# Revisiting Abnormalities in Brain Network Architecture Underlying Autism Using Topology-Inspired Statistical Inference

Sourabh Palande, Vipin Jose, Brandon Zielinski, Jeffrey Anderson, P. Thomas Fletcher, and Bei Wang

University of Utah

**Abstract**

A large body of evidence relates autism with abnormal structural and functional brain connectivity. Structural covariance MRI (scMRI) is a technique that maps brain regions with covarying gray matter density across subjects. It provides a way to probe the anatomical structure underlying intrinsic connectivity networks (ICNs) through analysis of gray matter signal covariance. In this paper, we apply topological data analysis in conjunction with scMRI to explore network-specific differences in the gray matter structure in subjects with autism versus age-, gender- and IQ-matched controls. Specifically, we investigate topological differences in gray matter structure captured by structural correlation graphs (SCGs) derived from three ICNs strongly implicated in autism, namely, the salience network (SN), the default mode network (DMN) and the executive control network (ECN). By combining topological data analysis with statistical inference, our results provide evidence of statistically significant network-specific structural abnormalities in autism.

## 1 Introduction

Autism is a complex developmental disorder characterized by impairment in social interactions, difficulty in verbal and nonverbal communications and repetitive behaviors. Although the exact mechanism of its development remains unclear, there is strong evidence relating autism to abnormal structural and functional connectivity between brain regions. Structural abnormalities can be identified using voxel-based morphometry by comparing gray matter and white matter volumes or densities, as well as using cortical thickness and their respective growth trajectories [Schumann et al., 2010, Zielinski et al., 2014] across diagnostic groups. Although the gross brain differences have been well-documented [Courchesne et al., 2007], investigations into specific regional abnormalities in brain structure have reported divergent results [Stigler et al., 2011].

These inconsistent findings, however, may reflect discrete abnormalities in any given brain network. Research has revealed a finite set of canonical domain-specific *resting state* or *intrinsic connectivity networks* (ICNs) that organize the brain function [Fox et al., 2005]. Many of the regions with reported abnormalities in autism lie within these ICNs. Network-specific differences could account for seemingly contradictory findings from previous studies.

Structural covariance MRI (scMRI) maps regions of gray matter that have a statistically significant correlation with a specific seed region of interest (ROI) across subjects. This suggests shared developmental or genetic influences between the gray matter region and the seed ROI [Zielinski et al., 2010]. Seeley et al. [Seeley et al., 2009] have used scMRI to demonstrate that specific adult dementias affect distinct ICNs and the corresponding gray matter regions. Using a

similar technique, Zielinski et al. [Zielinski et al., 2012] have shown that there are network-specific structural differences between autism and control groups which are consistent with clinical aspects of the disease and that reported functional abnormalities in autism may have a structural basis. Several recent studies have applied the scMRI technique to find evidence of network-specific structural abnormalities in other diseases such as Alzheimer's [Montembeault et al., 2016] and Huntington's [Minkova et al., 2016].

However, scMRI can only reveal shared influences between gray matter voxels and a specific seed ROI. Alternatively, we can model correlations across subjects, between all possible pairs of gray matter regions, as a network. Such a network represents structural relationships between regions which are not captured by conventional scMRI. Comparing these networks across diagnostic groups may provide information that is not given by direct comparisons anchored by network seeds.

Network comparison is not an easy problem, especially when the networks are weighted. Graph-theoretic measures have been proposed previously to compare networks [Bullmore and Sporns, 2009]. However, a major drawback of these measures is their reliance on a fixed topology. That is, these measures are typically based on a binary graph obtained by thresholding the connectivity matrix. The choice of threshold is crucial in such analyses. Different heuristics have been suggested to determine the threshold depending on which properties of the network are of interest. However, it is often not possible to determine a unique optimal threshold.

In this paper, we apply topological data analysis to graphs derived from three ICNs strongly implicated in autism; the default mode network (DMN), the salience network (SN) and the executive control network (ECN). These graphs are referred to as the structural correlation graphs (SCGs). Our method is based on a core technique from topological data analysis known as *persistent homology* [Edelsbrunner et al., 2002] and follows closely the work of Lee et al. [Lee et al., 2012]. We compare topological invariants of two graphs across all thresholds (obtained via persistent homology) to make statistical inference. By combining topological data analysis with statistical inference, our results provide evidence of statistically significant structural abnormalities underlying ICNs in autism. Our results are consistent with the observations of Zielinski et al. [Zielinski et al., 2012] and may offer new insights towards interpreting fine-scale network-specific structural differences.

# 2 Materials and Methods

## 2.1 Structural Correlation Graph

We first use scMRI to identify a set of brain regions, underlying a specific *intrinsic connectivity network* (ICN), according to methods published previously [Zielinski et al., 2010]. This is done by first determining a seed ROI known to anchor a specific ICN and then implementing a generalized linear model (GLM) to identify regions that have covarying gray matter densities with that seed ROI, across subjects. Specifically, for a given seed ROI, separate condition (group)-by-covariate analysis is performed for each voxel. The mean seed gray matter density is the covariate of interest and disease status is the grouping variable. Total brain volume (TBV) and age are included as a covariates-of-no-interest. This design enables us to determine the whole-brain patterns of seed-based structural covariance in each group. To identify regions with significant gray matter density covariance with the seed ROI across subjects in a diagnostic group, one-sample $t$-tests with family-wise error correction are performed.

The regions are identified based on their structural relationship with a specific seed ROI. We model the structural relationships between all pairs of regions as a graph. Correlations between gray matter densities across subjects, for all pairs of regions, are modeled as a weighted, undirected

graph $G(V, E, W)$. The vertices of the graph represent gray matter regions and the edge weights are given by magnitudes of pairwise correlations.

