A Bayesian Hierarchical Correlation Model for fMRI Cluster Analysis

Camille Gómez-Laberge, Andy Adler, Ian Cameron, Thanh Nguyen, and Matthew J. Hogan

Abstract—Data-driven cluster analysis is potentially suitable to search for, and discriminate between, distinct response signals in blood oxygenation level dependent functional magnetic resonance imaging (BOLD fMRI), which appear during cerebrovascular disease. In contrast to model-driven methods, which *test* for a particular BOLD signal whose shape must be given beforehand, data-driven methods *generate* a set of BOLD signals directly from the fMRI data by clustering voxels into groups with correlated time signals. Here we address the problem of selecting only the clusters that represent genuine responses to the experimental stimulus by modeling the correlation structure of the clustered data using a Bayesian hierarchical model. The model is empirically justified by demonstrating the hierarchical organization of the voxel correlations after cluster analysis. BOLD signal discrimination is demonstrated using i) simulations that contain multiple pathological BOLD response signals and ii) fMRI data acquired during an event-related motor task. These demonstrations are compared with results from a model-driven method based on the general linear model. Our simulations show that the data-driven method can discriminate between the BOLD response signals, while the model-driven method only finds one signal. For fMRI, the data-driven method distinguishes between the BOLD signals appearing in the sensorimotor cortex and those in basal ganglia and putamen, while the model-driven method combines these signals into one activation map. We conclude that the proposed data-driven method provides an objective framework to identify and discriminate between distinct BOLD response signals.

Index Terms—Functional magnetic resonance imaging, Bayesian hierarchical model, cluster analysis, stroke, cerebrovascular disease

I. INTRODUCTION

UNDERSTANDING the regulatory mechanisms of cerebral blood flow and its dysfunction in disease is a central topic in the neurological scientific literature. Ample evidence now exists indicating that factors related to age, disease, and medication can alter cerebrovascular regulation and, consequently, affect the blood oxygenation level-dependent (BOLD) signal measured by functional magnetic resonance imaging (fMRI) [1]–[5]. Conventional methods based on tests of variance, which require pre-defined hemodynamic models to detect a cerebrovascular response, may not be appropriate in this setting, since the diseased response may deviate considerably

from the assumed model. These facts have led to investigations of data-driven methods that search for distinct BOLD signal responses during an fMRI study [6]–[8]. These methods aim to identify *multiple distinct* hemodynamic responses to an event-related stimulus by, first, grouping together voxels having similar time series into clusters and, second, by identifying which of these clusters are significantly correlated to the stimulus paradigm.

In order to address the second aim, we propose a Bayesian hierarchical model [9] to describe the correlation structure of the observed voxel clusters. The model provides a joint posterior distribution of voxel and cluster correlation values in order to determine the statistical significance of the voxel correlations within each cluster with respect to the global correlation of the data set. We begin by justifying the choice of a Bayesian hierarchical model by empirically demonstrating the hierarchical organization of the correlation data. Next, the model is developed and applied to fMRI cluster analysis by computing estimates of the joint posterior distribution with a Markov chain Monte Carlo (MCMC) algorithm. We demonstrate the proposed method's capability for BOLD signal identification and discrimination by preparing fMRI simulations that contain three distinct BOLD response signals that have been observed in subjects with and without cerebrovascular disease (CVD) and by corrupting these signals with characteristic fMRI noise. The proposed method is also applied to real fMRI data acquired during resting and during an event-related hand motor task. The fMRI participants all gave informed consent in accordance with the Ottawa Hospital Research Ethics Board. A performance benchmark is provided for these experiments by comparing the results with those from a model-driven method based on a t statistic calculated using the general linear model (GLM) and the "canonical hemodynamic function."

It seems natural to use a multilevel statistical framework to describe the correlation structure of fMRI data. In fact, multilevel Bayesian frameworks such as Gaussian mixture models have also been used for model-driven fMRI analysis [10], [11]. The motivation for these frameworks is that fMRI data are strongly dependent in space and time. In addition to correlations between proximal voxels, a considerable degree of correlation is also expected between disparate regions, which can arise from general variations in blood pressure, electronic noise, magnetic field inhomogeneities, or body movement. Within a multilevel framework, these factors can be modeled appropriately at the voxel, regional, and global levels. Interestingly, the study of hierarchical models for fMRI analysis was also recently proposed during recent methodological debates
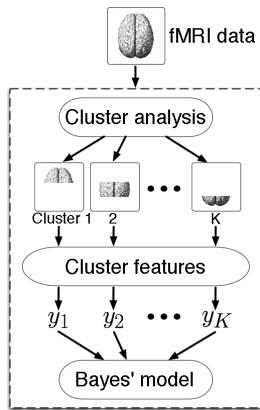
Fig. 1. Overview of the fMRI cluster analysis method. The acquired fMRI data are partitioned by a cluster analysis algorithm into a set of k clusters. The set of clusters is then parameterized by k features. The features y are then used in a Bayesian model to determine the statistical significance of each cluster. It is important to note that in this model, y does not represent the fMRI time-series data, but rather features of those data.

on the limitations of unilevel models employing multiple comparison correction strategies [12]–[14]. Nonetheless, while the aforementioned research suggests that multilevel frameworks might lead to analytical improvements over multiple comparisons correction factors, the aim of this article is not to compare statistical models but rather to demonstrate how data-driven fMRI cluster analysis can be used to search for, and discriminate between, distinct BOLD response signals.

II. METHODS

An overview of the analysis method is illustrated in Fig. 1. The fMRI data are acquired as a time series of voxels within the brain during an event-related experimental paradigm. In the *cluster analysis* stage, the data set is partitioned into k groups of voxels with correlated time series, such that each voxel has highest correlation with voxels in its group. This stage allows for the automatic detection and distinction of different responses without requiring hemodynamic or paradigm stimulus models. We used the fuzzy k-means algorithm [15], with the hyperbolic correlation distance [16]. This algorithm is appropriate for functional neuroimage data, since the optimal value k is unknown and can be determined heuristically from the algorithm's fuzzy membership values [17]. Our algorithm implementation is based on modifications of previously developed software [6]. In the *cluster feature* stage, the measured motor signal (e.g., the signal from a pressure transducer when the subject closes their hand) is compared to the BOLD signal from each cluster using the causal cross-correlation function, which we recently described in [7]. For each cluster i, this feature provides the correlation coefficient r with largest absolute value between the motor signal $m[t]$ and a BOLD signal $x[t+d]$, by optimizing the non-negative delay d constrained to be no greater than the shortest inter-stimulus duration. We calculate y_i and σ_i as the weighted average and weighted standard deviation, respectively, of the correlation coefficients from all voxels, where the weights are the membership values for cluster i. These data provide information of the distribution of correlation coefficients within

each cluster and are used to fit the Bayesian hierarchical model.

