Spatially adaptive mixture modeling for analysis of fMRI time series

Thomas Vincent, Laurent Risser and Philippe Ciuciu, Member, IEEE

Abstract—Within-subject analysis in fMRI essentially addresses two problems, the detection of brain regions eliciting evoked activity and the estimation of the underlying dynamics. In [1, 2], a detection-estimation framework has been proposed to tackle these problems jointly, since they are connected to one another. In the Bayesian formalism, detection is achieved by modeling activating and non-activating voxels through independent mixture models (IMM) within each region while hemodynamic response estimation is performed at a regional scale in a nonparametric way. Instead of IMM, in this paper we take advantage of spatial mixture models (SMM) for their non-linear spatial regularizing properties. The proposed method is unsupervised and spatially adaptive in the sense that the amount of spatial correlation is automatically tuned from the data and this setting automatically varies across brain regions. In addition, the level of regularization is specific to each experimental condition since both the signal-to-noise ratio and the activation pattern may vary across stimulus types in a given brain region. These aspects require the precise estimation of multiple partition functions of underlying Ising fields. This is addressed efficiently using first path sampling for a small subset of fields and then using a recently developed fast extrapolation technique for the large remaining set. Simulation results emphasize that detection relying on supervised SMM outperforms its IMM counterpart and that unsupervised spatial mixture models achieve similar results without any hand-tuning of the correlation parameter. On real datasets, the gain is illustrated in a localizer fMRI experiment: brain activations appear more spatially resolved using SMM in comparison with classical General Linear Model (GLM)-based approaches, while estimating a specific parcel-based HRF shape. Our approach therefore validates the treatment of unsmoothed fMRI data without fixed GLM definition at the subject level and makes also the classical strategy of spatial Gaussian filtering deprecated.

Index Terms—fMRI; estimation of partition function; Bayesian analysis; MCMC; Spatial mixture models; joint detection estimation; MRF; Ising field.

I. INTRODUCTION

Since the first report of the BOLD effect in human [3], functional MRI (fMRI) has represented a powerful tool to non-invasively study the relation between cognitive task and the hemodynamic (BOLD) response, which indirectly reflects evoked neuronal activity. Within-subject analysis in fMRI essentially addresses two problems. The first one is about the detection or localization of activated brain areas in response to given stimulus types or to behavioral tasks, while the second one concerns the estimation of the underlying temporal dynamics, usually referenced as the characterization of the Hemodynamic Response Function (HRF). The first question has been extensively treated in the context of General Linear Models (GLM) in the last decade; see for instance [4, 5]. The second question has been investigated in a parametric [6–10], semi-parametric [11–13] or nonparametric [14–18] setting. The main differences between these approaches lie in the underlying assumptions regarding the evoked responses: parametric methods treat the HRF as a deterministic function whose parameters have to be identified while nonparametric approaches define the hemodynamic filter as a stochastic object of finite length. In the same vein, a stochastic modeling of hemodynamic responses has been conducted in a state space formulation either relying on physiological models [19–24] or on less informed evolution schemes (e.g., ARX, ARMAX models) [25–27].

Physiological Balloon-like models [19–24] are attractive in their ability to jointly estimate the HRF properties and the neuronal events eliciting evoked responses that can be seen under a delayed and filtered form in fMRI BOLD signals. Nonetheless, their highly nonlinear structure induces identifiability problems especially when fitting their parameters from non-activating or evoked-free fMRI time series. Hence, they are not able to provide a reliable answer to the joint question of brain activity detection and hemodynamics estimation from every fMRI time series. In [1, 2], a combined approach has been proposed to reconcile both issues in a region-based analysis i.e., on a set of prespecified parcels, which are supposed to be functionally homogeneous. In [28–30], such parcels were identified using parcellation algorithms that resort either to statistical or agglomerative clustering under connectivity constraints. Given the parcellation result, a key idea is to assign to each parcel a single but unknown nonparametric HRF shape, making the definition of spatially varying general linear models feasible. The main motivation here was to adapt the GLM definition to the local microvascular network with a restricted number of regressors. Within a given parcel, the evoked response is supposed to be modulated in magnitude across experimental conditions (or stimulus types) and voxels to accommodate voxel-dependent and stimulus-specific signal fluctuations. Following [16, 17], the HRF was modeled as a random quantity and physiological prior information is introduced to recover a smooth time course. Recently, several works segregating neuro- and hemodynamics events from fMRI time series have been proposed in the context of linear dynamic systems [31, 32]. Regarding the dynamic aspect, these contributions extend previous works since they are able to...
recover the Neuronal Response (NR) or both the HRF and
NR. However, from a spatial point of view, these works remain
massively univariate and proceed either voxel by voxel or on
averaged profiles (e.g., profile of independent components in
ICA [33]).

To enhance the estimation of regression coefficients in
the GLM context, global stationary and more recently lo-
cal non-stationary spatial regularization models have been
introduced [13, 34–36]. These Bayesian spatial priors encode
similarity between neighboring voxels and estimate one or se-
veral degrees of smoothness of the parameter images through-}

out the brain. In this regard, they do not require data to
be smoothed prior to entering a statistical model. However,
since these model take place in the family of edge-preserving
regularization models [37, 38], posterior inference1 is not nu-
merically tractable at low computational cost without con-
idering a mean-field approximation [39] and the Variational
Bayesian (VB) formalism [40]. This kind of approximations
breaks down numerical complexity by assuming conditional
posterior independence, which means independence across
voxels in the fMRI context. It also provides closed form
routines for iterative estimation of parameters of interest
and hyper-parameters provided that conjugate priors have been
retained. Nonetheless, the literature in spatial regularization
offers alternative choices that make the definition of non-
stationary spatial models and their global optimization [41,
42] or simulation feasible [43–49], even if VB approximation
4

5 can also be derived [50]. These models enter in the class of
Spatial Mixture Models (SMM) in which the spatial structure
is embedded on hidden allocation variables that specify the
voxel states in the fMRI context (activating or non-activating)
through a discrete Markov Random Field (MRF) [46,48,
51,52]. In this regard, the detection problem becomes a
segmentation issue that can be addressed using a region-
based method [43,53] instead of an edge-preserving restora-
tion one [37,38,54].

In [1,2], the authors have investigated the choice of the
best prior mixture model that serves as prior distribution on
the evoked response magnitude, called the Neural Response
Level (NRL) hereafter. More precisely, it has been shown
that the use of inhomogenous IMM (e.g., gamma-Gaussian)
permits to better disentangle activating from non-activating
voxels at the expense of higher computational cost. Nonethe-
less, in case of very few activations in a given parcel, the IMM
model overestimates the false positive rate (specificity). Both
on synthetic and real fMRI datasets, it has been demonstrated
in [55] that the homogeneous supervised SMM (SSMM)
outperforms its IMM counterpart in terms of false positive
control and activation cluster recovery. In this regard, our
work [55] is close to previous contributions [46,48,50,52,
56,57] but more general in the sense that our purpose is not
limited to the segmentation of a statistical map.

Akin to [46,51], in the present paper, we jointly estimate
HRFs and perform activation detection. However, our HRF
estimation procedure proceeds parcelwise and not voxelwise

1 e.g., the computation of the maximum a posteriori or the posterior mean
estimates do not admit a closed form expression.

and does not rely on any parametric assumption. We also ge-
eralize contributions [46,51] in the following directions:

- We do not impose HRF parameters (delay, peak magni-
tude, width, ...) to be sampled from parametric distribu-
tions (eg, Gaussian or gamma densities).

- Whatever the type of event-related fMRI paradigm (ie

slow or fast), our approach can handle multiple expe-

imental conditions through the estimation of stimulus

varying NRLs.

- The trial-by-trial variability or the non-stationarity of

the BOLD response captured in [51] in the context of

slow event-related designs can also be efficiently

addressed in the proposed JDE framework irrespective

of the experimental design. For the sake of simplicity,

this paper assumes a stationary model of NRLs (constant

across trials) but non-stationary extensions already exist
to model the so-called repetition suppression effect in
cognitive neuroscience [58,59].

- Following [44, 48] but in contrast to [46, 51], we develop
an unsupervised extension that makes the segmentation of
brain activations fully automatic. The unsupervised aspect
means that the MRF coupling parameter is automatically

\textit{tuned from the data}. Our unsupervised regularization

is also spatially adaptive meaning that the amount of

regularization varies across parcels. This requires a pre-
cise estimation of the partition function of the MRF

defined over each parcel. A small subset of these partition
functions are estimated by path sampling [60–62] while

the large remaining part is accurately approximated using

fast and precise extrapolation techniques [63,64].