These derived patterns of distributed structural coherence are referred to as structural correlation graphs (SCGs). We compare SCGs derived from three intrinsic connectivity networks strongly implicated in autism, the SN, the ECN and the DMN. In the context of this paper, for simplicity (unless otherwise specified), we describe these SCGs by the name of their corresponding ICNs, namely, SN-SCG, ECN-SCG and DMN-SCG.

### 2.2 Graph Filtration

We extract topological features at multiple scales from a structural correlation graph $G$ by constructing a nested sequence of graphs from $G$, referred to as the graph filtration.

Let $V = \{v_i \mid i = 1, \ldots, n\}$ be the vertex set with $n$ vertices. Let $E$ denote the edge set and $W$ denote the set of edge weights. The edge between vertices $v_i$, $v_j$ is denoted by $e_{ij}$ and its weight is denoted by $w_{ij}$. $|E|$ denotes the number of edges. For a given threshold $\lambda$, we obtain a binary graph $G_\lambda$ by removing edges with weight $w_{ij} \leq \lambda$. The adjacency matrix $A_\lambda = (a_{ij}(\lambda))$ is given by:

$$a_{ij}(\lambda) = \begin{cases} 0 & w_{ij} \leq \lambda \\ 1 & o.w. \end{cases}$$

As $\lambda$ increases, more and more edges are removed from the graph. We can generate a sequence of thresholds, $\lambda_0 = 0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_q$, where $q \leq |E|$ by setting $\lambda_i$'s equal to edge weights arranged in ascending order.

Corresponding to the sequence of thresholds we get a nested sequence of binary graphs, referred to as a *graph filtration* $\mathbf{G}$ [Lee et al., 2012]:

$$G_{\lambda_0} \supseteq G_{\lambda_1} \supseteq G_{\lambda_2} \supseteq \cdots \supseteq G_{\lambda_q}.$$

We can measure the connectivity of a graph by its 0-th Betti number, $\beta_0$, which is the number of connected components in the graph. As the threshold $\lambda$ increases, $\beta_0(G_\lambda)$ of the corresponding graph also increases. The number of connected components of the graphs in filtration $\mathbf{G}$ form a monotonic sequence of integers,

$$\beta_0(G_{\lambda_0}) \leq \beta_0(G_{\lambda_1}) \leq \beta_0(G_{\lambda_2}) \leq \cdots \leq \beta_0(G_{\lambda_q}).$$

Suppose we start with a connected graph $G = G_{\lambda_0}$, we have $\beta_0(G_{\lambda_0}) = 1$ and $\beta_0(G_{\lambda_q}) = |V| = n$ by construction. The plot of $\beta_0(G_\lambda)$ as a function of threshold $\lambda$ is called the $\beta_0$ curve. Given a finite graph with $n$ nodes, there are at most $\binom{n}{2}$ unique edge weights. If we choose the set of all the unique edge weights, sorted in ascending order, to be the thresholds, then with finitely many threshold values, we can estimate the $\beta_0$ curve for all $\lambda$. Computing the $\beta_0$ curve for a given graph could follow the standard algorithm for persistent homology [Edelsbrunner et al., 2002]. In practice, a simpler algorithm relying on the notion of a minimum spanning tree can be used to capture the $\lambda$ values when we are only concerned with tracking the number of components (clusters) during the filtration.

### 2.3 Statistical Inference

Our data consists of subjects divided into two diagnostic groups (samples), autism (ASD) and typically developing controls (TDC). We would like to test whether these two samples come from the

same underlying distribution. More specifically, we want to test whether there are any statistically meaningful differences in the 0-dimensional topology (connectivity) of the SCGs derived from the two samples. We do this by examining the equivalence of the corresponding $\beta_0$ curves.

Let $G$ and $H$ represent the SCGs obtained from autism and control samples respectively with corresponding graph filtrations $\mathbf{G}$ and $\mathbf{H}$. We want to test the null hypothesis,

$$H_0 : \beta_0(G_\lambda) = \beta_0(H_\lambda) \quad \text{for all } \lambda,$$

against the alternative,

$$H_1 : \beta_0(G_\lambda) \neq \beta_0(H_\lambda) \quad \text{for some } \lambda.$$

Since we are dealing with finite graphs, a discrete version of the null hypothesis is stated as,

$$H_0 : \beta_0(G_{\lambda_i}) = \beta_0(H_{\lambda_i}) \quad \text{for all } \lambda_i,\ i = 1, 2, \ldots, q,$$

with the alternative being

$$H_1 : \beta_0(G_{\lambda_i}) \neq \beta_0(H_{\lambda_i}) \quad \text{for some } \lambda_i,\ i = 1, 2, \ldots, q.$$

The distance between two graph filtrations $\mathbf{G}$ and $\mathbf{H}$ with respect to the corresponding $\beta_0$ curves can be defined as :

$$D_q(\mathbf{G}, \mathbf{H}) = \sup_{0 \leq i \leq q} |\beta_0(G_{\lambda_i}) - \beta_0(H_{\lambda_i})|. \tag{1}$$

Intuitively, $D_q$ measures the largest gap between the two $\beta_0$ curves. The $p$-value is the probability that $D_q$ will take a value equal to or greater than the observed value under the null hypothesis. In order to determine this $p$-value, we need the distribution of $D_q$ under the null hypothesis.

### 2.3.1 Permutation Test

A permutation test provides a simple way to estimate the distribution of $D_q$ under the null hypothesis. Let $D_q^*$ denote the value computed from the two original samples. To estimate the sampling distribution of $D_q$, in each iteration, we pool subjects from both samples, randomly permute subject group labels in the pooled data and form two new samples. Then, using these two new samples, we:

1. Construct SCGs for the two samples separately;
2. Apply graph filtration to both SCGs and obtain their corresponding $\beta_0$ curves;
3. Compute the distance $D_q$ between the two curves.

