# JOINT DETECTION-ESTIMATION OF BRAIN ACTIVITY IN FMRI USING AN AUTOREGRESSIVE NOISE MODEL

*Salima Makni,*[1,3] *Philippe Ciuciu,*[1,3] *Jérôme Idier,*[2] *and Jean-Baptiste Poline*[1,3]

[1]Service Hospitalier Frédéric Joliot (CEA) 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] name@shfj.cea.fr, [2] Jerome.Idier@irccyn.ec-nantes.fr

## ABSTRACT

Different approaches have been considered so far to cope with the temporal correlation of fMRI data for brain activity detection. However, it has been reported that modeling this serial correlation has little influence on the estimate of the hemodynamic response function (HRF). In this paper, we examine this issue when performing a joint detection-estimation of brain activity in a given homogeneous region of interest (ROI). Following [1], we adopt a space-varying AR(1) temporal noise model and assess its influence, on both the estimation of the HRF and the detection of brain activity, using synthetic and real fMRI data. We show that this model yields a significant gain in detection specificity (lower false positive rate).

## 1. INTRODUCTION

Within-subject analysis in fMRI consists generally in answering two questions: first, which parts of the brain are activated by a given stimulus and second, what is the temporal dynamics of the brain response during activations. These two points are often addressed separately, while they clearly depend on each other. In [2], we have developed a Bayesian detection-estimation approach to perform these two tasks simultaneously in a region-based analysis.
It has been shown that the noise component of the fMRI time series is correlated in time [3]. Nonetheless, the authors of [4] have outlined that the noise correlation structure has little influence on the HRF estimation. Therefore in [2], we have considered a spatially varying Gaussian *white noise* (SVGWN) model.

In this work, we address the question of the impact of the noise correlation on our procedure and choose the most commonly adopted model [3, 5]. In fact, we consider a spatially varying *first-order autoregressive* noise (SVAR(1)N) model for the errors, and assess its influence on both *estimation and detection* of brain activity. The detection of brain activity is achieved using a mixture of two Gaussian distributions as a prior model on the "neural" response levels (NRLs), while the HRF is constrained to be smooth in the time domain with a Gaussian prior. All parameters of interest as well as hyperparameters are estimated from the joint posterior distribution using Gibbs sampling and posterior mean estimates (PMEs). We compare the performances of the SVGWN and SVAR(1)N models, both on synthetic and real fMRI datasets. This comparison demonstrates the necessity of modeling the temporal correlation of fMRI time series in the context of joint estimation-detection of brain activity. In fact, we first observed a strong influence in the detection results and second, unlike [4], we also noted some impact on the HRF shape estimation.

## 2. REGIONAL MODELING OF FMRI DATA

### 2.1. Problem formulation

Let us define $\boldsymbol{y}_j = (y_{j,t_n})_{n=1:N}$ as the fMRI time course measured in voxel $V_j$ at time $t_n$. Here, we consider that a functionnally homogeneous ROI $\mathcal{R} = (V_j)_{j=1:J}$ is characterized as a first approximation by a single HRF shape $\boldsymbol{h} = (h_{d\Delta t})_{d=0:D}$ and second by a task and voxel dependent magnitude adjustment of the HRF (called the NRL) described by parameter $a_j^m$ for voxel $V_j$ and condition $m$. Then in $\mathcal{R}$ the model reads:

$$\boldsymbol{y}_j = \sum_{m=1}^{M} a_j^m \boldsymbol{X}^m \boldsymbol{h} + \boldsymbol{P} l_j + \boldsymbol{b}_j, \quad \forall\, j = 1:J, \qquad (1)$$

where $\boldsymbol{X}^m = (x^m_{t_n - d\Delta t})_{n=1:N}^{d=0:D}$ is a binary matrix corresponding to the arrival times for the $m$th condition. Note that $\boldsymbol{P} l_j$ models the trend. In [2], for simplicity reason, we have just considered a SVGWN model. Here, a more sophisticated modeling is introduced using an autoregressive process to account for the serial correlation of fMRI time series. Following [3, 5], $\boldsymbol{b}_j$ is a SVAR(1)N process *i.e.,* $b_{j,t_n} = \rho_j\, b_{j,t_n-1} + \varepsilon_{j,t_n}$, $\forall j$, $t$, where $\boldsymbol{\varepsilon}_j \sim \mathcal{N}(0, \sigma^2_{\varepsilon_j})$.