A. Correlation structure of clustered data

In order to study the empirical distribution of $y = (y_1, \ldots, y_k)$, we consider two questions: 1) at the voxel level, how is the temporal correlation $r(x[t], m[t])$ between each voxel $x[t]$ in the image and the motor signal $m[t]$ distributed? and 2) at the cluster level, how is the correlation $r(x[t], v_i[t])$ between voxels x in cluster i and the cluster centroid v_i distributed? Histograms of these correlations were computed to answer these questions. Fig. 2(a) shows the empirical distribution of the correlation coefficients of $r(x[t], m[t])$ for each of ten resting fMRI sessions, where each curve represents a different session. Fig. 2(b) shows a color-scale histogram of correlation coefficients $r(x[t+d], m[t])$ along the y-axis, for each cluster $i = 1, \ldots, k$ arranged along the x-axis, from one session. The y-axis indicates the correlation coefficient between the optimally delayed BOLD signal data in a voxel $x[t+d]$ and the motor signal $m[t]$. The delay is selected so that the cross-correlation between $x[t]$ and $m[t]$ is maximum [7]. The color bar is used to illustrate the histogram indicating the relative quantity of voxels in a given cluster having correlation values along the y-axis. These plots reveal several phenomena.

First, at the voxel level, Fig. 2(a) shows that each session produces a Gaussian-shaped distribution of correlation coefficients $r(x[t], m[t])$ that are not always centered at zero but often have a positive or negative bias. Indeed, this bias has been demonstrated in previous neuroimage studies resulting in a stimulus-correlated average time signal, often referred to as the *global signal* [18], [19]. Hence, the global signal could be modeled by a Gaussian random variable with mean and variance conditioned by the empirical distribution of $r(x[t], m[t])$. Second, at the cluster level, Fig. 2(b) shows how the correlation coefficients $r(x[t+d], m[t])$ organized by cluster (i.e., column-wise) are symmetrically distributed about the most frequent correlation value, which coincides with the average correlation coefficient $r(v_i[t+d], m[t])$. Furthermore, the spread of these distributions is directly related to the correlation $r(x[t], v_i[t])$. Hence, within each cluster, it is reasonable to approximate these distributions by a Gaussian random variable whose mean and variance are conditioned by the empirical distributions of y_i and $r(x[t], v_i[t])$. Together, these observations indeed suggest a hierarchical influence at global, regional, and voxel levels, since correlations at each voxel appear to be governed by distributions both at the global and cluster levels.

Therefore, for all voxels from each cluster i, we set y_i and σ_i^2 as the sample average and variance of correlations $r(x[t+d], m[t])$.

B. Bayesian hierarchical correlation model

Modelling voxel correlations using a multilevel model has the advantage of representing the variability within each cluster as well as between them. Moreover, the intersubject variability shown in Fig. 2(a) also indicates the impact a biased
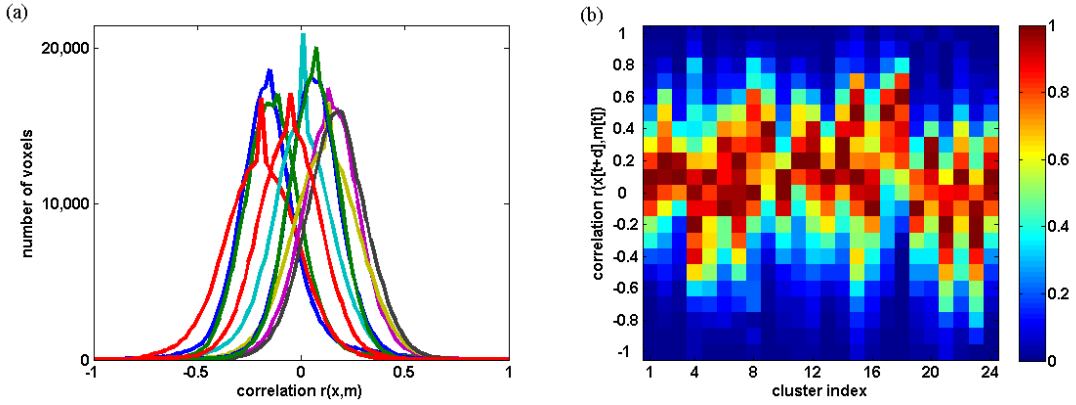
Fig. 2. Histograms demonstrating the hierarchical and Gaussian form of cluster voxel correlations. (a), Gaussian-shaped distribution of $r(x[t], m[t])$ correlations for each of ten resting fMRI sessions (each curve corresponds to one session) shows how the global signal can be biased in either direction. (b), Histograms (arranged in columns) of correlation coefficients $r(x[t + d], m[t])$ along the y-axis for each of the 24 clusters, along the x-axis, obtained from one session. The color bar quantifies the relative frequency of voxels in a given cluster having correlation values along the y-axis.
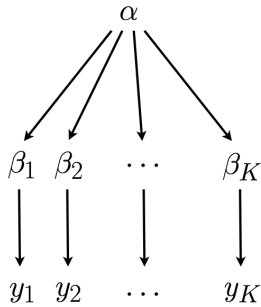

Fig. 3. Bayesian hierarchical model parameters (adapted from [9]) used in the final stage of the cluster analysis method. The model parameters (α, β) are fit to the data y, such that each cluster feature y_i is drawn from an underlying signal β_i. In turn, all underlying signals β are drawn from a global signal α.

global signal may have on determining cluster significance based on simple correlation thresholding.

A Bayesian approach can account for the bias introduced by the global signal by conditioning the model parameters based on the empirical observations y and their variance σ^2. Therefore, we propose a Bayesian hierarchical model to represent the correlative effects of the global signal and of each cluster by model parameters whose values are conditioned on the neuroimage data. The structure of the model parameters is illustrated in Fig. 3. From bottom to top, each of k clusters inputs a value y_i determined using the correlation data $r(x[t + d], m[t])$. The objective is to determine whether voxels forming cluster i are significantly affected by an underlying response or rather exhibit an unrelated signal. The underlying effect of each cluster is modeled as the parameter β_i (the true value that y_i is measuring). The apex of the structure is the hyperparameter α that models the global signal. The hyperparameter α controls this effect by modeling the influence of the global signal on each cluster via β_i. Our implementation closely follows the theoretical developments by Gelman *et al.* [9]; therefore, we only describe the details necessary to apply the model to the fuzzy k-means algorithm.

The joint posterior distribution $p(\beta, \alpha, \tau | y, \sigma)$ encapsulates the entire model by conditioning all model parameters

$\{\alpha, \beta, \tau\}$ with the given data $\{y, \sigma\}$, where the variance of α is estimated by the parameter τ^2.

We model the joint likelihood function $p(y, \sigma | \beta, \alpha, \tau)$ as the Gaussian distribution $\mathrm{N}(\beta_i, \sigma_i^2)$, since each y_i is the sample average of the member voxel correlations $r(x[t+d], m[t])$. For a large sample size of independent and identically distributed samples the central limit theorem guarantees an approach to normality. Although our samples do exhibit considerable interdependence, our empirical data in Fig. 2 show that their distribution of values remain unimodal and approximately Gaussian. The joint population distribution is modeled by assuming that the individual parameters β_i are exchangeable, since there is no a priori information to distinguish one cluster from another. Furthermore, we assume that the population and prior distributions have the Bayesian conjugate form of the likelihood function, since we do not have an explicit parametric form for these distribution functions. In the case of a Gaussian likelihood function, the conjugate forms for the population and prior distributions are also Gaussian. Therefore, the joint likelihood function and the joint population distribution are respectively written as

$$p(y, \sigma | \beta) = \prod_{i=1}^{k} \mathrm{N}(\beta_i, \sigma_i^2), \tag{1}$$

$$p(\beta | \alpha, \tau) = \prod_{i=1}^{k} \mathrm{N}(\alpha, \tau^2). \tag{2}$$

Finally, since nothing is known about α and τ a priori, we write the joint prior distribution in the non-informative form

$$p(\alpha, \tau) = p(\alpha | \tau) p(\tau) \propto 1. \tag{3}$$

The product of equations (1)–(3) describes the joint posterior distribution to within a constant factor.

