



# A Statistical Method to Distinguish Functional Brain Networks

André Fujita<sup>1\*</sup>, Maciel C. Vidal<sup>1</sup> and Daniel Y. Takahashi<sup>2</sup>

<sup>1</sup> Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, São Paulo, Brazil,

<sup>2</sup> Department of Psychology and Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

One major problem in neuroscience is the comparison of functional brain networks of different populations, e.g., distinguishing the networks of controls and patients. Traditional algorithms are based on search for isomorphism between networks, assuming that they are deterministic. However, biological networks present randomness that cannot be well modeled by those algorithms. For instance, functional brain networks of distinct subjects of the same population can be different due to individual characteristics. Moreover, networks of subjects from different populations can be generated through the same stochastic process. Thus, a better hypothesis is that networks are generated by random processes. In this case, subjects from the same group are samples from the same random process, whereas subjects from different groups are generated by distinct processes. Using this idea, we developed a statistical test called ANOGVA to test whether two or more populations of graphs are generated by the same random graph model. Our simulations' results demonstrate that we can precisely control the rate of false positives and that the test is powerful to discriminate random graphs generated by different models and parameters. The method also showed to be robust for unbalanced data. As an example, we applied ANOGVA to an fMRI dataset composed of controls and patients diagnosed with autism or Asperger. ANOGVA identified the cerebellar functional sub-network as statistically different between controls and autism ( $p < 0.001$ ).

**Keywords:** random graph, analysis of variance, graph spectrum, network science, functional connectivity, anogva

## OPEN ACCESS

### Edited by:

Jean-Baptiste Poline,  
University of California, Berkeley, USA

### Reviewed by:

Jason W. Bohland,  
Boston University, USA  
Amir Omidvarnia,  
Florey Institute of Neuroscience and  
Mental Health, Australia

### \*Correspondence:

André Fujita  
fujita@ime.usp.br

### Specialty section:

This article was submitted to  
Brain Imaging Methods,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 26 September 2016

**Accepted:** 30 January 2017

**Published:** 14 February 2017

### Citation:

Fujita A, Vidal MC and Takahashi DY  
(2017) A Statistical Method to  
Distinguish Functional Brain Networks.  
*Front. Neurosci.* 11:66.  
doi: 10.3389/fnins.2017.00066

## 1. INTRODUCTION

Graphs are widely used to represent interactions such as functional connectivity among brain regions (Bullmore and Sporns, 2009), social networks (Scott, 2012), and molecular interactions (Barabási and Oltvai, 2004). Once interaction graphs are obtained, a common problem is to verify if graphs of different populations are comparable or not. Standard approaches are based on algorithms to determine isomorphism—one-to-one correspondence—or how close to isomorphism are different graphs. For example, if the vertices of the graphs are labeled, one may count how many times a certain edge is present in each population. Otherwise, one may try to find an isomorphic sub-network, which problem is known to be NP-complete. Both strategies are not the most adequate given that real-world interaction graphs are heterogeneous and present intrinsic randomness. For example, functional brain networks of different individuals are structurally different, even belonging to the same group. Even networks from the same subject can change if measured on different times. Notice that in both examples, algorithms based on isomorphism will falsely discriminate graphs belonging to the same group or state. One solution for this problem is to assume that real-world graphs are generated by probabilistic processes (random graph models) and

then test whether populations of graphs are generated by the same random graph model. However, the model that generated the graph is rarely known in practice. Thus, the first step to discriminate random graphs is to identify highly distinctive features across different graph models.

The spectrum (set of eigenvalues) of the adjacency matrix describes several structural properties of a random graph, such as the number of walks, diameter, and cliques (Van Mieghem, 2010). Therefore, it is a natural candidate to distinguish graphs generated by different processes (Van Mieghem, 2010). Indeed, in general, the graph spectrum is a better and more general characterization of complex networks in comparison to other features, such as the number of edges, degree, and centrality measures (Takahashi et al., 2012). By analyzing the graph spectrum (Takahashi et al., 2012) defined the concept of graph spectral entropy and developed statistical methods on graphs for (i) model selection, (ii) parameter estimation, and (iii) a hypothesis test to discriminate whether two populations of graphs are generated by the same random graph model and parameters. These methods are important from a methodological viewpoint because it provided formal methods for the statistical inference using graph samples. From a practical perspective, these methods were essential to identify novel brain sub-networks associated with attention deficit hyperactivity disorder (ADHD) (Sato et al., 2013) and autism spectrum disorder (ASD) (Sato et al., 2015). However, the random graph comparison method introduced in Takahashi et al. (2012) cannot verify simultaneously whether three or more groups of graphs are generated by the same random graph model. For example, it is not possible to simultaneously test the equality of the functional brain networks of controls, autism, and Asperger subjects. One possible solution would be to compare the groups in a pairwise manner, nevertheless these methods in general give an inadequate control of type I error. Here, we introduce a statistical method to discriminate two or more populations of graphs simultaneously, namely ANOGVA (Analysis of Graph Structure Variability). Intuitively, if the original test proposed by Takahashi et al. (2012) is equivalent to a *t*-test, our proposed test is equivalent to the analysis of variance (ANOVA) (Fisher, 1918).

We illustrate the performance of ANOGVA through simulation studies and demonstrate the power of the test for identifying small differences in the parameters of the random graph models. We also applied our method to study the whole brain functional magnetic resonance imaging (fMRI) data of 908 controls and patients diagnosed with autism or Asperger.

## 2. MATERIALS AND METHODS

Let us first formalize our problem. Given  $k$  populations of graphs  $g_1, g_2, \dots, g_k$  where each population  $g_i (i = 1, \dots, k)$  is composed of  $|g_i|$  graphs, we would like to verify whether the graphs of the  $k$  populations were generated by the same probabilistic process, i.e., by the same random graph model [e.g., Erdős-Rényi (Erdős and Rényi, 1960), Watts-Strogatz (Watts and Strogatz, 1998), and Barabási-Albert (Barabasi and Albert, 1999) random graph models] and set of parameters. First we will describe the graph

spectrum. Based on the graph spectrum, we will define the Kullback-Leibler divergence between two random graphs that will be used to define the ANOGVA statistics.

### 2.1. Graphs and Graph Spectrum

A graph is a pair of sets  $G = (V, E)$  where  $V$  is a set of  $n$  vertices and  $E$  is a set of  $m$  edges that connect two vertices of  $V$ . A random graph  $g$  is a family of graphs, where the members of the family are generated by some probability law.

An undirected graph  $G$  with  $n$  vertices can be represented by its  $n \times n$  adjacency matrix  $\mathbf{A}$  where  $\mathbf{A}_{ij} = \mathbf{A}_{ji} = 1$  if vertices  $i$  and  $j$  are connected, and 0 otherwise. The spectrum of  $G$  is the set of eigenvalues ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ) of the adjacency matrix  $\mathbf{A}$ . For an undirected graph, the adjacency matrix is symmetric, and thus, its eigenvalues are real (Strang, 2011).