### 2.2. Likelihood

Let us denote $\theta_{0,j} = (\rho_j, \sigma^2_{\varepsilon_j})$ and $\boldsymbol{\theta}_0 = (\theta_{0,j})_{j=1:J}$. Since we assume the fMRI time series $\boldsymbol{y}$ to be statistically independent and identically distributed in space, the likelihood function reads:

$$p(\boldsymbol{y} \,|\, \boldsymbol{h}, \boldsymbol{a}, \boldsymbol{\ell}, \boldsymbol{\theta}_0) \prod_j |\boldsymbol{\Lambda}_j|^{1/2}\, \sigma_{\varepsilon_j}^{-N} \exp\Big(-\frac{\tilde{\boldsymbol{y}}_j^{\mathrm{t}} \boldsymbol{\Lambda}_j \tilde{\boldsymbol{y}}_j}{2\sigma^2_{\varepsilon_j}}\Big),$$

where $\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j - \sum_m a_j^m \boldsymbol{X}^m \boldsymbol{h} - \boldsymbol{P}\boldsymbol{\ell}_j$, and $\sigma_{\varepsilon_j}^{-2}\boldsymbol{\Lambda}_j$ denotes the inverse of the autocorrelation matrix of $\boldsymbol{b}_j$. According to [6], we have $|\boldsymbol{\Lambda}_j| = 1 - \rho_j^2$ and:
$(\boldsymbol{\Lambda}_j)_{1,1} = (\boldsymbol{\Lambda}_j)_{N,N} = 1$, $(\boldsymbol{\Lambda}_j)_{l,l} = 1 + \rho_j^2$, $\forall\, l = 2:N-1$,
$(\boldsymbol{\Lambda}_j)_{l+1,l} = (\boldsymbol{\Lambda}_j)_{l,l+1} = -\rho_j$, $\forall\, l = 1:N-1$,
$(\boldsymbol{\Lambda}_j)_{.,.} = 0$ otherwise.

## 3. THE DETECTION-ESTIMATION PROBLEM

Assuming that the given ROI $\mathcal{R}$ has homogeneous vasculature properties, we propose to estimate a single HRF shape $\boldsymbol{h}$, the corresponding voxel and condition dependent gain factors (NRLs) and simultaneously classify the voxels in $\mathcal{R}$ as either activated or not using the posterior distributions of the NRLs. Here, we propose an extension of [2] that accomodates the temporal correlation in fMRI time series and we study the gain brought by this novel joint detection-estimation approach. In what follows, we first present the prior information, derive the joint posterior distribution, and summarize our approach.

### 3.1. Prior information

**The HRF.** According to [7], the HRF can be characterized as a causal slow-varying function which returns to its baseline after about 25 sec. These assumptions lead to select a Gaussian prior on $\boldsymbol{h}\sim\mathcal{N}(\boldsymbol{0},\|\partial^2\boldsymbol{h}\|^2/2\sigma_{\boldsymbol{h}}^2)$, where $(\partial^2\boldsymbol{h})_{d\Delta t}\approx\frac{1}{(\Delta t)^2}(h_{(d+1)\Delta t}-2h_{d\Delta t}+h_{(d-1)\Delta t}),\ \forall d=1:D-1$.

**The NRLs.** We assume that different types of stimuli induce statistically independent neural responses: $p(\boldsymbol{a};\boldsymbol{\theta_a})=\prod p(\boldsymbol{a}^m;\boldsymbol{\theta}_m)$ with $\boldsymbol{a}=(\boldsymbol{a}^m)_{m=1:M}$, $\boldsymbol{a}^m=(a_j^m)_{j=1:J}$ and $\boldsymbol{\theta_a}=(\boldsymbol{\theta}_m)_{m=1:M}$, where vector $\boldsymbol{\theta}_m$ denotes the set of unknown hyperparameters related to the prior probability density function (pdf) of the NRLs for the $m$th stimulus type. Since only a few voxels of the ROI may be activated by a given condition, we introduce couples of random variables $z_j^m=(q_j^m,a_j^m)_{j=1:J}^{m=1:M}$ where $q_j^m$ is a binary random variable and indicates whether voxel $V_j$ is activated or not by condition $m$ (respectively $q_j^m$=1 or $q_j^m$=0). Conditionally to $q_j^m$, $a_j^m$ is modelled as a Gaussian random variable that represents the NRL for voxel $V_j$ and condition $m$. We therefore introduce a two-class Gaussian mixture prior distribution:
$p(a_j^m\,|\,\boldsymbol{\theta}_m)=\sum_{i=0,1}\Pr(q_j^m=i\,|\,\lambda_{i,m})\mathcal{N}(\mu_{i,m},\sigma_{i,m}^2)$,
with $\Pr(q_j^m=1)=\lambda_{1,m}$, $\Pr(q_j^m=0)=\lambda_{0,m}=1-\lambda_{1,m}$ and $\boldsymbol{\theta}_m=\big[\lambda_{1,m},\mu_{1,m},\sigma_{1,m}^2,\sigma_{0,m}^2\big]$. We set $\mu_{0,m}=0$ (as the mean of the NRLs for unactivated voxels).

