

## SOCIAL NETWORK MODELS FOR IDENTIFYING ACTIVE BRAIN REGIONS FROM FMRI DATA

BY ANA M. BARGO, ABHYUDAY MANDAL\*, LYNNE SEYMOUR,  
JENNIFER MCDOWELL†, AND NICOLE A. LAZAR

*University of Georgia*

Functional magnetic resonance imaging (fMRI) is an important tool for scientists studying brain function. FMRI data are complex in nature: they are massive in size and a low signal-to-noise ratio makes the elimination of some noise prior to model fitting desirable for improved identification of true brain activity. We propose two methods of reducing this noise: generalized indicator functional analysis and a hidden Markov model. Brain regions showing increased fMRI signal while subjects engaged in a visual/spatial motor task are identified using concepts from social network analysis and statistical mechanics. Conditional probabilities of activation given the degree to which pairs of voxels are related are modeled for three groups: people with schizophrenia, their asymptomatic relatives, and control subjects. We compare the conditional probability maps obtained for each group to evaluate for between-group differences in extent of task-related signal.

**1. Introduction.** Functional magnetic resonance imaging (fMRI) is a revolutionary, non-invasive technique for studying the activity of the human brain. The use of this functional neuroimaging procedure has led to advances in understanding the neural systems supporting basic perceptual and motor processes, as well as higher level cognitive processes. Published fMRI studies span a wide range of topics from mapping ocular dominance columns in visual cortex (Cheng, Waggoner and Tanaka, 2001) to development of trust occurring during the evolution of a money exchange game (King-Casas, Tomlin, Anen, Camerer, Quartz and Montague, 2005). The data presented in the current study relate to differences in brain function associated with a liability for developing the debilitating psychiatric illness known as schizophrenia.

Regardless of content area, fMRI data present numerous interesting challenges to statisticians and other data analysts, since they are enormous (hundreds of time points collected at a hundred thousand locations – *voxels* or *volume elements* of the brain), extremely noisy (the signal of interest

---

\*Supported by NSF grant DMS-0905731

†Supported by National Institute of Mental Health grants MH01852 and MH076998

*Keywords and phrases:* Functional Magnetic Resonance Imaging, Generalized Indicator Function Analysis, Hidden Markov Model, Markov Chain Monte Carlo, Probability Maps

is often only a small percentage of the total signal), and correlated both spatially and temporally in complex ways.

Historically, the complicated nature of the data meant that analysis was limited by computing speed and storage, in addition to methodology. As such, certain simplifying assumptions were made to conduct statistical analyses. Hence the earliest types of statistical analysis were quite simple –  $t$  tests at the voxel level, ignoring the spatial and temporal correlations; corrections for multiple testing (sometimes); and subjects analyzed individually. As the inadequacies of these approaches became evident, researchers shifted into more sophisticated analyses; in particular much effort has gone into accounting for the temporal dependence in fMRI data (Worsley and Friston, 1995) and into group level inference on multiple subjects (Lazar, Luna, Sweeney and Eddy, 2002; Lindquist, 2008).

Accounting for spatial correlation has been more difficult, in large part because the patterns of spatial dependence are complex, and do not easily fit the assumptions of traditional statistical spatial models. A simple example to illustrate the point is that the right and left hemispheres of the brain show some redundancy: a functional area in one hemisphere will often have a counterpart in the other hemisphere. That is, voxels that are located in different hemispheres, and hence are physically separated, could have very similar responses to a specific stimulus. More generally, voxels in different parts of the brain may react similarly to particular stimuli; similar patterns of activation are not predicated on physical proximity. Only in recent years have more realistic models for this type of behavior been developed. Recent spatial analyses of fMRI data sets have addressed the connectivity of neighboring voxels (Katanoda, Matsuda and Sugishita, 2002). The Bayesian framework also has been used to model the correlation between voxels which are not necessarily neighbors (Woolrich, Jenkinson, Brady and Smith, 2004; Gössl, Auer and Fahrmeir, 2001). Bowman and colleagues used hierarchical models in conjunction with spatial models to explore functional as well as physical connectivity (Bowman, Caffo, Bassett and Kilts, 2008). One feature of the Bowman et. al. (2008) approach is that it requires anatomical regions of interest to be identified prior to analysis, which may not always be feasible (especially in exploratory work involving single case studies, studies using clinical subject groups or those invoking new paradigms or complex cognitive processes).

In this paper, we take a different approach to the spatial modeling aspect of the problem, drawing on ideas from social networks and statistical mechanics (Geman and Geman, 1984; Wasserman and Faust, 1994). The model we propose accounts for both the task-related activation of the brain at the voxel level and for the functional connections between voxels in the brain. Unlike the model suggested in Bowman et al. (2008), regions of interest are not predefined; instead they are revealed by the results of the social network model.

The paper is organized as follows. Section 2 presents the fMRI paradigm, including visual stimuli, the associated motor responses, and the resulting data. Section 3 describes the statistical models and methods applied. Section 4 shows results of the data analyses. Section 5 concludes with the discussion of future research objectives.

**2. Scientific Background.** FMRI data were collected while subjects performed eye movements in response to visual stimuli (Camchong, Dyckman, Austin, Clementz and McDowell, 2008). The condition of interest is known as an “antisaccade task” and requires that a subject inhibit a glance (“saccade”) to a cue presented in the periphery and instead generate a glance (“antisaccade”) to an unmarked mirror image location (opposite side, same distance from center). An initial glance towards the cue constitutes an error and is typically construed as a failure of inhibition. This study used a standard block design (see Figure 1) in which the antisaccade condition of interest (8 trials each, 25.2 seconds duration) alternated with blocks of fixation trials (22.5 seconds) as a baseline condition.

FMRI data were collected using a GE Signa Horizon LX 1.5T MRI scanner (Milwaukee, WI) at the Athens Orthopedic Clinic MRI Center. Scanning began with the acquisition of a three-dimensional  $T1$ -weighted structural MRI scan for definition of anatomical structures within each brain (spoiled gradient-recall (SPGR) protocol:  $TE=2.8$  msec,  $TR=10.8$  msec, flip angle=20, number of excitations=2, matrix=256-by-256, field of view=24 (resulting in an in-plane resolution of 0.97-by-0.97 mm), slice thickness 1.5 mm, sagittal acquisition, 124 contiguous slices, scan time 5 min 41 sec). Following structural imaging,  $T2^*$ -weighted images were acquired during functional runs (spoiled-gradient pulse sequence (SPGR) with a spiral readout pattern in k-space, (matrix=64-by-64, field of view=24 resulting in an in-plane resolution of 3.75-by-3.75 mm, slice thickness=4 mm),  $TE=40$  msec,  $TR=1912$  msec with two interleaves resulting in an image acquisition time of 3.8 sec; flip angle=77, 24 supratentorial contiguous slices).

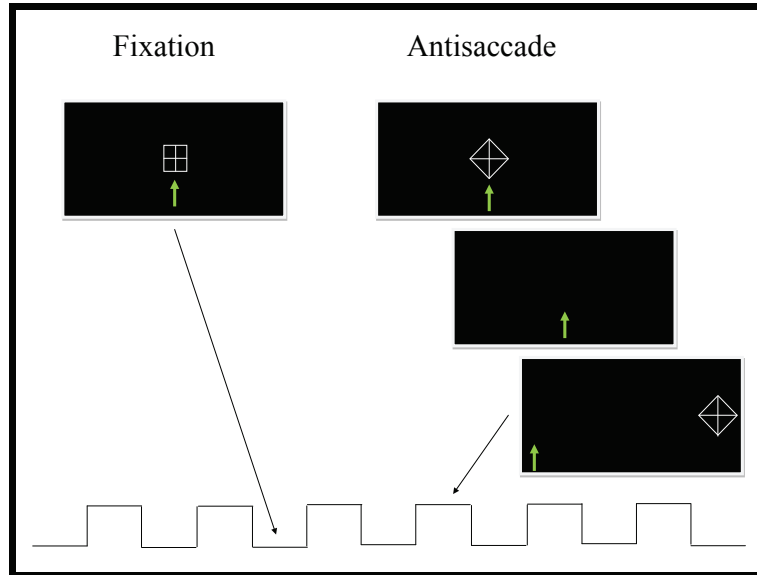


FIG 1. Stimuli used for antisaccade/fixation task. During the fixation condition, the center target was a cross with a bordering square. Participants were instructed to stare at the target for the duration of the block (22.5 sec). During the antisaccade condition, the target changed to a cross with a bordering diamond and started at the center (1700 ms). The central target went off, there was a gap (200 msec) and then a cross with a bordering diamond cue was presented in the periphery (1250 msec). The participants were instructed to look at the target while it was in the center, but when the cue arose in the periphery, to look at the mirror image location of that cue as quickly and accurately as possible, without looking at the cue itself. There were 8 antisaccade trial in each block. The green arrows show the correct eye location for each stimulus presentation (but were not present during the task).

