Community Detection in the Space of Functional Abnormalities Reveals both Heightened and Reduced Brain Synchrony in Autism

Archana Venkataraman[1], Daniel Y.-J. Yang[2], Kevin A. Pelphrey[2], and James S. Duncan[1,3]

[1]Department of Diagnostic Radiology, School of Medicine, Yale University, USA
[2]Center for Translational Developmental Neuroscience, Yale University, USA
[3]Department of Biomedical Engineering, Yale University, USA

Abstract. We propose a unified Bayesian framework to detect both hyper- and hypo-active communities within whole-brain fMRI data. Our model characterizes population-level differences in functional synchrony between a control and clinical group. We use a variational EM algorithm to solve for the latent posterior distributions and parameter estimates. We demonstrate that our method provides valuable insights into the neural mechanisms underlying social dysfunction in autism, as verified by the Neurosynth meta-analytic database. In contrast, both univariate testing and standard community detection via recursive edge elimination fail to identify stable functional communities associated with autism.

1 Introduction

Statistical analysis of fMRI activation patterns allow us to pinpoint functional differences induced by neurodevelopmental conditions, such as Autism Spectrum Disorder (ASD). While this approach informs us about localized cognitive deficits, there is increasing evidence that ASD reflects distributed impairments across multiple brain systems [1]. These findings underscore the importance of network-based methodologies for functional data. Here, we present a novel framework to extract both hyper- and hypo-active communities from task fMRI.

Community detection is the process of identifying highly interconnected subgraphs within a larger network; these nodes share common properties and are crucial to understanding the organization of complex systems [2, 3]. Most community detection algorithms are greedy and/or hierarchical procedures that optimize surrogate measures of cluster strength, such as modularity and information compression [3, 4]. Modularity-based approaches have recently been applied to fMRI data [5, 6]. Broadly, these works estimate a hierarchical community organization independently within each subject and use a post-hoc cluster matching procedure to perform group-level analysis. While promising, the results depend on preselected thresholds and subroutines (i.e., adjacency matrix construction and subgraph alignment). Moreover, since these methods do not impose topological constraints on group differences, the clinical findings are difficult to interpret.

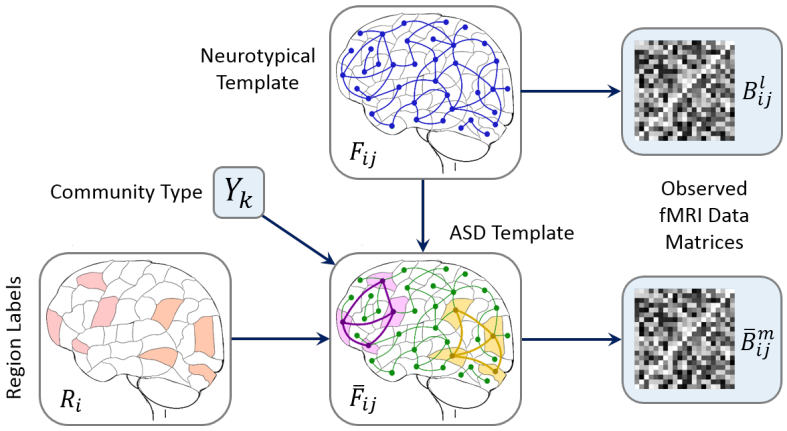

Fig. 1. Hierarchical model of community structure for $K = 2$. The label R_i indicates whether region i is healthy (white) or whether it belongs to one of the two abnormal communities (red). The binary variable Y_k denotes either a hyper-active (purple) or hypo-active (yellow) subgraph k. The neurotypical template $\{F_{ij}\}$ provides a baseline functional architecture for the brain, whereas the clinical template $\{\bar{F}_{ij}\}$ describes the latent organization of ASD. The green connections $\langle i, j \rangle$ are unchanged from baseline; the purple and yellow lines signify heightened and reduced synchrony, respectively. Each template generates a set of subject observations $\{B^l_{ij}\}$ and $\{\bar{B}^m_{ij}\}$ for the group.

Our approach differs from prior work in three crucial ways. First, we propose a unified generative model that describes the relationship between population templates and individual subject observations. Second, we perform community detection in the space of group-level functional *differences*, which allows us to identify both hyper- and hypo-active subgraphs. Finally, our framework can simultaneously detect multiple abnormal communities of varying type. We demonstrate our method on an fMRI study of social perception in autism.

