



Published in final edited form as:

*J Comp Neurol.* 2017 October 15; 525(15): 3251–3265. doi:10.1002/cne.24274.

## The significance of negative correlations in brain connectivity

Liang Zhan<sup>1</sup>, Lianne M. Jenkins<sup>2</sup>, Ori E. Wolfson<sup>3</sup>, Johnson Jonaris GadElkarim<sup>4</sup>, Kevin Nocito<sup>4</sup>, Paul M. Thompson<sup>5</sup>, Olusola A. Ajilore<sup>2</sup>, Moo K. Chung<sup>6</sup>, and Alex D. Leow<sup>2,3,4</sup>

<sup>1</sup>Computer Engineering Program, University of Wisconsin-Stout, Menomonie, Wisconsin

<sup>2</sup>Department of Psychiatry, University of Illinois, Chicago, Illinois

<sup>3</sup>Department of Computer Science, University of Illinois, Chicago, Illinois

<sup>4</sup>Department of Bioengineering, University of Illinois, Chicago, Illinois

<sup>5</sup>Imaging Genetics Center, and Institute for Neuroimaging and Informatics, Keck School of Medicine of USC, Marina del Rey, California

<sup>6</sup>Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, Wisconsin

### Abstract

Understanding the modularity of functional magnetic resonance imaging (fMRI)-derived brain networks or “connectomes” can inform the study of brain function organization. However, fMRI connectomes additionally involve negative edges, which may not be optimally accounted for by existing approaches to modularity that variably threshold, binarize, or arbitrarily weight these connections. Consequently, many existing Q maximization-based modularity algorithms yield variable modular structures. Here, we present an alternative complementary approach that exploits how frequent the blood-oxygen-level-dependent (BOLD) signal correlation between two nodes is negative. We validated this novel probability-based modularity approach on two independent publicly-available resting-state connectome data sets (the Human Connectome Project [HCP] and the 1,000 functional connectomes) and demonstrated that negative correlations alone are sufficient in understanding resting-state modularity. In fact, this approach (a) permits a dual formulation, leading to equivalent solutions regardless of whether one considers positive or negative edges; (b) is theoretically linked to the Ising model defined on the connectome, thus yielding modularity result that maximizes data likelihood. Additionally, we were able to detect novel and consistent sex differences in modularity in both data sets. As data sets like HCP become widely available for analysis by the neuroscience community at large, alternative and perhaps more advantageous computational tools to understand the neurobiological information of negative edges in fMRI connectomes are increasingly important.

---

**Correspondence** Alex Leow, Department of Psychiatry, University of Illinois, Chicago, IL 60607. aleow@psych.uic.edu, or, Liang Zhan, Computer Engineering Program, University of Wisconsin-Stout, Menomonie, WI 54751. zhan.liang@gmail.com.

#### SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

## Keywords

F1000; functional connectome; Human Connectome Project; modularity; negative correlations; resting state; RRID: SCR\_006942; RRID: SCR\_005361

---

## 1 | INTRODUCTION

Just as social networks can be divided into cliques that describe modes of association (e.g., family, school), the brain's connectome can be divided into modules or communities. Modules contain a series of nodes that are densely interconnected (via edges) with one another but weakly connected with nodes in other modules (Meunier, Lambiotte, & Bullmore, 2010). Thus, modularity or community structure best describes the intermediate scale of network organization, rather than the global or local scale. In many networks, modules can be divided into smaller sub-modules, thus can be said to demonstrate hierarchical modularity and *near decomposability* (the autonomy of modules from one another), a term first coined by Simon in 1962 (Meunier et al., 2010; Simon, 2002). Modules in fMRI-derived networks comprise anatomically and/or functionally related regions, and the presence of modularity in a network has several advantages, including greater adaptability and robustness of the function of the network. Understanding modularity of brain networks can inform the study of organization and mechanisms of brain function and dysfunction, thus potentially the treatment of neuropsychiatric diseases.