C. Gibbs sampling of the joint posterior distribution

This section provides a numerical approach to estimate the joint posterior distribution by simulating draws using the Gibbs sampler MCMC technique based on Gelman *et al.* [9].

This method has the advantage of implementing Bayesian hierarchical models that do not have closed-form solutions for their distribution functions.

First, the joint posterior distribution is factored into the alternate form

$$p(\beta, \alpha, \tau | y, \sigma) = p(\beta | \alpha, \tau, y, \sigma)\, p(\alpha | \tau, y, \sigma)\, p(\tau | y, \sigma) \quad (4)$$

so that realizations can be simulated from each term starting with the rightmost density to give τ, then α, and finally β. Simulated draws from equation (4) are obtained by iteratively performing batch realisations of $k + 2$ random variables $\{\beta_1, \ldots, \beta_k, \alpha, \tau\}$, where each batch simulates one point from $p(\beta, \alpha, \tau | y, \sigma)$.

We require the conditional density functions on the right-hand side of equation (4). To begin, we estimate the conditional posterior distribution of $\boldsymbol{\tau}^2$ using the *scaled inverse-χ^2 distribution* [9], [20] written as

$$\boldsymbol{\tau}^2 | \beta_i, \alpha, y \sim \text{Inv-}\chi^2(k - 1, \hat{\tau}^2), \quad (5)$$

where $\hat{\tau}^2$ is the unbiased variance estimate based on the realisations α and β_i given by

$$\hat{\tau}^2 = \frac{1}{k - 1} \sum_{i=1}^{K} (\beta_i - \alpha)^2. \quad (6)$$

Next, we seek the conditional posterior distribution of $\boldsymbol{\alpha}$. It is not hard to see that this distribution is simply the sample average and sample variance of the β_i terms

$$\boldsymbol{\alpha} | \beta, \tau \sim \text{N}\left(\frac{1}{k} \sum_{i=1}^{k} \beta_i, \frac{\tau^2}{k} \right). \quad (7)$$

Finally, the conditional posterior distribution of $\boldsymbol{\beta}_i$ is obtained based on the exchangeability assumption, since the multivariate distributions (2) and (1) have product forms. Thus, computing their product $p(\beta | \alpha, \tau)p(y, \sigma | \beta, \alpha, \tau)$ yields the Gibbs sampler estimate of each β_i as the Gaussian random variable

$$\boldsymbol{\beta}_i | \alpha, \tau, y_i, \sigma_i \sim \text{N}(\hat{\beta}_i, V_{\beta_i}) \quad (8)$$

with mean and variance

$$\hat{\beta}_i = \frac{y_i / \sigma_i^2 + \alpha / \tau^2}{1 / \sigma_i^2 + 1 / \tau^2}, \quad V_{\beta_i} = \frac{1}{1 / \sigma_i^2 + 1 / \tau^2}. \quad (9)$$

We used the following algorithm based on the Gibbs sampler equations (5)–(8) for all data in this study. At the end of each iteration, we include a means to check for MCMC convergence using the potential scale reduction threshold [9].

*Algorithm 1 (**Gibbs sampler**)*: Let (y_i, σ_i^2) for $i = 1, \ldots, k$ be a data set described by a Bayesian hierarchical model with random variables $\boldsymbol{\tau}$, $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$. Let $G = 10$ be the number of Gibbs samplers running in parallel, and $2N = 2,000$ the number of realisations from each sampler, and one small number, such as $\epsilon = 0.001$. Then a total of $GN = 10,000$ batch realisations are obtained by the doing the following for each of the G samplers:

1) Set the initial values for each $\beta_i = y_i$ and compute the initial value for α to be the sample average of all k

values β_i.
2) Simulate a realization of τ using equation (5), which can be implemented by first making random draws from a χ^2 random variable with $k - 1$ degrees of freedom, $X = \chi_{k-1}^2$, and then setting $\tau^2 = (k - 1)\hat{\tau}^2 / X$.
3) Simulate a realization of α using equation (7), with the latest realisations of τ and β.
4) Simulate a realization for each β_i using equation (8), with the latest realisations of α and τ.
5) Store the batch of latest realisations $\{\tau, \alpha, \beta_1, \ldots, \beta_k\}$.
6) Repeat steps 2-5 until $2N$ batches are obtained.
7) Discard the first N batches from each Gibbs sampler.
8) Compute the potential scale reduction threshold

$$\hat{R} = \sqrt{\frac{1}{W}\left(\frac{N-1}{N}W + \frac{1}{N}B \right)}, \quad (10)$$

where W and B are the within-sampler and between-sampler variance of the remaining GN realisations.
9) If the inequality $|\hat{R}| < 1 + \epsilon$ is satisfied, then these GN realisations are convergent and represent realisations from the true joint posterior distribution of the Bayesian hierarchical model.

Given a convergent MCMC sampling of the joint posterior distribution, we can calculate the significance of each cluster with respect to the global signal by computing Bayes' error between the distributions of the hyperparameter $\boldsymbol{\alpha}$ and the cluster parameters $\boldsymbol{\beta}_i$, for each $i = 1, \ldots, k$. Geometrically, it corresponds to the overlap area between the tails of both distributions. A summary graph for the area overlap between each $\boldsymbol{\beta}_i$ and $\boldsymbol{\alpha}$ can be obtained by generating a corresponding set of box plots arranged horizontally whose box and whiskers extend vertically and show the 5% and 95% quantiles of each distribution. When the box plot of a given β_i does not overlap with the box plot of α, then that cluster has a distribution of correlation values that are significantly different from the rest of the data set and potentially represent a stimulus-response cluster with probability of error $p < 0.05$.

D. Simulation of event-related BOLD fMRI data

Simulated data are used to test the sensitivity and specificity under various signal to noise ratio (SNR) levels. The temporal SNR for fMRI data is defined as the largest change in signal magnitude divided by the noise standard deviation. This SNR definition and range of $[0, 2]$ are common in the literature involving 1.5 Tesla MRI. The image volume contains 64^3 8-bit voxels over which an event-related design is executed for 160 T_R-cycles.