In this dataset, one full functional brain volume consisted of 24 axial slices acquired over 3.8 seconds. Each volume was collected 81 times; 40 volumes were taken when the subjects were performing the antisaccade task (denoted by  $n_A$ ), and 41 of the volumes were taken when the subjects were performing the fixation task (denoted by  $n_F$ ). The data were preprocessed using standard methods in AFNI (Cox, 1996): motion correction, slice timing correction, head motion correction, spatial and temporal filtering, and normalization into Talairach space (Talairach and Tournoux, 1988; Ward, 2000). This article focuses on slice number 28, which in a previous study with a large sample size ( $N=36$ ) was one among a number of slices reliably showing antisaccade related activity (Dyckman, Camchong, Clementz and McDowell, 2007; see Figure 2).

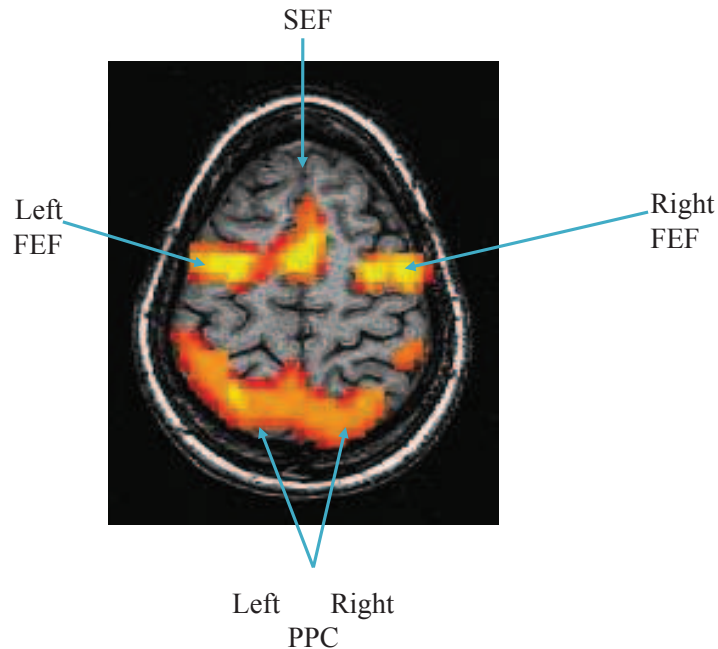


FIG 2. *Antisaccade-related activity in regions similar to those analyzed in the current study: frontal eye fields (FEF), supplementary eye fields (SEF) and posterior parietal cortex (PPC). Adapted from Dyckman et al., 2008.*

FMRI data during the antisaccade/fixation task were acquired for 3 groups of subjects: 1) 16 people diagnosed with schizophrenia (DSM-IV, First, Spitzer, Gibbon, and Williams, 1995), 2) 13 of their biological relatives (primarily unaffected siblings), and 3) 15 healthy control subjects recruited from the community. One goal of the study was to compare antisaccade related fMRI signal between groups.

People with schizophrenia and their biological relatives have deficits in executive control processes such as inhibition, which can be modeled successfully using cognitively complex saccadic eye movement tasks. Saccades exist in a hierarchy of increasingly complex behavior. At the lower end of the hierarchy are simple refixation saccades to a visual target, on which schizophrenia subjects perform normally (Clementz, McDowell and Zisook, 1994; Ettinger, Picchioni, Hall, Schulze, Touloupoulou, Landau, Crawford and Murray, 2006). At the higher end of that hierarchy are antisaccades, which require inhibition of a glance towards a cue and a generation of a glance to the mirror image of that cue (opposite side, same distance from the center), on which

the performance of schizophrenia participants is compromised. For correct antisaccade performance, participants must remember the requirement to generate a saccade to the peripheral cue's mirror image location (same amplitude, opposite direction), inhibit a reflexive saccade toward that cue upon presentation, and then program and generate a saccade to the cue's mirror image location. An error is defined as an initial glance toward the peripheral cue. Schizophrenia subjects make more antisaccade errors than control groups (Calkins, Iacono and Ones, 2008; McDowell, Myles-Worsley, Coon, Byerley and Clementz, 1999, Hutton, Cuthbert, Crawford, Kennard, Barnes and Joyce, 2001), and their relatives show an error rate that is typically intermediate between groups.

The generation of correct antisaccade responses requires a wide-spread neural system that supports attention, inhibitory control and generation of a saccade to a specific location in the absence of a visual target. In a number of studies, the basic saccadic system has been identified (see Figure 2 above; Dyckman et al., 2008; Sweeney et al., 2007). Generally the system shows activity of occipital (visual) cortex for target perception and registration, parietal cortex for spatial attention, the motor regions of frontal and supplementary eye fields for programming and execution of the saccade and prefrontal regions for inhibitory control. The current slice was selected for its good representation of parietal and motor regions. Given differences in antisaccade error rates between the three groups, it is expected that there will also be differences in extent of activity in some of the regions supporting this behavior, particularly between the schizophrenia and control subjects.

### 3. Statistical Methods.

3.1. *Statistical Background.* The model we propose is inspired by two areas of statistics: social networks and statistical mechanics. In each, identifying the dependence structure is the key feature of the model.

3.1.1. *Social Networks.* Social network models address dependence induced by social relations (Wasserman and Faust, 1994). Recent effort has focused on inference for the dependence parameter (Hunter, Goodreau and Handcock, 2005). In most social network models, the probability of a relational tie between two individuals increases as the characteristics of the individuals become more similar. We use this feature to apply social network models to fMRI data.

The social environment can be expressed as patterns or regularities in relationships among interacting units, or actors (Wasserman and Faust, 1994). Two types of variables are associated with network data: *structural* and *composition*. Structural variables are measured on pairs of individuals, or *actors*; composition variables are measurements of actor attributes. In our application of the social network model, voxels are the actors; the relationships between voxels in the brain are the structural variables, but we do not incorporate composition variables, which could include covariates about the subjects.

There are several schemes available for describing social network data. In this paper, we use the sociometric notation, where information about the actors is described via a matrix called the sociomatrix in the social network literature. Here we use the terminology co-activation matrix to emphasize the nature of relationships we exploit.

The model we use for the conditional probability of observing an active voxel given the activation status of the other voxels, as described in Equation (4) below, is a member of the  $p^*$  models (Strauss and Ikeda, 1990; Geyer and Thompson, 1992). These models are widely used in social network literature. Here the parameters are estimated via pseudo-likelihood estimation (Besag, 1974b) and Markov Chain Monte Carlo methods (Besag, 1974a; Geman and Graffigne, 1986; Casella and George, 1992).

3.1.2. *Statistical Mechanics.* The exponential family models used in the social networks literature have origins in statistical mechanics for modeling particle interactions on a lattice  $B$ . The Gibbs distribution gives joint probabilities for the particle interactions (Georgii, 1988). In the case of fMRI data, at a given time scan, the lattice  $B$  consists of all voxels in the acquisition matrix (the plane acquired by the scanner during an experiment). The set of all possible configurations at a slice in the system is defined to be  $\mathcal{B}$ . Here we consider a two-state system: a voxel is either active (represented by 1) or inactive (represented by 0). Although it is possible to represent more states in the model via the configuration set  $\mathcal{B}$ , we choose instead to let the matrix  $B$  denote the activation status, and represent different states via the co-activation matrix  $\mathcal{W}$ .

The product space,  $\Omega = \mathcal{B}^B$ , is called the configuration space (Georgii, 1988). Let  $\mathcal{H}^\phi$  be a Hamiltonian which assigns to each configuration  $y$  in  $\Omega$

a potential energy  $\mathcal{H}^\phi(y)$ . The potential for two voxels  $i$  and  $j$  is denoted by  $\phi_{ij}$  and is defined as  $y_j w_{ij}$ , where  $y_j \in \mathcal{B}$  is the activation level of voxel  $j$  and  $w_{ij}$  is the value of the associated co-activation matrix  $\mathcal{W}$ . The energy associated with any element  $y_i$  in the lattice  $B$  is therefore defined to be  $\mathcal{H}^\phi(y_i) = \sum_{j \in B} \phi_{ij} = \sum_{j \in B} y_j w_{ij}$ .

A given Gibbs specification can be defined in terms of many different potentials. Since the energy  $\mathcal{H}^\phi$  is dependent on the number of voxels (the more voxels there are, the higher the value will be), we appeal to the *equivalence of potentials* (Georgii, 1988) to ensure the numerical stability of all calculations. The equivalence of potentials states that any two potentials are equivalent to each other if their difference is constant. Therefore, since oftentimes the value of  $\mathcal{H}^\phi$  is so high that the Gibbs distribution in (1) is the same for all values of  $\phi$ , we reduce the value of all Hamiltonians by a constant.