Given a set of random graphs  $g$  generated by the same probability law, the set of eigenvalues  $\Lambda$  are random vectors. Let  $\delta$  be the Dirac delta function and the brackets " $\langle \cdot \rangle$ " indicate the expectation with respect to the probability law of the random graph, the spectral distribution of a random graph  $g$  is defined as:

$$\rho_g(\lambda) = \lim_{n \rightarrow \infty} \left\langle \frac{1}{n} \sum_{j=1}^n \delta(\lambda - \lambda_j / \sqrt{n}) \right\rangle. \quad (1)$$

The spectral distribution is directly associated with the structural features of the graphs (Albert and Barabási, 2002) and can be considered as a fingerprint of the random graph, where each random graph model is associated with a specific spectral distribution  $\rho_g$  (Van Mieghem, 2010).

### 2.2. Kullback-Leibler Divergence

Once the spectral distribution of a graph is defined, we can describe a measure of similarity between two spectral distributions. If two spectral distributions are different, then the respective graphs should be different.

Let  $\rho_{g_1}$  and  $\rho_{g_2}$  be the spectral distributions of random graphs  $g_1$  and  $g_2$ , respectively. If the support of  $\rho_{g_2}$  contains the support of  $\rho_{g_1}$ , the Kullback-Leibler (KL) divergence between two spectral distributions  $\rho_{g_1}$  and  $\rho_{g_2}$  is (Kullback and Leibler, 1951):

$$KL(\rho_{g_1} | \rho_{g_2}) = \int_{-\infty}^{+\infty} \rho_{g_1}(\lambda) \log \frac{\rho_{g_1}}{\rho_{g_2}} d\lambda, \quad (2)$$

otherwise,  $KL(\rho_{g_1} | \rho_{g_2}) = +\infty$  (we assume  $0 \log \frac{0}{0} = 0$ ).

For Equation (2),  $\rho_{g_2}$  is called the reference measure. The KL divergence is non-negative and zero if and only if  $\rho_{g_1}$  and  $\rho_{g_2}$  are equal.

### 2.3. Estimation of the Spectral Density

To estimate the spectral density ( $\hat{\rho}_g$ ), we use the same procedure described by Takahashi et al. (2012). First we compute the eigenvalues of the adjacency matrix of the graph. Then, we apply a Gaussian kernel regression using the Nadaraya-Watson estimator (Nadaraya, 1964) for regularization of the estimator. Finally, we normalize the density to obtain the integral below the curve equal to one. In general, smaller sample and/or graph

sizes require larger bandwidth and smaller bin numbers for the smoothing kernels. The opposite holds for larger sample and/or graph sizes. The exact bandwidth size and bin number that maximize the statistical power depends on the data and the alternative hypotheses, but some rules of thumb exist in the literature that have shown good performance in our simulations. The bandwidth of the kernel is chosen as  $\frac{\lambda_1 - \lambda_n}{\text{number of bins}}$  (Sain, 1996), where the number of bins is selected by using the Sturges' criterion (Sturges, 1926). Type I errors are controlled by our bootstrap procedure discussed below for any choice of bandwidth size and bin number.

## 2.4. Analysis of Graph Structure Variability—ANOGVA

We are now able to describe ANOGVA. Given  $k$  populations of graphs  $g_1, g_2, \dots, g_k$ , the test consists of verifying if all populations of graphs were generated by the same random graph model. For this, we test if all the spectral distributions are equal.

Let  $\hat{\rho}_{g_1}, \hat{\rho}_{g_2}, \dots, \hat{\rho}_{g_k}$  be the estimated spectral distributions of populations of graphs  $g_1, g_2, \dots, g_k$ , respectively, where  $\hat{\rho}_{g_i}$  ( $i = 1, \dots, k$ ) is the average of the graphs spectra in population  $g_i$ . Also, set  $\hat{\rho}_{g_M} = \frac{\sum_{i=1}^k \hat{\rho}_{g_i}}{k}$ . The support of  $\hat{\rho}_{g_M}$  includes the support of  $\hat{\rho}_{g_i}$  for any  $i$ . Formally, we test:

$H_0: KL(\rho_{g_1}, \rho_{g_M}) = KL(\rho_{g_2}, \rho_{g_M}) = \dots = KL(\rho_{g_k}, \rho_{g_M}) = 0$ , i.e., the graphs from  $g_1, g_2, \dots, g_k$  are generated by the same random graph model (the spectral distributions are equal).

$H_1$ : "At least one population of graphs is generated in a different manner".

We will use the statistic  $\Delta = \sum_{i=1}^k KL(\hat{\rho}_{g_i}, \hat{\rho}_{g_M})$  to build the test statistic. This statistic is the generalization of the Jensen-Shannon divergence (Jensen, 1906; Shannon, 1948) for  $k > 2$ . Under the null hypothesis, we expect small  $\Delta$ , while large  $\Delta$  suggests a rejection of the null hypothesis. The exact or asymptotic distribution of  $\Delta$  under the null hypothesis is not known; therefore, we use a computational procedure based on the permutation test to construct the empirical distribution. The steps for the permutation test is as follows:

1. Construct permuted samples  $g_i^*$ , for  $i = 1, \dots, k$  by resampling (without replacement)  $|g_i|$  graphs from the entire dataset  $\{g_1 \cup g_2 \cup \dots \cup g_k\}$ .
2. Calculate  $\hat{\rho}_{g_i^*}$  for each  $g_i^*$  ( $i = 1, \dots, k$ ).
3. Calculate  $\hat{\Delta}^* = \sum_{i=1}^k KL(\hat{\rho}_{g_i^*}, \hat{\rho}_{g_M})$ .
4. Repeat steps 1 to 4 until the desired number of replications is obtained.
5. The  $p$ -value for the observed statistic  $\hat{\Delta}$  is the fraction of times  $\hat{\Delta}^*$  obtained in the permuted dataset is at least as large as  $\hat{\Delta}$  estimated in the original dataset.

**Figure 1** illustrates the idea behind ANOGVA. In summary, the spectral distribution of each population is compared to the reference distribution (the average of the spectral distributions,  $\rho_{g_M}$ ). If the sum of the distances (KL divergence) is large, it means that at least one of the spectral distributions is different when compared to the reference ( $\rho_{g_M}$ ). In other words, at least one of the populations of graphs was generated by a different random graph model and/or set of parameters.

## 2.5. Graph Models

### 2.5.1. Erdős-Rényi Random Graph

Erdős-Rényi random graph (Erdős and Rényi, 1960) is defined as  $n$  labeled vertices where each pair of vertices ( $v_i, v_j$ ) is connected by an edge with a given probability  $p$ .

### 2.5.2. Geometric Random Graph

A geometric random graph (Penrose, 2003) is constructed by randomly placing  $n$  vertices in a space  $\mathbb{R}^d$  according to a specified probability distribution (usually, uniform distribution) and connecting two vertices by an edge if their distance (according to some metric) is smaller than a certain neighborhood radius  $r$ .