The normal BOLD signal can be divided into three post-stimulus phases: i) a delay before the response onset, ii) a positive peak width, and iii) a negative undershoot period after the response. Fig. 4(a) plots three BOLD signals normalised to unit extrema whose characteristics are based on those described in the literature. Signal 1 represents a typical BOLD response in normal subjects. The selected values for the onset delay (1.3 s), time-to-peak (5 s), peak width (8.3 s), and undershoot duration (11 s) are the average of previously reported values [21]–[23]. Signal 2 is a delayed response in
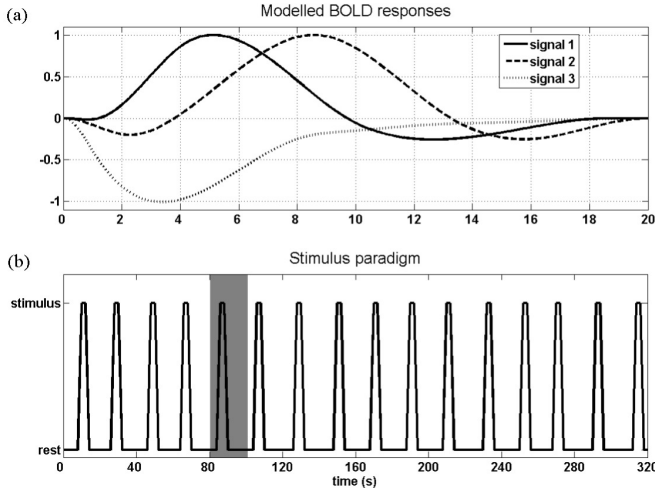
(a)

(b)

Fig. 4. Simulated fMRI data. (a), Modelled BOLD responses from normal subjects (signal 1) and CVD patients (signals 2,3). (b), Event-related paradigm with interstimulus period randomly varying between 16-20 s (gray area indicates simulated BOLD response duration).

the presence of CVD, where an average onset delay of 3.8 s and the appearance of an early trough before onset have been reported [24], [25] . Signal 3 is an advanced, negative peak response also from CVD patients, however, this time exhibiting a time-to-peak advancement of 2 s [25], [26].

The event-related paradigm time series is illustrated in Fig. 4(b). The signals are added to a central block of the image volume, each confined to one of three regions, and decrease linearly in SNR moving upwards along the y-axis (e.g., see Fig. 6(b)). The remainder of the image volume consists of a low intensity background and noise voxels without signal.

The noise for each voxel is simulated from an empirical analysis of water phantom data. Four sessions were completed using a standard echo-planar imaging sequence, whence a central 16^3-voxel region from each session was observed over 16 consecutive T_R cycles. Before analysis, these data underwent the same preprocessing as our human fMRI data. The average phantom session mean and standard deviation in arbitrary image intensity units (au; measured as 1517 ± 30.79 au, $n = 4$) are simulated as Gaussian noise with a 2.0% standard deviation from a mean intensity of 1500 au. A slight negative trend is also simulated, based on the observed phantom linear intensity decrease with slope -0.025 au/T_R ($r = -0.98$, $n = 4$). Spatial and temporal noise correlations are simulated by estimating the autocovariance function $\gamma_{ww}[m] = \mathcal{E}\{\mathbf{w}_n \mathbf{w}_{m+n}\} - \mathcal{E}\{\mathbf{w}_n\}^2$ of a stationary real random process $\mathbf{w}_n$. The phantom data $w[n]$ are treated as a realization of $\mathbf{w}_n$, where n is a multivariate index over space-time. The first term is estimated using the convolution property of the discrete Fourier transform (DFT): computing the product $\Phi_{ww}[k] = W[k]W[-k]$, where $W[k]$ is the DFT of $w[n]$, and then taking the inverse DFT of $\Phi_{ww}[k]$. The second term is estimated as the square of the sample mean of $w[n]$. The function values are plotted along the four axes in Fig. 5, where x, y, and z represent the MRI read encode, phase encode, and slice select orientations, respectively. The
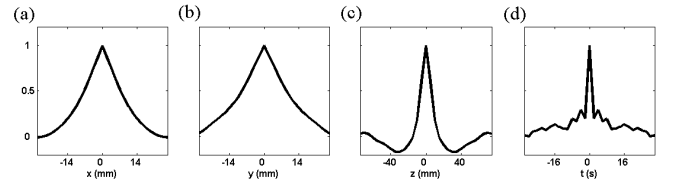
(a) (b) (c) (d)

Fig. 5. The autocovariance function estimate plotted along (a) the read-encode x-axis, (b) the phase-encode y-axis, (c) the slice-select z-axis, and (d) the time axis.

correlation coefficient estimates are $r_x = 0.865 \pm 0.008$, $r_y = 0.898 \pm 0.002$, $r_z = 0.636 \pm 0.024$, $r_t = 0.208 \pm 0.076$ ($n = 4$). These negative trend and autocovariance functions are used to simulate the fMRI noise corrupting the aforementioned BOLD response signals.

E. Acquisition of event-related BOLD fMRI data

To qualitatively demonstrate the performance of the proposed method during the performance of an event-related motor task, BOLD fMRI data were acquired from one healthy subject (36-year-old female). Whole brain imaging was performed on a 1.5 Tesla Magnetom Symphony MR scanner (Siemens, Erlangen, Germany) using an echo-planar imaging (EPI) sequence ($T_R/T_E = 2000/30$ ms; $70°$) and a voxel size of $1.72 \times 1.72 \times 5.00$ mm^3. The fMRI session produced a 160-image data set collected over 5 minutes, roughly. A stimulus-response apparatus provided a visual feedback hand-motor task whose events of two-second duration occurred irregularly to minimize habituation. During events, the subject was instructed to squeeze a palm-sized ball placed in her right hand in order to raise a computer-generated horizontal bar displayed on-screen to meet a target. The visual scene was kept as small as possible to avoid visual area stimulation. The only visible objects in the room were two horizontal bars whose vertical movement was limited to subtend no more than a 5-degree angle with respect to the subjects eyes. A detailed description of the task is given in [7].