Each permutation gives us a new value of $D_q$. The $p$-value is given by the fraction of $D_q$ values greater than or equal to $D_q^*$.

For an exact permutation test, we would have to compute $D_q$ for all possible permutations of the samples which is often computationally infeasible. Instead, we perform the test on a random subset of all possible permutations. This random permutation test is not exact, however with enough permutations, it can closely approximate the exact test. The $p$-value $\hat{p}$ obtained is an estimate of the true underlying $p$-value $p$. For each permutation, the test ($D_q \geq D_q^*$) is a Bernoulli trial with probability of success $p$. Permutations are independently sampled from a uniform distribution, and $\hat{p}$ is an unbiased estimator of $p$. The standard error of $\hat{p}$ can be approximated by $\sqrt{\hat{p}(1-\hat{p})/N}$, where $N$ is the number of permutations performed.

### 2.3.2 Bootstrap

Permutation tests are widely used to test hypotheses when the underlying distribution is unknown. The only assumption they make is that of exchangeability under null distribution. However, in complex cases like ours, even that assumption may be too strong or too difficult to verify. Bootstrap tests are commonly used to get confidence intervals and they generally have low power compared to permutation tests. But they also make much weaker assumptions about the underlying distribution. Here we will implement a version of bootstrap to estimate sampling distribution of $D_q$. We once again combine both samples to create one pool of subjects. Then we generate two new samples (same size as the original), by sampling subjects *with replacement* from this pool and ignoring their original group labels. With these new bootstrap samples, we proceed to compute the SCGs, the $\beta_0$ curves and the distance $D_q$ in the same way as the permutation test. The $p$-value is computed as the fraction of $D_q$ values obtained from the bootstrap samples that are greater than or equal to $D_q^*$ obtained from the original sample. Note that the only difference between the two tests implemented here is whether we re-sample from the pooled data with (bootstrap) or without (permutation) replacement.

## 2.4 Data Preprocessing

We derive our SCGs from the ICNs previously reported by Zielinski et al. [Zielinski et al., 2012, Zielinski et al., 2010]. Here, we review the preprocessing pipeline. 49 male subjects with autism (ASD), aged 3-22 years, are compared to 49 age-, gender- and IQ-matched typically developing control subjects (TDC). The group-wise mean age (standard deviation) is 13.27 (5.07) for ASD subjects and 13.67 (5.53) for TDC subjects. Images are acquired using a Siemens 3.0 Tesla MRI scanner. Whole brain isotropic MPRAGE image volumes are acquired in the sagittal plane using an 8-channel receive-only RF head coil, employing standard techniques (TR = 2300 ms, TE median = 3 ms, matrix median = $256 \times 256 \times 160$, flip angle = $12^\circ$, voxel resolution = 1 mm$^3$, acquisition time = 9 min 12 sec).

Customized image analysis templates are created by normalizing, segmenting and averaging T1 images using SPM8 according to the processing pipeline proposed in [Altaye et al., 2008, Wilke et al., 2008]. First, images are transformed into standard space using a 12-parameter affine-only linear transformation and segmented into three tissue classes representing gray matter, white matter and cerebrospinal fluid. Then smoothly varying intensity changes as well as artifactual intensity alterations resulting from the normalization step are corrected for using a standard modulation algorithm within SPM8. Finally, the resulting segmented maps are smoothed using a 12-mm full-width at half-maximum Gaussian kernel.

In performing the scMRI analysis, a two-pass procedure is utilized, wherein study-specific templates are first created by segmenting our sample using a canonical pediatric template. Then tissue-specific prior probability maps are created from our sample. The tissue compartments are then re-segmented using this sample-specific template, so that the age range of our sample precisely matches that of the template(s) upon which the ultimate segmentations are based.

## 2.5 Structural Correlation Graphs and Statistical Inference

We would like to construct SCGs that capture structural relationships, across subjects, between all pairs of gray matter regions from a predefined set of regions. We begin by constructing a *whole brain* SCG as follows: 1-mm spheres are placed at grid points of a uniform grid on the entire preprocessed image volume. After applying the gray matter mask, we obtain a set of 7266

regions. The *whole-brain* SCG (denoted Global-SCG) is constructed by computing correlations, across subjects, between all pairs of these regions.

To study network-specific structural relationships, 4-mm-radius spherical seed ROIs are selected within the right frontoinsular cortex (R FI) [Seeley et al., 2009], the right dorsolateral prefrontal cortex (R DLPC) [Seeley et al., 2007] and the right posterior cingulate cortex (R PCC) [Fair et al., 2008]. These regions anchor the salience network (SN), the executive control network (ECN) and the default mode network (DMN), respectively [Seeley et al., 2009, Fair et al., 2008].

For each diagnostic group and each seed ROI, we obtain the set of regions covarying with the seed ROI, following the process described in Section 2.1. The structural covariance maps corresponding to the seed ROI are shown in Fig. 1(a)-(c). Further comparisons in Fig. 2 show that the maps for two diagnostic groups do not completely overlap. Some regions present in the map for the control group are absent in the map for the autism group. Conversely, some regions are present only in the map for the autism group but not in the map for the control group. Table 1 lists the number of regions present in controls but not in autism, in autism but not in controls and in both as well as in either autism or control. A network specific set of ROIs is given by the union of all regions covarying with the corresponding seed ROI, in either the autism group map or the control group map.

| | Controls only | Autism only | Both | Either |
|---|---|---|---|---|
| R PCC (DMN-SCG) | 9 | 21 | 9 | 39 |
| R FI (SN-SCG) | 21 | 1 | 10 | 32 |
| R DLPC (ECN-SCG) | 22 | 5 | 12 | 39 |

Table 1: Number of ROIs identified from scMRI map for a given seed region. Last column shows the number of nodes in the corresponding SCG

Thus, we have one *Global* set of ROIs and three network specific sets of regions. For each set of ROIs, we compute the group level SCGs and perform the random permutation test as described in Section 2.3.1. We perform one million permutations in case of network-specific SCGs and ten thousand permutations in case of the Global-SCG (due to computational constraints). We also perform the bootstrap test as described in Section 2.3.2 with one million bootstrap samples in case of network SCGs and ten thousand in case of the Global-SCG.