**The low-frequency drift.** Vector $\boldsymbol{\ell}=(\ell_j)_{j=1:J}$ defines the unknown parameters of the orthonormal basis function $\boldsymbol{P}$. We assume that $\boldsymbol{\ell}$ is a random process independent of $\boldsymbol{h}$ such that $p(\boldsymbol{\ell};\eta^2)=\prod_j\mathcal{N}(\boldsymbol{0},\eta^{-2}\boldsymbol{I}_Q)$.

**The hyperparameters.** The set of hyperparameters is denoted $\boldsymbol{\Theta}=\big[\boldsymbol{\epsilon}^2,\boldsymbol{\rho},\sigma_{\boldsymbol{h}}^2,\eta,\boldsymbol{\theta_a}\big]$. For these parameters, we resort to the following uninformative priors:
$p(\sigma_{\boldsymbol{h}}^2,\eta^2,\epsilon_j^2)=(\sigma_{\boldsymbol{h}}\eta\epsilon_j)^{-1}$, $p(\rho_j)=\mathbb{1}_{(-1,1)}(\rho_j)$ and $p(\boldsymbol{\theta}_m)\propto\sigma_{1,m}^{-1}\sigma_{0,m}^{-1}\lambda_{1,m}^{-1/2}\lambda_{0,m}^{-1/2}$.

### 3.2. The joint posterior distribution

Considering the constructed model and assuming no further prior dependence between parameters, formal application of the chain rule yields:

$$p(\boldsymbol{h},\boldsymbol{a},\boldsymbol{\ell},\boldsymbol{\Theta}|\boldsymbol{y})\propto p(\boldsymbol{y}|\boldsymbol{h},\boldsymbol{a},\boldsymbol{\ell},\boldsymbol{\theta}_0)p(\boldsymbol{a}|\boldsymbol{\theta_a})p(\boldsymbol{h}|\sigma_{\boldsymbol{h}}^2)p(\boldsymbol{\ell}|\eta^2)p(\boldsymbol{\Theta})$$

$$p(\boldsymbol{h},\boldsymbol{a},\boldsymbol{\ell},\boldsymbol{\Theta}|\boldsymbol{y})\propto\sigma_{\boldsymbol{h}}^{-D}\eta^{-JQ}\prod_j\Big(\frac{\big(1-\rho_j^2\big)^{1/2}}{\sigma_{\tilde{\varepsilon}_j}^{N+1}}\,\mathbb{1}_{(-1,1)}(\rho_j)\Big)$$
$$\exp\Big(-\frac{\boldsymbol{h}^{\mathrm{t}}\boldsymbol{R}^{-1}\boldsymbol{h}}{2\sigma_{\boldsymbol{h}}^2}-\sum_j\Big(\frac{1}{2\sigma_{\tilde{\varepsilon}_j}^2}\tilde{\boldsymbol{y}}_j^{\mathrm{t}}\boldsymbol{\Lambda}_j\tilde{\boldsymbol{y}}_j+\frac{1}{2\eta^2}\big\|\boldsymbol{\ell}_j\big\|^2\Big)\Big)$$
$$\prod_m\Big(p(\boldsymbol{\theta}_m)\prod_j\sum_{i=0}^{1}\frac{\lambda_{i,m}^{1/2}\lambda_{1-i,m}^{-1/2}}{\sigma_{i,m}^2\sigma_{1-i,m}}\exp\Big(-\frac{(a_j^m-\mu_{i,m})^2}{2\sigma_{i,m}^2}\Big)\Big),\tag{2}$$

### 3.3. The Gibbs sampling algorithm

To draw realizations of the posterior pdf (2), we use a Gibbs sampler which consists in building a Markov chain deriving from the target distribution (2), by sequentially generating random samples from the full conditional pdfs of all the unknown parameters and hyperparameters. Finally, PMEs are computed from these realizations after discarding the first part of the Markov chain (a burn-in period = $I$).