ASD is characterized by impaired social reciprocity, communication deficits, and repetitive/stereotyped behavior. Task fMRI studies reveal an inconsistent pattern of altered connectivity between the cortex and limbic structures [7]. However, the analyses are restricted to predefined neural systems and do not reflect whole-brain information. At present, network analysis of ASD has focused on aggregate measures of degree and centrality [8], which fail to pinpoint a concrete etiological mechanism for the disorder. In contrast, our method robustly identifies hyper- and hypo-active *subnetworks* that map onto key functional domains, in which autistic individuals exhibit strengths and vulnerabilities, respectively.

2 Generative Model of Abnormal Communities

Our model assumes that altered whole-brain functional synchrony can be decomposed into K non-overlapping communities. Each community is associated with a binary label Y_k that indicates either a hyper-active ($Y_k = 1$) or a hypo-active

($Y_k = -1$) subgraph. In this work Y_k is specified *a priori* by the user for greater control over the results. This level of input is akin to setting the number of clusters or dictionary components in an unsupervised learning scenario; it further enables us to explore the network evolution in the absence of ground truth. If desired, it is straightforward to model Y_k as an unknown random variable.

We use multinomial variables F_{ij} and $\bar{F}_{ij}$ to denote the latent functional synchrony between regions i and j in the neurotypical and ASD groups, respectively. Empirically, we find that three states: low ($F_{ij} = 0$), medium ($F_{ij} = 1$), and high ($F_{ij} = 2$), best capture the dynamic range and variability of our fMRI data. The conditional relationships are defined such that $\bar{F}_{ij} > F_{ij}$ within a hyper-active community, and $\bar{F}_{ij} < F_{ij}$ within a hypo-active community.

The fMRI metrics B^l_{ij} for subject l and $\bar{B}^m_{ij}$ for subject m are noisy observations of the underlying latent structure. Fig. 1 outlines the generative process. While the variables are similar to [9], our framework identifies interconnected subgraphs, and it distinguishes between hyper- and hypo-active functional synchrony. These two properties describe a fundamentally different abnormal network topology, which has not been captured by prior Bayesian models.

Node Selection: The multinomial variable R_i indicates whether region i is healthy ($R_i = 0$), or whether it belongs to an abnormal community k. We assume an *i.i.d.* multinomial prior for R_i as follows:

$$P(R_i = k; \pi^r) = \pi^r_k, \qquad k = 0, \ldots, K \tag{1}$$

The unknown parameters π^r_k are shared by all regions in the network.

Latent Network Topology: The latent functional synchrony F_{ij} denotes the co-activation between regions i and j in the neurotypical template. As previously described, F_{ij} is modeled as a tri-state random variable. The multinomial prior is *i.i.d.* across all pairwise connections, i.e., $P(F_{ij} = s; \pi^f) = \pi^f_s$, $\forall s = 0, 1, 2$.

The clinical variable $\bar{F}_{ij}$ depends on the control template F_{ij} and the abnormal communities defined by the region labels R and subgraph type Y. To better understand the conditional distribution $P(\bar{F}_{ij}|F_{ij}, R, Y)$, it is convenient to define an auxiliary variable T_{ij}, which indicates a healthy ($T_{ij} = 0$) or abnormal ($T_{ij} = 1$) edge $\langle i, j \rangle$ between member nodes of a community. Mathematically,

$$P(T_{ij} = 1|F_{ij}, R_i, R_j, Y; \eta) = \begin{cases} \eta, & R_i = R_j > 0, Y_k = 1, F_{ij} < 2, \\ \eta, & R_i = R_j > 0, Y_k = -1, F_{ij} > 0, \\ 0, & \text{Otherwise.} \end{cases} \tag{2}$$

As seen, T_{ij} is Bernoulli with parameter η if both regions are in the same community, with the exception of border conditions $Y_k = 1, F_{ij} = 2$ and $Y_k = -1, F_{ij} = 0$, in which case $T_{ij} = 0$. Eq. (2) is a special case of the planted ℓ-partition model [3], where η controls the density of edges within the community.

Ideally, $\bar{F}_{ij} = F_{ij}$ for healthy edges. Conversely, if $T_{ij} = 1$, then $\bar{F}_{ij} > F_{ij}$ for $Y_k = 1$, and $\bar{F}_{ij} < F_{ij}$ for $Y_k = -1$. However, we assume a second level of noise, in which the parameter ϵ controls the probability of deviating from the

$R_i = R_j = k > 0$ and $Y_k = 1$

		$\bar{F}_{ij}$		
		0	1	2
F_{ij}	0	$(1-\eta)(1-\epsilon)+\eta\epsilon$	$(1-\eta)(\epsilon/2)+\eta\left(\frac{1-\epsilon}{2}\right)$	$(1-\eta)(\epsilon/2)+\eta\left(\frac{1-\epsilon}{2}\right)$
	1	$\epsilon/2$	$(1-\eta)(1-\epsilon)+\eta(\epsilon/2)$	$(1-\eta)(\epsilon/2)+\eta(1-\epsilon)$
	2	$\epsilon/2$	$\epsilon/2$	$1-\epsilon$