### 2.6 Type I Error Rate for Permutation test

The number of permutations in an exact permutation test depends on the number of subjects in the sample. Our random permutations test closely approximates this exact test. However, it should be noted that the test does not take into account the number of ROIs. To check whether our test is robust to increasing number of ROIs, we have implemented the following simulation.

We generate two $m \times n$ data matrices ($m$ is the number of subjects and $n$ is the number of ROIs), where each element is randomly generated from a standard normal distribution so that both the samples (data matrices) come from the same underlying distribution. We compute the SCGs from these data matrices and perform the random permutations test as described in section 4.1. Since the underlying distribution of the samples is the same, rejecting the null hypothesis would be a type I error.

We keep the number of subjects $m = 100$ fixed through the whole simulation. Then we set the number of ROIs $n = 10, 50, 100, 500, 1000$ respectively, and for each value of $n$, we perform 100 random permutation tests as described in the previous paragraph to find the proportion of tests

resulting in type I error. Table 2 shows the findings. At 0.05 level of significance, our test is more conservative meaning it is much less likely to get false positives.

| number of ROIs | 10 | 50 | 100 | 500 | 1000 |
|---|---|---|---|---|---|
| Error rate | 0.01 | 0.02 | 0.03 | 0.05 | 0.03 |

Table 2: Type I error rate for increasing number of ROIs

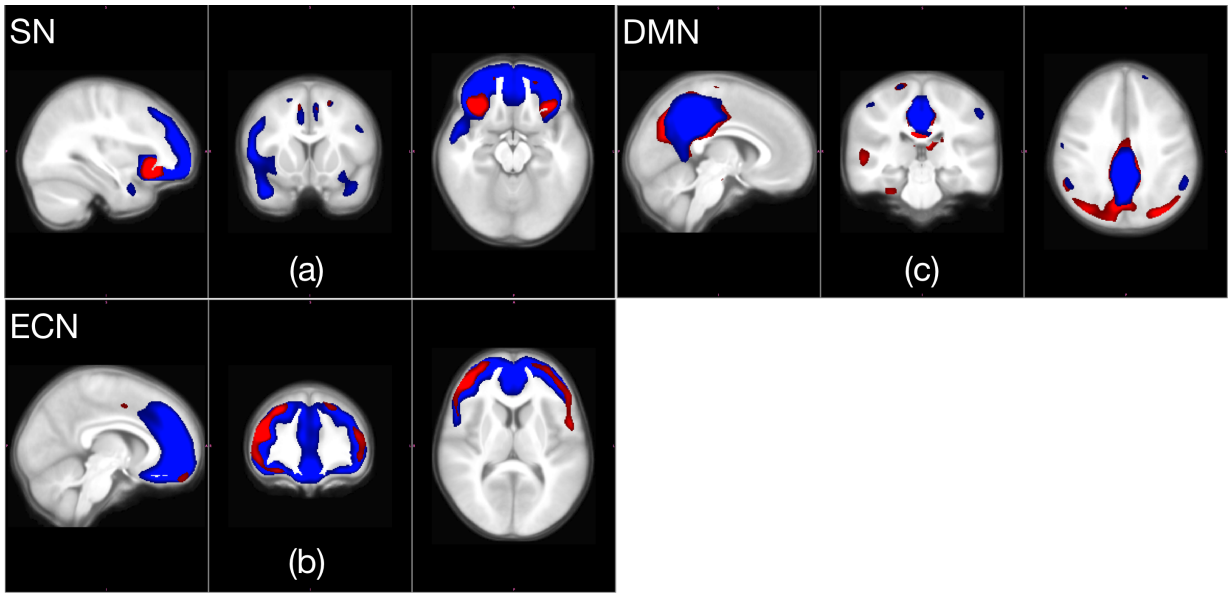


Figure 1: (a)-(c) Structural covariance maps with seed in R FI, R DLPC and R PCC, anchoring SN, ECN and DMN, respectively. Red represents the autism group map, blue represents the control group map.

## 3 Results

We apply both the random permutation test and the bootstrap test to compare SCGs across groups of subjects with autism and typically developing control subjects. We begin by comparing the global SCGs composed of 7266 gray matter regions in the preprocessed images. Then, for a closer analysis, we compare the SCGs generated with seed ROIs anchoring the three ICNs (SN, ECN and DMN), referred to as SN-SCG, ECN-SCG and DMN-SCG. These are much smaller SCGs, composed of 32, 39, 39 ROIs respectively (See Table 1). Recall that the structural covariance maps for the autism and the control groups overlap in very few regions. We construct and compare SCGs derived from sets of regions that are present in either controls or in autism.

The $\beta_0$ curves corresponding to the global SCGs and the seed-specific SCGs are shown in Fig. 3. Table 3 lists the estimated $p$-values obtained from the random permutation tests along with the standard errors in the estimation. Note that the number of permutations is sufficiently large in each case so that the estimation standard errors are small and not likely to affect the outcomes of the permutation tests. Table 4 lists the $p$-values obtained from the bootstrap test. Note that the conclusion from both tests is comparable at 0.05 level of significance. By combining topological data analysis with statistical inference, our results provide evidence of statistically significant network-specific structural abnormalities in autism SN-SCGs.