The Gibbs distribution is defined to be

$$(1) \quad P(Y = y) = \frac{\exp[-\mathcal{H}^\phi(y)]}{\sum_{x \in \Omega} \exp[-\mathcal{H}^\phi(x)]}.$$

**3.2. Building the Sociomatrix.** FMRI data have low signal-to-noise ratio, where the signal is variation in the data that is explained by the specific task that subjects are performing; noise is the variation that is not explained by the task. The low signal-to-noise ratio implies that task-related activation is relatively hard to discern. Here we separate the signal from the noise using Generalized Indicator Function Analysis (GIFA) and the Hidden Markov Model (HMM) before constructing the co-activation matrix  $\mathcal{W}$ .

The co-activation matrix  $\mathcal{W}$  is the key part of our proposed model. We are interested in the strengths of the relationships between pairs of voxels in the brain and therefore we deviate from the standard practice of using a sociomatrix in which the elements are simply indicators of a relationship. Instead, we build a sociomatrix that summarizes the strength of the relationship between voxels.

There are already several methods in common use for separating the signal from the noise in fMRI data (Lazar, 2008). Two of the most popular methods are principal component analysis (PCA) (Johnson and Wichern, 2002; Ulfarsson and Solo, 2006) and independent component analysis (ICA) (Hy-



varinen and Oja, 2000; Guo and Pagnoni, 2008). PCA assumes that the data are correlated and attempts to transform them into uncorrelated orthogonal components, while ICA assumes that the data are an additive mixture of an unknown number of mixed signals, and tries to separate their origin without any knowledge of the source signals. An important distinction between ICA and PCA is that the former assumes non-normality and independence of the signals. While both methods are widely used in analyzing fMRI data, we found that PCA and ICA do not provide an appropriate basis for building a sociomatrix that can be used with the social network model (this is demonstrated later in Figure 3 below).

The failure of PCA and ICA for our specific purpose provides the motivation for considering alternative noise removal techniques; here, as already mentioned, we look at GIFA and HMM. GIFA is a model-free method for analyzing imaging data which has the advantage of using the stimulus associated with each scan to separate signal from noise. This allows one to use the knowledge that certain images were acquired during fixation, whereas others were acquired during antisaccade performance. GIFA has never been used on fMRI data before, though it has been previously used in other imaging modalities, such as PET and MRI (Yokoo, 2001). HMM assumes that the data are a Markov process with a number of unobservable different signal levels. This is useful in constructing the co-activation matrix: different levels of activation are of interest when analyzing functional relationships in the brain. Here we would like to find the important signal levels associated with the difference matrix. This application of HMM is novel in fMRI data, though it has been used in other imaging techniques, such as quantitative MRI analysis (Zhang, 2008).

In our application, the difference matrix  $X$  has a total of  $40 \times 48 = 1920$  voxels. Any one acquisition matrix in any time point is denoted by  $B$ . Let a fixation scan be denoted by  $F$ , and an antisaccade scan be denoted by  $A$ . The number of images is 81 per subject. The activation vector,  $y$ , is a binary vector of activation states (0 for an inactive voxel, 1 for an active voxel). This vector is an indicator of activity based on the difference matrix  $X$ . Voxels with a low measured signal (500 or less for our dataset) are considered inactive and assigned the value of 0, whereas the other (active) voxels are labeled as 1. Note that this is just a crude screening to yield an initial classification of voxels into one of two states; the social network model that we apply to the sociomatrix will depict in more detail the differing activation levels of the voxels.

3.2.1. *Construction of  $\mathcal{W}$  Using PCA and ICA.* PCA decomposes the covariance matrix of a multivariate dataset in order to transform it into a small number of uncorrelated variables, or principal components (Johnson and Wichern, 2002). ICA (Hyvarinen and Oja, 2000) defines an additive model for multivariate data, assuming that these data are nongaussian and mutually independent. We first applied these two commonly-used methods to our data in an attempt to identify the signal derived from the antisaccade task and separate it from the unrelated noise.

Selecting the number of signal components which adequately explain the variability of the data matrix is not a trivial problem for PCA, ICA (or, for that matter, GIFA). For PCA, we only use the first principal component, since it explains approximately 80% of the variance. We considered using further components from PCA. However, we found that subsequent components include variation due to fluctuation in scans outside of the brain, as well as between-subject variation, thereby introducing noise of a different type.

When performing ICA, the order of components is unimportant, as they are statistically independent and a blind method of separation is used (Hyvarinen and Oja, 2000). The goal for ICA, as opposed to PCA, is not dimension reduction, but to decompose a given signal into independent components. Therefore, we use all  $n$  components to ensure all signals are used.

Once we isolate the PCA and ICA signals, we derive the co-activation matrix  $\mathcal{W}$  for each case using  $\gamma_1 X$  (for PCA) and  $\sum_i \gamma_i X$  (for ICA). The second and third rows of Figure 3 illustrate the resulting difference matrix using PCA and ICA. Specifically with controls and schizophrenia subjects, neither PCA nor ICA capture any activation inside the brain. Both methods rather capture boundary activation (most likely due to differences in brain sizes even after the transformation into Talairach coordinates). Therefore, Figure 3 shows that PCA and ICA oversmooth the data for some groups, indicating that these two approaches are inappropriate for further analysis. We do not fit the social network model of Section 3.3 with a co-activation matrix based on PCA or ICA.

3.2.2. *Construction of  $\mathcal{W}$  by Generalized Indicator Function Analysis.* GIFA separates the signal from the noise in multivariate data using the weighted difference of signal and noise as a statistic (Yokoo, Knight and Sirovich, 2001). With this method, we decompose each scan for a slice into

two parts much like a standard regression: the signal and the noise. Once the noise is separated from the signal, we find the structure of the co-activation matrix.

In GIFA, we perform a modified PCA on  $X$  using the signal and noise covariance matrices (defined below) instead of the regular variance-covariance matrix. Let  $\bar{B}$  be  $\frac{1}{n_{img}}(n_A\bar{A} + n_F\bar{F})$ . Here,  $n_{img}$  denotes the total number of images in a group (multiple subjects each with multiple scans). Hence,  $\bar{B}$  denotes the average over all the subjects in a particular group. The *signal covariance matrix* for any group is given by

$$K_S = (\bar{F} - \bar{B})^T(\bar{F} - \bar{B}) + (\bar{A} - \bar{B})^T(\bar{A} - \bar{B}),$$

which is the overall variation of fMRI values attributed to fixation and antisaccade. The *noise operator* covariance matrix for any group is

$$K_N = \frac{1}{n_F + n_A - 1} \left( \sum_{i=1}^{n_F} (F_i - \bar{F})^T (F_i - \bar{F}) + \sum_{i=1}^{n_A} (A_i - \bar{A})^T (A_i - \bar{A}) \right),$$

which is the individual variation of individual fixation and antisaccade values (Yokoo et. al., 2001).

Maximizing the difference between signal and noise is recast as the problem of solving the system of equations  $(K_S - \delta K_N)t = \gamma t$ . Here  $\delta$  acts as a threshold to measure the quality of the signal. The solution is the matrix of eigenvectors of the newly defined covariance matrix,  $K_t - \delta K_N$  (Yokoo et. al., 2001). The solutions to that system of equations are  $(\gamma_k, t_k)$  for  $k = 1, \dots, n=1920$ . Here we choose  $\delta = 4$  as it captures approximately 95% of the signal in the first solution ( $k = 1$ ); it captures the signal components which are two standard deviations or higher from the mean signal. Therefore, we use the first order projection of the data,  $\gamma_1 X$ , to represent the denoised difference matrix  $X$ .

Once the activation due to task-related differences has been identified, we derive the co-activation matrix using  $\gamma_1 X$ . The continuous values of the signal matrix,  $\gamma_1 X$ , are further rescaled to the values,  $0, 1, \dots, 5$ , indicating different levels of activation (the closer the value is to 5, the higher the activation level of the voxels). Preliminary studies show that the results of applying the model to different co-activation matrices are quite robust to different discretizations. Next we reshape  $\gamma_1 X$  into the vector  $v = [v_i]$ ,  $i = 1, \dots, n$ . Finally, given a value for  $\alpha$ , where  $\alpha = 1, 2, \dots, 5$  and the signal

vector  $v$ , the  $(i, j)$  element of the co-activation matrix is calculated as

$$(2) \quad w_{ij} = (v_i + v_j)^\alpha,$$

where  $i, j = 1, \dots, n$ .