### 2.5.3. K-regular Random Graph

A  $k$ -regular random graph (Meringer, 1999) is a graph where every vertex has the same degree (number of adjacent vertices). A  $k$ -regular random graph with degree  $deg$  is called a  $deg$ -regular random graph or regular random graph of degree  $deg$ .

### 2.5.4. Watts-Strogatz Random Graph

Watts-Strogatz random graph (Watts and Strogatz, 1998) presents small-world properties (short average path lengths) and a higher transitivity (clustering coefficient) than Erdős-Rényi random graphs.

The algorithm for constructing a Watts-Strogatz random graph is as follows:

Input: Let  $n$ ,  $nei$ , and  $p_w$  be the number of vertices, mean degree, and the rewiring probability, respectively.

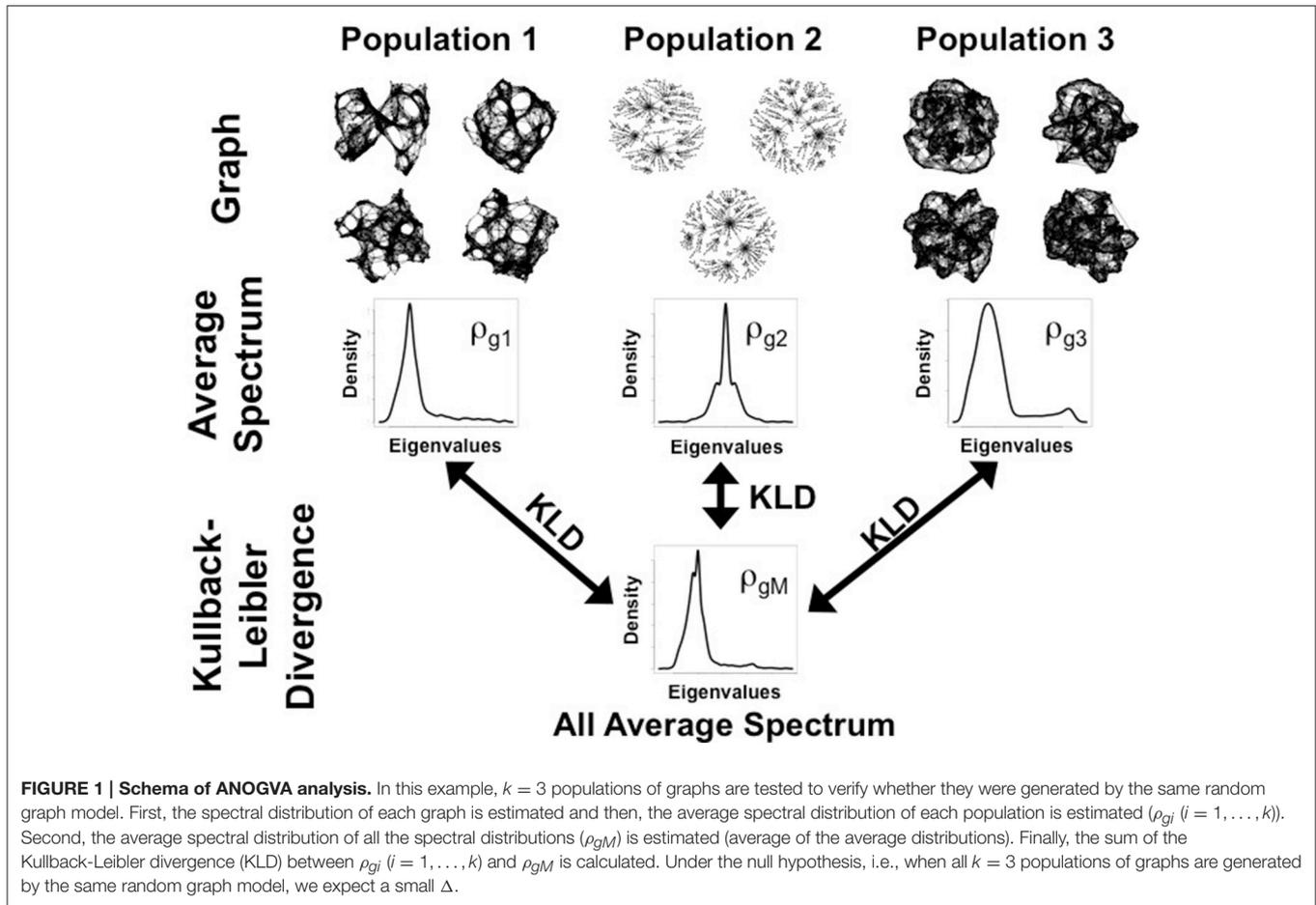
1. construct a ring lattice with  $n$  vertices, in which every vertex is connected to its first  $nei$  neighbors ( $\frac{nei}{2}$  on either side);
2. choose a vertex and the edge that connects it to its nearest neighbor in a clockwise sense. With probability  $p_w$ , reconnect this edge to a vertex chosen uniformly at random over the entire ring. This process is repeated by moving clockwise around the ring, considering each vertex in turn until one lap is completed. Next, the edges that connect vertices to their second-nearest neighbors clockwise are considered. As in the previous step, each edge is randomly rewired with probability  $p_w$ ; continue this process, circulating around the ring and proceeding outward to more distant neighbors after each lap, until each edge in the original lattice has been considered once.

**Output:** the Watts-Strogatz random graph.

### 2.5.5. Barabási-Albert Random Graph

Barabási-Albert random graph (Barabasi and Albert, 1999) has a power-law degree distribution due to preferential attachment of vertices (the more connected a vertex is, the more likely it is to receive new edges).

Barabasi and Albert (1999) proposed the following construction: start with a small number of ( $n_0$ ) vertices. At each iteration, add a new vertex with  $m_1$  ( $m_1 \leq n_0$ ) edges that connect the new vertex to  $m_1$  different vertices already present in the graph. To select which vertices the new vertex will connect, assume that the probability that a new vertex will be connected to vertex  $v_i$  is proportional to the degree of vertex  $v_i$  and the



scaling exponent  $p_s$  ( $P(v_i) \sim degree(v_i)^{p_s}$ , where  $degree(v_i)$  is the number of adjacent edges of vertex  $v_i$  in the current iteration).

### 2.6. Simulations Description

For the simulation studies, we analyzed five random graph models, namely Erdős-Rényi (Erdős and Rényi, 1960), geometric (Penrose, 2003),  $k$ -regular (Meringer, 1999), Watts-Strogatz (Watts and Strogatz, 1998), and Barabási-Albert (Barabasi and Albert, 1999). All simulations were carried out in R using the package `igraph`. The name of the functions and respective parameters analyzed in our study are: function `erdos.renyi.game` - parameter `p`; function `grg.game` - parameter `radius`; function `k.regular.game` - parameter `k`; function `watts.strogatz.game` - parameter `p`; and function `barabasi.game` - parameter `power`. For the Watts-Strogatz random graph model, we selected the parameter `p` as the varying parameter in our simulations because it changes the graph structure without altering the number of edges.