Mathematical techniques derived from graph theory (Fornito, Zalesky, & Breakspear, 2013) have been developed to measure and describe the modular organization of neural connectomes (Bullmore & Sporns, 2009; Sporns & Betzel, 2016). Different methods for module detection have been applied in network neuroscience, and offer different strengths and weaknesses (reviewed in Sporns & Betzel, 2016). Optimization algorithms are typically used to maximize the Q modularity metric or its variants (Danon, Diaz-Guilera, Duch, & Arenas, 2005). These algorithms vary in accuracy as there are tradeoffs made with computational speed (Rubinov & Sporns, 2010). Simulated annealing (e.g., Guimera & Amaral, 2005; Guimera, Sales-Pardo, & Amaral, 2004) is a slower, more accurate method for smaller networks, however, could be computationally expensive with larger networks (Danon et al., 2005). The Newman method (Newman, 2006; Newman & Girvan, 2004) reformulates modularity with consideration of the spectral properties of the network, and is also considered fairly accurate with adequate speed for smaller networks (Rubinov & Sporns, 2010). More recently, the Louvain method (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) has been developed for large networks (millions of nodes and billions of edges). Its rapid computation and ability to detect modular hierarchy (Rubinov & Sporns, 2010) has led to it becoming one of the most widely utilized methods for detecting communities in large networks. Comparisons with other modularity optimization methods have found that the Louvain method outperforms numerous other similar methods (Aynaud, Blondel, Guillaume, & Lambiotte, 2013; Lancichinetti & Fortunato, 2009).

However, these existing methods were mostly originally developed for networks with only positive connections and may additionally suffer from suboptimal reproducibility (Butts,

2003; Fortunato & Barthe-lemy, 2007; Guimera & Sales-Pardo, 2009). With the advent of connectomics, they also have been heuristically applied to fMRI brain networks, in which we have the additional complication of negative correlations. To this end, some methods largely ignore fMRI networks' negative edges (Fornito et al., 2013), only considering the right tail of the correlation histogram, that is, the positive edges (Schwarz & McGonigle, 2011). However, in functional neuroimaging, negative edges may be neurobiologically relevant (Sporns & Betzel, 2016), depending on factors such as data preprocessing steps, particularly the removal of potentially confounding signal such as head motion, global white-matter or whole-brain average signal, before calculation of the correlation matrix, because removal of such signal could result in detection of anticorrelations that were not present in the original data (Schwarz & McGonigle, 2011). Ignoring negative edges is achieved with binarization of a network (so-called "hard thresholding"), by selecting a threshold then replacing edge values below this threshold with zeros, and replacing supra-threshold values with ones (van den Heuvel et al., 2017). Some researchers retain the weights of the supra-threshold edge values, which has the effect of compressing the positive edges, however, the negative edges remain suppressed (Schwarz & McGonigle, 2011). Choice of threshold is important as more severe thresholds increase the contributions from the strongest edges, but can result in excessive disconnection of nodes within networks, in comparison to less stringent thresholds. Rather than binarizing networks, some researchers choose a "soft thresholding" approach that replaces thresholding with a continuous mapping of correlation values into edge weights, which had the effect of suppressing, rather than removing weaker connections (Schwarz & McGonigle, 2011). Linear and nonlinear adjacency functions can be employed, and the choice can be made to retain the valence of the edge weights, when appropriate.

An alternative to optimization methods discussed above, independent components analysis (ICA) has been applied to functional neuroimaging data (Beckmann, DeLuca, Devlin, & Smith, 2005). This method assumes that voxel time series are linear combinations of subsets of representative time series (Sporns & Betzel, 2016). Patterns of voxels load onto spatially independent components (modules). Unlike optimization methods, ICA allows for overlapping communities (Sporns & Betzel, 2016), although the number of ICA components needs to be pre-specified.

Utilizing a distance-based approach, recently, a new technique for investigating the hierarchical modularity of structural brain networks has been developed (GadElkarim et al., 2012, 2014). Rather than maximizing Q, the path length associated community estimation (PLACE) uses a unique metric that measures the difference in path length between versus within modules, to both maximize within-module integration and between-module separation (GadElkarim et al., 2014). It utilizes a hierarchically iterative procedure to compute global-to-local bifurcating trees (i.e., dendrograms), each of which represents a collection of nodes that form a module.

In this study, we developed a related novel method for functional brain networks—probability associated community estimation (PACE), that uses probability, not thresholds or the magnitude of BOLD signal correlations. We conducted experiments using this method, as well as six different implementations within the widely used brain connectivity toolbox

(BCT) (<http://www.brain-connectivity-toolbox.net/>) using data from the freely accessible 1,000 functional connectomes or F1000 project data set (F1000, RRID:SCR\_005361) (Biswal et al., 2010) and the Human Connectome Project (HCP, RRID:SCR\_006942) (Van Essen et al., 2012, 2013), and further examined differences in resting-state functional connectome's modularity (i.e., the resting-state networks [RSNs]) between males and females.