- Our spatial regularization varies across experimental con-
ditions in every parcel since both the Signal-to-Noise
(RSNR) and the activation pattern may fluctuate from

one condition to another according to the involvement of

the given parcel in the experimental paradigm.

- Our approach is also able to account for deactivations or

negative BOLD responses [65] occurring for instance in

epilepsy, by introducing a 3-state Potts MRF as spatial

model instead of an Ising one; see [66] for details.

Hence, the proposed methodology introduces spatially adap-
tive levels of regularization at a reasonable computational cost
in adopting a fully \textit{exact} Bayesian inference of brain activity.

Parcel-based parameters of interest (HRF shape, NRLs) as well
as hyper-parameters are estimated in the Posterior Mean (PM)
sense from unsmoothed fMRI time series after convergence of
a hybrid Metropolis within Gibbs sampling procedure [67].

The rest of this paper is structured as follows: Section II
summarizes our parcel-based approach for whole brain anal-
ysis of fMRI time series by extending the Joint Detection-
Estimation (JDE) framework proposed in [1,2] to spatial
mixture models. The extension of this methodology to the
whole brain analysis is addressed in Section III where the
specific concerns of unsupervised and spatially adaptive regu-
larization (USMM) are dealt with. As shown, they respectively
require specific numerical schemes for estimating single and
then multiple partition functions (PF) of Ising fields. The
construction of the brain parcellation is then discussed and
illustrated in Section IV. Section V then shows the interest of USMM in comparison with IMM and SSMM on artificial datasets. Simulations results also demonstrate the importance of accounting for spatial fluctuations of the HRF shape. Then, our approach is illustrated on real fMRI datasets in Section VI in order to demonstrate that there exist relevant alternative strategies to spatial filtering of fMRI time series, specifically for contrasts of interest associated to low SNR. The last section finally summarizes our contributions and opens the discussion about the pros and cons of our method and its future improvements and extensions.

II. THE SPATIALLY REGULARIZED JDE APPROACH

A. Notations

We denote vectors and matrices with bold lower and upper case letters, respectively (e.g., \( \mathbf{y} \) and \( \mathbf{P} \)). A vector is by convention a column vector. Scalars are denoted with non-bold lower case letters. The transpose is denoted by \(^\top\). Unless stated otherwise, subscripts \( i,j,m,n \) and \( \gamma \) are respectively indexes over mixture components, voxels, stimulus types and time points. The probability distribution functions (pdf) are denoted using calligraphic letters (eg, \( \mathcal{N} \) and \( \mathcal{G} \) for the Gaussian and gamma distributions).

The JDE framework proposed in [2, 55] relies on a prior parcellation of the brain into \( \mathcal{P} = (\mathcal{P}_\gamma)_{\gamma=1:\Gamma} \) functionally homogeneous and connected parcels [28], where typically \( \Gamma \approx 500 \) to cover the whole brain (see Subsection IV-A and [68, 69] for their computation and the assessment of \( \Gamma \)); see also Fig. 2 for illustrations. Every parcel \( \mathcal{P}_\gamma \), comprising voxels \( (V_j)_{j=1:J_\gamma} \) is characterized by a specific model of the BOLD signal, which consists in estimating a single HRF shape \( \mathbf{h}_\gamma \) whatever the number of experimental conditions involved in the paradigm. In what follows, we first summarize the parcel-based modeling of the BOLD signal before specifying the priors involved over each parcel. In other words, our Bayesian parcel-based inference of brain activity is independently repeated over \( \Gamma \) different parcels with different sets of parameters on the corresponding fMRI time series. For notational simplicity, in what follows we drop off the \( \gamma \) index that makes reference to the parcel except for \( \mathbf{h}_\gamma \) and \( J_\gamma \).

B. Forward parcel-based model of the BOLD signal

The forward bilinear model of the BOLD signal introduced in [1] and extended in [2] to account for serial correlation of fMRI time series is a time-invariant model that characterizes each and every parcel \( \mathcal{P}_\gamma \) by a single neurovascular impulse response \( \mathbf{h}_\gamma \) and a NRI for each voxel and stimulus type. As illustrated in Fig. 1, this means that the HRF shape \( \mathbf{h}_\gamma \) is assumed constant within \( \mathcal{P}_\gamma \), while its magnitude \( a_m^\gamma \) can vary in space \( (j = 1 : J_\gamma) \) and across experimental conditions \( (m = 1 : M) \), where \( M \) is the total number of stimulus types. Then, the generative BOLD model reads:

\[
\forall j = 1 : J_\gamma, \quad \mathbf{y}_j = \sum_{m=1}^{M} a_m^\gamma \mathbf{X}^m \mathbf{h}_\gamma + \mathbf{P}\mathbf{\ell}_j + \mathbf{b}_j \tag{1}
\]

where:

- \( \mathbf{y}_j = (y_{j,n})_{n=1:N} \) denotes the fMRI signal measured in voxel \( V_j \) at times \( n = 1 : N \) (\( N \) is the number of scans).
- \( \mathbf{X}^m = (x_n^{m-d\Delta t})_{n=1:N,d=0:D} \) is a \( N \times (D + 1) \) binary matrix coding for the occurrences of the \( m \) stimulus type, with \( \Delta t \) is the sampling period of the unknown HRF \( \mathbf{h}_\gamma = (h_{d\Delta t}^{\gamma})_{d=0:D} \) in \( \mathcal{P}_\gamma \).
- \( a_m^\gamma \) stands for the NRI in voxel \( V_j \) for condition \( m \).
- \( \mathbf{P} \) is a low frequency orthogonal matrix of size \( N \times Q \). To each voxel is attached an unknown weighting vector \( \mathbf{\ell}_j \) to estimate the trend in \( V_j \). We denote \( \mathbf{L} = [\mathbf{\ell}_1 | \ldots | \mathbf{\ell}_{J_\gamma}] \) the set of low frequency drifts involved in \( \mathcal{P}_\gamma \).
- \( \mathbf{b}_j \in \mathbb{R}^N \) is the noise in \( V_j \) and follows a first-order autoregressive process: \( \mathbf{b}_j \sim \mathcal{N}(0, \sigma_b^2 \mathbf{I}) \) where \( \mathbf{A} = \begin{bmatrix} a_1 & \ldots & a_M \end{bmatrix} \) be the whole NRI matrix in \( \mathcal{P}_\gamma \) where \( a_m^\gamma = (a_m^\gamma)_{j=1:J_\gamma} \).

Although the noise structure is correlated in space [70, 71], we neglect such spatial dependency and consider the fMRI time series \( \mathbf{Y} = [y_1 | \ldots | y_N] \) independent in space but not identically distributed. The reason is twofold: first, neglecting the spatial dependencies of noise is tenable when the BOLD signal model itself is flexible enough to account for HRF shape fluctuations. Indeed, part of the usually observed spatial correlation of the noise is due to a misspecification of the BOLD signal model. Second, the noise correlation is much lower than that of the evoked BOLD response. When its modeling introduces additional computational complexity, there is no evidence in the literature that ignoring this correlation induces strong bias on the sought parameters. Hence, our likelihood reads:

\[
p(\mathbf{Y} | \mathbf{h}_\gamma, \mathbf{A}, \mathbf{L}, \theta_0) \propto \prod_{j=1}^{J_\gamma} |A_j|^{1/2} \sigma_j^{-N} \exp \left( -\frac{\mathbf{y}_j^\top A_j \mathbf{y}_j}{2\sigma_j^2} \right) \tag{2}\]

where \( \theta_{0,j} = (\rho_j, \sigma_j^2) \), \( \theta_0 = (\theta_{0,j})_{j=1:J_\gamma} \) and \( \tilde{\mathbf{y}}_j = y_j - \sum_m a_m^\gamma \mathbf{X}^m \mathbf{h}_\gamma - \mathbf{P}\mathbf{\ell}_j \).

Note that model (1) is bilinear in the sense that Eq. (1) linearly depends on \( \mathbf{h}_\gamma \) when \( \mathbf{A} \) is fixed and vice-versa. This means that the Maximum Likelihood (ML) solution \( (\mathbf{h}^*_\gamma, \mathbf{A}^*) \) cannot be distinguished from any other pair \( (h_s^\gamma / \mathbf{A}, \mathbf{A} x \mathbf{s}) \) whatever the scale parameter \( s > 0 \). The Bayesian formalism is helpful to get rid of such identifiability problems and define a reference scale. In the next subsection, we introduce priors on \( \mathbf{h}_\gamma \) and \( \mathbf{A} \), which are helpful to fix this scale to an arbitrary value \( c \). However, this value is not necessarily optimal for exploring the posterior distribution. Hence, instead of normalizing \( \mathbf{h}_\gamma \) deterministically \( (h_s^\gamma = c) \), it is shown in Appendix C how this reference scale \( s \) can be selected to speed up the exploration of the posterior distribution.