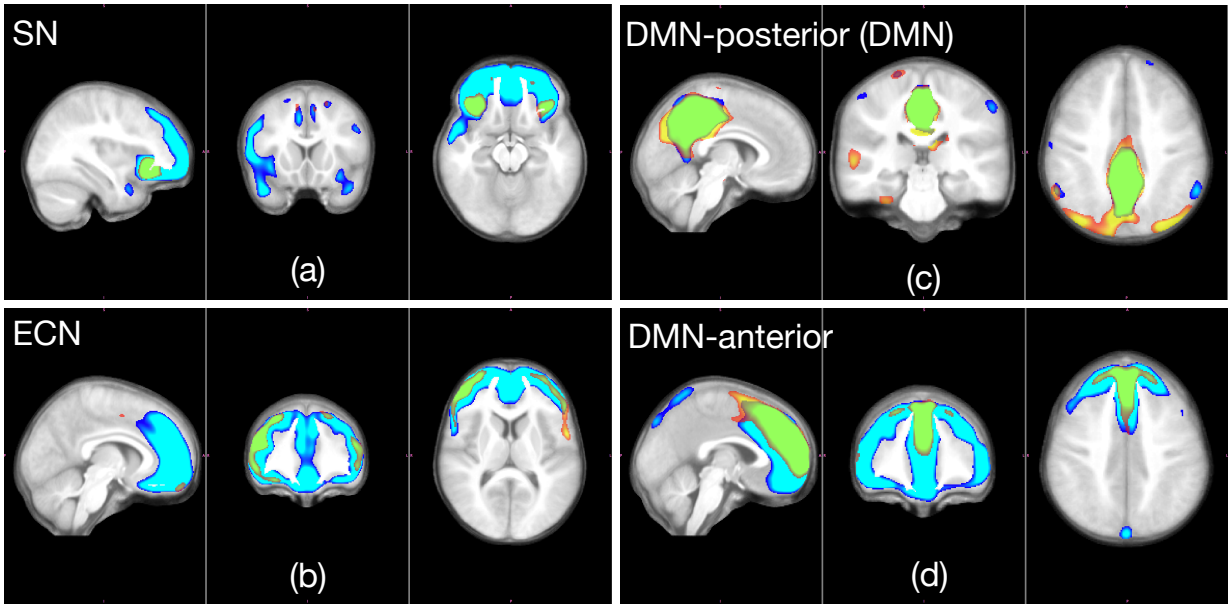


Figure 2: scMRI maps are further illustrated here with red to yellow (autism) and dark blue to light blue (control) color maps. The color gradation indicates increasing statistical significance. The overlapping regions among the autism and control groups are highlighted in green. Note for (c) and (d): Our data consists of subjects with an average age of about 13 years. The underlying structure of the DMN is not fully developed at this age. We include two DMN maps with different seeds to show that the posterior part (c) is not yet integrated with the anterior part (d). In our analysis, we use the posterior covariance map (c) which corresponds to the most common seed for DMN in adults (R PCC).

Figure 4 shows a comparative visualization of the two SN-SCGs that correspond to ASD and TDC groups respectively at the threshold $\lambda = 0.8174$. This is the threshold at which the gap between the corresponding $\beta_0$ curves is the largest ($D_q = 21$, with $\beta_0$ being 6 and 27 for ASD and TDC respectively). The visualization shows that the gray matter densities are much less correlated across the ASD subjects compared to the TDC subjects. The most likely reason behind these differences is that there is more idiosyncrasy in the autism cohort in which regions show greater or lesser cortical thickness or subtle differences in the position of gyri or sulci. There is a growing sense that this idiosyncrasy might be the most reliable difference in autism, meaning that each ASD case is unique; different from group mean values in some way.

| | Global-SCG | DMN-SCG | SN-SCG | ECN-SCG |
|---|---|---|---|---|
| $p$-value estimate | 0.3985 | 0.3658 | **0.00614** | 0.1118 |
| standard error | 0.004895 | 0.000481 | $\mathbf{7.81172 \times 10^{-5}}$ | 0.000315 |

Table 3: Estimated $p$-values for random permutation test on SCGs and corresponding standard errors.

| | Global-SCG | DMN-SCG | SN-SCG | ECN-SCG |
|---|---|---|---|---|
| $p$-value | 0.3685 | 0.3970 | **0.0073** | 0.0890 |