#### 2.6.1. Simulation 1

To verify the controls of types I and II errors of ANOGVA in inferring whether populations of graphs are equally generated, we constructed four scenarios.

1. Scenario 1—under the null hypothesis: We constructed three populations of graphs ( $g_1, g_2, g_3$ ) with  $n = 300$ . Graphs were generated by an Erdős-Rényi (ER) random graph model (Erdős and Rényi, 1960). The parameters  $p_j$ 's (the probability of inclusion of an edge) of the ER random graph models were generated by a truncated normal distribution with lower bound, upper bound, mean and variance set as 0, 10, 1, and 1, respectively, for each graph, and then linearly normalized to the interval  $[0; 1]$  (i.e., let  $p_j$ ,  $j = 1, \dots, |g_1|+|g_2|+|g_3|$ , be the random numbers generated by the truncated normal distribution. Then, we linearly normalize each  $p_j$  to the interval  $[0; 1]$  by dividing it by the upper bound). Notice that this normalization is necessary because the parameter of the ER random graph model is the probability of inclusion of an edge. The parameters were generated in the same manner for all graphs of the three populations. This scenario was constructed to evaluate the control of the rate of false positives under the null hypothesis (all graphs were generated by the same random process).
2. Scenario 2—under the alternative hypothesis: Populations of graphs  $g_1$  and  $g_3$  were constructed as described in scenario 1. The parameters of the graphs of population  $g_2$  were generated by a truncated normal distribution with mean 1.5 and unit variance, and then linearly normalized to the interval  $[0; 1]$ .

- This scenario was constructed to evaluate the power of the test when the parameters of the graphs are different.
3. Scenario 3—under the alternative hypothesis: Populations of graphs  $g_1$  and  $g_3$  were constructed as described in scenario 1. The parameters of the graphs of population  $g_2$  were generated by a truncated normal distribution with mean and variance set as 1.5, and 1, respectively, and then linearly normalized to the interval [0; 1]. The number of graphs are set as  $|g_1| = |g_3| = 125$  and  $|g_2| = 50, 75, 100$ . This scenario was constructed to evaluate the power of the test when the datasets are not balanced.
  4. Scenario 4—under the alternative hypothesis: Populations of graphs  $g_1$  and  $g_3$  were constructed as described in scenario 1. The population  $g_2$  is generated by a Watts-Strogatz random graph model (Watts and Strogatz, 1998) with the dimension of the starting lattice equal to one and the neighborhood within which the vertices of the lattice is connected equal to four. The rewiring probability is generated by a truncated normal distribution with lower bound, upper bound, mean and variance set as 0, 10, 1, and 1, respectively, and then linearly normalized to the interval [0; 1]. The purpose of this scenario is to evaluate the power of the test when one of the populations of graphs is generated by a different random graph model. To make this scenario more realistic, we misclassified the labels of the graphs in rates of 0, 20, 30, and 40%.

For scenarios 1, 2, and 4, the number of graphs varied ( $|g_1| = |g_2| = |g_3| = 50, 75, 100, 125$ ). For each number of graphs, the experiment was repeated 1,000 times. Then, we constructed receiver operating characteristic (ROC) curves to evaluate the control of the rate of false positives and the power of the proposed statistical test.

## 2.6.2. Simulation 2

One alternative to verify whether graphs are generated by the same probabilistic process is the application of ANOVA on the features of the graphs (e.g., the betweenness centrality). To compare the performance of ANOVA with a simple application of ANOVA on the features, we constructed three populations of graphs ( $g_1, g_2, g_3$ ), each one composed of  $|g_1| = |g_2| = |g_3| = 100$  graphs. The number of vertices was set to  $n = 300$ . The graphs were generated by Erdős-Rényi (Erdős and Rényi, 1960), geometric (Penrose, 2003),  $k$ -regular (Meringer, 1999), Watts-Strogatz (Watts and Strogatz, 1998), and Barabási-Albert (Barabasi and Albert, 1999) random graph models. The parameter of the random graph models were generated by normal distributions with mean one for  $g_1$  and  $g_3$  and mean 1.5 for  $g_2$ , all of them with unit variance. Then, they were linearly normalized to the interval [0; 1]. The parameters for these five graph models are: the probability of adding an edge for Erdős-Rényi, the radius for the geometric, i.e., a vertex is connected to all vertices at distance smaller than the radius (we set the space dimension where the vertices are located to two), the vertex degree for the  $k$ -regular, the rewiring probability for Watts-Strogatz, and the power of the preferential attachment probability for the Barabási-Albert. Since the parameters for  $k$ -regular and

Barabási-Albert random graph models are integers, we took the floor of the parameter normalized between zero and one and multiplied it by 10. Then, we measured five features that are commonly considered in the literature, namely the number of edges, the average betweenness centrality (Freeman, 1977) (number of shortest paths from all vertices to all others that pass through that vertex), the average closeness centrality (Bavelas, 1950) (one divided by the sum of the distances from one vertex to all other vertices), assortativity (Newman, 2002) (preference of a vertex to attach to others in terms of degree), and transitivity (Wasserman and Faust, 1994) (relative number of triangles in the graph, compared to total number of connected triples of vertices) for each graph. Finally, we applied ANOVA on these features and compared the statistical power with ANOVA. A  $p$ -value cutoff of 0.05 was set to determine whether the test rejected the hypothesis that the three populations were generated by the same random graph model. This experiment was repeated 1,000 times and the proportion of rejected null hypothesis was calculated.

## 2.7. Application to the ABIDE Dataset

### 2.7.1. Dataset Description

A large resting state fMRI dataset initially composed of 908 individuals comprising controls and subjects diagnosed with autism and Asperger was downloaded from the ABIDE I Consortium website ([http://fcon\\_1000.projects.nitrc.org/indi/abide/](http://fcon_1000.projects.nitrc.org/indi/abide/)). The ABIDE I dataset is fully anonymized in compliance with the Health Insurance Portability and Accountability (HIPAA) Privacy Rules and the 1,000 Functional Connectomes Project/INDI protocols. Protected health information identifiers and face information from structural images are not included in this dataset. For further details, refer to Di Martino et al. (2014).

To pre-process the brain imaging data, we carried out the Athena pipeline downloaded from (<http://www.nitrc.org/plugins/mwiki/index.php/neurobureau:AthenaPipeline>) which can be summarized as follows: exclusion of the first four scans; slice timing correction; deoblique dataset; correction for head movements; masking the volumes to exclude non-brain regions; co-registration of mean image to the respective anatomic image of the subject; spatial normalization to MNI space ( $4 \times 4 \times 4$  mm resolution); extraction of BOLD time series from white matter and cerebrospinal-fluid; removing effects of white matter, cerebrospinal-fluid, motion and trend using linear multiple regression; temporal band-pass filter ( $0.009 < f < 0.08$  Hz); and spatial smoothing the filtered data using a Gaussian filter (FWHM = 6 mm). We used the CC400 atlas (Craddock et al., 2012) to define the 351 regions of interest (ROIs). Then, we removed 35 ROIs including the ventricles (identified by using the MNI atlas), resulting 316 ROIs (vertices of the graph) for the construction of functional brain networks. The average time series within the ROIs were considered as to be the region representatives. The head movement during magnetic resonance scanning was treated by using the “scrubbing” procedure described by Power et al. (2012). Individuals with a number of adequate scans less than 100 after the “scrubbing” procedure were discarded. It resulted in 896 subjects for subsequent analyses. Thus, the dataset used in this study is composed of 529 controls (430 males, mean age  $\pm$  standard deviation,  $17.47 \pm 7.81$

years), 285 autistic patients (255 males,  $17.53 \pm 7.13$  years), and 82 Asperger patients (70 males,  $19.97 \pm 11.37$  years). For further details, see **Table 1**.