C. Bayesian inference

The Bayesian approach developed in [2] introduces proper priors on the unknown parameters \( (\mathbf{h}_\gamma, \mathbf{A}) \) in order to recover...
a robust estimate of brain activity (localization and activation profile). Akin to [2, 17], the prior density for the HRF remains Gaussian, \( h_m \sim \mathcal{N}(0, \nu^2 R) \) with \( R = (D_1^2 D_2^2)^{-1} \), which allows us to estimate a smooth HRF shape since \( D_2 \) is 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 [17].

1) Neural Response Levels: Regarding NRLs, according to the maximum entropy principle we assume that different types of stimulus induce statistically independent NRLs i.e., \( p(A | \theta_m, a) = \prod_m p(a^m | \theta^m) \) with \( \theta_m = (\theta^m)_{m=1:M} \). Vector \( \theta^m \) denotes the set of unknown hyperparameters related to the \( m \)th stimulus type. However, for certain class of paradigms (e.g., attention and motor dual tasks; priming effects, ...) this assumption may potentially be unlikely. In such cases, a between-condition prior covariance matrix could be introduced, the difficulty lying in the choice of a relevant and know correlation model to limit the computational complexity.

Mixture models are introduced to segregate activating voxels from non-activating ones. To this end, let \( q^m \) be the allocation variable that states whether voxel \( V_j \) is activating \((q^m_j = 1)\) or not \((q^m_j = 0)\) in response to stimulus \( m \). The NRLs still remain independent conditionally upon \( q^m \). This means that \( p(a^m | q^m, \theta^m) = \prod_m p(a^m_j | q^m_j, \theta^m) \) for every condition \( m \). In the case of an IMM [1, 2], the marginal density of the NRLs factorizes over voxels and reads:

\[
p(a^m | \theta^m) = \prod_{j=1}^{J} \prod_{i=0}^{1} p(a^m_j | q^m_j, \theta^m) \Pr(q^m_j = i | \theta^m). \tag{3}
\]

An important feature of IMM lies in the definition of the mixing probability (or weight) \( \Pr(q^m_j = i) \) which is constant over voxels and so independent of \( j \). This means that IMM explicitly estimates the proportions of voxels in the activating and non-activating classes.

Instead, spatial mixture models are introduced here to favor clustered activations and the mixing probabilities become space-varying. Hereafter, the marginal density no longer factorizes over voxels and reads:

\[
p(a^m | \theta^m) = \sum_{q^m} \prod_{j=1}^{J} p(a^m_j | q^m_j, \theta^m) \Pr(q^m_j = i | \theta^m) \]  \tag{4}

Spatial correlation is directly incorporated in the probabilities of activation through a hidden Ising MRF on the allocation variables \( q^m \), as already done in image analysis [44,45] or in neuroimaging [46, 57]. Here, as opposed to IMM, the proportions of voxels for the different classes is not explicit. The prior density on the allocation variables reads:

\[
\Pr(q^m | \beta_m) = Z(\beta_m)^{-1} \exp(\beta_m U(q^m)) \tag{5}
\]

with \( U(q^m) = \sum_{j < k} I(q^m_j = q^m_k) \) and \( I(A) = 1 \) if \( A \) is true and \( I(A) = 0 \) otherwise. The notation \( j \sim k \) means that the sum extends over all pairs \((j, k)\) of neighbouring sites. The neighbouring system can be defined either in 3D in the brain volume intersecting region \( P_o \) or in 2D along the cortical surface. In this paper, we only consider the 3D case using 6-connectivity. Extensions to 18 or 26 neighborhood system are straightforward. The Ising MRF in (5) has no external field. Previous works have shown that anatomical prior information can be embedded in an external field to increase activation probability in the grey matter [46]. The parameter \( \beta_m > 0 \) in (5) controls the amount of spatial regularization: large values of \( \beta_m \) associate higher probabilities to configurations containing clusters of like-valued neighboring binary variables. Since activation patterns within parcel \( P_o \) should be different from one stimulus type to another, different parameters \( \beta_m \) are considered across different stimuli types. The normalization constant of the MRF, also called the...
partition function \(Z(\cdot)\) reads:
\[
Z(\beta_m) = \sum_{q^m \in \{0,1\}^i} \exp(\beta_m U(q^m)) \tag{6}
\]
and guarantees that the MRF defines a pdf.

In what follows, we assume that \((a^m_j | q^m_j = i) \sim \mathcal{N}(\mu_{1,m}, \sigma^2_{1,m})\), with \(i = 0, 1\). We impose \(p_{0,m} = 0\) for all the marginals of the MRFs, leading to \(\theta^m = \{p_{0,m}, \mu_{1,m}, \sigma^2_{1,m} \}. Note that a Bernoulli-Gaussian formulation has also been tested in fMRI in [46]. This modeling corresponds to a degenerated mixture \((v_{0,m}) = \emptyset\).

2) Noise and drift parameters: To complete the Bayesian model, priors are required for all the remaining parameters. The noise and drift parameters, \(\theta_0\) and \(L\) respectively, are assumed independent in space:
\[
P(\theta_0, L | v) = \prod_{i} p(\theta_{0,i}) p(\ell_j | v_i)\quad \text{and without informative prior knowledge, the following priors are chosen:} \quad \ell_j \sim \mathcal{N}(0, v_i I)\quad \text{and} \quad p(\rho_j, \sigma^2_j) = \sigma^2_j^{-1} I(\rho_j < 1)\quad \text{to ensure stability of the AR(1) noise process}\. Note that the mean of the NRLs in non-activating voxels, leading to \(\text{Pr}(\rho^m_j = 1 | v_i) = \frac{1}{2}\) for all \(v_i \).

Interestingly, such proper prior defined over \(\mathcal{N}\) allows us to easily compute using field PF at discrete \(\beta\)-values defined over the same range; see Section III.

3) Mixture parameters: Similarly, the prior considered for \(v_{0,m}\) is \(p(v_{0,m}) = \frac{1}{2}\) because \(\text{we do not expect non-activating voxels in any parcel. Hence, class 0 should never be empty a priori.})\ If this assumption is not tenable, we could introduce a conjugate prior (an inverse Gamma law) denoted as \(\mathcal{IG}(\alpha_{0,m}, \beta_{0,m})\), as already done for the class of activating voxels. We would then avoid degeneracy problem that could prevent its sampling. In the same way, a proper prior \(N(\alpha_{1,m}, \beta_{1,m})\) is chosen for \(\mu_{1,m}\) \((\mu_{1,m} = x \text{ and } \beta_{1,m} = y)\), the prior on \(\beta = (\beta_m)_{m=1:M}\) is independent and identically distributed (iid) across conditions and follows a uniform pdf over fixed range: \(p(\beta_m) = U([0, \beta_{\text{max}}])\) with \(\beta_{\text{max}} = 1.6\). Interestingly, such proper prior defined over \([0, \beta_{\text{max}}]\) allows us to easily compute using field PF at discrete \(\beta\)-values defined over the same range; see Section III.