**The HRF and its scale.** $\boldsymbol{h}$ is $\mathcal{N}(\boldsymbol{\mu_h},\boldsymbol{\Sigma_h})$-distributed. $\boldsymbol{\Sigma}_{\boldsymbol{h}}^{-1}=\sigma_{\boldsymbol{h}}^{-2}\boldsymbol{R}^{-1}+\sum_j\sigma_{\varepsilon_j}^{-2}\boldsymbol{S}_j^{\mathrm{t}}\boldsymbol{\Lambda}_j\boldsymbol{S}_j$ (with $\boldsymbol{S}_j=\sum_m a_j^m\boldsymbol{X}^m$) and $\boldsymbol{\mu_h}=\boldsymbol{\Sigma_h}\sum_j\sigma_{\varepsilon_j}^{-2}\boldsymbol{S}_j^{\mathrm{t}}\boldsymbol{\Lambda}_j(\boldsymbol{y}_j-\boldsymbol{P}\boldsymbol{\ell}_j)$.
$\boldsymbol{\sigma}_{\boldsymbol{h}}^2$ is sampled according to an inverse Gamma pdf: $p(\boldsymbol{\sigma}_{\boldsymbol{h}}^2\,|\,\boldsymbol{h})\sim\mathcal{IG}\big(D/2,\boldsymbol{h}^{\mathrm{t}}\boldsymbol{R}^{-1}\boldsymbol{h}/2\big)$.

**The nuisance variables and the scale $\eta^2$.** Vectors $\ell_j$ are independent, their sampling is achieved in parallel according to $\mathcal{N}(\boldsymbol{\mu}_{\ell_j},\boldsymbol{\Sigma}_{\ell_j})$ with: $\boldsymbol{\Sigma}_{\ell_j}^{-1}=\eta^{-2}\boldsymbol{I}_Q+\sigma_{\varepsilon_j}^{-2}\boldsymbol{P}^{\mathrm{t}}\boldsymbol{\Lambda}_j\boldsymbol{P}$ and $\boldsymbol{\mu}_{\ell_j}=\sigma_{\varepsilon_j}^{-2}\boldsymbol{\Sigma}_{\ell_j}\boldsymbol{P}^{\mathrm{t}}\boldsymbol{\Lambda}_j(\boldsymbol{y}_j-\boldsymbol{S}_j\boldsymbol{h})$.
$\eta^2$ is simulated according to: $\mathcal{IG}((QJ+1)/2,\sum_j\big\|\ell_j\big\|^2/2)$.

**The NRLs.** Sampling the mixture is done sequentially for each voxel $V_j$ and condition $m$ using two nested loops. Since the likelihood and the prior are Gaussian with respect to the NRLs, it can be shown that $p(a_j^m|rest=$ remaining variables $)$ is a Gaussian mixture as well:

$$p(a_j^m\,|\,rest)\sim\sum_{i=0,1}\lambda_{i,j}^m\mathcal{N}\big(\mu_{i,j}^m,(\sigma_{i,j}^m)^2\big),$$