Table 4: $p$-values for bootstrap test on SCGs

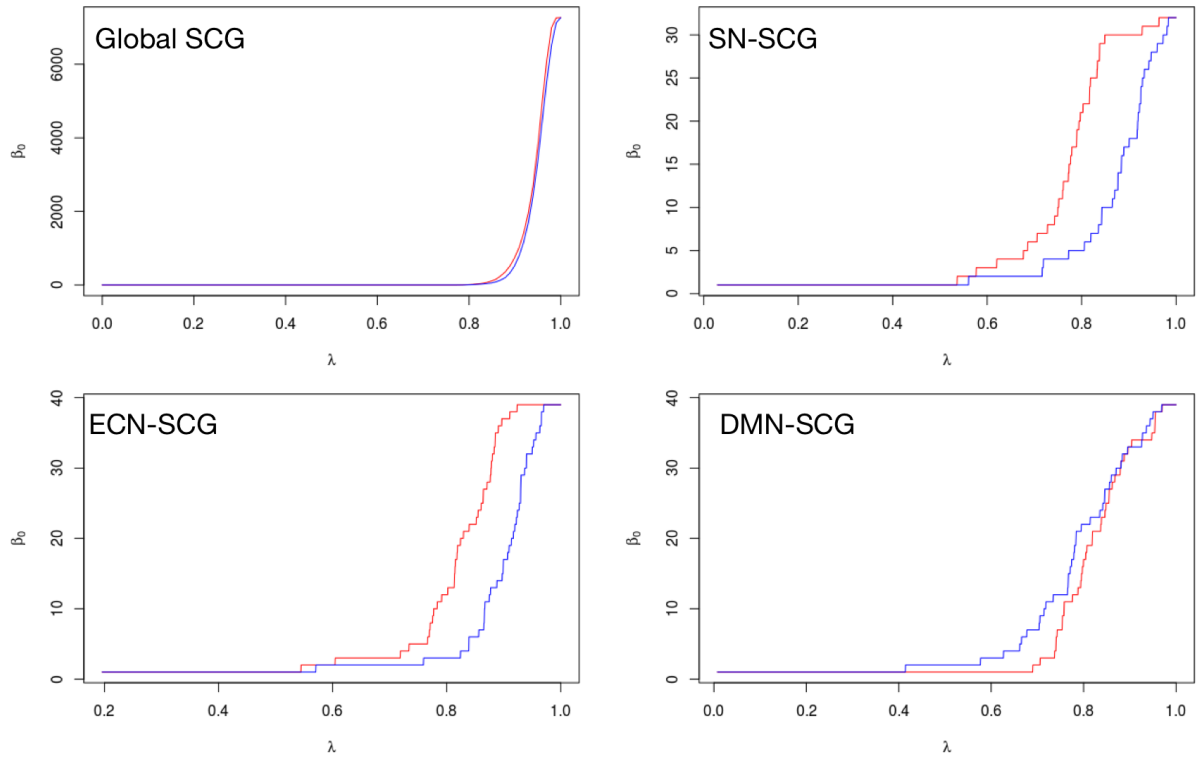


Figure 3: $\beta_0$ curves from Global SCGs as well as SN-SCGs, ECN-SCGs and DMN-SCGs, generated from regions present in either autism (red) or controls (blue) respectively.

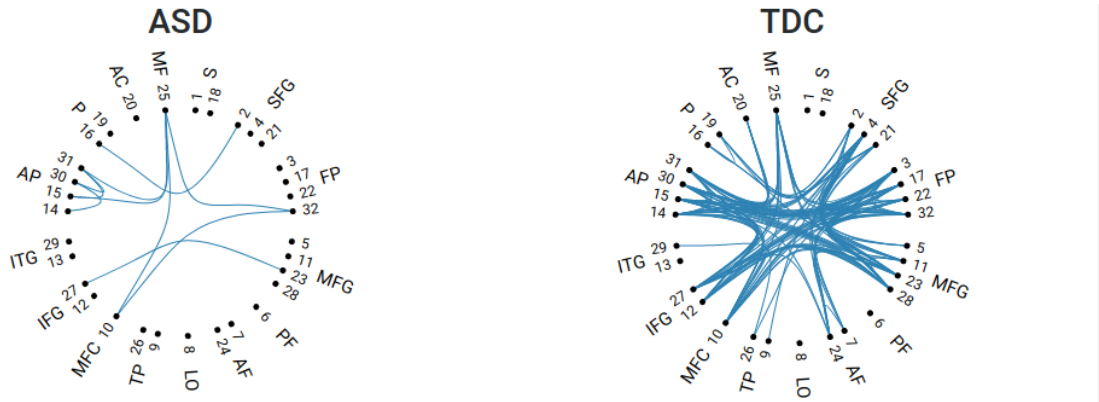


Figure 4: SN-SCG at threshold $\lambda = 0.8174$ that corresponds to $D_q^*$ for both ASD (left) and TDC (right) groups. The ROIs (4mm spheres) are grouped by anatomical regions they are placed in as follows : S - SMA, SFG - sup frontal gyrus, FP - frontal pole, MFG - middle frontal gyrus, PF - post fusiform, AF - anterior Fronto-Insular, LO - Lateral Occipital, TP - Temporal pole, MFC - medial frontal cortex, IFG - inferior frontal gyrus, ITG - inferior temporal gyrus, AP - anterior paracingulate, P - paracingulate, AC - anterior cingulate, MF - medial frontal (ventromedial prefrontal cortex). Image courtesy of Yiliang Shi.

# 4 Conclusion and Discussion

## 4.1 Exact Statistical Inference

In an earlier effort [Palande et al., 2017], used an exact statistical inference method, originally proposed by [Chung et al., 2017], to test the significance of differences in the connectivity of SCGs. Chung et al. have given a combinatorial construction for the exact $p$-value along with an expression for the asymptotic $p$-value in [Chung et al., 2017, Theorem 2]. This construction is the same as the combinatorial construction of the two sample Kolmogorov-Smirnov (KS) test, and uses the fact that the $\beta_0$ sequences are monotonic increasing sequences. Here we discuss why this exact inference method is, in fact, not appropriate in our setting. Recall that our data is represented as two $m \times n$ matrices, $m$ being the number of ASD or TDC subjects and $n$ being the number of ROIs.

1. In [Chung et al., 2017, Theorem 2], the $p$-value for the exact inference obtained combinatorially only depends on $n$, the number of connected components (i.e. the number of ROIs). Therefore the $p$-value computation from [Chung et al., 2017, Theorem 2] is independent of the sample size. However, for a hypothesis test of group differences, we would expect the $p$-values to be dependent on the sample size, therefore [Chung et al., 2017, Theorem 2] is not applicable in our setting.