4) Posterior distribution: Considering the constructed model and assuming no further prior dependence between parameters, Bayes’ rule gives us:
\[
p(h_{\gamma}, A, L, \Theta | Y) \propto p(Y | h_{\gamma}, A, L, \Theta_0) p(A | \theta_A) p(h_{\gamma} | v_h) p(L | v_L) p(\theta_0) p(\theta_A) p(v_h, v_L) \tag{7}
\]
\[
\propto v_h \frac{\pi}{2} v_L \prod_{j=1}^{J_L} \frac{-x_j}{(1 - p_j^2)^{1/2}} \prod_{j=1}^{J_T} \frac{(1 - p_j^2)}{\sigma_j^{2+1}} \left(1 \left(\rho_j < 1\right) \right)
\]
\[
\exp\left(-\frac{1}{2v_h} \sum_{j=1}^{J_L} \frac{y_j A_j y_j^T}{\sigma_j^2} \right) + \frac{1}{2v_L} \|\ell_j\|^2 \right) \]
\[
\exp\left(-\frac{\sum_{j=1}^{J_T} h_j^T R^{-1} h_j}{2v_h} \right) \prod_{m=1}^{M} p(\theta^m | \alpha^m | \theta^m). \]
Note that the $B_m^{(n)}$ term handles the Gaussian truncation since the transition kernel is not symmetric. If $\beta_m^{(n)}$ is not accepted, then $\beta_m^{(n+1)} = \beta_m^{(n)}$. As shown in Eq. (9), an accurate evaluation of $A_m^{(n)}$ relies on a precise estimation of the ratio of partition functions $Z(\beta)$. Exact evaluations of $Z(\beta)$ is a reasonable task when they are impossible except for very small grids since it requires the evaluation of $U(q_m)$ for all possible combinations of $q_m$ whose number is exponentially related to the grid dimensions. Fast and robust estimation of $Z(\beta)$ is thus a key issue to get an unsupervised approach. Robustness is achieved by avoiding numerical overflow, thus considering Eq. (9) under its logarithmic form. Akin to [44, 57], efficiency is obtained here adopting path sampling to perform this estimation on a discrete grid of $\beta$-values, as detailed in what follows. Combined with interpolation of $Z(\beta)$ outside the grid points, this strategy makes the MH sampler of $\beta$ feasible.

### B. Estimation of single partition function

It is worth noting that in the special case where $\beta_m = 0$, the elements of $q_m$ are spatially independent, and $Z$ has a closed form expression $Z(0) = 2^{L}$. This property is fundamental for path sampling since this algorithm essentially relies on relative evaluations of $Z(\beta_m)$ for distinct values of $\beta_m$ [60–62, 74]. Path-sampling is actually an extension of importance sampling for estimating ratios of normalizing constants, by considering series of easy-to-sample unnormalized intermediate densities. Such a strategy was proven efficient to tabulate the PF for the Ising case; see [73] for details.

Let $f_1(q)$ and $f_2(q)$ be two densities with respective support $\Omega_1$ and $\Omega_2$ such that $\Omega_1 \subseteq \Omega_2$. Each density is known up to a normalizing constant, $f_s(q) = p_s(q)/Z_s$. The importance sampling identity is meaningful in order to estimate ratios of normalizing constants [60]:

$$Z_1/Z_2 = \mathbb{E}_2 \left[ \frac{p_1(q)}{p_2(q)} \right],$$

(10)

where $\mathbb{E}_s$ denotes the expectation with respect to $f_s$. Applying Eq. (10) to the context of Ising fields and considering a Monte Carlo approximation, we get:

$$\forall \beta_m, \tilde{Z}(\beta_m) \simeq \frac{Z(\beta_m)}{L} \sum_{l=1}^{L} \exp \left( (\beta_m - \beta_m^{(l)}) U(q_m^{(l)}) \right).$$

(11)

where $Z(\beta_m)$ is assumed to be known and $\{q_m^{(1)}, \cdots, q_m^{(L)}\}$ are distributed according to $\operatorname{Pr}(q_m^{(l)} | \beta_m)$. To avoid numerical overflow, Eq. (11) is usually handled in logarithmic scale:

$$\log \tilde{Z}(\beta_m) \simeq \log \frac{Z(\beta_m)}{L} + \log \sum_{l=1}^{L} \exp \left( (\beta_m - \beta_m^{(l)}) U(q_m^{(l)}) \right).$$

To estimate $Z$ around $\beta_m^0$, $L$ realizations $q_m^l$ are thus generated at $\beta_m = \beta_m^0$ before computing $U(q_m^l), \forall l = 1 : L$. Starting with $\beta_m^0 = 0$, the overall scheme consists in computing $\log Z(\beta_m)$ on a relatively dense $\beta$-grid (i.e., $\beta_m^0 = g \beta_{\max}/G$, $g \in \{0, 1, \cdots, G\}$ with e.g., $G = 10^0$) using Eq. (11) and then interpolating $\log \tilde{Z}(\beta_m)$ for $\beta_m$-values outside the grid points. This incremental procedure makes the estimation of $Z(\beta_m^0)$ relatively robust since it only relies on $\tilde{Z}(\beta_m^0)$. Interpolation between grid points is robust even if it is linearly performed since $Z(\beta_m)$ is an increasing function of $\beta_m$ and does not present strong irregularities. Furthermore, results strongly depend on our ability to generate draws $q_m$ that cover the whole support of the field.

While the Gibbs sampler is practically efficient at small $\beta$-values i.e., in little correlated configurations, it becomes impractical in highly correlated fields (large $\beta$-values) because it moves too slowly from one configuration to another. In such cases, a more efficient algorithm is the iterative Swendsen-Wang (SW) sampling scheme [75] provided that the Ising MRF to be sampled is symmetric. For non-symmetric models embedding external fields that could be useful to account for anatomical-based priors in the neuroimaging context [46, 52, 57, 76], several generalizations have been proposed [44, 49, 77, 78]. However in this paper, we only focus on the symmetric case.

### IV. Whole brain analysis: spatially adaptive USMM

#### A. Derivation of brain parcellation from fMRI data

As outlined in Section II, our BOLD signal modeling is parcel-dependent and spatial regularization differs over each parcel $P_i$. The methodology proposed in Section III to make this regularization unsupervised is therefore applicable to all parcels separately making our whole brain analysis of fMRI data fully spatially adaptive. This makes sense given that the stimulus-specific SNRs also vary in space. Of course, the critical issue is to exhibit such functionally homogeneous parcellation of brain. To this end, several algorithms have been proposed [28, 29, 68, 69]: they segregate the brain into connected and functionally homogeneous regions by minimizing a criterion reflecting both the spatial and functional structures of the dataset. The functional part of this criterion can be computed either from the raw fMRI time series or from voxel-based hemodynamic features (time-to-peak and time-to-undershoot, peak and undershoot magnitudes, ...) [79], which can be extracted from nonparametric HRF estimates [17, 18].

Since all these parcels have a specific geometry, we proceed to the PF estimation of 3D Ising MRF for each out of them. Still, the computation cost increase is reasonable as several hundreds of parcels are typically necessary to cover the entire brain. For an averaged-size parcel of 250 voxels, the PF estimation requires around 10 seconds which results in an increase of about 30 minutes for a whole brain analysis, compared to a time of 1.5 hour for the complete analysis, it yields a large increase of 33%. Hence, in the next part, we introduce fast numerical alternatives based on efficient approximations of 3D Ising partition functions, so that the overall cost would be quite negligible.

#### B. Multiple 3D Ising MRF partition function estimation

Algorithms with polynomial time complexity [39, 80] provide efficient alternatives to a single PF estimation. However, except for large grids, around the critical value $\beta_m$ where the phase transition occurs.
TABLE I

<table>
<thead>
<tr>
<th>Steps of the unsupervised JDE algorithm in a given parcel $P$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Estimate $Z(\cdot)$ using Eq. (11) for $\beta^0 = g/\max /G$, $g = 0 : G$</td>
</tr>
<tr>
<td>- Interpolate linearly $Z(\cdot)$ outside the grid points</td>
</tr>
<tr>
<td>- Setting up: choose $h_0^0$, $A^0(\cdot)$, $\sigma^2$, $\theta_0^0$</td>
</tr>
<tr>
<td>- Iteration $t$: draw samples $h_t^0$, $A_t^0(\cdot)$, $Q_t^0(\cdot)$, $(\sigma^2)$, $\theta_t^0$ from the full conditional posterior pdfs:</td>
</tr>
<tr>
<td>- HRF: $h_t^0 \sim p(h_t^0</td>
</tr>
<tr>
<td>- HRF variance: $v_t^{h_t} \sim p(h_t^0</td>
</tr>
<tr>
<td>- Drift: $\nu_j^{(t)} = 1 : J_s$, $J_s, J_s \sim p(\nu_j^{(t)}</td>
</tr>
<tr>
<td>- Drift variance: $\nu_j^{(t)} \sim p(\nu_j^{(t)}</td>
</tr>
<tr>
<td>- NRLs: External loop over stimuli $m$ ; Internal loop over voxels $j$:</td>
</tr>
<tr>
<td>- Identify $\lambda_t^{j}, N(\nu_t^{m_t}, v_t^{m_t})$ using Eqs. (14)-(16).</td>
</tr>
<tr>
<td>- $(a_t^{m_t}) \sim U(0,1]$ if $(a_t^{m_t}) \leq \lambda_t^{j}$</td>
</tr>
<tr>
<td>- $(\sigma^2)^{(t)} = 0$, $(q_t^{m_t}) = 1$ otherwise $(q_t^{m_t}) = 0$.</td>
</tr>
<tr>
<td>- $(\sigma^2)^{(t)}</td>
</tr>
<tr>
<td>- Update $\nu_j^{(t)}$ accordingly.</td>
</tr>
<tr>
<td>- Noise parameters: $\forall j = 1 : J_s$, $(\sigma^2, \rho^2) \sim \pi((\sigma^2, \rho^2)</td>
</tr>
<tr>
<td>- Mixture parameters: for every condition $m$,</td>
</tr>
<tr>
<td>- Variance of NRLs for nonactivated voxels: $(\nu_{m_t})^{(t)} \sim p(\nu_{m_t}, v_{m_t})$, cf. [1, Eq. (12)].</td>
</tr>
<tr>
<td>- Variance of NRLs for activated voxels: $(\nu_{m_t})^{(t)} \sim p(v_{m_t}, v_{m_t})$, cf. [1, Eq. (13)].</td>
</tr>
<tr>
<td>- Mean of NRLs for activated voxels: $(\mu_{m_t})^{(t)} \sim p(v_{m_t}, v_{m_t})$, cf. [1, Eq. (14)].</td>
</tr>
<tr>
<td>- Regularization level $\beta_t^0$ for every $m$: cf. Eq. (9).</td>
</tr>
<tr>
<td>- Iterate until convergence is achieved and compute posterior mean estimates (Discard $x^t = \ldots , x^{T_0}$, as burn-in period): $x^{T_0} = \frac{1}{T_0 - T_0} \sum_{t=T_0+1}^{T_0} x_t^{(t)}, \forall x \in {h_t, A_t, Q_t, \sigma^2, \rho^2, \theta_0^0}$</td>
</tr>
</tbody>
</table>