III. RESULTS

A. Simulated BOLD fMRI data

Summary results from the proposed method are shown in Fig. 6. The joint posterior distribution $p(\beta, \alpha, \tau | y, \sigma)$ is illustrated in Fig. 6(a) by a set of box plots representing the MCMC distributions of the parameters $\alpha, \beta_1, \ldots, \beta_k$ arranged horizontally. A low Bayes' error corresponds to a cluster whose box plot does not overlap with that of α shown in gray. Boxes marked with an asterisk have $p < 0.05$. The figure shows that the simulated data are partitioned into 24 clusters, where β corresponding to clusters 3, 4, and 5 are significantly different from α. The three response-signals are well distinguished by the voxel maps in Fig. 6(b). The averaged voxel time series are shown in Fig. 6(c). The correlations and response delay pairs (r_i, d_i) for clusters 3, 4, and 5, are $(0.78, 8\,\text{s})$, $(0.78, 6\,\text{s})$, $(-0.87, 4\,\text{s})$, respectively. Due to the inclusion of high SNR values, these correlations may seem
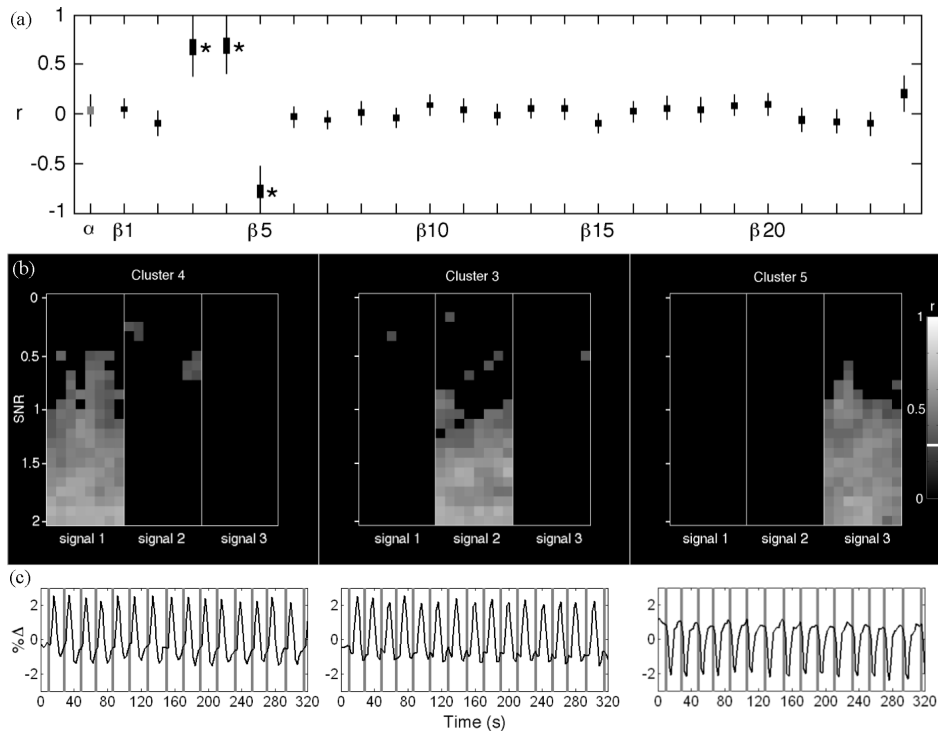
Fig. 6.   Simulation results from the proposed method. (a), Box plots of paradigm-cluster correlation values from the entire data set α and each cluster β_i ($i = 1, \ldots, k$). The * symbol denotes clusters that are significantly different from α with $p < 0.05$. (b), Significant voxel-cluster correlation maps for clusters 3-5 (membership threshold indicated by white bar on scale). (c), Averaged voxel time series (events indicated by vertical gray bars).
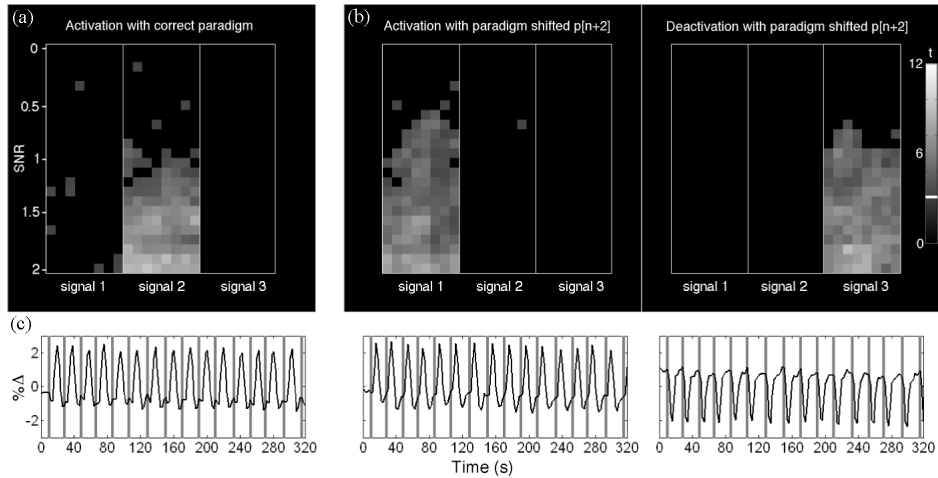

Fig. 7.   Simulation results from the GLM-t method. (a), Significant voxel-paradigm t-maps obtained from the correct paradigm. (b), Significant voxel-paradigm t-maps obtained from the paradigm advanced by 4 seconds (t-threshold indicated by white bar on scale). (c), Averaged voxel time series (events indicated by vertical gray bars).

large for fMRI; however, the detection of noisy signals are also appreciable in the figure. Moreover, the delays match those seen in Fig. 4(a). The stability of the proposed method was verified by three repeated analyses on three simulated data sets. In all cases the same three clusters were identified from a range of $[24, 30]$ clusters whose correlation coefficient estimate differed on average by 3.8×10^{-4} within a data set, and by 0.034 between data sets.

Results using the GLM-t testing for the normal hemodynamic response are computed for comparison. Two, one-tailed tests are done to detect a significant increase or decrease in voxel intensity, referred to as "activation" or "deactivation," respectively. Summary results from the GLM-t method are shown in Fig. 7. Using the paradigm that generated the data shown in Fig. 4(b), only signal 2 is identified as an activation as shown in Fig. 7(a). The deactivation finds no signal voxels. Signals 1 and 3, however, are identified by advancing the paradigm by 4 seconds ($p[n] \rightarrow p[n + 2]$) as suggested by the time-to-peak of the simulated signals in Fig. 4(a). The significance threshold for the GLM-t is $p < 0.001$. This value was chosen ad hoc to approximate the significance threshold of the proposed method so that

the voxel maps in Figs. 6 and 7 can be visually compared across the SNR range. Indeed, the structure of the GLM t-maps are similar to the r-maps from the proposed method in Fig. 6, as expected from previous correlation analyses [16]. Thus, similar results between methods are observed as their respective significance thresholds, indicated by the white bar across the gray scales, increase proportionately. Both methods have comparable sensitivity. This can be seen by recording the SNR of the weakest signal voxel correctly identified in each of the 24 image slices. The weakest signal voxel correctly identified by the proposed method has median SNR $= 0.51$ and worst case SNR $= 0.78$; similarly, the GLM-t method has median SNR $= 0.56$ and worst case SNR $= 0.83$.

The major difference seen between the methods is related to specificity, namely, the ability to distinguish between the three response signals. In general, the proposed method is capable of distinguishing various signals automatically, while the GLM-t requires informed adjustments of the paradigm to achieve similar results. Even with the necessary adjustments, the GLM-t method does not discern signals 1 and 2 as well as the proposed method. In this simulation, the proposed method seldom misclassifies these signal voxels. The worst case misclassification by the proposed method occurred in a voxel with SNR $= 0.78$. The strongest signal voxels misclassified by the GLM-t method, however, occur at higher signal strengths in voxels with a median SNR $= 0.56$, and worst case SNR $= 2.00$. This difference is apparent when comparing the SNR level of the confounded voxels in cluster 3 of Fig. 6(b) with those of Fig. 7(a). Both methods identified signal 3 with perfect specificity.

Finally, we applied the proposed method to a set of ten resting fMRI data sessions to determine if clusters could be falsely selected. To do this, we applied the event-related stimulus sequence shown in Fig. 4(b) to these data, and the Bayesian model correctly rejected all clusters. The largest stimulus-related centroid correlation observed in these clusters had a coefficient of 0.18, which led to a probability of being significantly different from the global signal of correlation $\alpha = 0.06 \pm 0.14$ (mean $\pm$ SD) with corresponding Bayes' error $p = 0.21$.

B. Acquired BOLD fMRI data

Summary results from the proposed method and the GLM-t method are shown in Fig. 8(a) and (b), respectively. Since the subject performed a visual feedback controlled motor task with her right hand, the BOLD response signals expected according to previous neuroimage research should be localized in the left sensorimotor cortex near the "hand area" and the supplementary motor area [27]–[29]. Secondly, responses involving the basal ganglia and thalamus are also possible, since they have also been implicated in sensory and motor function [30]. In particular, the putamen and caudate nucleus of the basal ganglia has previously been identified in two neuroimage studies utilizing a visual feedback controlled motor task [29], [31]. We do not expect to observe the deactivation of a default network, since these were not identified in the aforementioned studies. Lastly, although the task is visually cued, activations

in the visual area are not expected because the cue represented only a small proportion of the visual field as was described in section II. E.