$R_i = R_j = k > 0$ and $Y_k = -1$

		$\bar{F}_{ij}$		
		0	1	2
F_{ij}	0	$1-\epsilon$	$\epsilon/2$	$\epsilon/2$
	1	$(1-\eta)(\epsilon/2)+\eta(1-\epsilon)$	$(1-\eta)(1-\epsilon)+\eta(\epsilon/2)$	$\epsilon/2$
	2	$(1-\eta)(\epsilon/2)+\eta\left(\frac{1-\epsilon}{2}\right)$	$(1-\eta)(\epsilon/2)+\eta\left(\frac{1-\epsilon}{2}\right)$	$(1-\eta)(1-\epsilon)+\eta\epsilon$

Table 1. Conditional density tables $P(\bar{F}_{ij}|F_{ij}, R, Y; \eta, \epsilon)$ for hypo-active (top) and hyper-active (bottom) communities.

above rules. Marginalizing out the graph $\{T_{ij}\}$ to better estimate the community assignments R_i yields the conditional relationships in Table 1.

Finally, if regions i and j are not in the same community, then $T_{ij} = 0$ according to Eq. (2). Representing F_{ij} and $\bar{F}_{ij}$ as length three indicator vectors, such that $F_{ij}^{\mathrm{T}}\bar{F}_{ij} \in \{0, 1\}$, we obtain the following baseline distribution:

$$P(\bar{F}_{ij}|F_{ij}, \{R_i \neq R_j | R_i = R_j = 0\}; \epsilon) = (1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \tag{3}$$

Data Likelihood: The fMRI correlation B_{ij}^l for subject l is a noisy observation of the functional template F_{ij}. We assume a conditional Gaussian distribution, with mean and variance controlled by the latent synchrony: $P(B_{ij}^l|F_{ij} = s; \{\mu, \sigma^2\}) = \mathcal{N}\left(B_{ij}^l; \mu_s, \sigma_s^2\right)$. The likelihood for an ASD patient $\bar{B}_{ij}^m$ has the same functional form and parameter values but relies on the clinical template $\bar{F}_{ij}$. Notice that we are modeling the data observations as a *mixture* of Gaussians, which can be used to represent non-Gaussian correlation values [10].

Variational Inference: Let $\Theta = \{\pi^r, \pi^f, \eta, \epsilon, \mu, \sigma^2\}$ denote the collection of non-random model parameters, and recall that the subgraph types Y_k are given. We combine the prior and likelihood terms to obtain the joint density of latent and observed variables: $P(R, F, \bar{F}, B, \bar{B}|Y; \Theta)$. The community assignments R_i induce a complex coupling across pairwise connections $\langle i, j \rangle$. Therefore, we employ a variational EM framework [11] to infer the approximate posterior $Q(\cdot)$ and corresponding parameters Θ that minimize the variational free energy

$$\mathcal{F} = -E_Q[\log P(R, F, \bar{F}, B, \bar{B}|Y; \Theta)] - \mathcal{H}(Q). \tag{4}$$

Notice that $-\mathcal{F}$ is a lower bound to the marginal log-likelihood $P(B, \bar{B}|Y; \Theta)$.

Our approximate posterior assumes the following factorized form:

$$Q(R, F, \bar{F}) = \prod_{i=1}^{N} q_i^r(R_i; \tilde{\beta}_i) \prod_{edges\ \langle i,j \rangle} q_{ij}^c(F_{ij}, \bar{F}_{ij}; \tilde{\nu}_{ij}), \tag{5}$$

where $q_i^r(\cdot)$ is a multinomial distribution with $K+1$ states, parameterized by $\tilde{\beta}_i$. Likewise, $q_{ij}^c(\cdot)$ is a multinomial distribution with 9 states, parameterized by $\tilde{\nu}_{ij}$, accounting for the 9 configurations in Table 1. Eq. (5) preserves the dependency between F_{ij} and $\bar{F}_{ij}$ and is scalable to accommodate a large number of regions N.

During the E-step, we fix Θ and iteratively update the elements of $Q(\cdot)$ to minimize the variational free energy. The iterations for $\tilde{\nu}_{ij}$ can be expressed in closed form given $\tilde{\beta}$. However, the updates for $\tilde{\beta}_i$ are coupled. Therefore, we perform an inner fixed-point iteration until convergence of the region posterior distribution. In the M-step we estimate the model parameters Θ given $Q(R, F, \bar{F})$. The updates for $\{\pi^r, \pi^f, \mu, \sigma^2\}$ parallel those of a Gaussian mixture model. We jointly update the concentration η and noise ϵ via Newton's method.