## 2 | METHODS

The popular  $Q$ -based modular structure (Blondel et al., 2008; Reichardt & Bornholdt, 2006; Ronhovde & Nussinov, 2009; Rubinov & Sporns, 2011; Sun, Danila, Josic, & Bassler, 2009) is extracted by finding the set of nonoverlapping modules that maximizes the modularity metric  $Q$ .

$$Q(G) = \frac{1}{2m} \sum_{i \neq j} \left( A_{ij} - \frac{k_i k_j}{2m} \right) \delta(i, j)$$

For a binary graph  $G$ ,  $m$  is the total number of edges,  $A_{ij} = 1$  if an edge links nodes  $i$  and  $j$  and 0 otherwise,  $\delta(i, j)$  if nodes  $i$  and  $j$  are in the same community and 0 otherwise, and  $k_i$  is the node degree of  $i$  (i.e., its number of edges). For weighted graphs,  $m$  is the sum of the weights of all edges while  $A_{ij}$  becomes the weight of the edge that links nodes  $i$  and  $j$ , and  $k_i$  is the sum of all weights for node  $i$ .

Approaches based on  $Q$ -maximization are naturally suitable for understanding the modularity of structural connectome where all edges are nonnegative. As an alternative to  $Q$  maximization, we previously developed a graph distance (shortest path length) based modularity approach for the structural connectome. By exploiting the structural connectome's *hierarchical modularity*, this *path length associated community estimation technique* (PLACE) is designed to extract global-to-local hierarchical modular structure in the form of bifurcating dendrograms (GadElkarim et al., 2012). PLACE has potential advantages over  $Q$  (GadElkarim et al., 2014), as it is hierarchically regular and *scalable* by design. Here, the degree to which nodes are separated is measured using graph distances (Dijkstra, 1959) and the PLACE benefit function is the  $\Psi^{\text{PL}}$  metric, defined at each bifurcation as the difference between the mean inter- and intra-modular graph distances. Thus, maximizing  $\Psi^{\text{PL}}$  is equivalent to searching for a partition with stronger intra-community integration and stronger between-community separation (GadElkarim et al., 2012, 2014; Lamar et al., 2016; Ye et al., 2015; Zhang et al., 2016).

### 2.1. | PACE for functional connectomes

Here, let us describe the PACE-based modularity of a functional connectome mathematically represented as an undirected graph  $\text{FC}(V, E)$ , where  $V$  is a set of vertices (i.e., nodes) and  $E$  is a set of edges (indexed by considering all pairs of vertices). Each edge of  $E$  is associated with a weight that can be either positive or negative.

Given a collection of functional connectomes  $\mathcal{S}$  on the same set of nodes  $V$  (but having edges with different weights), we can define the following *aggregation graph*  $G(V, E)$ . For each edge  $e_{ij}$  in  $E$  connecting node  $i$  and node  $j$ , we consider  $P^-_{ij}$ , the probability of observing a negative value at this edge in  $\mathcal{S}$  (i.e., nodes  $i$  and  $j$  exhibit an ‘anti-activating’ functional relationship). In the case of HCP, for example,  $\mathcal{S}$  thus consists of all healthy subjects’ resting-state functional connectome and this probability can simply be estimated using the ratio between the number of connectomes in  $\mathcal{S}$  having the edge  $e_{ij} < 0$  and the total number of connectomes in  $\mathcal{S}$ . Similarly, we define the probability of an edge in  $E$  being nonnegative as  $P^+_{ij}$  (i.e., nodes  $i$  and  $j$  exhibit a co-activating relationship). Naturally, the  $P^- - P^+$  pair satisfies the following relationship:

$$P^-_{i,j} + P^+_{i,j} = 1, \quad \forall (i, j), \quad i \neq j$$

Then, given  $C^1, C^2, \dots, C^N$  that are  $N$  subsets (or communities) of  $V$ , we define the mean intra-community edge positivity or negativity  $\overline{P^\pm(C^n)}$  for the  $n$ -th community  $C^n$  as:

$$\overline{P^\pm(C^n)} = \frac{\sum_{i, j \in C^n, i < j} P^\pm_{i,j}}{|C^n|(|C^n| - 1)/2}$$

Here,  $|C^n|$  represents the size (i.e., number of nodes) of the  $n$ -th community. Similarly, we could define the mean inter-community edge positivity and negativity (between communities  $C^n$  and  $C^m$ ) as:

$$\overline{P^\pm(C^n, C^m)} = \overline{P^\pm(C^m, C^n)} = \frac{\sum_{i \in C^n, j \in C^m} P^\pm_{i,j}}{|C^n||C^m|}$$

Here, the first equality holds as correlation-based functional connectomes are undirected. The intuition of PACE for fMRI connectomes is that edges that are most frequently anticorrelations should be placed across communities.