In Fig. 8(a), the Bayesian hierarchical model from the proposed method identified clusters 8, 17, and 24 as being significantly different from the global signal. Their correlation distributions are marked with the symbol * in the box plot. Indeed, the expected brain regions for the performed task are identified, and the responses in the cerebral cortex (clusters 8 and 24) are distinguished from those of the basal ganglia and thalamus (cluster 17). The cluster centroid plots superimposed over the motor responses (gray vertical bars) also reveal the BOLD signal differences between the identified clusters. These results are compared with the GLM-t analysis in Fig. 8(b), which shows the maximum intensity projection of the activation t-map. The expected regions identified by the proposed method also appear in the GLM-t analysis and are indicated by the anatomical overlay sections A, B, and C corresponding to clusters 8, 24, and 17, respectively. No significant voxels were identified in the deactivation map.

Although the expected regions were identified by both methods, differences in other regions can be seen. In particular, an unexpected bilateral response in the parietal cortex seen in section A of the GLM-t map yet only faintly appears in cluster 8 of the proposed method. To compare these regions, we computed their voxel's time sequences and measured their correlation to the pressure signal generated during the hand-grip events. The sensorimotor response signal had a higher correlation coefficient $r(\text{sensorimotor}, \text{hand}) = 0.59$ than the parietal response $r(\text{parietal}, \text{hand}) = 0.31$. Nevertheless, both regions were remarkably correlated to the hand signal. The second notable difference is that the basal ganglia response seen in cluster 17 of the proposed method does not appear in the GLM-t map, which only faintly shows the thalamic response. In light of the simulation results, which reveal how the standard GLM can overlook responses that differ from the assumed model, it is plausible that these differences stem from the variability of the hemodynamic response throughout the brain, as previously demonstrated [8], [32]. Moreover, the notable differences of the BOLD signals plotted in Fig. 8(a) from cortex in cluster 8, and basal ganglia in cluster 17 also support this explanation.

IV. DISCUSSION

Data-driven fMRI cluster analysis seems beneficial for the study of the complex regulatory mechanisms of the cerebrovasculature, since in CVD, unanticipated changes can take place in the BOLD response signal [1]–[5]. However, these exploratory techniques are hindered by the difficulty in determining the statistical significance of the identified clusters. For example, the fuzzy k-means offers a unique advantage over the standard k-means or ISODATA algorithm, since its fuzzy membership functions can be used to determine the optimal k number [6], [17]. Unfortunately, the task of *directly* integrating fuzzy logic algorithms into a probability space remains unresolved [33]–[35]. Here we addressed this problem by using a Bayesian hierarchical model for the correlation
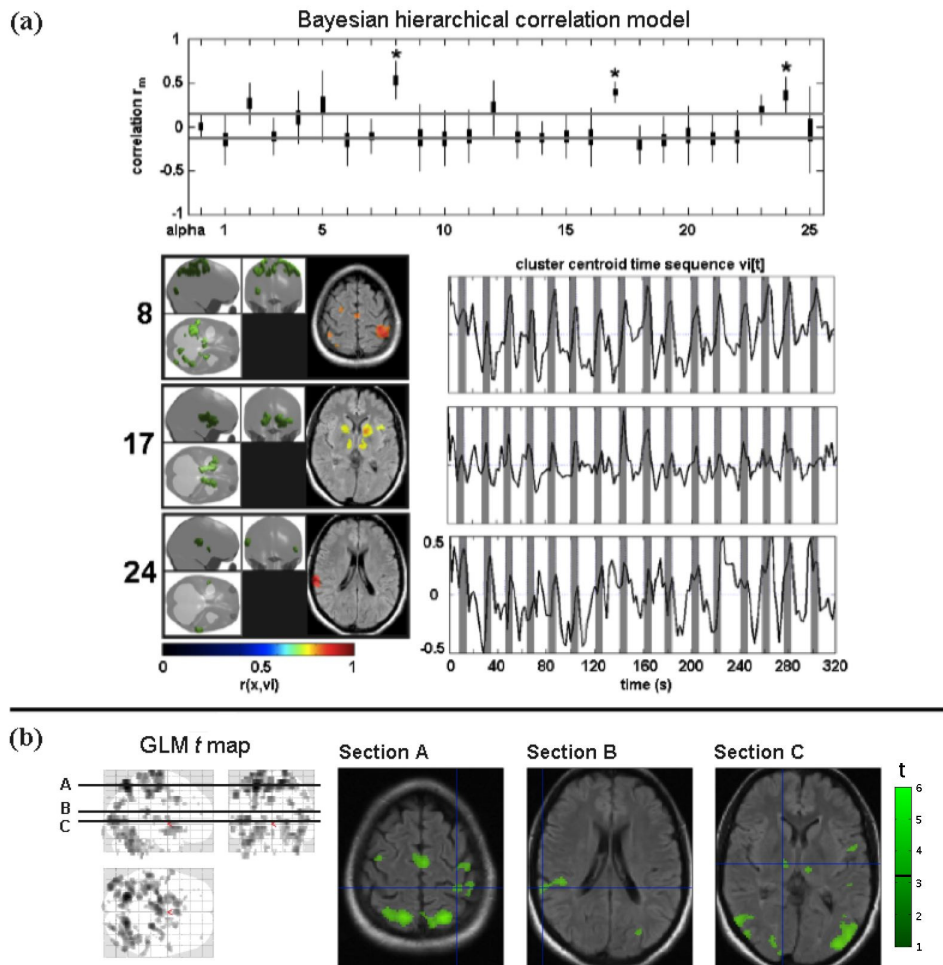
Fig. 8. Results from fMRI data acquired during an event-related right-hand motor task. (a), Proposed method: box plots of paradigm-cluster correlation values from the entire data set α and each cluster β_i ($i = 1, \ldots, k$). The * symbol denotes clusters that are significantly different from α with $p < 0.05$. The significant cluster voxels are shown in green from three computer rendered views, and the anatomical overlay, in radiological convention. Centroid plots for all clusters are shown with the same scale. (b), GLM t method: significant voxel-paradigm t-map shown in three maximum intensity projections. Anatomical overlay sections A, B, and C reveal the brain regions (blue cross-hairs) identified in agreement with clusters 8, 24, and 17, respectively, from the proposed method (t-threshold indicated by black bar on scale).

structure of the clustered data in order to determine whether any one cluster possessed a correlation distribution significantly different from the others. Our simulation and fMRI results show that the proposed data-driven method is capable of identifying and discriminating distinct BOLD response signals with a detection performance that is comparable to model-driven approaches based on the GLM. We emphasize, however, that this comparison should not be viewed as an indication to replace model-based methods. On the contrary, our results suggest that data-driven and model-driven analysis are perhaps best utilized as complementary tools.

In response to methodological concerns about fMRI analysis [12], [13], a statistical commentary about the multiple comparisons framework and the Bayesian hierarchical model has been recently published by Lindquist and Gelman [14]. The authors argue that the multiple comparisons framework in essence can only provide inference at the voxel level. Since voxel interactions do indeed exist in fMRI data, corrections of the p-value threshold are essential to obtain a reasonable estimate the type I error rate of the tested hypothesis. Considerable

research has addressed this issue using the GLM framework [36]–[38], perhaps most elegantly appealing to the theory of smooth random fields [39]–[41]. Nonetheless, on their own, these corrections provide a scalar that uniformly adjusts the confidence interval around a point estimate, i.e., keeping all point estimates stationary as though they were statistically independent. On the other hand, hierarchical models can represent inter-voxel dependence directly by "shrinking" the confidence interval of each point estimate toward the regional and global means. Moreover, global signal correction strategies (e.g., [18], [42], [43]) are not required. Nevertheless, the efficacy of the Bayesian hierarchical model remains to be determined for BOLD fMRI signal detection and discrimination problems.