We emphasize that given Y, and hence the number of communities K, both the posterior distribution and the model parameters are estimated directly from the observed data. We do not tune any auxiliary parameters or thresholds.

Estimating the Abnormal Networks: The idealized graph of functional differences $\{T_{ij}\}$ provides valuable insights into the topological properties of each abnormal community. We can retrospectively approximate these variables based on the maximum *a posteriori* (MAP) solution for R and the parameter estimates $\hat{\Theta}$. Specifically, given R and Y, our model decouples by pairwise connection, so we can assign each T_{ij} independently. Furthermore, Eq. (2) implies that $T_{ij} = 0$ for $R_i \neq R_j$ and $R_i = R_j = 0$. For the remaining case ($R_i = R_j > 0$), we can select the value $T_{ij} \in \{0, 1\}$ according to the following optimization problem:

$$\hat{T}_{ij} = \arg\max_{T} E_Q \left[\log P(T|F_{ij}, \bar{F}_{ij}, \hat{R}^{MAP}, Y; \hat{\eta}, \hat{\epsilon}) \right]. \tag{6}$$

The right-hand side of Eq. (6) can be computed by multiplying the natural logarithm of entries in Table 1 with the corresponding variational posterior parameter $\tilde{\nu}_{ij}$ and summing across all nine configurations of $\{F_{ij}, \bar{F}_{ij}\}$. Eq. (6) weighs the density of edges in the idealized communities η against the latent noise ϵ, which encourages deviation of the functional templates.

Model Evaluation: The marginal distribution $q_i^r(R_i; \tilde{\beta}_i)$ informs us whether region i is healthy or if it belongs to one of the K abnormal communities. We evaluate the robustness of these region assignments via bootstrapping. In particular, we fit the model to random subsets of the data, such that the ratio of neurotypical controls to ASD patients is preserved. Each subset contains only 80% of the total number subjects. We re-sample the data subsets 50 times and average the region posterior probability estimates across the runs.

The Neurosynth database (`www.neurosynth.org`) provides an unbiased and comprehensive evaluation of the functionality supported by each hyper- and hypo-active community. Given a distributed activation pattern, the meta-analytic framework enables us to perform a reverse inference by computing the posterior probability $P(Feature|Activation)$ for individual psychological features [12]. For each of the detected communities, we computed the region-wise Pearson correlation with the statical maps obtained for each of the 3,099 Neurosyth feature terms; words with correlation above the default threshold ($r > 0.001$) were retained. The corresponding topics were automatically generated by the Neurosynth inference engine via Latent Dirichlet Allocation (LDA).

Finally, we present results from a simulated experiment, which demonstrates that our variational EM algorithm recovers the ground truth region labels. Here, the subject correlations are generated according to Fig. 1 using the noise statistics and community organization observed in our fMRI dataset.

Baseline Methods: We compare our generative model with two alternative techniques: statistical testing and a well-known community detection procedure, which prunes edges from the original network. Such analyses evaluate the benefit of explicitly modeling both hyper- and hypo-active communities.

The two sample t-test confirms or rejects the null hypothesis that the group means of a given task fMRI correlation are equal. The corresponding p-value is computed via the Student-t distribution and quantifies the statistical significance of the result. In this experiment, we vary the significance threshold to determine whether the univariate functional differences naturally exhibit an underlying community structure, and whether these connections organize by type.

The seminal method of [13] iteratively removes edges from the network in order to discover community structure. The edges are selected according to a *current-flow betweenness* measure, which favors connections that lie between communities rather than ones that span a single community. Mathematically, consider the electrical circuit created by placing a unit resistance on each edge of the network and a unit current source and unit current sink at a given pair of vertices (s, t). Let $\mathbf{A}$ denote the binary adjacency matrix for the network, such that $\mathbf{A}_{ij} = 1$ if nodes i and j are connected by an edge and $\mathbf{A}_{ij} = 0$ otherwise. The diagonal matrix $\mathbf{D}$ captures the vertex degrees: $\mathbf{D}_{ii} = \sum_j \mathbf{A}_{ij}$. Let $\mathbf{v}_i$ be the voltage at node i. Then, by Kirchhoff's laws we have

$$\sum_j \mathbf{A}_{ij}(\mathbf{v}_i - \mathbf{v}_j) = \delta_s(i) - \delta_t(i) \tag{7}$$

where nodes s and t correspond to the source and sink, respectively, and $\delta_v(\cdot)$ is a unit impulse with mass at vertex v. Eq. (7) can be written in matrix form as

$$(\mathbf{D} - \mathbf{A})\mathbf{v} = \mathbf{z} \tag{8}$$

such that $\mathbf{z}_i = 1$ if $i = s$ (source), $\mathbf{z}_i = -1$ if $i = t$ (sink), and $\mathbf{z}_i = 0$ otherwise. We can invert the graph Laplacian $(\mathbf{D} - \mathbf{A})$ in Eq. (8) by selecting a ground vertex g and removing the corresponding row and column ($\mathbf{v}_g = 0$ by default).