Fig. 2 Parcellation of the brain: (a) using the Voronoi based method - (b) as obtained with an optimal anatomical-functional parcellation [28].

None of them is able to perform multiple PF estimation at the same time, a problem that arises in our whole brain parcel-based analysis. Since several hundreds of grids of variable size and shape are manipulated in our fMRI application as illustrated in Fig. 2, fast estimation of multiple PF is necessary. To this end, in [63, 64] we have proposed several hybrid schemes which consist in resorting to path sampling to get log-scale estimates $(\log Z_p(\beta))_{p=1}^P$ in a small subset of reference graphs $(G_p)_{p=1}^P$ and then in using extrapolation formulas to obtain $\log Z(\beta)$ for the large remaining set of brain regions to be analyzed, referenced here by a test graph $T$.

In what follows, we summarize the main ideas and underlying assumptions of this extrapolation techniques.

1) Linear/Bilinear extrapolation schemes: In [73], the authors have proposed a linear regression procedure to estimate $(\log Z_p(\beta))_{p=1}^P$ as a function of the number of cliques in the grids $(G_p)_{p=1}^P$. Estimates of $\log Z(\beta)$ are then linearly computed using the estimated regression coefficients and the number of cliques in $T$: $\log Z(\beta_k) = \hat{a}_k e_T$ at each $\beta_k$ regularization level ($\beta_k = k \Delta \beta$). A bilinear extension of this technique, which also takes the number of sites in the grid into account, has been developed in [63]. This procedure was shown to be efficient to estimate log-PFs in small and irregular grids such as those appearing in our fMRI application. However, the accuracy of linear/bilinear PF extrapolations strongly depends on the homogeneity and the number of reference grids: the less homogeneous the reference set, the larger the approximation error. These reasons motivate the derivation of a more reliable and versatile approach to grid inhomogeneities as requested by our fMRI application.

2) Min-max extrapolation scheme: We have therefore proposed a more robust 2-step technique in [64]:

a) Akin to [73], reference log-PFs $\log Z_p(\beta_k)$ are estimated using path sampling on inhomogeneous reference grids $(G_p)_{p=1}^P$.

b) For any test graph $T$, the quantity $\log Z_T(\beta)$ is approximated from a single reference log-PF estimate out of $(\log Z_p(\beta))_{p=1}^P$ selected by a grid homogeneity criterion $\mathcal{L}(G_p)$ as well as an approximation error criterion $\mathcal{A}(\beta, G_p)$ defined by:

$$\mathcal{A}(\beta, G_p) = || \log Z_T(\beta) - \log Z_p(\beta, G_p) ||^2 / || \log Z_T(\beta) ||^2$$

with

$$\log Z(\beta, G_p) = \left( \frac{T}{G_p} \log Z_p(\beta) - \log 2 \right) + \log 2, \quad (12)$$

and where $(c_T, c_{G_p})$ and $(n_T, n_{G_p})$ are the number of cliques and sites of the Ising MRFs defined over $T$ and $G_p$, respectively. The theoretical justification of this construction can be found in [64].

In short, the extrapolation formula (12) is built up according to two principles: i) an unbiased asymptotic approximation error and ii) an exact approximation of $(\log Z_T(\beta))'$ for $\beta \to 0^+$. The reference grid $G_{ref}$ is exhibited using a min-max principle, which consists in minimizing wrt all reference grids $(G_p)_{p=1}^P$ the maximal approximation error $\mathcal{A}(\beta, G_p)$ given by (see [64]):

$$\mathcal{A}(0, G_p) \overset{\triangle}{=} \left\| (n_T - 1) - c_T (n_{G_p} - n_T) / c_{G_p} \right\|^2 / n_T^2. \quad (13)$$

Once $G_{ref}$ has been identified, the log-PF estimate in $T$ is thus given by $\log Z_T(\beta, G_{ref})$ according to Eq. (12).
V. SIMULATION RESULTS

A. Generation of Artificial fMRI datasets

Artificial fMRI datasets were simulated according to the forward BOLD model defined by 1 and illustrated in Fig. 1, considering two experimental conditions (M = 2). Simulation setup at low Signal-to-Noise Ratio (SNR) is reported in Table II which is the general context to examine JDE behaviour on artificial data either deriving from the true Ising prior for the labels (subsection V.D) or from hand drawn labels maps (subsections V. B, C, E). The considered map size is 400 voxels, which is close to the size a parcel obtained from a brain parcellation on real data. Convergence to simulated values is theoretically guaranteed for spatial variables when considering an asymptotic situation, i.e., very large size. Here, we rather focus on a small problem in order to grasp our model behaviour on a more realistic situation.

![Diagram](image-url)

Fig. 3. From top to bottom rows: simulated labels \( Q^{\text{true}} \), NRLs \( A^{\text{true}} \) and examples of artificial time series \((M = 2)\). On the bottom row, the solid line corresponds to the stimulus induced signal, the dashed line to the artificial data and vertical blue and green lines correspond to stimulus onsets for the two conditions. Simulated signals on the left side depict an active voxel for both conditions with \( a_1^1 = 2.25 \) and \( a_1^2 = 1.9 \). Signals on the right side depict an inactive voxel with \( a_1^1 = -0.1 \) and \( a_1^2 = -0.3 \).

<p>| TABLE II |</p>
<table>
<thead>
<tr>
<th>Setup for the generation of artificial fMRI datasets.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Paradigm:</strong> 30 trials, ( \forall m ), and fast event-related: ( \mu_{\text{ ISI }} = 5s, \sigma_{\text{ ISI }} = 2.9s )</td>
</tr>
<tr>
<td><strong>NRLs:</strong> ( (a_j^m</td>
</tr>
<tr>
<td><strong>labels:</strong> 20x20 maps</td>
</tr>
<tr>
<td><strong>Drift:</strong> ( \xi_j \sim N(0, 0.3 I_k) )</td>
</tr>
<tr>
<td><strong>Noise:</strong> ( \epsilon_j = 2 )</td>
</tr>
<tr>
<td><strong>HRF:</strong> canonical shape duration of 25 seconds</td>
</tr>
</tbody>
</table>

We have compared the efficiency of different models: IMM (independent MM, which do not model any spatial correlation), SSMM (supervised SMM, where \( \beta \) is set to a fixed value) and USMM (unsupervised SMM, where \( \beta \) is estimated). SSMM was iterated over a grid of \( \beta \) values ranging in \([0, 2]\) to manually search for optimal detection results. The USMM version relied on a path sampling estimate of \( \log Z(\beta_m) \), over the considered 20x20 map. Our first assessment step dealt with the detection performance when label maps are composed of realistic activation clusters and so do not follow the Ising prior, therefore there is no obvious \( \beta \) value associated to these maps. Moreover, in order to illustrate the need for adaptive \( \beta \) estimation, cluster configuration differs from one condition to another. As shown in Fig. 3, for \( m = 1 \), clusters are numerous with variable sizes, whereas for \( m = 2 \) there is only one cluster with a stretched shape.