with: $(\sigma_{i,j}^m)^{-2}=\sigma_{i,m}^{-2}+\sigma_{\varepsilon_j}^{-2}\boldsymbol{g}_m^{\mathrm{t}}\boldsymbol{\Lambda}_j\boldsymbol{g}_m$,
$\mu_{i,j}^m=(\sigma_{i,j}^m)^2\big(\sigma_{\varepsilon_j}^{-2}\boldsymbol{g}_m^{\mathrm{t}}\boldsymbol{\Lambda}_j\boldsymbol{e}_{m,j}+i\,\mu_{i,m}\sigma_{i,m}^{-2}\big)$,
$\lambda_{i,j}^m=\big(1+\tilde{\lambda}_{1-i,j}^m/\tilde{\lambda}_{i,j}^m\big)^{-1}$,
$\tilde{\lambda}_{i,j}^m=\lambda_{i,m}^2\lambda_{1-i,m}\sigma_{i,m}^{-1}\sigma_{i,j}^m\,\exp\big((\mu_{i,j}^m)^2/2(\sigma_{i,j}^m)^2\big)$,
where $\boldsymbol{g}_m=\boldsymbol{X}^m\boldsymbol{h}$ and $\boldsymbol{e}_{m,j}=\boldsymbol{y}_j-\boldsymbol{P}\boldsymbol{\ell}_j-\sum_{m'\neq m}a_j^{m'}\boldsymbol{g}_{m'}$. We first sample the binary label $q_j^m$ which consists in generating $u_j^m$ from the uniform pdf $\mathcal{U}([0,1])$ and then in applying the following rule: $u_j^m\leq\lambda_{1,j}^m\Longrightarrow q_j^m=1$, otherwise $q_j^m=0$. Then, we sample the NRL $a_j^m$ conditionnally to $q_j^m=i$ according to $\mathcal{N}(\mu_{i,j}^m,(\sigma_{i,j}^m)^2)$.

**The noise variances.** Sampling $\sigma_{\tilde{\varepsilon}_j}^2$ amounts to drawing variables according to $\mathcal{IG}((N+1)/2,\|\tilde{\boldsymbol{y}}_j\|_{\boldsymbol{\Lambda}_j}^2/2)$.

**The AR parameters.** For each voxel $V_j$, we have:

$$p(\rho_j\,|\,rest)\propto(1-\rho_j^2)^{1/2}\exp(-\frac{A_j}{2\sigma_{\tilde{\varepsilon}_j}^2}(\rho_j-\frac{B_j}{A_j})^2)\ \mathbb{1}_{(-1,1)}(\rho_j),\tag{3}$$

where $A_j = \sum_{n=2}^{N-1} \tilde{y}_{j,n}^2$ and $B_j = \sum_{n=1}^{N-1} \tilde{y}_{j,n}\tilde{y}_{j,n+1}$. Unfortunately, (3) is not a referenced pdf. We therefore propose to approximate $p(\rho_j \,|\, rest)$ by a Beta distribution $g(\rho_j) \sim \mathcal{B}(\alpha_j, \beta_j)$ defined over $(-1,1)$[1], whose parameters have to be set in an appropriate way. Hence, $g$ is of the form:

$$g(\rho_j) \propto (1+\rho_j)^{\alpha_j - 1}(1-\rho_j)^{\beta_j - 1}, \qquad \forall |\rho_j| < 1.$$

Calculating this approximation amounts to deriving a Beta pdf, say $\mathcal{B}(a_j, b_j)$, close to the exponential term in (3). For doing so, we calculate second order Taylor expansions of $\log(1+\rho_j)$ and $\log(1-\rho_j)$ about $m_j = B_j/A_j$:

$$\log(1 \pm \rho_j) \approx \log(1 \pm m_j) \pm \frac{\rho_j - m_j}{1 \pm m_j} - \frac{1}{2}\frac{(\rho_j - m_j)^2}{(1 \pm m_j)^2},$$

Then, let us find some scalars $(a_j, b_j, c_j)$ such that

$$a_j \log(1+\rho_j) + b_j \log(1-\rho_j) + c_j \approx -\frac{A_j}{2\sigma_{\varepsilon_j}^2}\big(\rho_j - m_j\big)^2,$$

After some straightforward calculations, we are able to identify the pair $(a_j, b_j)$: $a_j = A_j\big(1-m_j^2\big)\,(1+m_j)\,/2\sigma_{\varepsilon_j}^2$ and $b_j = A_j\big(1-m_j^2\big)\,(1-m_j)\,/2\sigma_{\varepsilon_j}^2$. Then, we have:

$$g(\rho_j) \propto \big(1-\rho_j^2\big)^{1/2}\,(1+\rho_j)^{a_j}(1-\rho_j)^{b_j}\,\mathbb{1}_{(-1,1)}(\rho_j)$$

with $\alpha_j = a_j + 3/2$ and $\beta_j = b_j + 3/2$.
We finally resort to a Metropolis-Hastings algorithm [8] that uses $g(\rho_j)$ as instrumental distribution.