**3.2.3. Construction of  $\mathcal{W}$  Using Hidden Markov Model.** The hidden Markov model (HMM) is typically used when the data to be analyzed originate from a medium with excessive noise (Rabiner and Juang, 1986). Here we use HMM for separating the task-related signal from noise. In this approach we consider that each voxel has an unknown label that represents the level of contrast of antisaccade to fixation. We assume that voxels are conditionally independent, identically distributed following a normal distribution with mean and standard deviation conditional on the label. The unknown labels in turn follow a nearest-neighbor structure, where neighboring voxels tend to share the same label. We use two algorithms (Hoshen-Kopelman and K-means) in order to determine an appropriate number of labels for each group separately (Hoshen, Berry, and Minser, 1997; Hoshen, 1998). The HMM takes into account the noisiness of the data in the lower level of the hierarchical structure (individual voxels), whereas the higher level of the hierarchy (the labels associated with the voxels) captures contrast related to the antisaccade versus fixation task.

If two voxels  $z$  and  $z^*$  share an edge, we call them *neighbors*. The set of neighbors of voxel  $z$  is denoted by  $\mathcal{N}_z$ . Each voxel  $z$  has an observed difference intensity,  $x_z$ , and a hidden label  $l_z$ . The unknown matrix of labels  $L$  is assumed to create an  $M$ -state Markov random field (MRF) on the state space  $\mathcal{S} = \{1, \dots, M\}$ .

We fit a two-level hierarchical model. The higher level describes the structure of the MRF with probability distribution function:

$$P(L = l|\beta) = \frac{\exp\left\{-\beta l \sum_{z^* \in \mathcal{N}_z} x_{z^*}\right\}}{g(\beta)} \quad \text{for all } z \in \Lambda_n.$$

The normalizing constant  $g(\beta) = \sum_x \exp\left\{-\beta \sum_{z^* \in \mathcal{N}_z} x_{z^*}\right\}$  has  $M^n$  summands, and is not computationally tractable. Given the unknown labels  $L = l$ , the elements of the difference matrix  $X$  are conditionally independent and normally distributed, following the distribution:

$$P(x_z|l_z) = \frac{1}{\sigma_{l_z}} \exp\left\{-\frac{(x_z - \mu_{l_z})^2}{2\sigma_{l_z}^2}\right\}.$$

Write  $\mu = (\mu_1, \dots, \mu_M)^T$  and  $\sigma^2 = (\sigma_1^2, \dots, \sigma_M^2)^T$ , and then  $\theta^T = (\mu, \sigma^2, \beta)^T$ . Here, we estimate  $\mu$ ,  $\sigma^2$ , and  $\beta$  using pseudo-likelihood and Markov Chain Monte Carlo (MCMC). Label identification is done via the Hoshen-Kopelman cluster counting algorithm (Hoshen et. al., 1997; Hoshen, 1998) for initial cluster identification and the K-means algorithm (Hartigan, 1975) for cluster verification.

Once the activation due to task-related differences has been identified, we derive the co-activation matrix using the label mean matrix,  $L_M$ , represented by the means  $\mu_k$ , where  $k = 1, \dots, M$ . The continuous values of  $L_M$  are rescaled to the values of  $0, 1, \dots, 5$  and we reshape  $L_M$  into the vector  $u = [u_i]$ , for  $i = 1, \dots, n$ . Finally, again, similar to equation (2), the co-activation matrix is calculated as

$$(3) \quad w_{ij} = (u_i + u_j)^\alpha,$$

where  $i, j = 1, \dots, n$ . Again, the above discretization is somewhat arbitrary, but as before, the results are similar for different discretizations.

*3.2.4. Results of Denoising the Signal for Construction of  $\mathcal{W}$ .* Figure 3 shows the performance of PCA, ICA, GIFA, and HMM applied to  $X$  from each group. In Figure 3, as well as all other figures, signal is shown on a single axial plane of acquired data, with brain space centered in roughly the center of that area. The brain is oriented so that the front of the head is at the top of the images, the back of the head is at the bottom of the image, and the left and right sides of the head are on the left and right sides of the image, respectively (neurological orientation). The regions that would be expected to show signal are shown in Figure 2 (above) - bilateral PPC, bilateral FEF and SEF. Compared to the raw difference matrices, major improvements are evident using all four methods – the noise outside the brain is virtually eliminated with every method used. The most notable aspect of the data is the high signal fluctuation on the lower left corner of the raw difference matrix (especially for the control group), which is completely smoothed once the first GIFA signal is applied to the difference. Compared to PCA and ICA, GIFA and HMM retain significant activation inside the brain but reduce more noise.

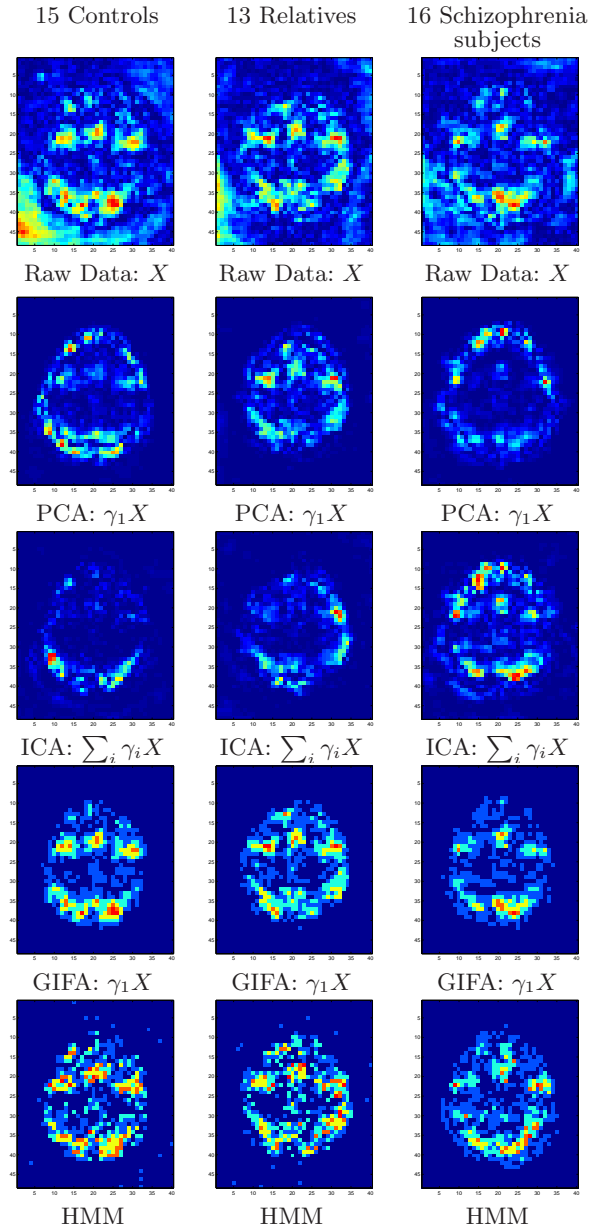


FIG 3. Comparison of PCA, ICA, GIFA, and HMM to Raw Average Difference:  $X$ . For all three groups, PCA and ICA oversmooth the data everywhere. GIFA and HMM, on the other hand, remove noise outside of the brain while keeping BOLD signal inside of the brain intact. This shows a single axial slice with the back of the head at the bottom of the images, and the left and right sides of the head on the left and right sides of the images (neurological orientation).

GIFA appears to work better than ICA and PCA in this case because GIFA takes into account the prior knowledge of when the stimuli were presented, thus yielding better results than a blind approach. Differences are evident in Figure 3, where it is clear that the noise outside of the brain is removed by all methods, but both PCA and ICA appear to oversmooth the data, whereas GIFA and HMM retain the differences of activation inside the brain.

The parameter  $\alpha$  is a tuning parameter which determines how much weight will be given to the values of the co-activation matrix  $\mathcal{W}$ . The higher the tuning parameter, the more weight is given to the higher valued relationships between any two voxels. If two voxels  $i$  and  $j$  are highly active so that the values of  $v_i$  and  $v_j$  (or  $u_i$  and  $u_j$ ) are large, then the value of the co-activation matrix  $\mathcal{W} = [[w_{ij}]]$  is large, and we say that the two voxels coactivate. If neither of the two voxels are activated, then the value of  $w_{ij}$  will be very small; and if one voxel is active and the other is inactive, then the value of  $w_{ij}$  will be small.

*3.3. The Social Network Model.* Once the co-activation matrix is estimated, we fit a social network model to estimate the conditional probability of activation for each voxel. Smith and Fahrmeir (2007) use the traditional Ising model as a prior to spatially smooth the indicator variables in their Bayesian variable selection model. The HMM in Section 3.2.3 uses the ordinary Ising approach at each level of activation for a label  $l$ , which allows us to retain different levels of activation.