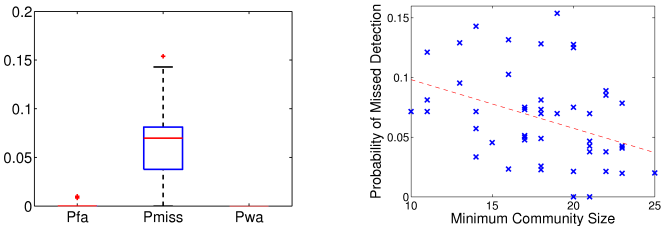

Fig. 2. Results across 50 synthetic trials. **Left:** Errors in region assignment corresponding to the probability of false alarm (Pfa), the probability of a miss (Pmiss) and the probability of an abnormal region being assigned to the wrong community (Pwa). **Right:** Inverse relationship between Pmiss and the minimum community size.

Our final betweenness score is the absolute difference in voltage along each edge, as summed over all source/sink pairs in the network. During each iteration, we remove the edge $\langle i, j \rangle$ with the highest current-flow betweenness and then recalculate these measures based on the reduced graph.

The above community detection procedure requires a sparse binary input graph $\mathbf{A}$ that describes the group-wise functional differences. Rather than averaging and thresholding the fMRI correlation values, we construct $\mathbf{A}$ via the Bayesian connection model described in [10]. Specifically, $\mathbf{A}_{ij} = 1$ if the posterior probability that the latent functional connectivity templates differ exceeds 0.5. We take this approach because data thresholding introduces nonlinear dependencies into the analysis, which are known to bias network results [14]. Moreover, since the fMRI likelihood in [10] is the same as in this work, we can directly evaluate the gain from our community priors in Eq. (1) and Table 1.

3 Social Perception in Autism

Synthetic Validation: We generate subject observations from our Bayesian model that have the same signal-to-noise ratio and community properties as the maximum likelihood solution from the clinical experiment. In particular, we set $(\mu_1 - \mu_0) = 0.13$, $(\mu_2 - \mu_1) = 0.2$, $\sigma_0^2 = 0.07$ and $\sigma_1^2 = \sigma_2^2 = 0.06$ for the data likelihood. Notice that there is significant overlap in the data distributions for the three latent connectivity states, i.e, the standard deviation of each Gaussian is 2-3 times greater than the separation between the means. We mimic the organization of our clinical dataset by specifying an underlying network with 150 regions and sampling 50 subjects in each population. The latent concentration and noise parameters are fixed at $\eta = 0.5$ and $\epsilon = 0.03$, respectively.

We assume one hyper-active and one hypo-active community ($K = 2$). During each trial, we uniformly sample the region prior π^r such that $11 - 16\%$ of regions belong to each abnormal cluster. We generate the random variables in the following order $R \rightarrow F \rightarrow \bar{F} \rightarrow B \rightarrow \bar{B}$ and use the variational EM algorithm to infer the region posterior distribution $q_i^r(R_i; \tilde{\beta}_i)$. Fig. 2 (left) illustrates the

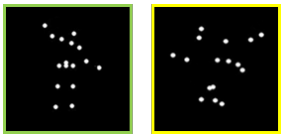

Fig. 3. Snapshot of the coherent biological motion (left) and scrambled motion (right) visual stimuli.

average errors across 50 random trials. As seen, the probability of a false alarm or wrong assignment is negligible. While the probability of a missed detection is slightly higher, the median error is only 7%, and the 75^{th} percentile error is under 9%. This error can likely be attributed to the data overlap. Fig. 2 (right) suggests an inverse relationship between the detection error and the minimum community size. Intuitively, larger communities have more abnormal edges and are easier to identify, whereas smaller communities can be overwhelmed by noise.

Biopoint Dataset: We demonstrate our methods on a clinical study of 72 ASD children and 43 age-matched ($p > 0.124$) and IQ-matched ($p > 0.122$) neurotypical controls. For each subject, a T1-weighted scan (MPRAGE, $TR = 1900ms$, $TE = 2.96ms$, flip angle = $9°$, $res = 1mm^3$) and a task fMRI scan (BOLD, $TR = 2000ms$, $TE = 25ms$, flip angle = $60°$, $res = 3.44 \times 3.44 \times 4mm$, 164 volumes) were acquired on a Siemens MAGNETOM Trio TIM 3T scanner.