While multilevel models have already been applied to fMRI analysis [10], [11], to our knowledge, this is the first implementation of a Bayesian hierarchical model for data-driven fuzzy cluster analysis. We note that a probabilistic, data-driven framework has already been developed for independent component analysis (PICA) for fMRI by Beckmann and Smith

[44]. However, the differences between PICA and the proposed method are considerable. For example, PICA decomposes the fMRI time series data into a set of orthogonal signals, while the proposed method uses fuzzy k-means, which does not have an orthogonality constraint. Secondly, PICA infers voxel significance based on comparisons between each detected signal and the residual Gaussian noise model, while the proposed method infers significance by assessing the salience of each detected signal relative to the global signal. Finally, PICA is not a hierarchical model.

In our implementation, we used a fuzzy clustering algorithm to systematically generate regions of interest based on voxel time series correlation. At the voxel level, our simulation results show comparable behavior between the proposed method and the GLM-t over a wide range of physiologically plausible BOLD signals at various SNR levels. Although we only used known response signals in our simulation study, the proposed method is not signal dependent and may be used to discover arbitrary response signals. This is especially important in two applications: i) event-related fMRI because short stimuli (under 4 seconds) are known to evoke responses which are non-linearly related to the sequence of stimuli and, hence, difficult to model a priori [45], [46]; and ii) clinical studies of CVD patients because of the possible alterations in cerebral blood flow caused by age, disease, and medication [1]–[5].

The proposed method remains in a preliminary stage and is not without limitations. There is currently no approach to make statistical comparisons between two particular clusters. While such future work would be beneficial to confirm cluster differences, here we began by empirically verifying that the clustering algorithm appropriately separates voxels according to the different signals utilized in fMRI data simulation. The apparent differences in the BOLD signals identified within the acquired fMRI data also motivate the development of intra-cluster comparisons. Second, while the fuzzy k-means cluster analysis does not require a hemodynamic response model *a priori*, it nevertheless makes implicit assumptions about the BOLD time series data such as temporal independence. An improvement would be to account for this dependence between measurements by using autoregressive models to account for the temporal autocorrelation of time series as previously described within [47] or before [48] Bayesian analysis. Third, refinements of the parametric density models of Bayesian hierarchical model could also lead to improvements. For example, we empirically chose to model the likelihood of mean Pearson correlation coefficients as a Gaussian function, since since we did not observe the mean correlation approaching ± 1 in our data, nor exhibiting marked skew. In general, however, the distribution of Pearson correlation coefficients can take on more complicated forms [49]. Additionally, one could modify the likelihood function to account for the prominent spikes appearing near the mean value of the histograms in Fig. 2(a), which is also pointed out by Lindquist and Gelman [14]. For this reason, the Gibbs sampling technique was implemented so that more detailed likelihood and prior functions can be evaluated without requiring a closed-form formula for the joint posterior distribution. Although cluster analysis seems to offer unique advantages in neuroimage analysis, it remains to be

shown how its assumptions affect signal detection performance in comparison to those based on hypothesis-testing statistics.

Extending the proposed method to group and population level analyses remains a challenging problem for future work. A major issue in performing group analysis with the proposed method is how to account for subject differences in the shape and location of identified regions. A simple approach might aggregate the identified regions from all subjects onto a single activation map in order to determine which voxels are most frequently activated. However, this approach would yield a group analysis where all cluster-based membership is lost. Perhaps a comprehensive clustering approach for group and population analysis could utilize a second hierarchical model to estimate a group template for the clusters identified in the individual subjects. Along this line, Kim *et al.* [11] have recently proposed an approach to tackle this problem in two dimensions by using a Bayesian hierarchical model to analyze cluster location and shape. Their approach could be instrumental in providing a common ground for group and population inference based on individual subject cluster shape, location, and hemodynamic signal features, such as those addressed here.

In conclusion, the proposed Bayesian hierarchical model for data-driven fMRI cluster analysis is encouraging and represents a significant step toward the development of an objective framework to identify, and estimate the significance of, distinct hemodynamic responses within the cerebrovascular system.

REFERENCES

[1] D. Attwell and C. Iadecola, "The neural basis of functional brain imaging signals," *Trends Neurosci.*, vol. 25, no. 4, pp. 621–625, 2002.

[2] M. D'Esposito, L. Y. Deouell, and A. Gazzaley, "Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging," *Nature Rev. Neurosci.*, vol. 4, pp. 863–872, 2003.

[3] M. Burke and C. Bührle, "BOLD response during uncoupling of neuronal activity and CBF," *NeuroImage*, vol. 32, pp. 1–8, 2006.

[4] H. Girouard and C. Iadecola, "Neurovascular coupling in the normal brain and in hypertension, stroke, and alzheimer disease," *J. Appl. Physiol.*, vol. 100, pp. 328–335, 2006.

[5] A. Kocharyan, P. Fernandes, X.-K. Tong, E. Vaucher, and E. Hamel, "Specific subtypes of cortical GABA interneurons contribute to the neurovascular coupling response to basal forebrain stimulation," *J. Cereb. Blood Flow Metab.*, vol. 28, no. 2, pp. 221–31, 2008.

[6] M. Jarmasz and R. L. Somorjai, "Exploring regions of interest with clustering analysis (EROICA) using a spectral peak statistic for selecting and testing the significance of fMRI activation time series," *Artif. Intell. Med.*, vol. 25, pp. 45–67, 2002.

[7] C. Gómez-Laberge, A. Adler, I. Cameron, T. Nguyen, and M. J. Hogan, "Selection criteria for the analysis of data-driven clusters in cerebral fMRI," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 10, pp. 2372–2380, 2008.

[8] J. Zhou and J. C. Rajapakse, "Fuzzy approach to incorporate hemo-dynamic variability and contextual information for detection of brain activation," *Neurocomputing*, vol. 71, pp. 3184–3192, 2008.

[9] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. Boca Raton, USA: Chapman & Hall/CRC Press, 2004.

[10] W. Penny and K. Friston, "Mixtures of general linear models for functional neuroimaging," *IEEE Trans Med Imaging*, vol. 22, pp. 504–514, 2003.

[11] S. Kim, P. Smyth, and H. Stern, "A Bayesian mixture approach to modeling spatial activation patterns in multisite fMRI data," *IEEE Trans Med Imaging*, vol. 29, pp. 1260–1274, 2010.

[12] E. Vul, C. Harris, P. Winkielman, and H. Pashler, "Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition," *Persp. Psych. Sci.*, vol. 4, pp. 274–290, 2009.

[13] M. D. Lieberman, E. T. Berman, and T. D. Wagner, "Correlations in social neuroscience aren't voodoo: Commentary on vul et al. (2009)," *Persp. Psych. Sci.*, vol. 4, pp. 299–307, 2009.