One might try to use the traditional Ising model to test for brain activation. However, as is shown in the Appendix, this model simply separates brain from non-brain in the image, and hence is useless for detecting activation within the brain.

Instead of the ordinary Ising model, we use the so-called fully connected Ising model (Strauss and Ikeda, 1990; Wasserman and Faust, 1994; Hoff, 2002; Handcock, 2003; Hunter et. al., 2005; Snijders, Pattison, Robins, Handcock, 2006):

$$(4) \quad P(Y_i = 1 | y_j, j \neq i) = \frac{\exp\{-\beta \sum_{j \neq i} y_j w_{ij}\}}{1 + \exp\{-\beta \sum_{j \neq i} y_j w_{ij}\}},$$

where  $Y_i = 1$  denotes an active voxel. The traditional Ising model incorporates the effect of the nearest neighbors on the conditional probability of

activation of a single voxel. The fully connected Ising model, on the other hand, incorporates the effect of all other voxels in the activation matrix on the conditional probability of activation based on the co-activation of those voxels.

**Pseudo-Likelihood:** The pseudo-likelihood approach (Besag, 1974b) involves multiplying the values of  $P(Y_i = y_i | Y_j = y_j, j \neq i)$  over all  $i = 1, \dots, n$  and maximizing this with respect to  $\beta$ . We call the resulting estimate  $\hat{\beta}_{MPLE}$ . The pseudo-likelihood is:

$$P\mathcal{L}(y, \beta) = \prod_{i=1}^n P(Y_i = y_i | y_j, j \neq i) = \prod_{i=1}^n \frac{\exp\{-\beta y_i \sum_{j \neq i} y_j w_{ij}\}}{1 + \exp\{-\beta \sum_{j \neq i} y_j w_{ij}\}}.$$

Although the pseudo-likelihood approach yields consistent results under certain conditions (Besag, 1974b), the variance of  $\hat{\beta}$  cannot be estimated. However,  $\hat{\beta}_{MPLE}$  is a good starting point for an MCMC chain, from which variance estimates can be obtained.

**Markov Chain Monte Carlo:** The log likelihood is:

$$l(y, \beta) = -\beta \sum_{i=1}^n y_i \left( \sum_{j \neq i} y_j w_{ij} \right) - \log [g(\beta)].$$

The calculation of  $-\log[g(\beta)]$  is numerically prohibitive (Seymour, 1993; Geyer and Thompson, 1992). We use a Markov chain of length  $L$  to approximate the value of  $r_L(\beta)$  and thereby the log-likelihood of  $l(y, \beta)$ , where

$$r_L(\beta) = \frac{1}{L} \sum_{l=1}^L \exp \left\{ -(\beta - \theta) \sum_{i=1}^n y_i^{(l)} \left( \sum_{j \neq i} y_j^{(l)} w_{ij} \right) \right\}.$$

The approximation to the log-likelihood thus is:

$$l_L(y; \beta) = \beta \sum_{i=1}^n y_i \left( \sum_{j \neq i} y_j w_{ij} \right) - \log r_L(\beta) + \log L.$$



#### 4. Results of the Social Network Model.

4.1. *Parameter Interpretation and Model Comparison.* Table 1 shows GIFA and HMM estimates of the dependence parameter  $\beta$  for both MPLE and MCMC. Recall that the tuning parameter,  $\alpha$ , determines how much weight will be given to the values of the co-activation  $\mathcal{W}$  in the distribution. The higher the value of the tuning parameter, the more weight is given to highly active voxels. Overall, the MPLE estimates are fairly close to the MCMC estimates across all three groups and with both methods of calculating the co-activation matrix  $\mathcal{W}$ , when  $\alpha$  is closer to 1. The MCMC parameter is normally distributed (Geyer and Thompson, 1992), and due to this normality, the pseudo-likelihood estimate does fall within the 95% confidence interval of the MCMC parameter for each group of interest. As  $\alpha$  increases, the MCMC and MPLE estimates drift farther apart for all three groups. The MCMC variance estimates also increase as the values of  $\alpha$  increases.

TABLE 1  
GIFA and HMM fits, with MPLE and MCMC estimates, for controls, relatives of schizophrenia subjects and schizophrenia subjects. The chain size used for each of the three groups was 10,000.

		GIFA			HMM		
		Controls	Relatives	Schizo- phrenia	Controls	Relatives	Schizo- phrenia
$\alpha = 1$	$\hat{\beta}_{MPLE}$	-0.2228	-0.2343	-0.2304	-0.1861	-0.2077	-0.1849
	$\hat{\beta}_{MCMC}$	-0.2385	-0.2398	-0.2481	-0.1745	-0.2123	-0.1930
	$\hat{\sigma}_{MCMC}$	0.0111	0.0127	0.0115	0.0209	0.0136	0.0220
$\alpha = 2$	$\hat{\beta}_{MPLE}$	-0.1816	-0.1785	-0.1803	-0.1821	-0.1936	-0.1595
	$\hat{\beta}_{MCMC}$	-0.2385	-0.2398	-0.2481	-0.1959	-0.2171	-0.1945
	$\hat{\sigma}_{MCMC}$	0.0190	0.0236	0.0150	0.0131	0.0217	0.0217
$\alpha = 3$	$\hat{\beta}_{MPLE}$	-0.1653	-0.1602	-0.1652	-0.1744	-0.1797	-0.1455
	$\hat{\beta}_{MCMC}$	-0.1541	-0.1443	-0.1713	-0.1599	-0.1975	-0.1697
	$\hat{\sigma}_{MCMC}$	0.0264	0.0289	0.0148	0.0117	0.0197	0.0200
$\alpha = 4$	$\hat{\beta}_{MPLE}$	-0.1582	-0.1526	-0.1595	-0.1677	-0.1702	-0.1382
	$\hat{\beta}_{MCMC}$	-0.1626	-0.1883	-0.1365	-0.2012	-0.2054	-0.1573
	$\hat{\sigma}_{MCMC}$	0.0286	0.0315	0.0175	0.0154	0.0200	0.0226
$\alpha = 5$	$\hat{\beta}_{MPLE}$	-0.1545	-0.1487	-0.1567	-0.1624	-0.1639	-0.1341
	$\hat{\beta}_{MCMC}$	-0.2443	-0.2560	-0.2362	-0.2031	-0.1859	-0.1543
	$\hat{\sigma}_{MCMC}$	0.0295	0.0308	0.0241	0.0225	0.0219	0.0306

In this model, as with the traditional Ising model, the parameters measure the level of dependence found in the GIFA and HMM matrices, respectively.

The closer this parameter is to zero, the less dependent the data are. In all cases, the traditional test for dependency which has as a null hypothesis that for a group  $g$  there is no dependence (or in other words,  $H_0 : \beta_g = 0$ ) fails, since 0 does not fall within 2 standard deviations of  $\hat{\beta}$ . Thus, the model identifies structure in these data, and further examination of the estimated conditional probabilities is needed.

*4.2. Probability Maps.* Figures 4 and 5 show maps of the conditional probabilities of activation for controls, relatives and schizophrenia patients for GIFA and HMM, respectively. The tuning parameter  $\alpha$  adjusts the emphasis of the signal intensity of the most active voxels in the co-activation matrix. We focus most of the remaining discussion on the conditional probability maps when  $\alpha = 1$ . Table 2 presents the number of voxels with conditional probabilities greater than 0.9 for  $\alpha = 1$ .

TABLE 2  
*Frequency of voxels with conditional probabilities greater than 0.9 for controls, relatives, and schizophrenia subjects when  $\alpha = 1$  using GIFA and HMM.*

GIFA						
	Left Frontal Eye Field (Left FEF)	Right Frontal Eye Field (Right FEF)	Supplementary Eye Field (SEF)	Left Posterior Parietal Cortex (Left PPC)	Right Posterior Parietal Cortex (Right PPC)	Non-Area of Interest
Controls	8	10	7	16	12	0
Relatives	7	7	10	14	10	0
Schizophrenia	1	2	5	7	12	0
HMM						
	Left Frontal Eye Field (Left FEF)	Right Frontal Eye Field (Right FEF)	Supplementary Eye Field (SEF)	Left Posterior Parietal Cortex (Left PPC)	Right Posterior Parietal Cortex (Right PPC)	Non-Area of Interest
Controls	8	8	10	8	12	0
Relatives	4	9	7	1	12	0
Schizophrenia	1	3	3	2	11	0

A number of broad conclusions can be drawn from Table 2 and Figures 4 and 5. First, visual inspection of the between group data shows that in a general sense, where controls show a high conditional probability value, the schizophrenia participants show lower conditional probability values. The relatives range from midway between controls and schizophrenia to normal values in most regions. Second, visual inspection across techniques shows that all three groups appear to have high conditional probability values in right PPC. For the other regions (SEF and right FEF), there is general consistency in the relatives showing more similarity to the control participants, or at least less similarity to the participants with schizophrenia.