A good synthetic criterion for detection assessment is the Area Under the Receiver Operating Characteristic curve (AUROC) which represents the sensitivity-specificity compromise. Results for this criterion are shown in Fig. 4. In terms of detection quality, IMM is comparable to using \( \beta = 0 \) with SSMM, whatever the condition of interest. When varying \( \beta \) with SSMM, too low \( \beta \) yields not enough spatial regularization whereas too high \( \beta \) values yield too much correlation and poor detection results, so that there is an “optimal” range of \( \beta \) values for a given label map. This optimal range varies across conditions: [0.4, 1] for \( m = 1 \) and [0.8, 1.5] for \( m = 2 \), because of different activation patterns. This illustrates the need for an adaptive estimation of \( \beta \). USMM fulfills this role by providing \( \beta \) estimates within the good ranges, i.e., \( \beta_m^\text{PM} = 0.86 \) and \( \beta_m^\text{USMM} = 0.99 \). USMM is able to adapt across condition and adjust \( \beta_m^\text{USMM} \) to the actual level of correlation within the field.

To go one step further in the mixture model behaviour, we then present results on label maps estimates, NRLs estimates and prior mixture components estimates. Regarding SSMM, we present these results for relevant \( \beta \) values extracted from the AUROC curves in Fig. 4: \( \beta = 0.2 \) - few spatial regularization, \( \beta = 0.8 \) - optimal setting for \( m = 1 \), \( \beta = 1.1 \) - optimal setting for \( m = 2 \) and \( \beta = 1.8 \) - very high spatial regularization level.

![Graph](image-url)

Fig. 4. Comparison of different supervised settings of \( \beta \) (SSMM) versus USMM with respect to the area under ROC (AUROC) for conditions \( m = 1 \) (left) and \( m = 2 \) (right). AUROC for IMM is not displayed, but is almost the same as SSMM-\( \beta = 0 \). The solid line shows the evolution of the AUROC index for supervised JDE runs. The diamond point indicates the AUROC index for the unsupervised setting.

B. IMM, SSMM, USMM comparison
Label map estimates are depicted in Fig. 5. First, we can state that SSMM is not beneficial over IMM if $\beta$ is not tuned to a right value. Indeed Fig. 5-$\beta = 0.2$ and Fig. 5-$\beta = 1.8$ show close results and Fig. 5-$\beta = 1.8$ illustrates a misspecification of $\beta$ where a too high value yields too much spatial regularization and results worse than IMM.

The optimal $\beta$ setting regarding SSMM for $m = 1$ is $\beta_1 = 0.8$ but this value is not suitable for $m = 2$. Conversely, the optimal setting for $m = 2$ is not suitable for $m = 1$. The unsupervised approach (see Fig. 5-USMM) is able to recover these nearly optimal $\beta$ settings: $(\tilde{q}^m)^{PM}$ maps for USMM and SSMM-[$\beta = 0.8$] are quite identical and so are $(\tilde{q}^m)^{PM}$ maps for USMM and SSMM-[$\beta = 1.1$].

The same conclusions can be drawn from results on $\tilde{a}^{PM}$ depicted in Fig. 6, where we can state the impact of a right regularization with USMM, SSMM ($\beta_1 = 0.8$ and SSMM-$\beta_2 = 1.1$) whereas IMM and SSMM with misspecified $\beta$-values lead to less sensitive results.

Considering the estimation of the prior mixture model components $\tilde{\theta}$, as shown in Fig.7, IMM leads to very overlapped estimated components with a high variance for the activating class and so does SSMM-[$\beta = 0.2$]. A misspecification of $\beta$ as shown with SSMM-[$\beta = 1.8$] leads to degenerate mixtures where component variances are very high. Again, the optimal setting for one condition may not be suitable for the other condition, see Fig.7-SSMM-[$\beta = 0.8$]. USMM yields consistent estimates of mixture components for all conditions.

C. Robustness to SNR decrease

We are interested here in the impact of signal degradation by noise. To this end, we varied the noise variance in simulations such as $\sigma_j^2 \in [0.5, 5]$ and we compared the efficiency of USMM versus IMM. Fig. 8 depicts the evolution of the AUROC index for the two conditions in terms of SNR computed for each simulation, according to: $\text{SNR} = \frac{1}{2} \sum_j 10 \log_{10} \frac{\| \sum_k a^m X \cdot h_k \|}{\sigma_j^2}$. For $m = 1$, where activations are more scattered, the USMM asset only holds for $\text{SNR} > 0.2$. Indeed, below 0.2 the degradation is too important and there is not enough regularity within the data to enable a proper recovery. For $m = 2$, where activations are more clustered, USMM clearly outperforms IMM whatever the noise level and still maintains a good AUROC index around 0.9 in the worst situation. Similarly, estimated $\hat{\beta}$ is stable against SNR decrease when there is enough signal in the data and then strongly decreases when there is too much noise (for $\text{SNR} < 0.2$).

D. Reproducibility of USMM results

Previous simulation results tested the USMM JDE procedure only on two particular configurations. We focus here on result reproducibility for the estimation of $\beta$. For doing so with a hidden field, we need to simulate multiple realizations of label maps and the most convenient way is to generate label maps following the Ising model, each of them being associated to a $\beta$ value. Accordingly, each simulation data set was generated as explained in Table II, with label maps $q^1$ and $q^2$ deriving here from the Ising prior at a given $\beta$ value and we applied the JDE procedure to these data. This simulation step was repeated over 100 random Ising field realizations and over a $\beta$-grid defined as $\beta_{true} \in \{0, 0.1, \ldots, 1\}$ leading to a Monte Carlo validation scheme where the recovery of true $\beta$-values can be assessed as well as variability across estimations. Fig. 9 shows the results of this procedure and some typical maps of Ising fields at increasing $\beta$-values with detection results in terms of false positives/negatives. Note that we only present results for $m = 1$ since they are the same for $m = 2$. The overall estimation of the regularization level $\beta$ is accurate since the regression $\beta^{true}_{m,1} \approx \beta^{true}_{m,2}$ does not strongly deviate from the first bisector. Optimal recovering appears for $\beta^{true}_{1,2} > 0.2$ with a small bias, while highest errors occur for very uncorrelated situations. This quite imperfect adequation of $\beta^{true}_{m,2}$ estimates regarding simulated quantities is first due to the small grid dimensions: convergence of hyper-parameters are only guaranteed in the asymptotic situation, which means in the case of very large grids and second, due to a low SNR, which can favor misclassification and induce biased estimates $\hat{\beta}_{m,2}$, suitable for real data analysis: salt-and-pepper situations are not realistic in the fMRI context where activations are expected to be clustered. We also emphasize the fact that Ising fields are hidden here so that hyperparameter estimation is a much more delicate issue than in the case of observed fields. We tested the latter situation using our PF extrapolation scheme and results were indeed very accurate (not shown).

Finally, note that dynamics estimation variability regarding this Monte Carlo validation is very little. For any simulated activation field, $\hat{h}_{\alpha}^{m}$ was very close to the simulated $h_{\alpha}$ (results not shown). Indeed, the HRF temporal reproducibility was not affected by varying activation localizations.

E. Hemodynamic Response Function shape variability

Spatial modeling is performed at the parcel level through SMM as was illustrated in previous sections V-B to V-D. At
Fig. 5. Comparison of different supervised settings of $\beta$ (SSMM) versus IMM and USMM with respect to the PM label maps $(\hat{q}^1)^{\text{PM}}$ (a) and $(\hat{q}^2)^{\text{PM}}$ (c), true activating clusters are rounded in green. Histograms on (b) and (d), for conditions $m = 1$ and $m = 2$ respectively, illustrate the segregation of true activations (red) and true inactivations (blue) in the corresponding maps. Y-axis shows densities. Note: purple regions stand for histogram overlaps.

Fig. 6. Comparison of different supervised settings of $\beta$ (SSMM) versus IMM and USMM wrt NRL estimates $(\hat{a}^1)^{\text{PM}}$ (top) and $(\hat{a}^2)^{\text{PM}}$ (bottom).