### 2.7.2. Functional Brain Networks

A functional brain network can be modeled as a graph, i.e., a pair of sets  $G = (V, E)$ , in which  $V$  is the set of regions of interest—ROIs (vertices), and  $E$  is the set of functional connectivity (edges) among ROIs. In the current study, the functional connectivity between two ROIs was obtained by calculating the Spearman's correlation coefficient between ROIs  $i$  and  $j$  ( $i, j = 1, \dots, 316$ ) for each individual  $q = 1, \dots, 896$ . Thus, a functional brain network  $G^q$  with 316 ROIs can be represented by its adjacency matrix  $A^q$  with  $316 \times 316$  elements  $A_{ij}^q$  containing the association between the ROIs  $i$  and  $j$  ( $i, j = 1, \dots, 316$ ;  $q = 1, \dots, 896$ ). Site, gender, age effects and the proportion of removed volumes by the “scrubbing” were modeled with a generalized linear model (GLM) with the strength of association ( $z$ -value associated with the Spearman correlation coefficient) as the response variable and the effects as covariates. The residuals of the model were considered as the connectivity filtered by these effects. Then,  $p$ -values for each Spearman's correlation coefficient (without site, gender, age effects) between ROIs  $i$  and  $j$  were calculated and corrected for the false discovery rate (FDR) (Benjamini and Hochberg, 1995). The choice for the Spearman's correlation is based on the fact that it is robust to outliers and is also able to

identify non-linear monotonic relationships (de Siqueira Santos et al., 2014).

Functional sub-networks were defined as the same as defined by Sato et al. (2015), namely somatomotor, visual, default-mode, cerebellar, and fronto-parietal.

## 3. RESULTS

### 3.1. Simulations

We evaluated the controls of types I and II errors of ANOVA in verifying whether graphs are generated by the same random graph model.

**Figures 2A–D** illustrate, respectively, the ROC curves obtained by simulating scenarios 1, 2, 3, and 4 described in Section 2.6.1. The x-axis represents the  $p$ -value threshold and the y-axis represents the proportion of rejected null hypothesis given a  $p$ -value threshold. The ROC curve under the null hypothesis lies in the diagonal. Under the alternative hypothesis, we expect to obtain a curve above the diagonal. In our case, the nominal  $p$ -value is on the x-axis and the proportion of rejected null hypotheses is on the y-axis.

The further is the curve above the diagonal, the higher is the power of the test. By analyzing the ROC curves, it is possible to notice that (i) **Figure 2A**—the ROC curves under the null hypothesis lie in the diagonal as expected, i.e., the statistical test is effectively controlling the rate of false positives (the proportion of rejected null hypothesis is as expected by the  $p$ -value threshold); (ii) **Figure 2B**—the power of the test increases as the number of graphs increases under the alternative hypothesis (the parameter of one of the populations is generated in a different manner); (iii) **Figure 2C**—the power of the test increases as the datasets are more balanced; and (iv) **Figure 2D**—the case the graph samples are generated by different models, the power of the test increases as the number of misclassified graphs decreases. When the number of misclassification is zero, it is possible to notice that the power of the method is higher than when only the parameter is different (**Figure 2C**).

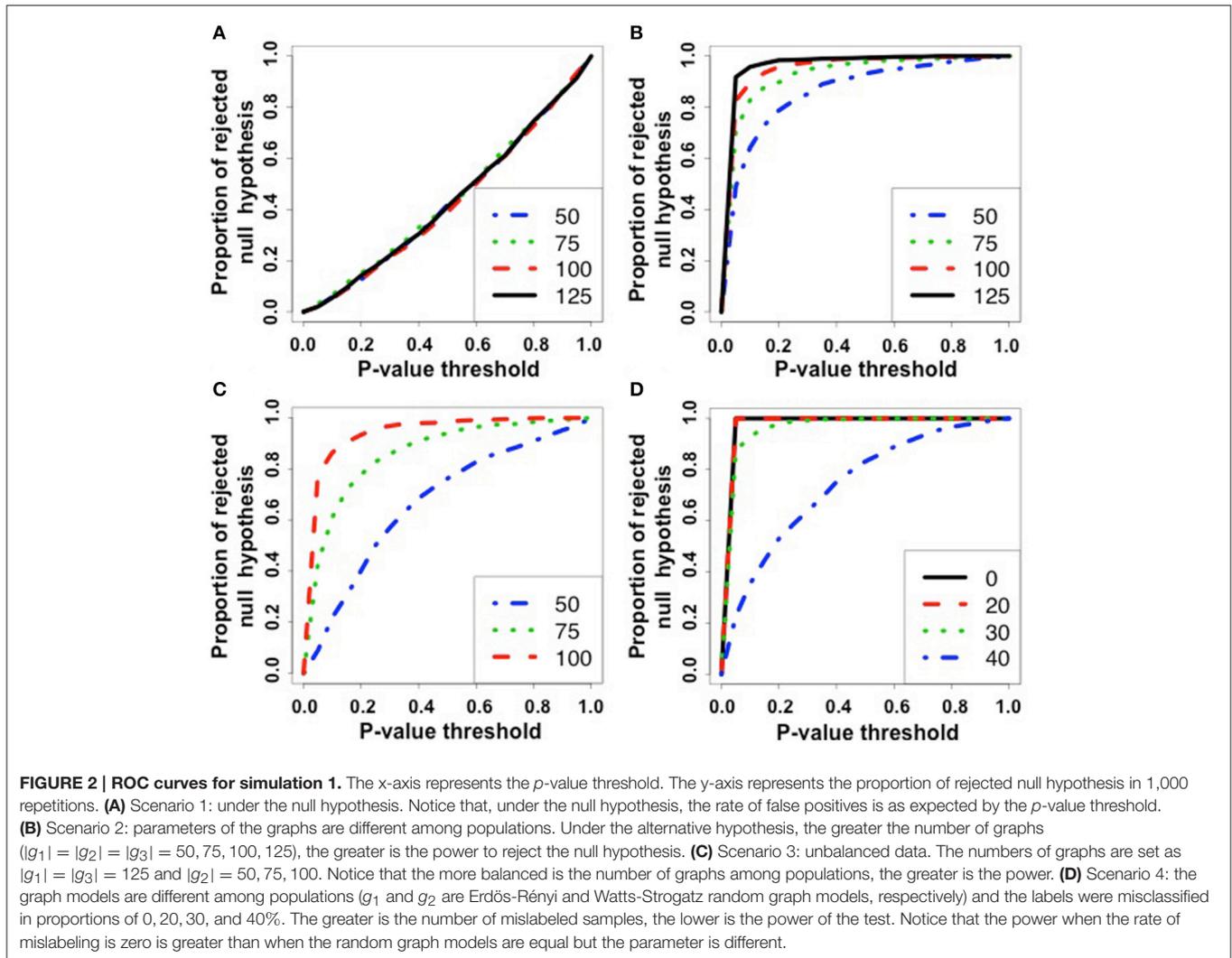
We compared ANOVA with the application of ANOVA on other features, such as the number of edges, betweenness centrality, closeness centrality, assortativity, and transitivity in five random graph models namely, Erdős-Rényi, geometric,  $k$ -regular, Watts-Strogatz, and Barabási-Albert. **Figure 3** describes the proportion of rejected null hypotheses in these graph models. By analyzing **Figure 3**, we notice that only ANOVA followed by transitivity are able to discriminate all random graph models with different parameters, including the Watts-Strogatz random graph model. Betweenness centrality presented low power for geometric,  $k$ -regular, Watts-Strogatz, and Barabási-Albert random graph models. Closeness centrality was not able to identify differences between Watts-Strogatz random graph models. Assortativity coefficient presented low power for Erdős-Rényi and Watts-Strogatz random graph models.