**The weighting probabilities.** Sampling the probabilities $\boldsymbol{\lambda} = (\lambda_{1,m})_{m=1:M}$ can be parallelized. Drawing a realization of $\lambda_{1,m}$ consists in sampling from a Beta pdf: $\mathcal{B}(J_{1,m}+3/2, J_{0,m}+3/2)$,. Note that $C_{1,m}$ and $C_{0,m}$ stand for the sets of activated and unactivated voxels, respectively, for condition $m$. $J_{i,m} = \text{Card}\,[C_{i,m}]$, with $\sum_i J_{i,m} = J$.

**The mixture parameters.** For every condition $m$:
$p(\sigma_{0,m}^2 \,|\, \boldsymbol{z}^m) \sim \mathcal{IG}\big((J_{0,m}-1)/2, \nu_{0,m}/2\big)$, where
$\nu_{0,m} = \sum_{j \in C_{0,m}} (a_j^m - \eta_{0,m})^2$ and $\eta_{0,m} = J_{0,m}^{-1} \sum_{j \in C_{0,m}} a_j^m$.
$p(\sigma_{1,m}^2 \,|\, \boldsymbol{z}^m) \sim \mathcal{IG}\big((J_{1,m}-1)/2, \nu_{1,m}/2\big)$, where
$\nu_{1,m} = \sum_{j \in C_{1,m}} (a_j^m - \eta_{1,m})^2$ and $\eta_{1,m} = J_{1,m}^{-1} \sum_{j \in C_{1,m}} a_j^m$.
$p(\mu_{1,m} \,|\, \sigma_{1,m}^2, \boldsymbol{z}^m) \sim \mathcal{N}\big(\eta_{1,m}, J_{1,m}^{-1}\sigma_{1,m}^2\big)$.

## 4. RESULTS

### 4.1. Simulation results

Two sets of trials were generated, each of them corresponding to a specific stimulus type ($M$=2). The ROI $\mathcal{R}$ consisted of $J$=60 voxels, with $J_{1,1}$=22 (activated voxels) and $J_{0,1}$=38 for condition 1 and $J_{1,2}$=$J_{0,2}$=30 for condition 2. The distributions of the NRLs were set as follows: $\boldsymbol{a}^1_{j \in C_{1,m=1}} \sim \mathcal{N}(\mu_{1,1}$=10$, \sigma^2_{1,1}$=3), $\boldsymbol{a}^1_{j \in C_{0,m=1}} \sim \mathcal{N}(0, \sigma^2_{0,1}$=1), $\boldsymbol{a}^2_{j \in C_{1,m=2}} \sim \mathcal{N}(\mu_{1,2} = 2, \sigma^2_{1,2} = 0.3)$, $\boldsymbol{a}^2_{j \in C_{0,m=2}} \sim \mathcal{N}(0, \sigma^2_{0,2}$=0.4).
These settings provided us with various realistic combination of signal, noise and signal to noise ratio. For all voxels, the binary stimulus sequence was convolved with the canonical HRF $\boldsymbol{h}_c$[2], whose exact shape appears in Fig. 1(a) and (b) in ■-line. A space varying low frequency drifts $\boldsymbol{P}\ell_j$ and a SVAR(1)N $\boldsymbol{b}_j$ were then added to the stimulus-induced signal $\sum_m a_j^m \boldsymbol{X}^m \boldsymbol{h}$ in every voxel $V_j$. All AR parameters were set to the same value: $(\rho_j)_{j=1:J}$=0.4, which is compatible with the serial correlation observed on actual fMRI time series. The PMEs of the HRF yielded by our Gibbs sampler are plotted in Fig. 1(a)-(b) for the SVGWN and SVAR(1)N models, respectively. The corresponding NRL estimates are depicted in Fig. 1(c)-(d) for condition 1 and in Fig 1(e)-(f) for condition 2, respectively. Our simulation results first show that the correlated noise model provides a HRF estimate closer to the true shape. Second, the AR coefficients estimates are close to the true value (0.4) and third, we obtained more accurate estimates of the NRLs with smaller error bars when modeling the temporal correlation (compare first Fig. 1(c)-(d) and second Fig. 1 (e)-(f)).