2. Applying [Chung et al., 2017, Theorem 2] directly in our setting would imply an enumeration of permutations over ROIs with $\binom{2n}{n}$ possibilities. Our standard permutation test, on the other hand, has $\binom{2m}{m}$ possible permutations since each sample consists of $m$ subjects. In our experiment, the set of ROIs for each SCG is selected because of their high covariance with a particular seed region. These ROIs are meant to represent specific attributes of the ASD and TDC populations. Permuting ROIs would result in sets of ROIs that do not represent either of these populations. Again, the exact inference procedure [Chung et al., 2017, Theorem 2] is clearly unrelated to the standard permutation test in our setting.

In summary, the exact inference method proposed by [Chung et al., 2017] is not applicable here to make inferences about population differences between the ASD and TDC subjects.

## 4.2 Highlights, Limitations and Future Directions

Similar covariance of gray matter density across a population of individuals is thought to be mediated by shared genetic or developmental factors [Zielinski et al., 2012]. Structural covariance maps help us identify brain regions highly covarying with a specific seed region. These maps exhibit similar architecture to more familiar intrinsic connectivity networks derived from functional MRI images [Fox and Raichle, 2007], but also exhibit differences in architecture, with extensive regions of the frontal lobe showing greater alignment with salience (SN) and executive control network (ECN) rather than with the default mode network (DMN) [Zielinski et al., 2010]. These differences suggest that while these frontal lobe regions may show greater synchrony with DMN regions such as posterior cingulate cortex and temporoparietal junction, the developmental influences across subjects may be more related to changes in brain attentional networks.

The work in [Zielinski et al., 2012] has performed direct comparisons of the structural covariance maps. The regions in these maps are assigned significance measures based on their covariance with respect to the specified seed region. Using these maps, [Zielinski et al., 2012] have shown the existence of multidimensional, structure-function relationships. Our work helps summarize these relationships using a more robust topological data analysis model. The SCGs encode all pairwise

associations among the ROIs, where the extent of an association is measured by the magnitude of correlations across subjects. Our experiments provide evidence of statistically significant differences in the 0-dimensional topological features of SCGs derived from SN (SN-SCGs).

Our results confirm the significant differences in structural covariance in autism; therefore being consistent with the findings of [Zielinski et al., 2012]. The most striking findings in our results show decreased structural covariance among individuals with autism in the integration of frontal lobe regions with SN hubs in the frontoinsula. Our findings of decreased integration of salience and executive networks, with increased integration of default regions within the frontal lobe align with results investigating functionally-defined intrinsic connectivity networks [Abbott et al., 2016] and suggest that shared developmental influences may underlie the particular specificity of SN connectivity abnormalities in autism [Uddin et al., 2013]

The DMN has also been associated with atypical connectivity in autism [Anderson et al., 2011, Assaf et al., 2010, Chen et al., 2015, Lynch et al., 2013]). Specifically, decreased within-network integration and increased between-network connectivity for the DMN has been reported in autism; see [Padmanabhan et al., 2017]. The increased structural covariance between posterior cingulate cortex and default regions, but decreased anterior DMN integration may suggest that DMN connectivity differences in autism may be related to altered development of frontal lobe components of the network with more idiosyncratic development of frontal network architecture in autism.

Our experiments fail to capture any statistically significant differences in the topology of SCGs derived from DMN. It is possible that the aforementioned idiosyncratic development results in more complex topological differences that are not captured by the pairwise interactions between ROIs. The $\beta_0$ curves only encode 0-order topological features that correspond to the evolution of the number of connected components. Analyzing three-way or four-way interactions between ROIs, capturing higher-order topological features such as tunnels and voids and focusing on specific nodes and edges directly involved in merging components in the graph filtration may provide further insights into the differences in DMN architecture.

Lastly, it should be noted that the SCGs reported here are abstract networks and do not represent physical connectivity between the regions. This limits the interpretability of our results to some extent and deeper analysis is needed in order to quantify and better interpret the differences suggested by the statistical inference.

**Acknowledgements.** This work was supported in part by NSF grant IIS-1513616, NIH grants R01EB022876, K08MH100609 and R01MH080826. Figure 4 was generated using a comparative visualization tool developed by Yiliang Shi, University of Utah.

**Disclosure Statement.** No competing financial interests exist.

## References

[Abbott et al., 2016] Abbott, A. E., Nair, A., Keown, C. L., Datko, M., Jahedi, A., Fishman, I., and Müller, R. A. (2016). Patterns of atypical functional connectivity and behavioral links in autism differ between default, salience, and executive networks. *Cerebral Cortex*, 26(10):4034–4045.

[Altaye et al., 2008] Altaye, M., Holland, S. K., Wilke, M., and Gaser, C. (2008). Infant brain probability templates for MRI segmentation and normalization. *NeuroImage*, 43(4):721–730.

[Anderson et al., 2011] Anderson, J. S., Nielsen, J. A., Froehlich, A. L., Dubray, M. B., Druzgal, T. J., Cariello, A. N., Cooperrider, J. R., Zielinski, B. A., Ravichandran, C., Fletcher, P. T.,

Alexander, A. L., Bigler, E. D., Lange, N., and Lainhart, J. E. (2011). Functional connectivity magnetic resonance imaging classification of autism. *Brain*, 134(12):3742–3754.

[Assaf et al., 2010] Assaf, M., Jagannathan, K., Calhoun, V. D., Miller, L., Stevens, M. C., Sahl, R., O'Boyle, J. G., Schultz, R. T., and Pearlson, G. D. (2010). Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. *NeuroImage*, 53(1):247–256.

[Bullmore and Sporns, 2009] Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198.

[Chen et al., 2015] Chen, C. P., Keown, C. L., Jahedi, A., Nair, A., Pflieger, M. E., Bailey, B. A., and Müller, R. A. (2015). Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism. *NeuroImage: Clinical*, 8:238–245.