The experimental paradigm features coherent and scrambled point-light animations created from motion capture data. The coherent biological motion depicts an adult male actor performing movements relevant to early childhood experiences [15]. The scrambled animations combine the trajectories of 16 randomly selected points from the coherent displays. Snapshots of these stimuli are illustrated in Fig. 3. Six biological motion clips and six scrambled motion clips were presented without audio in an alternating-block design ($24s$ per block).

We segment the anatomical images into 150 cortical and sub-cortical regions based on the Destrieux atlas in Freesurfer [16]. The fMRI data was preprocessed using FSL [17]. The pipeline consisted of: 1) motion correction using MCFLIRT, 2) interleaved slice timing correction, 3) BET brain extraction, 4) spatial smoothing with FWHM 5mm, and 5) high-pass temporal filtering. The functional and anatomical data were registered to the MNI152 standard brain. The fMRI measure B^l_{ij} is computed as the Pearson correlation coefficient between the mean time courses of regions i and j. For regularization, we center the correlation distribution of each subject and fix μ_1 in subsequent analysis. We randomly initialize the variational EM algorithm 10 times and select the optimal $\mathcal{F}$ solution.

Abnormal Communities: Due to space constraints, we focus on a single configuration of subnetwork types $\{Y_k\}$. Fig. 4 presents two abnormal communities inferred by our model. We have specified one hyper-active cluster ($Y_1 = 1$) and one hypo-active cluster ($Y_2 = -1$). Recall that these subnetworks reflect atypical functional synchrony across the entire stimulus paradigm.

The hyper-active community is depicted in red and concentrates in the left superior temporal sulcus (STS), the visual cortex, and the somatosesory cortex. Conversely, the blue regions indicate hypo-synchrony. As seen, they localize to

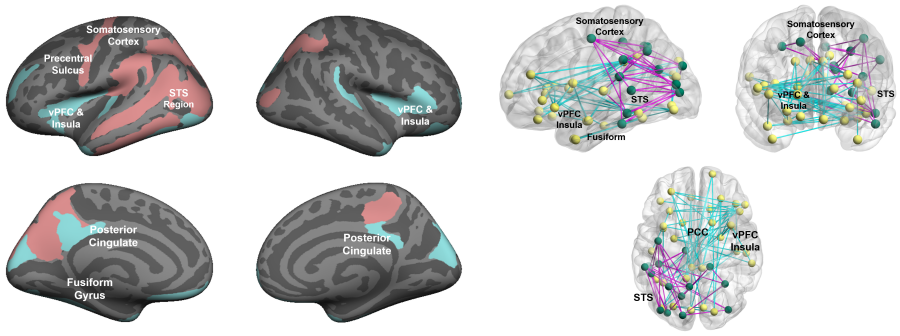

Fig. 4. Abnormal communities inferred by our Bayesian model for $K = 2$, assuming $Y_1 = 1$ and $Y_2 = -1$. **Left:** Region membership in each community. Red indicates hyper-synchrony, and blue areas delineate the hypo-active cluster. **Right:** Estimated network of abnormal functional synchrony. Nodes correspond to regions within each community. Blue lines signify reduced functional synchrony in ASD across the paradigm; magenta lines denote increased functional synchrony in ASD.

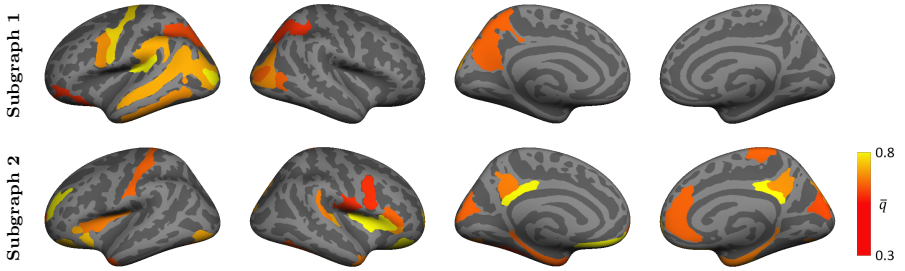

Fig. 5. Average marginal posterior probability $q_i^r(\cdot)$ for each community across 50 random samplings of the fMRI data. Each subset includes 80% of the subjects, such that the ratio of ASD patients to neurotypical controls is preserved. The colorbar denotes the average posterior probability $\bar{q}_i^r$ for each region i.

the bilateral ventral prefrontal cortex (vPFC), insula, and posterior cingulate. The hypo-active network also includes subcortical activations in the caudate and amgydala, which are not shown on the cortical surface plots. The corresponding network diagrams [18] reveal increased synchrony between temporal and occipital nodes (magenta lines), along with reduced synchrony between the frontal and parietal nodes (blue lines). These results support a view of impaired social communication in ASD. Consistent with neurocognitive findings, Fig. 4 suggests hyper-connectivity, and perhaps hyper-functionality, in a visual perception network. Hypo-active circuits include two networks that are well known for their role in social cognition and the high-level interpretation of social stimuli [1].