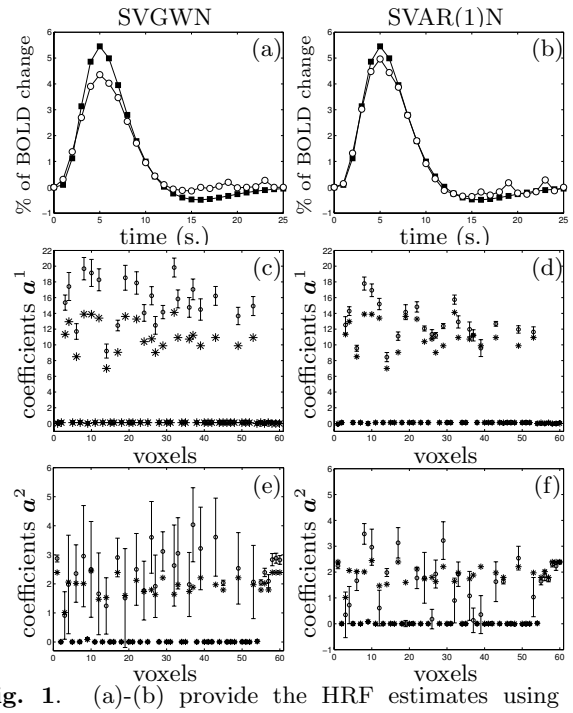


**Fig. 1**. (a)-(b) provide the HRF estimates using the SVGWN and SVAR(1)N models, respectively. Symbols ■ and ∘ represent the true $\boldsymbol{h}_c$ and the corresponding HRF estimate, respectively; (c)-(d) provide the NRL estimates for the first condition using the SVGWN and SVAR(1)N models, respectively; (e)-(f) provide the same results for the second condition. Symbols ∗ and ∘ represent true and estimated NRLs, respectively. The error bars correspond to ± the sampled posterior standard deviations.

### 4.2. Experimental results

The experiment is a fast event-related paradigm, performed on a 3-T whole-body system (Bruker, Germany), which consisted of a single session of $N$=125 scans lasting $TR$=2.4 $s$ each. The main goal of this experiment was to quickly map several brain functions such as motor, visual and auditory responses, as well as higher cognitive functions like computation. Here, we focus on two stimulus types: audio and video. The chosen ROI is a SPM cluster obtained from $t$ maps (thresholded at $P$=0.001 corrected for multiple comparisons) based on standard SPM activation detection using

[1] If $x \in (0,1)$ and $x \sim \mathcal{B}(\alpha, \beta)$ then $\rho = 2x - 1$ is said $\mathcal{B}(\alpha, \beta)$-distributed over $(-1, 1)$.

[2] used in SPM2 `www.fil.ion.ucl.ac.uk/spm/`

a canonical model, least squares estimation and inference on relevant contrasts (these results were obtained with SPM2). The 632 voxels of $\mathcal{R}_1$ are located around the voxel of Talairach coordinates in millimeters: ($X$=−63, $Y$=−21, $Z$=3). For the SVGWN model, the HRF estimate corresponding to $\mathcal{R}_1$ is plotted in Fig. 2(a). For the same ROI, the corresponding HRF estimate yielded by the SVAR(1)N model is depicted in Fig. 2(b). In $\mathcal{R}_1$, Fig. 2(c)-(d) show the maps of the NRLs difference between the SVGWN and SVAR(1)N models for the audio and video conditions, respectively. These figures suggest that the white noise (WN) model may lead to over estimated NRLs, a phenomenon also observed in figure 1(c)-(d). Note that the absolute value for the audio and video conditions reflect the choice of a region responding more to audio stimulus compared to video one. As a matter of fact, we expect little activity in the temporal gyrus due to visual condition. Our Gibbs sampler provides us with a PME ($\bar{p}_j^m = \frac{1}{K_0 - I + 1} \sum_{k=I}^{K_0} (q_j^m)^k$, $K_0$ is the number of iterations). We use this criterion for brain activation detection and classify voxel as active if it is more likely to belong to the class 1 distribution, as inactive otherwise. Note that this does not control for the type I risk of error, but simply inform on the most likely label for each voxel.

These maps are reported for the audio stimulus in Fig. 2(e)-(f) and for the video stimulus in Fig. 2(g)-(h). In all these panels, we compare the classification results using both noise models. The comparison of Figs. 2(e)-(f) show only little difference between the two noise models. Nevertheless, (f) does show more voxels labelled "active" suggesting that our SVAR(1)N model leads to increased sensitivity.