[14] M. A. Lindquist and A. Gelman, "Correlations and multiple comparisons in functional imaging," *Persp. Psych. Sci.*, vol. 4, pp. 310–313, 2009.

[15] J. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern*, vol. 3, pp. 32–57, 1973.

[16] X. Golay, S. Kollias, G. Stoll, D. Meier, A. Valavanis, and P. Boesiger, "A new correlation-based fuzzy logic clustering algorithm for fMRI," *Magn. Reson. Med.*, vol. 40, pp. 249–260, 1998.

[17] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* New York, USA: Plenum, 1981.

[18] E. Zarahn, G. K. Aguirre, and M. D'Esposito, "Empirical analyses of BOLD fMRI statistics," *NeuroImage*, vol. 5, pp. 179–197, 1997.

[19] K. Murphy, R. M. Birn, D. A. Handwerker, T. B. Jones, and P. A. Bandettini, "The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?" *NeuroImage*, vol. 44, pp. 893–905, 2009.

[20] H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory.* Boston, USA: Harvard Business School, 1961.

[21] K. J. Friston, P. Fletcher, O. Josephs, A. Holmes, M. D. Rugg, and R. Turner, "Event-related fMRI: characterizing differential responses," *NeuroImage*, vol. 7, pp. 30–40, 1998.

[22] S. A. Huettel, J. D. Singerman, and G. McCarthy, "The effets of aging upon the hemodynamic response measured by functional MRI," *NeuroImage*, vol. 13, pp. 161–175, 2001.

[23] R. B. Buxton, K. Uludağ, D. J. Dubowitz, and T. T. Liu, "Modeling the hemodynamic response to brain activation," *NeuroImage*, vol. 23, pp. S220–S233, 2004.

[24] L. M. Carusone, J. Srinivasan, D. R. Gitelman, M. M. Mesulam, and T. B. Parrish, "Hemodynamic response changes in cerebrovascular disease: implications for functional MR imaging," *Am. J. Neuroradiol.*, vol. 23, pp. 1222–1228, 2002.

[25] A. C. Roc, J. Wang, B. M. Ances, D. S. Liebeskind, S. E. Kasner, and J. A. Detre, "Altered hemodynamics and regional cerebral blood flow in patients with hemodynamically significant stenoses," *Stroke*, vol. 37, pp. 382–387, 2006.

[26] J. Röther, R. Knab, F. Hamzei, J. Fiehler, J. R. Reichenbach, C. Büchel, and C. Weiller, "Negative dip in BOLD fMRI is caused by blood flow-oxygen consumption uncoupling in humans," *NeuroImage*, vol. 15, pp. 98–102, 2002.

[27] S. C. Cramer, S. P. Finklestein, J. D. Schaechter, G. Bush, and B. R. Rosen, "Activation of distinct motor cortex regions during ipsilateral and contralateral finger movements," *J. Neurophysiol.*, vol. 81, pp. 383–387, 1999.

[28] P. Nachev, C. Kennard, and M. Husain, "Functional role of the supplementary and pre-supplementary motor areas," *Nat. Rev. Neurosci.*, vol. 9, pp. 856–869, 2008.

[29] A. Sterr, S. Shen, C. Kranczioch, A. J. Szameitat, W. Hou, and B. Sorger, "fmri effects of task demand and feedback accuracy on grip force tracking," *Neurosci Lett*, vol. 457, pp. 61–65, 2009.

[30] E. R. Kandel, J. H. Schwartz, and T. M. Jessel, *Principles of Neural Science*, 4th ed. New York, USA: McGraw Hill, 2000.

[31] R. S. Turner, M. Desmurget, J. Grethe, M. D. Crutcher, and S. T. Grafton, "Motor subcircuits mediating the control of movement extent and speed," *J. Neurophysiol.*, vol. 90, pp. 3958–3966, 2003.

[32] M. W. Woolrich, M. Jenkinson, J. M. Brady, and S. M. Smith, "Fully bayesian spatio-temporal modeling of fmri data," *IEEE Trans Med Imaging*, vol. 23, no. 2, pp. 213–31, Feb 2004.

[33] J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition: Methods That Search for Structure in Data.* New York, USA: IEEE, 1992.

[34] M. Bouguessa, S. Wang, and H. Sun, "An objective approach to cluster validation," *Patt. Recog. Lett.*, vol. 27, p. 1419, 2006.

[35] R. J. Hathaway, J. C. Bezdek, and J. M. Huband, "Scalable visual assessment of cluster tendency for large data sets," *Patt. Recog.*, vol. 39, p. 1315, 2006.

[36] K. J. Friston, A. P. Holmes, K. J. Worsley, J. P. Poline, C. D. Frith, and R. S. J. Frackowiak, "Statistical parametric maps in functional imaging: a general linear approach," *Hum. Brain Mapp.*, vol. 2, pp. 189–210, 1995.

[37] K. J. Friston, A. P. Holmes, and J. B. Poline, "Analysis of fMRI time-series revisited," *NeuroImage*, vol. 2, pp. 45–53, 1995.

[38] K. J. Worsley and K. J. Friston, "Analysis of fMRI time-series revisited—again," *NeuroImage*, vol. 2, pp. 173–181, 1995.

[39] K. J. Worsley, A. C. Evans, S. Marrett, and P. Neelin, "A three-dimensional statistical analysis for CBF activation studies in human brain," *J. Cereb. Blood Flow Metab.*, vol. 12, pp. 900–918, 1992.

[40] R. J. Adler and J. E. Taylor, *Random Fields and Geometry.* New York, USA: Springer, 2007.

[41] J. R. Chumbley and K. J. Friston, "False discovery rate revisited: FDR and topological inference using Gaussian random fields," *NeuroImage*, vol. 44, pp. 62–70, 2009.

[42] A. E. Desjardins, K. A. Kiehl, and P. F. Liddle, "Removal of confounding effects of global signal in functional MRI analyses," *NeuroImage*, vol. 13, pp. 751–758, 2001.

[43] P. M. Macey, K. E. Macey, R. Kumar, and R. M. Harper, "A method for removal of global effects from fMRI time series," *NeuroImage*, vol. 22, pp. 360–366, 2004.

[44] C. F. Beckmann and S. M. Smith, "Probabilistic independent component analysis for functional magnetic resonance imaging," *IEEE Trans Med Imaging*, vol. 23, pp. 137–152, 2004.

[45] G. M. Boynton, S. A. Engel, G. H. Glover, and D. J. Heeger, "Linear systems analysis of functional magnetic resonance imaging in human V1," *J. Neurosci.*, vol. 16, pp. 4207–4221, 1996.

[46] A. L. Vazquez and D. C. Noll, "Nonlinear aspects of the BOLD response in functional MRI," *NeuroImage*, vol. 7, pp. 108–118, 1998.

[47] M. F. Ramoni, P. Sebastiani, and I. S. Kohane, "Cluster analysis of gene expression dynamics," *Proc. Natl. Acad. Sci.*, vol. 99, pp. 9121–9126, 2002.

[48] F. de Pasquale, C. D. Gratta, and G. L. Romani, "Empirical Markov chain Monte Carlo Bayesian analysis of fMRI data," *NeuroImage*, vol. 42, pp. 99–111, 2008.

[49] J. F. Kenney and E. S. Keeping, *Mathematics of Statistics*, 2nd ed. Princeton, USA: Van Nostrand, 1951, vol. 2.