Fig. 7. Comparison of different supervised settings of $\beta$ (SSMM) versus IMM and USMM wrt prior mixture model component estimates $\hat{\theta}^m$ for $m = 1$ (top) and $m = 2$ (bottom). Dashed line densities correspond to the Gaussian pdf associated to inactivating voxels in the mixture while continuous line densities define the Gaussian pdf for activating voxels.
a higher regional scale, the JDE approach enables the HRF shape to vary across parcels, thus also performing spatial modeling, which is the focus of the current section. The same setup presented in V-A was used and we also kept the label maps depicted in Fig. 3 but these data were equally split vertically into two ROIs. Two different HRFs that differ from the canonical shape were simulated, as depicted in Fig. 10(a).

As we wanted to test the impact of the spatial relaxation of the HRF, we compared detection results for three situations:

(i) fixing the HRF to its canonical version, which is analogous to the classical GLM situation.

(ii) estimating one HRF over the concatenation of the two ROIs, which can be compared to GLM using a Finite Impulse Response (FIR) model.

(iii) estimating one HRF for each ROI.

Apart from the HRF modeling, we applied the USMM approach in all cases.

Fig. 11 shows how HRF modeling impacts detection. The most inflexible case, in Fig. 11(a) (one single fixed canonical HRF) yields the least sensitive results. The simulated HRF for the left ROI is close to the canonical shape (Fig. 10(a)), therefore we get more sensitive results in this part. However, for the right ROI, simulation strongly departs from the canonical shape, especially for the time-to-peak value so that sensitivity is dramatically decreased.

A more flexible situation is estimating one HRF for the whole data, results of which are depicted in Fig. 11(b). Compared to the previous situation, sensitivity is decreased in the left ROI and increased in the right one. Indeed, the estimated HRF is a compromise (or average) between the two simulated ones, as shown in Fig. 10(b-purple line), so that the estimated HRF is a worse solution than the canonical version for the left ROI, but a better one for the right ROI. However, sensitivity is globally enhanced.

Finally, the most flexible situation, in Fig. 11(c), allows a proper recovery of response levels and HRF are well estimated, see Fig. 10(b-orange and blue lines). In another respect, the case (ii) (one HRF for two ROIs) is equivalent to using an inadequate parcellation whereas case (iii) resorts to the true parcellation. Hence, the worse detection results for case (ii) compared to case (iii) also bring up the question of the sensitivity to the input parcellation. This issue is tackled in section VII.

VI. RESULTS ON REAL FMRI DATASETS

We applied the JDE procedure to 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 consisted of a single session of $N = 125$ scans lasting $TR = 2.4$ s each, yielding 3-D volumes composed of $64 \times 64 \times 32$ voxels. The paradigm was a fast event-related design comprising sixty auditory, visual and motor stimuli, defined in ten experimental conditions (auditory and visual sentences, auditory and visual calculations, left/right auditory and visual clicks, horizontal and vertical checkerboards).

We still compare the three versions of the JDE procedure: IMM, SSMM ($\beta = 0.8$) and USMM, in order to assess the impact of the adaptive spatial correlation model. Posterior mean estimates $\hat{A}$ have been computed over 5000 realizations of the Gibbs sampler after a burn-in period of $10^3$ iterations. Fig. 12 shows normalized contrasts maps of auditory computation (AC) versus auditory sentence (AS), where the modeling of spatial correlation seems to lead more sensitive results as activations in the parietal cortex are highlighted with SSMM and USMM whereas they are not with IMM. Moreover, these activations are coherent with the anatomy since they seem to follow the posterior part of the cingulate sulcus, which implication in numbers processing has been identified [81]. In another respect, Fig. 13 shows normalized contrast maps of auditory induced right click (RAC) versus auditory induced left click (LAC). As expected, the activations lie in the contralateral left motor cortex. Here, only USMM is more sensitive and we illustrate the advantage of an adaptive
spatial correlation model. Indeed, estimated $\beta_{\text{SSMM}}$ with USMM for the left auditory click was 0.56 so that the supervised setting of SSMM with $\beta = 0.8$ leads to too much correlation and less sensitive results.

Interestingly, Figs. 12-13 also depict the parcel-dependent maps of $\beta_{\text{SSMM}}$ estimates for the RAC and LAC experimental conditions. The gain in sensitivity in the USMM contrast map ($\hat{\alpha}_{\text{RAC}} - \hat{\alpha}_{\text{LAC}}$) results from a difference in the amount of spatial regularization introduced between the two conditions involved in the contrast. In parcels located in the left motor cortex, the BOLD signal is known to be stronger for the RAC than for the LAC condition. This is a possible interpretation of the lower regularization level estimated for RAC compared to LAC ($\hat{\beta}_{\text{LAC}} \approx 0.5$ vs. $\hat{\beta}_{\text{RAC}} \approx 0.75$) in the activating region outlined in Fig. 13.

Finally, following an illustrative purpose, Figs. 12-13 show estimated HRF shapes in the most active regions. They do not strongly depart from the canonical shape, expect for $\hat{\beta}_{20}$ which time-to-peak is shifted by around 2 seconds. The bumped tail of $\hat{\beta}_{232}$ can be explained by several hypothesis. There may be a sort of periodic scheme in the stimulus involved by the concerned activating regions. This may also be a particular behaviour of the local vascular system which responds to the undershoot. In any case, we would have to resort to a specific paradigm for a precise investigation. This concern goes beyond the scope of this paper and would be an interesting focus for future works centered on the neuro-vascular dynamics study, which is enabled by our approach.

VII. DISCUSSION AND CONCLUSION

In the present paper, unsupervised and spatially adaptive regularization has been integrated in a Bayesian approach that performs a conjoint detection-estimation of brain activity from fMRI time series. This extension has been achieved by estimating parcel-dependent regularization parameters $\beta$, which requires a precise estimation of the underlying 3D Ising field partition function. To this end, an extrapolation algorithm based on path-sampling has been systematically used over each parcel. This algorithm can be carried out prior to cycling over random samples drawn through Gibbs sampling since the partition functions only depend on the topological parcel configurations within the pre-computed parcellation.

Our approach first relies on the definition of an appropriate spatial scale which should be small enough so that the HRF shape invariance can be assumed and big enough so that we benefit from enough HRF reproductibility. This optimal spatial trade-off that impacts temporal modeling is fulfilled by the parcellation scheme we rely on. If we resort to such partitioning for HRF modeling, one could also think of taking the whole brain as the spatial support for the detection part, so that we would consider only one SMM to model response levels. However, this would obviously prevent us from making regularization spatially adaptive. Even if parcellation is well justified for HRF modeling from a physiological point of view, it is still questionable for response levels modeling. Indeed, two regions may be well suited to explain different vascular system properties so that we consider two different HRFs but they may not be suited for underlying activations which may span these two regions, for example. From a practical point of view, we did not observe any impact of parcel boundaries that would prevent activation clusters to span different regions. Moreover the parcellation optimality has been assessed by a sensitivity analysis in [82], where the parcellation method yields the same detection quality as a heavy Monte Carlo scheme which aims at marginalizing the parcellation. Ideally, the parcellation should be jointly exhibited so that parcel labels become random variables in our setting. In this respect, multivariate Gaussian mixture modeling (MSGMM) was applied in [79] to perform the required parcellation. Even if it settled a proof of concept as to merge parcellation with JDE, MSGMM alone suffers from a heavy computation burden.

One could also produce a meaningful functional parcellation by estimating for each parcel a label associated to a class of HRF. After treating an over-segmented parcellation (more parcels than HRF classes), spatially connected parcels which share the same labels could then be merged together and then be subsequently used in another JDE procedure to yield final reliable results. In another respect, reckoning on the second level analysis, parcellation at the group level can also be performed to account for between-subject variability [83]. Indeed, linking together brain parcels rather than voxels allows us to overcome the drawbacks of normalization due to spatial variability.

As a validation purpose, the right estimation of $\beta$ was assessed on simulations that followed prior modeling. Worst cases arised with highly correlated configurations, which are not expected to be interesting for dealing with real data. Simulations focusing on activation maps that do not follow prior Ising modeling confirmed the asset of SSMM versus IMM but also highlighted that a misspecification of a fixed $\beta$ value (SSMM) can lead to wrongly estimated activation label maps so that spatial correlation modeling does not bring any advantage compared to IMM. On the same simulated data, this limitation of SSMM was cleared by the USMM approach where the optimal setting of $\beta$ was correctly recovered. On real fMRI datasets, the latter situation is even worse for the SSMM setting. Indeed, the optimal setting of $\beta$ varies when considering different regions of the brain. In this respect, we identified regions where SSMM-$\beta = 0.8$ provides the same effect maps as IMM whereas USMM was more sensitive ($\beta = 0.56$). Hence, our claim is that our approach is not only unsupervised but also an adaptive spatial regularization scheme. As a remark, the sensitivity gain compared to IMM was mainly observed for low contrasts between conditions. For contrast involving higher response levels (for example auditory vs. visual conditions), there were no noticeable gain in resorting to USMM. Indeed, in this case there is enough information within data and spatial regularization may be useless.