[Chung et al., 2017] Chung, M. K., Villalta-Gil, V., Lee, H., Rathouz, P. J., Lahey, B. B., and Zald, D. H. (2017). Exact topological inference for paired brain networks via persistent homology. In Niethammer, M., Styner, M., Aylward, S., Zhu, H., Oguz, I., Yap, P.-T., and Shen, D., editors, *Information Processing in Medical Imaging, Lecture Notes in Computer Science*, volume 10265, pages 299–310. Springer, Cham.

[Courchesne et al., 2007] Courchesne, E., Pierce, K., Schumann, C. M., Redcay, E., Buckwalter, J. A., Kennedy, D. P., and Morgan, J. (2007). Mapping early brain development in autism. *Neuron*, 56(2):399–413.

[Edelsbrunner et al., 2002] Edelsbrunner, H., Letscher, D., and Zomorodian, A. J. (2002). Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533.

[Fair et al., 2008] Fair, D. A., Cohen, A. L., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Barch, D. M., Raichle, M. E., Petersen, S. E., and Schlaggar, B. L. (2008). The maturing architecture of the brain's default network. *Proceedings of the National Academy of Sciences*, 105(10):4028–4032.

[Fox and Raichle, 2007] Fox, M. D. and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9):700–711.

[Fox et al., 2005] Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9673–9678.

[Lee et al., 2012] Lee, H., Kang, H., Chung, M. K., Kim, B. N., and Lee, D. S. (2012). Persistent brain network homology from the perspective of dendrogram. *IEEE Transactions on Medical Imaging*, 31(12):2267–2277.

[Lynch et al., 2013] Lynch, C. J., Uddin, L. Q., Supekar, K., Khouzam, A., Phillips, J., and Menon, V. (2013). Default mode network in childhood autism: Posteromedial cortex heterogeneity and relationship with social deficits. *Biological Psychiatry*, 74(3):212–219.

[Minkova et al., 2016] Minkova, L., Eickhoff, S. B., Abdulkadir, A., Kaller, C. P., Peter, J., Scheller, E., Lahr, J., Roos, R. A., Durr, A., Leavitt, B. R., Tabrizi, S. J., Klöppel, S., and Investigators, T.-H. (2016). Large-scale brain network abnormalities in Huntington's disease revealed by structural covariance. *Human Brain Mapping*, 37(1):67–80.

[Montembeault et al., 2016] Montembeault, M., Rouleau, I., Provost, J.-S., and Brambati, S. M. (2016). Altered gray matter structural covariance networks in early stages of Alzheimer's disease. *Cerebral Cortex*, 26(6):2650.

[Padmanabhan et al., 2017] Padmanabhan, A., Lynch, C. J., Schaer, M., and Menon, V. (2017). The default mode network in autism. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 2(6):476–486.

[Palande et al., 2017] Palande, S., Jose, V., Zielinski, B., Anderson, J., Fletcher, P. T., and Wang, B. (2017). Revisiting abnormalities in brain network architecture underlying autism using topology-inspired statistical inference. In Wu, G., Laurienti, P., Bonilha, L., and Munsell, B., editors, *Connectomics in NeuroImaging (CNI), Lecture Notes in Computer Science*, volume 10511, pages 98–107. Springer, Cham.

[Schumann et al., 2010] Schumann, C. M., Bloss, C. S., Barnes, C. C., Wideman, G. M., Carper, R. A., Akshoomoff, N., Pierce, K., Hagler, D., Schork, N., Lord, C., and Courchesne, E. (2010). Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *Journal of Neuroscience*, 30(12):4419–4427.

[Seeley et al., 2009] Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L., and Greicius, M. D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron*, 62(1):42–52.

[Seeley et al., 2007] Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., and Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9):2349–2356.

[Stigler et al., 2011] Stigler, K. A., McDonald, B. C., Anand, A., Saykin, A. J., and McDougle, C. J. (2011). Structural and functional magnetic resonance imaging of autism spectrum disorders. *Brain Research*, 1380:146–161.

[Uddin et al., 2013] Uddin, L. Q., Supekar, K., Lynch, C. J., Khouzam, A., Phillips, J., Feinstein, C., Ryali, S., and Menon, V. (2013). Salience network-based classification and prediction of symptom severity in children with autism. *JAMA Psychiatry*, 70(8):869–879.

[Wilke et al., 2008] Wilke, M., Holland, S. K., Altaye, M., and Gaser, C. (2008). Template-O-Matic: A toolbox for creating customized pediatric templates. *NeuroImage*, 41(3):903–913.

[Zielinski et al., 2012] Zielinski, B. A., Anderson, J. S., Froehlich, A. L., Prigge, M. B. D., Nielsen, J. A., Cooperrider, J. R., Cariello, A. N., Fletcher, P. T., Alexander, A. L., Lange, N., Bigler, E. D., and Lainhart, J. E. (2012). scMRI reveals large-scale brain network abnormalities in autism. *PLOS ONE*, 7(11):1–14.

[Zielinski et al., 2010] Zielinski, B. A., Gennatas, E. D., Zhou, J., and Seeley, W. W. (2010). Network-level structural covariance in the developing brain. *Proceedings of the National Academy of Sciences*, 107(42):18191–18196.

[Zielinski et al., 2014] Zielinski, B. A., Prigge, M. B. D., Nielsen, J. A., Froehlich, A. L., Abildskov, T. J., Anderson, J. S., Fletcher, P. T., Zygmunt, K. M., Travers, B. G., Lange, N., Alexander, A. L., Bigler, E. D., and Lainhart, J. E. (2014). Longitudinal changes in cortical thickness in autism and typical development. *Brain*, 137(6):1799–1812.