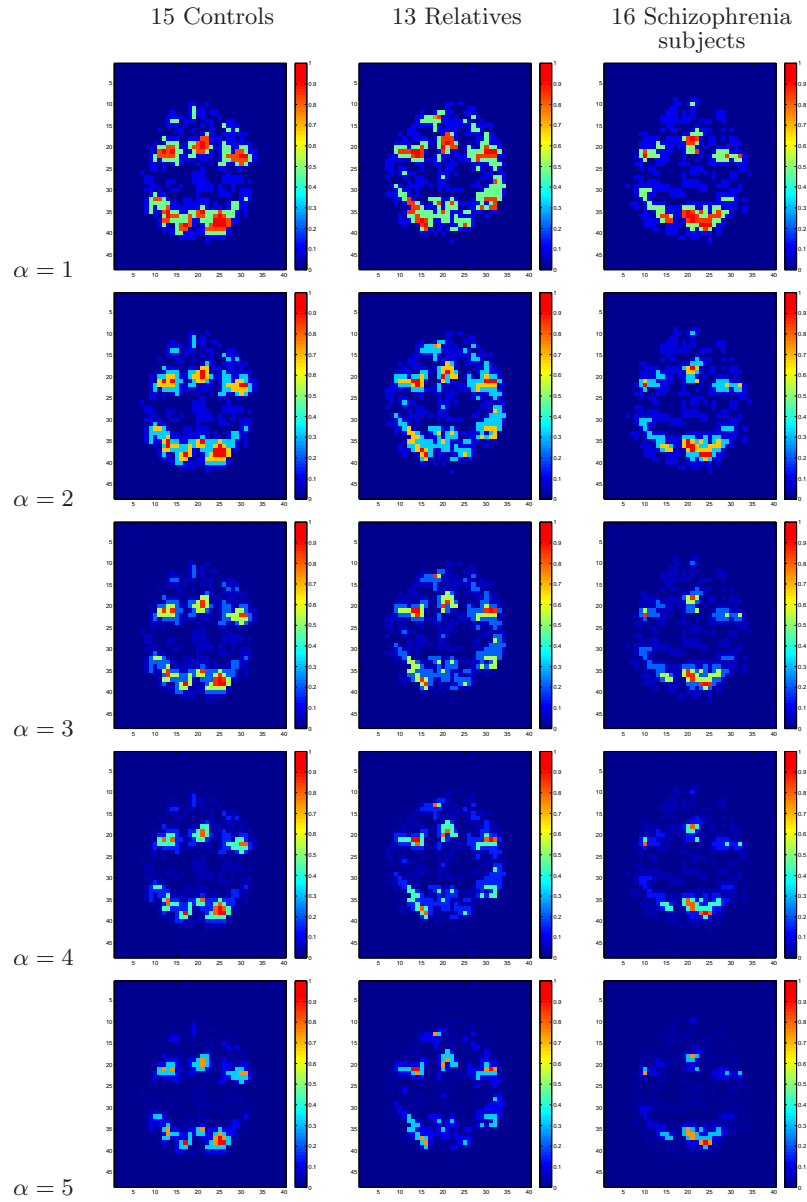


FIG 4. Conditional probabilities of activation for controls, relatives of schizophrenia subjects, and schizophrenia subjects using GIFA to construct the co-activation matrix. The conditional probabilities of activation in the FEF, SEF, and PPC are very close to 1, especially in the controls. The tuning parameter,  $\alpha$ , governs the sparseness of the conditional probabilities in the brain. As  $\alpha$  increases, the conditional probability levels are less similar across the groups.

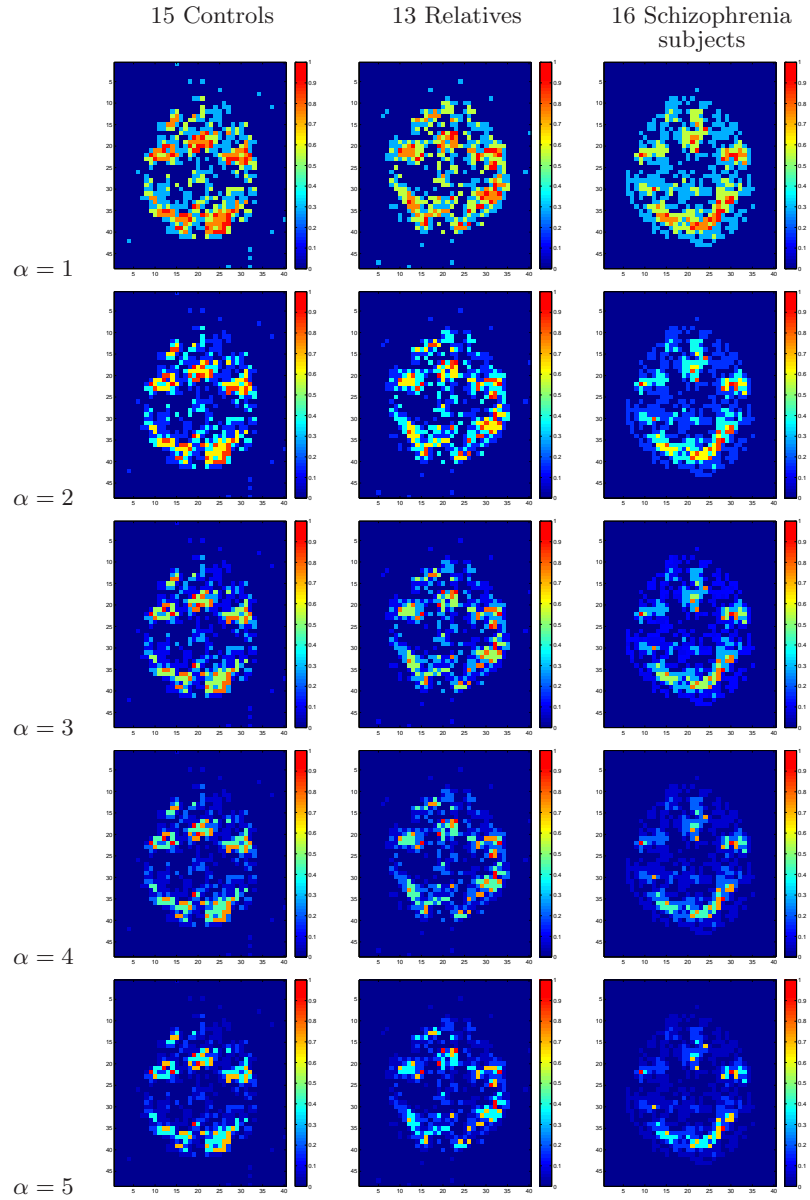


FIG 5. Conditional probabilities of activation for controls, relatives of schizophrenia subjects, and schizophrenia subjects using HMM to construct the co-activation matrix. The conditional probabilities of activation in the FEF, SEF, and PPC are very close to 1, especially in the controls. The tuning parameter,  $\alpha$ , governs the sparseness of the conditional probabilities in the brain. As  $\alpha$  increases, the conditional probability levels are less similar across the groups.

Importantly, however, for left PPC (and maybe the left FEF) the technique used determines that the relatives are considered to be more like controls (GIFA) rather than participants with schizophrenia (HMM). GIFA conditional probabilities appear to be more clustered than HMM conditional probabilities, which appear to be more scattered. Finally, no active voxels are identified outside of the areas of interest.

In the conditional probability maps in Figures 4 and 5, as  $\alpha$  increases from 1 to 5, the effect of the co-activation matrix is weighted more heavily. Thus, where the voxels are found to be more active and the co-activation matrix indicates a strong relationship between multiple pairs of voxels (the sum of the product of the co-activation matrix terms with other voxels is high), the conditional probabilities of activation are high (greater than 0.9). Conversely, where the voxels are found to be active but the co-activation matrix indicates a weak, or nonexistent relationship between multiple pairs of voxels, the conditional probabilities of activation are small (less than 0.05). The values of the co-activation matrix range from zero to  $10^\alpha$ ; for higher values of  $\alpha$ , the values of the co-activation matrix are more sparsely spread. Therefore, the conditional probabilities will be more swayed by the extremely high (and low) values of the co-activation matrix.

**5. Discussion.** The model we implement takes into account the relative strength of interactivity between voxels, regardless of physical proximity. The two methods we examine yield somewhat similar results, but with a very different set of assumptions. GIFA is model free, while HMM is model dependent. The results obtained from each method are also somewhat visually different, with GIFA conditional probabilities appearing to be more clustered than HMM conditional probabilities.