Fig. 5 reports the average posterior probability $\bar{q}_i^r(\cdot)$ of each region across 50 resamplings of the data. We have displayed only the regions for which $\bar{q}_i^r > 0.3$,

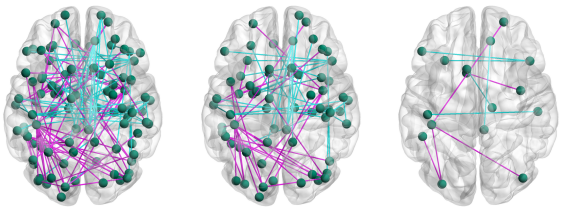

Fig. 6. Population differences in functional connectivity when varying the significance threshold. Blue lines denote a lower average correlation across ASD patient than for neurotypical controls. Magenta lines signify the opposite relationship.

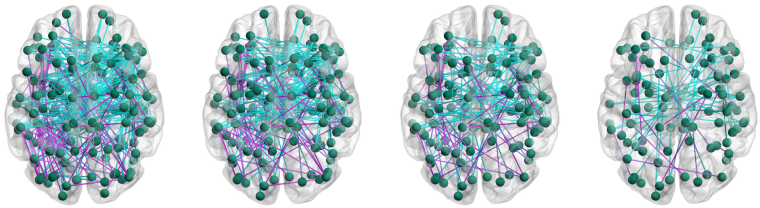

Fig. 7. Community detection based on current-flow betweenness. Left to right depicts the original network of pairwise functional differences (294 edges) and the iterative removal of 80, 160 and 240 edges. Blue lines correspond to reduced functional synchrony in ASD; magenta lines denote increased functional synchrony in ASD.

thereby emphasizing the most prominent patterns. Favorably, our bootstrapping analysis consistently recovers both the red hyper-active and blue hypo-active communities in Fig 4. Additionally, subgraph 2 uncovers medial activations in the right hemisphere that are not present in Fig. 4. This behavior suggests that while some regions are weakly implicated by the data, they are not included in the full variational EM solution. Despite the inconsistency, the level of reproducibility in Fig. 5 has not been demonstrated in prior fMRI studies of ASD.

Baseline Comparison: Figs. 6 and 7 illustrate the connectivity differences based on the two sample t-tests and the current-flow betweenness algorithm, respectively. For ease of visualization, we have colored each pairwise connection such that blue corresponds to reduced synchrony in ASD, and magenta denotes increased synchrony in ASD. However, we do not treat the high and low connections differently when applying the baseline methods.

In Fig. 6 we threshold the (uncorrected) p-value to enforce different levels of statistical significance. As an aside, none of the connections survive a false discovery rate (FDR) correction for multiple comparisons, which suggests that the univariate functional differences are relatively weak in our dataset. Similar to Fig. 4, we observe interconnected magenta lines in STS region and a high

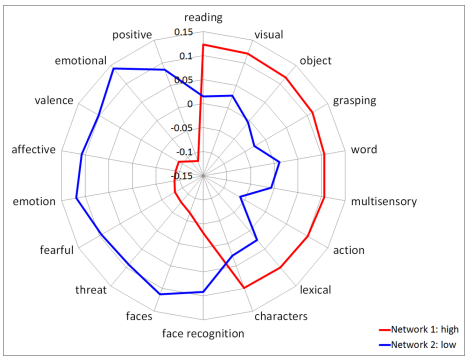

Fig. 8. Correlation values of the top 9 features for each network based on the Neurosynth database. These topics represent the specificity of neurocognitive functions derived from meta-analytic decoding. Red corresponds to the hyper-active community, and blue denotes the hypo-active subgraph. These networks map onto functional domains that highlight both the relative strengths (basic visual processing) and vulnerabilities (interpreting social cues and emotional processing) of ASD, respectively.

concentration of blue lines in the default mode and precentral cortices for $p < 0.01$. There is also a second magenta cluster in the left frontal regions. Reducing the threshold to $p < 0.005$ prunes some of the connections; however, the frontal areas no longer exhibit a clique-like organization. Finally, the result for $p < 0.001$ contains isolated groups of vertices rather than a unified network.