The activity detection maps for the video condition (Figs. 2(g)-(h)) appear less homogeneous in space, with "patchy" cluster of unactivated voxels (class 0 in black). As expected, it is clear that modeling the correlated noise increases the classification scheme specificity. This is inferred since this region should show very little "class 1" voxels under the video condition. This number is drastically reduced when the AR(1) model is used (panel (h)) compared to the WN model (panel (g)), and much less voxels appear in white (class 1) in (h) than in (g). Incidentally, this results also suggest that the classification scheme using the common autoregressive noise model, although performing much better than the WN model, is still likely to overestimate the number of false positive. This is the topic of future research.

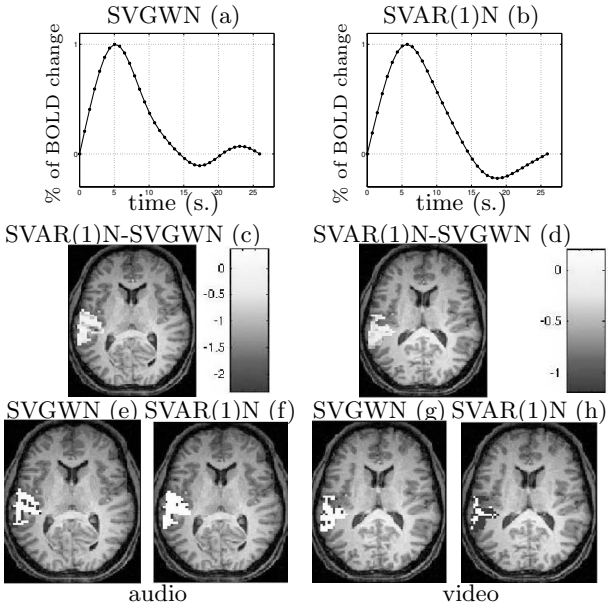


**Fig. 2**. Estimation results for the (audio minus video) contrast using raw data. (a)-(b): HRF estimate for $\mathcal{R}_1$ using the SVGWN and the SVAR(1)N model, respectively. (c)-(d): the audio (resp. video) voxels difference between the NRL estimates using the SVAR(1)N model and those of the SVGWN model. (e)-(f)/(g)-(h): classification results for the audio (resp. video) stimulus (black: class 0, white: class 1). Values are only plotted for voxels in $\mathcal{R}_1$.

## 5. CONCLUSION

We have shown that the SVAR(1)N model brings some improvements in both sensitivity and specificity at the expense of some computational load, since AR coefficient as well as drift parameters are now sampled. Our simulations have illustrated that this noise model has an influence over the PMEs of the NRL and their error bars, as well as on the HRF shape estimation. We have checked that these conclusions still hold on real fMRI data.

## 6. REFERENCES


[1] E. T. Bullmore, M. Brammer, S. C. Williams, S. Rabe-Hesketh, N. Janot, A. David, J. Mellers, R. Howard, and P. Sham, "Statistical methods of estimation and inference for functional MR image analysis," *Magn. Reson. Med.*, vol. 35, pp. 261–277, 1996.

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

[3] M. Woolrich, B. Ripley, M. Brady, and S. Smith, "Temporal autocorrelation in univariate linear modelling of fMRI data", *Neuroimage*, vol.14(6), pp.1370-1386, 2001.

[4] G. Marrelec, H. Benali, P. Ciuciu, and J.-B. Poline, "Bayesian estimation of the hemodynamic response function in functional MRI", in *Bayesian Inference and Maximum Entropy Methods*, R. Fry, Ed., Baltimore, MD, 2001, MaxEnt Workshops.

[5] K. Worsley, C. Liao, J. Aston, V. Petre, G. Duncan, F. Morales, and A. Evans, "A general statistical analysis for fMRI data," *Neuroimage*, vol.15, no.1, pp.1-15, 2002.

[6] S. M. Kay, *Modern Spectral Estimation*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[7] G. Marrelec, P. Ciuciu, M. Pélégrini-Issac, and H. Benali, "Estimation of the hemodynamic response function in event-related functional MRI: Bayesian networks as a framework for efficient bayesian modeling and inference," *IEEE TMI*, vol.23, no.8, pp.959-967, 2004.

[8] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97, Jan. 1970.