To summarize, our approach enables a finer recovering of subtle contrasts by adapting the spatial regularization to varying contrast-to-noise ratios as well as various underlying activation patterns.

From a perspective point of view, the current work can
be generalized to 3-class USMM, thus involving Potts fields, as the detection of an additional deactivation voxel state is relevant in several fMRI contexts, like in epilepsy. The path-sampling step extension to Potts model is straightforward and NRL sampling demands little adaptation. However, model selection strategies have to be developed to avoid overfitting or interpretation issues when resorting to 3-class modeling in a region where 2-class modeling is more adequate, i.e., when there is no deactivation. Finally, we intend to validate the current method on group analysis to appraise the impact of our spatial regularization scheme to the sensitivity of group level effect maps.

**APPENDIX**

A. Generalized Inverse Gaussian distribution

The density of the $GIG$ distribution reads [84]:

$$f_{GIG}(z) = \frac{(\alpha, \beta)}{2^{\frac{1}{2}} K_\frac{1}{2}(\sqrt{\alpha\beta})} z^\frac{\lambda - 1}{2} \exp\left(-\frac{1}{2}(\beta z^{-1} + \alpha z)\right),$$

where $z > 0, \lambda \in \mathbb{R}$, and $K_\frac{1}{2}(.)$ is the third-type Bessel function. Efficient accept-reject simulation techniques do really exist for generating samples distributed according to this pdf.

B. The “neural” response levels

The prior on the NRLs ($A$) being a Gaussian mixture and the likelihood being Gaussian when $h_e$ is fixed, the full conditional posterior density of $A$ is also a Gaussian mixture. From Eq. (7), it can be shown that each $a_i^m \in A$ is obtained by sampling a 2-class posterior spatial Gaussian mixture in voxel $V_j$ for the $m$th stimulus type:

$$p(a_i^m | y_j, a_i^m \neq m, q_{k<j}, \ldots) = \sum_{i=0,1} \lambda_i^m \mathcal{N}(\mu_i^m, \sigma_i^m),$$

which can be decomposed in three steps: (i) Identify the posterior parameters $(\lambda_i^m, \mu_i^m, \sigma_i^m)$; (ii) Sample the binary label $q_i^m$ according to $\lambda_i^m$ and (iii) Sample the NRL $a_i^m | q_i^m$ according to $\mathcal{N}(\mu_i^m, \sigma_i^m)$. As detailed in [2], we have for $i = 0, 1$:

$$v_{i,j}^m = \left(\frac{1}{v_{i,m}} + \frac{g_i^m \Lambda_i g_m}{\sigma_i^m}\right)^{-1},$$

$$\mu_i^m v_{i,j} + \mu_i^m v_{i,m} = \frac{g_i^m \Lambda_i e_{m,j}}{\sigma_i^m} + \sum_{1 < \lambda \neq m} q_i^m \mu_i^m \sigma_i^m v_{i,m}$$

where $g_m = X^m h_{\gamma}$ and $e_{m,j} = y_j - P\hat{\theta}_j - \sum_{2 < \lambda \neq m} a_{i}^m g_i^m = \tilde{y}_j + g_m$. The posterior probability $\lambda_i^m$ of the event $q_i^m = i$ reads:

$$\lambda_i^m = \left(1 + \sum_{1 < \lambda \neq m} \pi_i^m \right)^{-1}$$

with $r_i^m = (v_i^m / v_{i,m})^{1/2} \exp\left((\mu_i^m)^2 / v_i^m - i(\mu_i^m)^2 / v_{i,m}^2\right)$ and $\pi_i^m = \Pr(q_i^m = i | q_{k<j}^m, \beta_m)$. To compute (16), we only need to evaluate

$$\pi_i^m \rightarrow \pi_i^m \exp\left(2\beta_m (2i - 1) \sum_{k<j} (2q_k^m - 1)\right),$$

which only depends on the labels in the neighborhood of $V_j$.

C. Sampling of the reference scale $s$

This new step reads as follows: $[A^\text{new}, \mu^\text{new}] = [A^\text{old}, \mu^\text{old}] / s$ and $h_e^\text{new} = h_e^\text{old} \times s$, where $s$ stands for the scale of the problem that will be drawn along each iteration of the Gibbs sampler. Straightforward application of
the principles exposed in [85, § 3] allows us to derive the full conditional distribution of \( s \) from the full posterior pdf:

\[
f_S(s \mid \text{rest}) = \left| s \right|^{P-2-M(J_n+1)} f \left( \frac{\mu_1^{\text{old}}}{s}, \frac{\mu_2^{\text{old}}}{s}, s h_{\gamma}^{\text{old}}(Q, \Theta) \right) \quad \text{exp} \left( -\frac{s^2 (h_{\gamma}^{\text{old}})^T R^{-1} h_{\gamma}^{\text{old}}}{2 \sigma^2} \right)
\]

\[
p \left( \mu_1^{\text{old}} \right) \prod_{j=1}^{J_n} \mathcal{N} \left( a_j^{\text{old}}, \frac{\mu_2^{\text{old}}}{s}, v_j \right)
\]

(17)

where \( (\mu_1^{\text{old}}, v_j) \) are mean and variance vectors for all stimulus types. Interestingly, we can note that the full conditional pdf (17) of \( s \) is independent of the data \( Y \). This is of course due to the scale ambiguity present in our bilinear observation model in Eq. (1). Then, from Eq. (17) it can be deduced that \( s^2 \) follows a Generalized Inverse Gaussian pdf (see Appendix A), denoted as \( \mathcal{GIG}(\lambda, \alpha, \beta) \) where:

\[
\lambda = (P - M(J_n + 1))/2, \quad \alpha = (h_{\gamma}^{\text{old}})^T R^{-1} h_{\gamma}^{\text{old}} \sigma^2 \quad (18)
\]

\[
\beta = \frac{\| \mu_{\gamma}^{\text{old}} \|^2}{2 \sigma^2} + \sum_{j=1}^{J_n} (a_j^{\text{old}})^T \Sigma_j^{-1} a_j^{\text{old}} \quad (19)
\]

with \( \Sigma_j = \text{diag} \{ v_j^{m, \text{old}} \} \).

The global sampling scheme of the full posterior distribution attached to a 2-class Gaussian SSM and AR(1) noise becomes:

a) Sample first \( Q^{(1)} \) and then \( (A^{\text{old}}(Q^{(1)}) \) given that \( a_j^{m, \text{old}} = \sum_{\ell=0}^{\gamma} \lambda_{j,\ell} \mathcal{N}(\mu_j^{m, \text{old}}, v_j^{m, \text{old}}), \forall j, m \)

b) Sample \( h_{\gamma}^{\text{old}} \) according to \( \mathcal{N}(\mu_h, \Sigma_h) \);

c) Sample the drift \( (\ell_j) \) and hyper-parameters \( \Theta \) especially \( \mu_1^{\text{old}} \);

d) Simulate \( (s^2)^{(1)} \sim \mathcal{GIG}(\lambda, \alpha, \beta) \) with parameters given in (18)-(19) then take its square root \( s^{(1)} \);

e) Apply the stochastic normalization: \( \left[ A^{(1)}, \mu_1^{(1)} \right] \left/ s^{(1)} \right. \) and \( h_{\gamma}^{(1)} = h_{\gamma}^{\text{old}} \times s^{(1)} \).

As seen in Fig. 14(a-b), normalizing the HRF to constant values during Gibbs sampling may have a dramatic impact and disable the recovery of the HRF shape. In contrast, the proposed stochastic sampling step enables this recovery, up to a multiplicative constant which is inherent to the scale ambiguity. Still, for \( \| h_x^{\text{old}} \| > 10^{-2} \) we do not observe any HRF estimation degradation with the deterministic approach, see Fig. 14(c-d). However, this optimal setting for the normalizing constant depends on the data so that we actually have to automatically adapt the HRF scale and rather consider the stochastic approach.

ACKNOWLEDGMENT

The authors thank the Région ile de France and the ANR for financial support (grant number ANR-005-MMSA-0014-03, OPTIMED project). We are also grateful to Jérôme Idier and Jean-François Giovannelli for fruitful discussions and comments.

REFERENCES