The original network in Fig. 7 consists of 294 edges, which corresponds to connection-wise latent functional differences, as computed via [10]. We notice that this graph is a superset of the connections in Figs. 4, 6. However, pruning the edges according to the current-flow betweenness measure does not reveal interconnected communities. For example, despite removing 80% of the edges in the rightmost figure, the overall number of nodes is fairly similar to the original network. Hence, we cannot identify a natural stopping point for the algorithm.

Neurosynth Results: The polar plot in Fig. 8 illustrates the top nine meta-analytic constructs implied by the regions in each community, as inferred by our generative model. These mental states expose a clear functional distinction between the two networks. The hyper-active network maps onto visual perception and pattern recognition domains; individuals with ASD tend to outperform their neurotypical peers in such tasks. The hypo-active network implicates areas of clear deficit in ASD, such as emotion regulation and social communication, perception and reward. One hypothesis for these results is that ASD children are more stimulated by the point-light animations but do not interpret them.

4 Conclusion

We have demonstrated that a unified Bayesian framework for abnormal community detection retrieves key networks associated with ASD. Specifically, we observe heightened synchrony between visual processing areas of the brain, which corresponds to a relative strength in autistic individuals. Concurrently, we detect reduced synchrony within social information processing and social motivation networks; this finding relates to the hallmark socioemotional impairments of ASD [1]. Unlike prior methods, we explicitly model *differences* in the global functional organization of the brain, from latent region properties to the observed

fMRI measures. We use bootstrapping to verify the robustness of our region assignments. Subsequent decoding via the Neurosynth meta-analytic database confirms the clinical validity of our results within the neuroimaging literature.

Acknowledgments. This work was supported in part by R01 NS035193 (NINDS) and by R01 MH100028 (NIMH). D. Yang. is also supported by the Autism Speaks Meixner Postdoctoral Fellowship in Translational Research (#9284).

References

1. Pelphrey, K., et al.: Building a social neuroscience of autism spectrum disorder. Current Topics in Behavioral Neuroscience **16** (2014) 215–233
2. Fortunato, S.: Community detection in graphs. Physics Reports **486** (2010) 75–174
3. Lancichinetti, A., Fortunato, S.: Community detection algorithms: a comparative analysis. Physical Review **80** (2010) 1–12
4. Girvan, M., Newman, M.: Community structure in social and biological networks. PNAS **99** (2002) 7821–7826
5. Alexander-Bloch, A., et al.: The discovery of population differences in network community structure: New methods and applications to brain functional networks in schizophrenia. NeuroImage **59** (2012) 3889–3900
6. Mumford, J., et al.: Detecting network modules in fmri time series: A weighted network analysis approach. NeuroImage **52** (2010) 1465–1476
7. Koshino, H., et al.: Functional connectivity in an fmri working memory task in high-functioning autism. Neuroimage **24** (2005) 810–821
8. Paakki, J., et al.: Alterations in regional homogeneity of resting-state brain activity in autism spectrum disorders. Brain Research **1321** (2010) 169–179
9. Venkataraman, A., Kubicki, M., Golland, P.: From brain connectivity models to region labels: Identifying foci of a neurological disorder. IEEE Transactions on Medical Imaging **32** (2013) 2078–2098
10. Venkataraman, A., Rathi, Y., Kubicki, M., Westin, C.F., Golland, P.: Joint modeling of anatomical and functional connectivity for population studies. IEEE Transactions on Medical Imaging **31** (2012) 164–182
11. Jordan, M., Ghahramani, Z., Jaakkola, T.S., Saul, L.K.: An introduction to variational methods for graphical models. Machine Learning **37** (1999) 183–233
12. Yarkoni, T., et al.: Large-scale automated synthesis of human functional neuroimaging data. Nature Methods **8** (2011) 665–670
13. Newman, M., Girvan, M.: Finding and evaluating community structure in networks. Physical Reviews **69** (2004) arXiv:cond–mat/0308217
14. Scheinost, D., et al.: The intrinsic connectivity distribution: a novel contrast measure reflecting voxel level functional connectivity. NeuroImage **62** (2012) 1510–1519
15. Klin, A., et al.: Two-year-olds with autism orient to non-social contingencies rather than biological motion. Nature **459** (2009) 257–261
16. Fischl, B., et al.: Sequence-independent segmentation of magnetic resonance images. NeuroImage **23** (2004) 69–84
17. Smith, S.M., et al.: Advances in functional and structural mr image analysis and implementation as fsl. NeuroImage **23** (2004) 208–219
18. Xia, M., et al.: Brainnet viewer: A network visualization tool for human brain connectomics. PLoS ONE **8** (2013) e68910