### 3.2. Autism Spectrum Disorder

Functional brain networks were constructed as described in Section 2.7.2. The five sub-networks depicted in **Figure 4** are

**TABLE 1 | Description of the ABIDE data set.**

Site	TR (ms)	TE (ms)	Voxel-size (mm)	Scanner
Caltech	2,000	30	3.50×3.50×3.50	SIEMENS MAGNETOM TrioTim syngo MR B17
CMU	2,000	30	3.00×3.00×3.00	SIEMENS MAGNETOM Verio syngo MR B17
KKI	2,500	30	3.05×3.15×3.00	PHILIPS Achieva 3T
Leuven	1,667	33	3.59×3.59×4.00	PHILIPS INTERA 3T
MaxMun	3,000	30	3.00×3.00×4.00	SIEMENS MAGNETOM Verio syngo MR B17
NYU	2,000	15	3.00×3.00×4.00	SIEMENS MAGNETOM Allegra syngo MR 2004A
Olin	1,500	27	3.40×3.40×4.00	SIEMENS MAGNETOM Allegra syngo MR 2004A
Pitt	1,500	25	3.10×3.10×4.00	SIEMENS MAGNETOM Allegra syngo MR A30
SBL	2,200	30	2.75×2.75×2.72	PHILIPS INTERA 3T
SDSU	2,000	30	3.43×3.43×3.40	GE 3T MR750
Stanford	2,000	30	3.12×3.12×4.50	GE SIGNA 3T
Trinity	2,000	28	3.00×3.00×3.50	PHILIPS INTERA 3T (conferir)
UCLA	3,000	28	3.00×3.00×4.00	SIEMENS MAGNETOM TrioTim syngo MR B15
UM	2,000	30	3.44×3.44×3.00	GE SIGNA 3T
USM	2,000	28	3.40×3.40×3.00	SIEMENS MAGNETOM TrioTim syngo MR B17
Yale	2,000	25	3.40×3.40×4.00	SIEMENS MAGNETOM TrioTim syngo MR B17



based on sub-networks defined by Sato et al. (2015), namely somatomotor, visual, default-mode, cerebellar, and fronto-parietal.

