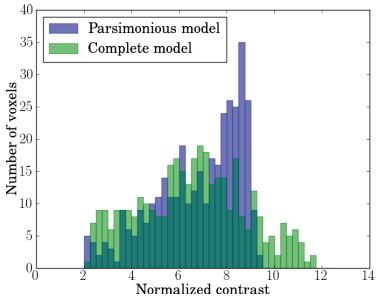


Fig. 4. Superimposed normalized contrast histograms for contrasts greater than 2 for the *complete* and *parsimonious* models (the darkest color shows the overlap). Higher contrasts are observed for the *complete model* but the majority of the contrasts show smaller values than in the parsimonious case.

5. CONCLUSION AND FUTURE WORK

We proposed to go beyond the standard event-related fMRI data analysis which assumes that all delivered stimuli induce a BOLD response and models them as effects of interest in the GLM context. Using a Bayesian hierarchical approach we further explored the variable selection principle, used previously to detect evoked brain activity [6], as a tool to perform relevant condition selection. Experiments on real data suggested the ability of our model to accurately select and exploit the most relevant stimulus types. Our *parsimonious model* improves the statistical significance of the detected voxels and limits false positive detection compared to the *complete model* while preserving the HRF estimation. Inference was performed using an MCMC procedure derived as a simple Gibbs sampler whose mixing and convergence properties have not been thoroughly investigated in this paper. The Bayesian variable selection literature is rich with a number of tools to improve the MCMC estimation part. Our goal is then to study in more detail the properties of the possible chains so as to design one that optimally fits our stimulus type selection goal. Faster alternatives to simulation intensive procedures are also of interest such as variational Bayes implementations that have shown good performance in particular in the JDE framework [11]. Another potentially important feature of our framework is the automatic setting of parameter τ_2. Its precise value does not seem to be so crucial but assessing the extent of its impact would deserve more careful investigations. Eventually, further real data analysis would be necessary for an extended study with a particular emphasis on the group-level impact of parcel-wise adaptive definition of parsimonious models [12].

6. REFERENCES

[1] N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann, "Neurophysiological investigation of the basis of the fMRI signal," *Nature*, vol. 412, no. 6843, pp. 150–157, Jul. 2001.

[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] T.T. Liu, "The development of event-related fMRI designs," *Neuroimage*, 2011, In Press, avaible online.

[4] S. Donnet, M. Lavielle, P. Ciuciu, and J.-B. Poline, "Selection of temporal models for event-related fMRI," in *2th Proc. IEEE ISBI*, Arlington, VA, Apr. 2004, pp. 992–995.

[5] R. B. O. Hara and M. J. Sillanpaa, "A review of Bayesian variable selection methods: What, how and which," *Bayesian Analysis*, vol. 4, no. 1, pp. 85–118, 2009.

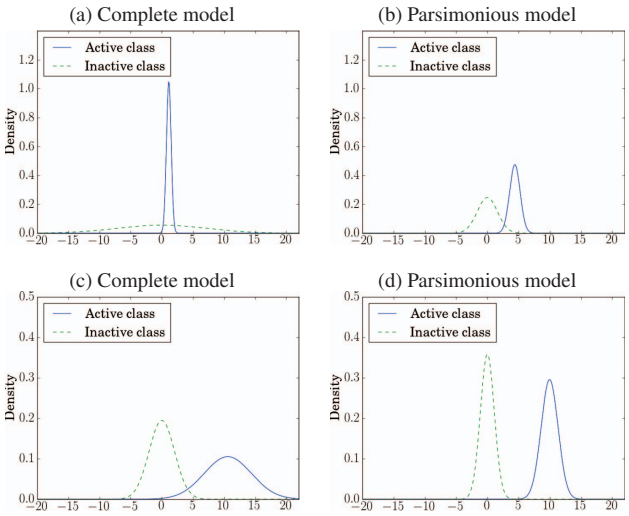


Fig. 5. Estimated prior Gaussian mixture model components for two stimulus types, Auditory calculation (first row) and Vertical checkboard (second row) and both models.

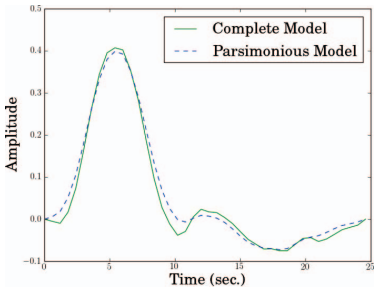


Fig. 6. Estimated HRFs in the considered visual region for the *complete* and *parsimonious* models. The two estimates show similar features.

[6] M. Smith, B. Putz, D. Auer, and L. Fahrmeir, "Assessing brain activity through spatial Bayesian variable selection," *Neuroimage*, vol. 20, pp. 802–815, 2003.

[7] M. Smith and L. Fahrmeir, "Spatial Bayesian variable selection with application to functional Magnetic Resonance Imaging," *J. Stat. Amer. Assoc.*, vol. 102, no. 478, pp. 417–431, 2007.

[8] L. Risser, T. Vincent, F. Forbes, J. Idier, and P. Ciuciu, "Min-max extrapolation scheme for fast estimation of 3D Potts field partition functions. application to the joint detection-estimation of brain activity in fMRI.," *Journal of Signal Processing Systems*, vol. 65, no. 3, pp. 325–338, 2011.

[9] S. Makni, P. Ciuciu, J. Idier, and J.-B. Poline, "Joint detection-estimation of brain activity in functional MRI: a multichannel deconvolution solution," *IEEE Tran. Sig. Process.*, vol. 53, no. 9, pp. 3488–3502, Sep. 2005.

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

[11] L. Chaari, F. Forbes, T. Vincent, M. Dojat, and P. Ciuciu, "Variational solution to the joint detection estimation of brain activity in fMRI," in *14thProc. MICCAI'11*, Toronto, Canada, Sep. 2011.

[12] S. Badillo, T. Vincent, and P. Ciuciu, "Impact of the joint detection-estimation approach on random effects group studies in fMRI," in *7th International Symposium on Biomedical Imaging*, Chicago, IL, Apr. 2011, pp. 376–380.