The current study shows that these techniques can be used to reveal the expected pattern of activation of the neural system supporting antisaccade performance. At the level of the slice selected for analysis, expected regions included bilateral PPC, bilateral FEF and SEF. Of course, in studies like these, scientists are often interested in whether there are group differences. In the current study, one region (right PPC) was similar across groups, although for all other regions the controls showed higher levels of signal than did the schizophrenia patients. Relatives generally ranged from between schizophrenia and normal values to normal values. This pattern of decreased activity in schizophrenia (that is not global) with intermediate values in schizophrenia is expected from the literature (Camchong et al,

2008).

In neuroimaging, we are usually working with a small signal-to-noise-ratio, so any statistical methods which can increase this ratio will have applications in the analysis of fMRI data. In this paper, we illustrate how GIFA and HMM can efficiently remove the noise while retaining the salient features of the signal. This research focuses only on one slice of the brain, but our proposed methodology can be extended to data sets pertaining to the entire brain. This is a topic of our future research. Currently, the computational (as well as storage) costs are prohibitive of straightforward analysis, but we expect to solve these problems by partitioning the brain efficiently. Another aim of future work is to quantify and make statistically rigorous the differences observed among the three groups of subjects. Incorporating covariates, such as a factor for group, might also be a way to answer questions about group differences.

#### APPENDIX: ISING MODEL RESULTS

The Ising model is the most basic model in statistical mechanics (Besag, 1974b; Besag, 1974a; Seymour, 1993). This basic model describes a collection of states on a lattice wherein each state can take one of two values. In our context, each voxel is considered an element of the lattice (the acquisition matrix of a given slice of the fMRI scan) and is declared either active or inactive. The conditional probability given the four nearest neighbors is:

$$P(Y_i = y_i | y_j, j \in N_j) = \frac{\exp \left\{ -\beta y_i \sum_{j \in N_j} y_j \right\}}{\exp \left\{ \beta \sum_{j \in N_j} y_j \right\} + \exp \left\{ -\beta \sum_{j \in N_j} y_j \right\}},$$

where  $y_i = \pm 1$ ,  $N_j$  is the collection of the four nearest neighbors of  $y_i$ , and  $\beta$  is the dependence parameter. The closer  $\beta$  is to zero, the less dependence there is in the data.

Applying this model to the raw difference  $X$  and the GIFA difference results in masking the brain from the background. In other words, higher probabilities of activation are assigned to areas within the brain. This can be seen in Figure 6. In the raw difference matrix  $X$ , most of the areas inside the brain have a conditional probability of one, and the areas outside the brain where

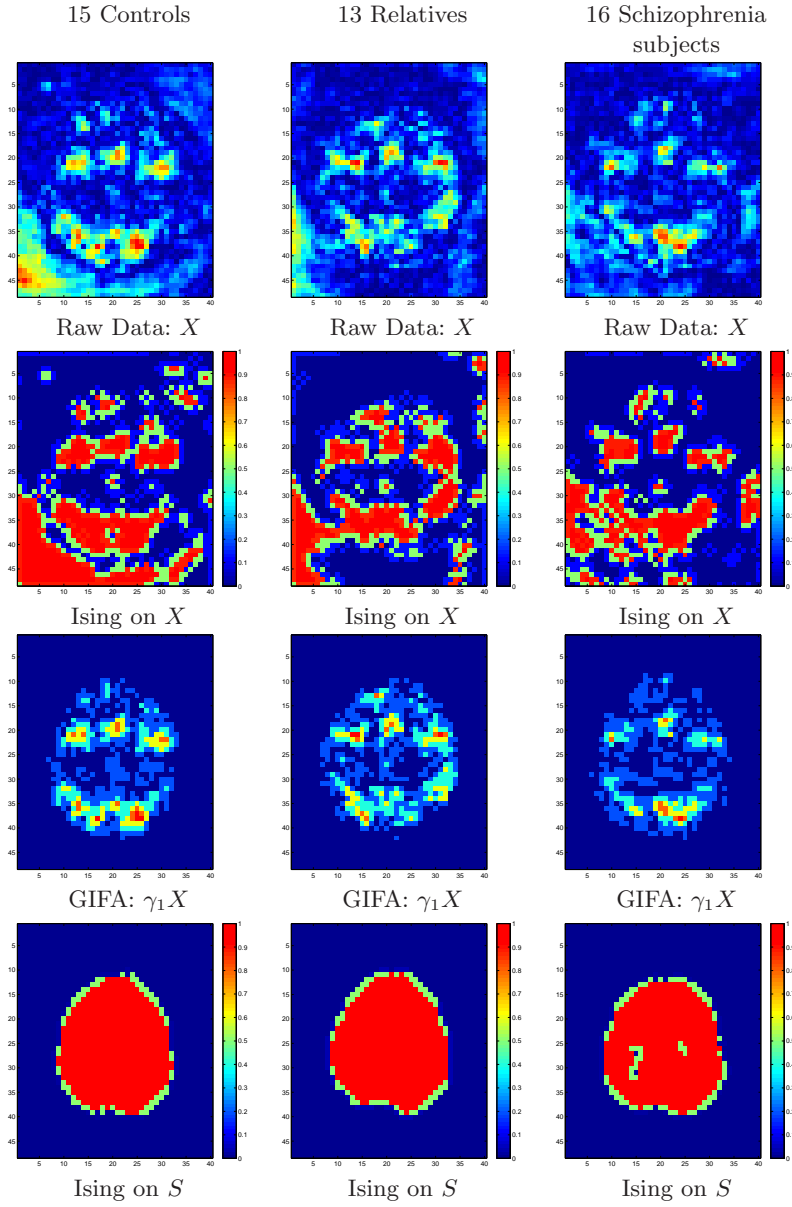


FIG 6. Comparison of Ising Model. Ising model applied to raw data and GIFA. The simplicity of this model yields probability maps which only have two probability values, much like the two states: “on” or “off”.

lower levels of activation are found are assigned a conditional probability of 0.5. The exception to this is noise that is found outside of the brain (along the left bottom area of the acquisition matrix), which is also assigned a conditional probability of one. In the GIFA matrix, all of the areas inside the brain are assigned a conditional probability of one, while the areas outside the brain are assigned a conditional probability of zero. The conditional probabilities around the boundaries of activation in both the raw data and GIFA are slightly larger than zero.

The main advantage of this model is its simplicity. However, since this is a two-state model, where the voxels are either activated or not, it cannot adequately capture the different levels of activity in the brain. In addition, it is a spatial nearest-neighbor model, and thus ignores the fact that an area of the brain may directly or indirectly influence the activation of a physically remote but functionally-related area.

The parameter estimates for  $\beta$  using the Ising model are somewhere between -1 and 0, thus indicating the data may have some dependence structure. However, since the conditional probabilities only take very few values (zero, one, and a very small value close to zero), it is difficult to interpret any other structure in the above images, as the only apparent distinction this model makes is between the brain and the background in the acquisition matrix.

TABLE 3  
*Ising MPLE estimates for controls, schizophrenia subjects and their relatives*

	Raw		
	Controls	Relatives	Schizophrenia
$\widehat{\beta}_{MPLE}$	-0.5298	-0.4819	-0.4994
	GIFA		
	Controls	Relatives	Schizophrenia
$\widehat{\beta}_{MPLE}$	-0.7535	-0.7408	-0.7713

### Acknowledgement

The authors would like to thank Dr. Jazmin Camchong, Dr. Kara Dyckman and Benjamin Austin for conducting the study and collecting the data used in this paper. We would further like to thank Qingyang Li, Cynthia Krafft and Michael Amlung for their roles in restoring and preparing the data and creating a figure. We would also like to thank the fMRI study group of University of Georgia for many helpful discussions.