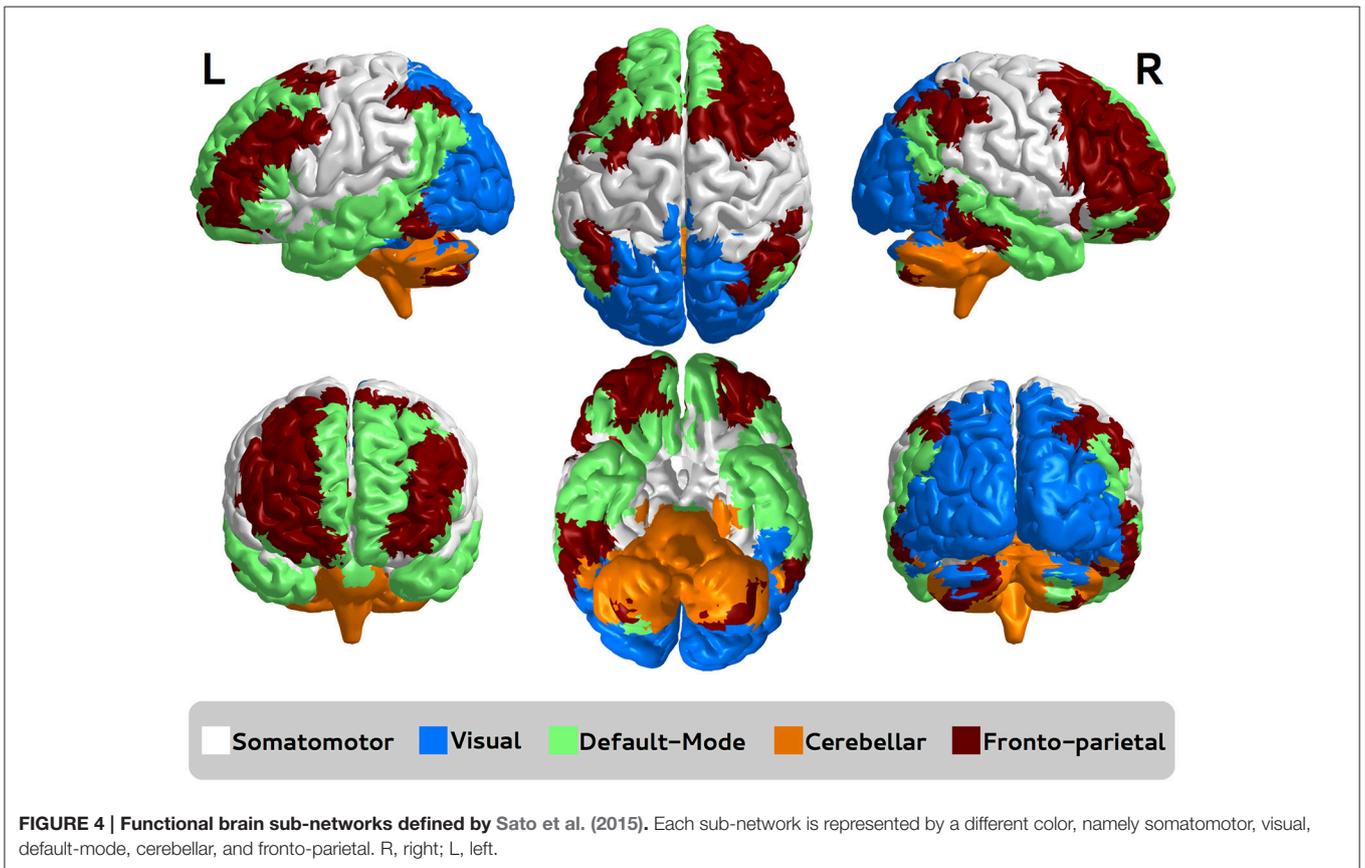
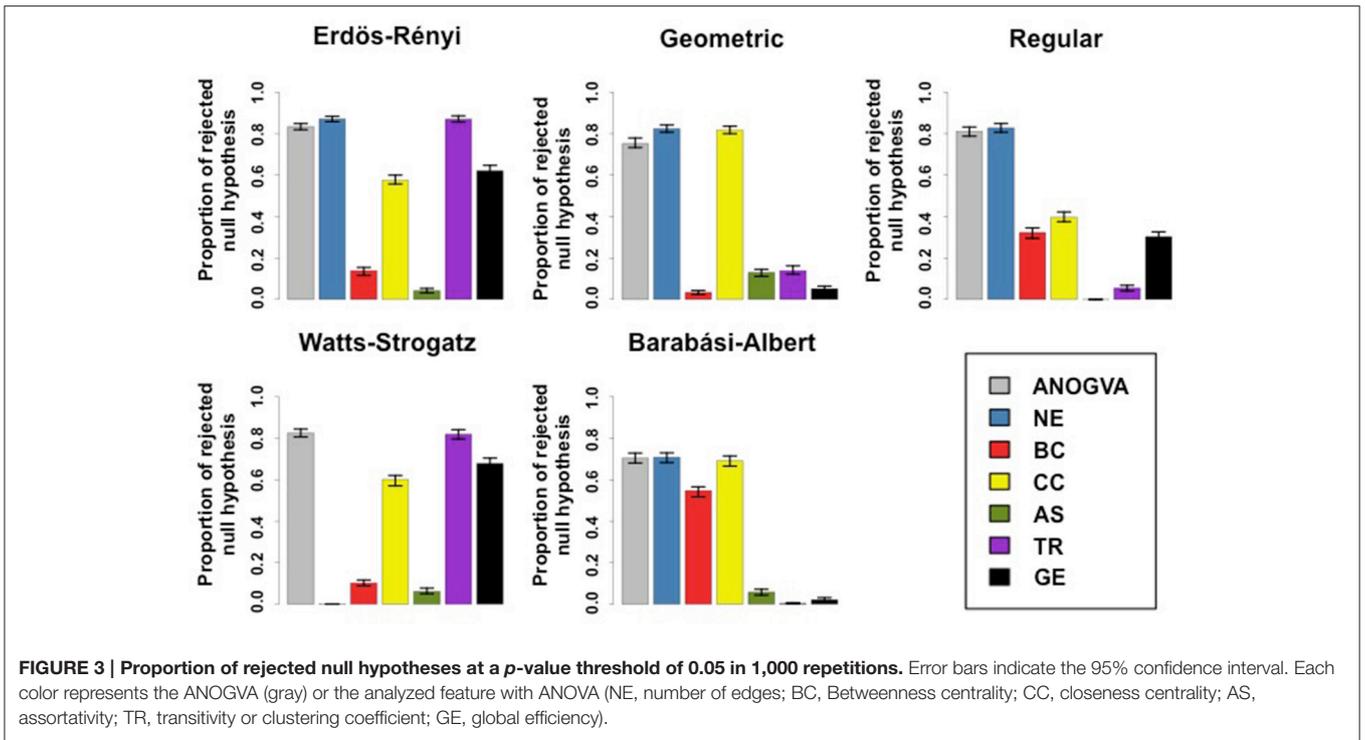
Here, we focused on the identification of which sub-network is associated with autism and Asperger by using ANOGVA. The number of permutations was set to 1,000. First, we compared the three groups (controls vs. autism vs. Asperger) for each sub-network to verify if there is at least one population that differs from the others. The test indicated a significant difference at a  $p$ -value threshold of 0.05 for cerebellar ( $p = 0.04$ ), and not for somatomotor ( $p = 0.13$ ), visual ( $p = 0.95$ ), default-mode ( $p = 0.17$ ), and fronto-parietal ( $p = 0.21$ ). These results suggest that the structure of the cerebellar functional sub-network is different at least in one of the populations among controls, autistic and Asperger patients. To identify which population is not equally generated in the cerebellar cluster, we carried out pairwise comparisons among the groups. Results indicate that there is no statistical evidence to discriminate controls vs. Asperger ( $p = 0.99$ ) and autism vs. Asperger ( $p = 0.13$ ), but there is significant difference between controls and autism ( $p < 0.001$ ).

This result indicates that the random process underlying the group control and autism are different.

## DISCUSSIONS

We introduced a method that is able to test two or more groups of graphs simultaneously. Our permutation-based test allows the estimation of  $p$ -values even for datasets with unknown probability distributions.

In a simulation study, we showed that the proposed method can indeed discriminate samples of graphs generated by different models and parameters. Similar to ANOGVA, the application of ANOVA on the number of edges can also discriminate a wide range of graph models. However, when the number of edges do not change, such as in the case of the Watts-Strogatz random graph model (the parameter is the rewiring probability), the ANOVA on the number of edges fails. Notice that the parameter of the Watts-Strogatz model only changes the structure of the graph and, in general, graph features are associated with the number of edges.



In the illustrative fMRI application example, we found that the cerebellar is associated with autism spectrum disorder. This result is consistent with other findings reported in the literature (Fatemi et al., 2012; Becker and Stoodley, 2013). Sato et al. (2015) showed that the network entropy of the cerebellar system is lower in autism than controls. Other fMRI studies reported reduced connectivity within cerebro-cerebellar motor networks during finger sequence tapping (Mostofsky et al., 2009) and also related to verb generation (Verly et al., 2014). Mosconi et al. (2015) showed that feedforward and feedback motor control abnormalities implicate cerebellar dysfunctions. We didn't find statistical differences between control and Asperger groups and between autism and Asperger groups. These results do not mean that there are no differences in graph structures between these groups as there is evidence in the literature that the brain activity and structure between these groups can be distinguished (McAlonan et al., 2002; Welchew et al., 2005). This result is most probably a result in lack of statistical power that can be improved by increasing the sample size, improving the classification criteria of the groups, and/or changing the experimental setup.

One limitation of our study is that the data collection protocols are heterogeneous among labs belonging to the ABIDE consortium. This issue was addressed by including the site as covariate in the GLM. Another solution would be to select the lab presenting the greater number of samples. However, since the number of samples decrease considerably when compared to analyzing the entire ABIDE dataset, the power of the test also decreases considerably. Another limitation is the fact that

the results are based on a resting-state fMRI protocol, and caution has to be taken to extend to other cognitive states. Regarding the method, ANOGVA can only be applied to undirected graphs. For directed graphs, more studies of the spectrum are necessary. Since the statistical test is based on a permutation procedure, one supposition is that the parameters of the model are sampled from probability distributions with finite variance.