## REFERENCES

- [1] J. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of the Royal Statistical Society B*, **36**, 192-235, 1974a.
- [2] J. Besag. Statistical analysis of non-lattice data. *The Statistician*, **24**, 179-195, 1974b.
- [3] F. D. Bowman, B. S. Caffo, S. S. Bassett and C. Kilts. A Bayesian hierarchical framework for spatial modeling of fMRI data. *NeuroImage*, **39**, 146-156, 2008.
- [4] M. E. Calkins, W. G. Iacono and D. S. Ones. Eye movement dysfunction in first-degree relatives of patients with schizophrenia: a meta-analytic evaluation of candidate endophenotypes. *Brain and Cognition*, **68**, 436-461, 2008.
- [5] J. Camchong, K. A. Dyckman, B. P. Austin, B. A. Clementz and J. E. McDowell. Common neural circuitry supporting volitional saccades and its disruption in schizophrenia patients and relatives. *Biological Psychiatry*, **64**, 1024-1050, 2008.
- [6] G. Casella and E. I. George. Explaining the Gibbs sampler. *American Statistician*, **46**, 167-174, 1992.
- [7] K. Cheng, R. A. Wagoner and K. Tanaka. Human ocular dominance columns are revealed by high-field functional MRI. *Neuron*, **32**, 359-374, 2001.
- [8] B. A. Clementz, J. E. McDowell and S. Zisook. Saccadic system functioning among schizophrenic patients and their first-degree biological relatives. *Journal of Abnormal Psychology*, **103**, 277-287, 1994.
- [9] R. W. Cox. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, **29**, 162-173, 1996.
- [10] K. Dyckman, J. Camchong, B. A. Clementz and J. E. McDowell. An effect of context on saccade-related behavior and brain activity. *NeuroImage*, **36**, 774-784, 2007.
- [11] U. Ettinger, M. Picchioni, M. H. Hall, K. Schulze, T. Touloupoulou, S. Landau, T. J. Crawford and R. M. Murray. Antisaccade performance in monozygotic twins discordant for schizophrenia: The Maudsley twin study. *American Journal of Psychiatry*, **163**, 543-545, 2006.
- [12] M. B. First, R. L. Spitzer, M. Gibbon and J. B. W. Williams. *Structured clinical interview for DSM-IV Axis I disorders (SCID-I)*. American Psychiatric Association, Inc., Washington, 1995.
- [13] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In *IEEE-PAMI6*, 721-741, 1984.
- [14] S. Geman and C. Graffigne. Markov random field image models and their application to computer vision. In *Proceedings of the International Congress of Mathematics, Mathematical Society*, Berkeley, 1986.
- [15] C. R. Genovese, N. A. Lazar and T. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, **15**, 870-878, 2002.
- [16] H. Georgii. *Gibbs Measures and Phase Transitions*. Walter de Gruyter & Co., Berlin, 1988.
- [17] C. Geyer and E. Thompson. Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society, Series B*, **54**, 657-699, 1992.
- [18] C. Gössl, D. Auer and L. Fahrmeir. Bayesian spatiotemporal inference in functional magnetic resonance imaging. *Biometrics*, **57**, 554-562, 2001.
- [19] Y. Guo and G. Pagnoni. A unified framework for group independent component analysis for multi-subject fMRI data. *NeuroImage*, **42**, 1078-1093, 2008.

- [20] M. Handcock. Degeneracy and inference for social network models. Working Paper no. 39, Center for Statistics and the Social Sciences, University of Washington, 2003. Available from <http://www.csss.washington.edu/Papers/>.
- [21] J. A. Hartigan. *Clustering Algorithms*. Wiley, New York, 1975.
- [22] P. Hoff. Latent space approaches to social network analysis. *Journal of the American Statistical Association*, **97**, 1090-1098, 2002.
- [23] J. Hoshen, M. W. Berry and K. S. Minser. Percolation and cluster structure parameters: The enhanced Hoshen-Kopelman algorithm. *Physical Review B*, **56**(2), 1455-1460, 1997.
- [24] J. Hoshen. On the application of the enhanced Hoshen-Kopelman algorithm for image analysis. *Pattern Recognition Letters*, **12**, 575-584, 1998.
- [25] D. Hunter, S. Goodreau and M. Handcock. Goodness of fit of social network models. Working Paper no. 47, Center for Statistics and the Social Sciences, University of Washington, 2005. Available from <http://www.csss.washington.edu/Papers/>.
- [26] A. Hyvarinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, **13**, 411-430, 2000.
- [27] S. B. Hutton, I. Cuthbert, T. J. Crawford, C. Kennard, T. R. E. Barnes and E. M. Joyce. Saccadic hypometria in drug naive and drug-treated schizophrenic patients: A working memory deficit? *Psychophysiology*, **38**, 125-132, 2001.
- [28] R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis, Fifth Edition*. Prentice Hall, New Jersey, 2002.
- [29] K. Katanoda, Y. Matsuda and M. Sugishita. Spatio-temporal regression model for the analysis of functional MRI data. *NeuroImage*, **17**, 1415-1428, 2002.
- [30] B. King-Casas, D. Tomlin, C. Anen, C. F. Camerer, S. R. Quartz and P. R. Montague. Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, **308**, 78-83, 2005.
- [31] N. A. Lazar. *The Statistical Analysis of Functional MRI Data*. Springer, New York, 2008.
- [32] N. A. Lazar, B. Luna, J. A. Sweeney and W. F. Eddy. Combining brains: A survey of methods for statistical pooling of information. *NeuroImage*, **16**, 538-550, 2002.
- [33] M. A. Lindquist. The statistical analysis of fMRI data. *Statistical Science*, **23**, 439-464, 2008.
- [34] J. E. McDowell, M. Myles-Worsley, H. Coon, W. Byerley and B. A. Clementz. Measuring liability for schizophrenia using optimized antisaccade stimulus parameters. *Psychophysiology*, **36**, 138-141, 1999.
- [35] L. R. Rabiner and B. H. Juang. An introduction to hidden Markov models. *Proceedings of the IEEE*, **77**, 257-286, 1986.
- [36] L. Seymour. Parameter estimation and model selection in image analysis using Gibbs-Markov random fields. Dissertation, University of North Carolina Chapel Hill, 1993.
- [37] M. Smith and L. Fahrmeir. Spatial Bayesian variable selection with application to functional magnetic resonance imaging. *Journal of the American Statistical Association*, **102**, 417-431, 2007.
- [38] T. Snijders, P. Pattison, G. Robins and M. Handcock. New specifications for exponential random graph models. *Sociological Methodology*, **36**, 99-153, 2006.
- [39] D. Strauss and M. Ikeda. Pseudo-likelihood estimation for social networks. *Journal of the American Statistical Association*, **85**, 204-212, 1990.

- [40] J. Sweeney, B. Luna, S. K. Keedy, J. E. McDowell and B. A. Clementz. FMRI studies of eye movement control: Investigating the interaction of cognitive and sensorimotor brain systems. *NeuroImage*, **36**, T54-T60, 2007.
- [41] J. Talairach and P. Tournoux. *Co-Planar Stereotactic Atlas of the Human Brain*. Thieme, New York, 1988.
- [42] M. O. Ulfarsson and V. Solo. Spatially local and temporally smooth PCA for fMRI. *Proceedings of the International Conference on Image Processing, IEEE*, 2853-2856, 2006.
- [43] B.D. Ward. Simultaneous Inference for fMRI Data. AlphaSim program documentation for AFNI. Medical College of Wisconsin, Milwaukee, 2000. <http://afni.nimh.nih.gov/pub/dist/doc/manual/AlphaSim.pdf>.
- [44] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, 1994.
- [45] M. Woolrich, M. Jenkinson, J. Brady and S. Smith. Fully Bayesian spatio-temporal modeling of fMRI data. *IEEE Transactions on Medical Imaging*, **23**, 213-231, 2004.
- [46] K. J. Worsley and K. J. Friston. Analysis of fMRI time-series revisited - again. *NeuroImage*, **2**, 173-181, 1995.
- [47] T. Yokoo, B. W. Knight and L. Sirovich. An optimization approach to signal extraction from noisy multivariate data. *NeuroImage*, **14**, 1309-1326, 2001.
- [48] X. Zhang, T. D. Johnson, R. J. A. Little and Y. Cao. Quantitative magnetic resonance image analysis via the EM algorithm with stochastic variation. *The Annals of Applied Statistics*, **2**, 736-755, 2008.

ANA M. BARGO  
 DEPARTMENT OF STATISTICS  
 UNIVERSITY OF GEORGIA  
 ATHENS, GA 30602-7952, USA  
 E-MAIL: amoura@uga.edu

ABHYUDAY MANDAL  
 DEPARTMENT OF STATISTICS  
 UNIVERSITY OF GEORGIA  
 ATHENS, GA 30602-7952, USA  
 E-MAIL: amandal@stat.uga.edu

LYNNE SEYMOUR  
 DEPARTMENT OF STATISTICS  
 UNIVERSITY OF GEORGIA  
 ATHENS, GA 30602-7952, USA  
 E-MAIL: seymour@stat.uga.edu

JENNIFER MCDOWELL  
 DEPARTMENT OF PSYCHOLOGY  
 UNIVERSITY OF GEORGIA  
 ATHENS, GA 30602-7952, USA  
 E-MAIL: jemcd@uga.edu

NICOLE A. LAZAR  
 DEPARTMENT OF STATISTICS  
 UNIVERSITY OF GEORGIA  
 ATHENS, GA 30602-7952, USA  
 E-MAIL: nlazar@stat.uga.edu