ANOGVA is implemented in R (R Core Team, 2014) and is available in the package `statGraph` (<http://www.ime.usp.br/~fujita/software.html>).

## AUTHOR CONTRIBUTIONS

AF and DT conceived the experiments, analyzed the results, and drafted the manuscript. AF conducted the experiments. MV pre-processed the fMRI data. All authors gave the final approval for publication.

## FUNDING

AF was partially supported by São Paulo Research Foundation (FAPESP 2013/01715-3, 2013/03447-6, 2013/07375-0, 2015/01587-0, and 2016-13422-9), CNPq (306319/2010-1), and NAP eSciencePRPUSP. MV was partially supported by CAPES fellowship. DT was partially supported by Pew Latin American Fellowship and Ciência Sem Fronteiras Fellowship (CNPq 246778/2012-1).

## REFERENCES

- Albert, R., and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.* 74:47. doi: 10.1103/RevModPhys.74.47
- Barabasi, A. L., and Albert, R. (1999). Emergence of scaling in random networks. *Science* 286, 509–512.
- Barabási, A. L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272
- Bavelas, A. (1950). Communication patterns in task-oriented groups. *J. Acoust. Soc. Am.* 22, 725–730.
- Becker, E. B., and Stoodley, C. J. (2013). Autism spectrum disorder and the cerebellum. *Int. Rev. Neurobiol.* 113, 1–34. doi: 10.1016/B978-0-12-418700-9.00001-0
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B (Methodol.)* 57, 289–300.
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Craddock, R. C., James, G. A., Holtzheimer, P. E. III., Hu, X. P., and Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* 33, 1914–1928. doi: 10.1002/hbm.21333
- de Siqueira Santos, S., Takahashi, D. Y., Nakata, A., and Fujita, A. (2014). A comparative study of statistical methods used to identify dependencies between gene expression signals. *Brief. Bioinform.* 15, 906–918. doi: 10.1093/bib/bbt051
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659–667. doi: 10.1038/mp.2013.78
- Erdős, P., and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* 5, 17–61.
- Fatemi, S. H., Aldinger, K. A., Ashwood, P., Bauman, M. L., Blaha, C. D., Blatt, G. J., et al. (2012). Consensus paper: pathological role of the cerebellum in autism. *Cerebellum* 11, 777–807. doi: 10.1007/s12311-012-0355-9
- Fisher, R. (1918). Studies in crop variation. I. an examination of the yield of dressed grain from broadbalk. *J. Agric. Sci.* 11, 107–135. doi: 10.1017/S0021859600003750
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry* 40, 35–41. doi: 10.2307/3033543
- Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.* 30, 175–193. doi: 10.1007/BF02418571
- Kullback, S., and Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Stat.* 22, 79–86. doi: 10.1214/aoms/1177729694
- McAlonan, G. M., Daly, E., Kumari, V., Critchley, H. D., van Amelsvoort, T., Suckling, J., et al. (2002). Brain anatomy and sensorimotor gating in aspergers syndrome. *Brain* 125, 1594–1606. doi: 10.1093/brain/awf150
- Meringer, M. (1999). Fast generation of regular graphs and construction of cages. *J. Graph Theor.* 30, 137–146. doi: 10.1002/(SICI)1097-0118(199902)30:2<137::AID-JGT7>3.0.CO;2-G
- Mosconi, M. W., Mohanty, S., Greene, R. K., Cook, E. H., Vaillancourt, D. E., and Sweeney, J. A. (2015). Feedforward and feedback motor control abnormalities implicate cerebellar dysfunctions in autism spectrum disorder. *J. Neurosci.* 35, 2015–2025. doi: 10.1523/JNEUROSCI.2731-14.2015
- Mostofsky, S. H., Powell, S. K., Simmonds, D. J., Goldberg, M. C., Caffo, B., and Pekar, J. J. (2009). Decreased connectivity and cerebellar activity in autism during motor task performance. *Brain* 132, 2413–2425. doi: 10.1093/brain/awp088
- Nadaraya, E. A. (1964). On estimating regression. *Theor. Prob. Appl.* 9, 141–142. doi: 10.1137/1109020

- Newman, M. E. (2002). Assortative mixing in networks. *Phys. Rev. Lett.* 89:208701. doi: 10.1103/PhysRevLett.89.208701
- Penrose, M. (2003). *Random Geometric Graphs*, Vol. 5. Oxford: Oxford University Press.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity mri networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Sain, S. R., and Scott, D. W. (1996). On locally adaptive density estimation. *J. Am. Stat. Assoc.* 91, 1925–1934. doi: 10.1080/01621459.1996.10476720
- Sato, J. R., Takahashi, D. Y., Hoexter, M. Q., Massirer, K. B., and Fujita, A. (2013). Measuring network's entropy in adhd: a new approach to investigate neuropsychiatric disorders. *Neuroimage* 77, 44–51. doi: 10.1016/j.neuroimage.2013.03.035
- Sato, J. R., Vidal, M., de Siqueira Santos, S., Massirer, K. B., and Fujita, A. (2015). Complex network measures in autism spectrum disorders. *IEEE ACM Trans. Comput. Biol. Bioinform.* doi: 10.1109/TCBB.2015.2476787. [Epub ahead of print].
- Scott, J. (2012). *Social Network Analysis*. Los Angeles, CA: Sage.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
- Strang, G. (2011). *Introduction to Linear Algebra*. Wellesley: Wellesley-Cambridge Press.
- Sturges, H. A. (1926). The choice of a class interval. *J. Am. Stat. Assoc.* 21, 65–66. doi: 10.1080/01621459.1926.10502161
- Takahashi, D. Y., Sato, J. R., Ferreira, C. E., and Fujita, A. (2012). Discriminating different classes of biological networks by analyzing the graphs spectra distribution. *PLoS ONE* 7:e49949. doi: 10.1371/journal.pone.0049949
- Van Mieghem, P. (2010). *Graph Spectra for Complex Networks*. Cambridge: Cambridge University Press.
- Verly, M., Verhoeven, J., Zink, I., Mantini, D., Van Oudenhove, L., Lagae, L., et al. (2014). Structural and functional underconnectivity as a negative predictor for language in autism. *Hum. Brain Mapp.* 35, 3602–3615. doi: 10.1002/hbm.22424
- Wasserman, S., and Faust, K. (1994). *Social Network Analysis: Methods and Applications*, Vol. 8. Cambridge: Cambridge University Press.
- Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature* 393, 440–442.
- Welchew, D. E., Ashwin, C., Berkouk, K., Salvador, R., Suckling, J., Baron-Cohen, S., et al. (2005). Functional disconnection of the medial temporal lobe in aspergers syndrome. *Biol. Psychiatry* 57, 991–998. doi: 10.1016/j.biopsych.2005.01.028

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Fujita, Vidal and Takahashi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.