PACE operates as follows. Given a collection of functional connectomes  $\mathcal{S}$ , PACE identifies a natural number  $N$  and a partition of  $V$ ,  $C^1 \cup C^2 \cup \dots \cup C^N = V$ , ( $C^i \cap C^j = \emptyset$  for all  $i \neq j$ ) which maximizes the PACE benefit function  $\Psi$ . Intuitively,  $\Psi$  computes the difference between mean inter-community edge negativity and mean intracommunity edge negativity. Moreover, considering the duality between  $P^-$  and  $P^+$ , our optimization problem thus permits an equivalent dual form.

Formally,  $\Psi =$

$$C^1 \cup C^2 \cup \dots \cup C^N = V, C^i \cap C^j = \emptyset \text{ for all } i \neq j \left\{ \begin{array}{l} \text{argmax} \\ \frac{\sum_{1 \leq n < m \leq N} \overline{P^-(C^n, C^m)}}{N(N-1)/2} - \frac{\sum_{1 \leq n \leq N} \overline{P^-(C^n)}}{N} \end{array} \right\} =$$

$$C^1 \cup C^2 \cup \dots \cup C^N = V, C^i \cap C^j = \emptyset \text{ for all } i \neq j \left\{ \begin{array}{l} \text{argmax} \\ \frac{\sum_{1 \leq n \leq N} \overline{P^+(C^n)}}{N} - \frac{\sum_{1 \leq n < m \leq N} \overline{P^+(C^n, C^m)}}{N(N-1)/2} \end{array} \right\}$$

To solve the above PACE optimization problem, we adopt a PLACE-like algorithm, which has been extensively validated (Ajilore et al., 2013; GadElkarim et al., 2012, 2014; Lamar et al., 2016; Ye et al., 2015; Zhang et al., 2016), and computed global-to-local four-level bifurcating trees (yielding a total of 16 communities at the fourth level; please refer to GadElkarim et al. (2014) for implementation details).

## 2.2 || Theoretical link between PACE and the Ising model

Here, let us further explore the relationship between PACE and the Ising model using a mean-field approximation approach. When defined on the human connectome, the Ising model consists of assigning atomic spins  $\sigma$  to each brain region or node to one of two states (+1 or -1). Given a specific ensemble spin configuration  $\sigma$  over the entire brain and assuming the absence of external magnetic field, the corresponding Hamiltonian is thus defined as:

$$H(\sigma) = - \sum_{(i, j) \in E} J_{ij} \sigma_i \sigma_j$$

Here,  $(i, j) \in E$  indicates that there is an edge connecting nodes  $i$  and  $j$ . In classic thermodynamics, the Hamiltonian relates a configuration to its probability via the following Boltzmann distribution equation:

$$P(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z}$$

Where  $\beta$  is the inverse temperature and the normalizing constant  $Z$  is often called the partition function  $Z = \sum_{\sigma} e^{-\beta H(\sigma)}$ . Note,  $J_{ij}$  is positive when the interaction is ferromagnetic  $J_{ij} > 0$ , and antiferromagnetic when  $J_{ij} < 0$ .

Given this general set-up, we are now ready to show the general equivalence between PACE and maximizing the joint likelihood of the observed resting state fMRI (rs-fMRI) connectome data over some unknown ferromagnetic/antiferromagnetic ensemble interaction  $J$  defined on the connectome.

First, as one subject's rs-fMRI connectome is independent of other subjects', the joint likelihood over  $S$  subjects can be computed by forming the product:

$$\text{Likelihood (observed rs-fMRI data } |J) = \frac{e^{-\beta \sum_{s=1}^S H(\rho^s)}}{Z} = \frac{e^{-\beta S \sum_{s=1}^S \overline{H(\rho)}}}{Z}$$

For convenience of notations, let us work with the negative mean Hamiltonian

$$-\overline{H(\rho)} = -\frac{\sum_{s=1}^S H(\rho^s)}{S} = \sum_{(i,j) \in E} J_{ij} \frac{\sum_{s=1}^S \sigma_i^s \sigma_j^s}{S} = \sum_{(i,j) \in E} J_{ij} \overline{\sigma_i \sigma_j}$$

While the spin  $\sigma_i^s$  at a node  $i$  for any subject  $s$  is unknown, with PACE we nevertheless could proceed to estimate the expected value of spin product (across all  $S$  subjects)  $\overline{\sigma_i \sigma_j}$  by noting that PACE assigns  $\sigma_i \sigma_j = 1$  with probability  $P_{ij}^+$  and  $-1$  with probability  $P_{ij}^-$ . Thus

$$\begin{aligned} \overline{\sigma_i \sigma_j} &= 1 \cdot P_{i,j}^+ + (-1) \cdot P_{i,j}^- = 1 \cdot P_{i,j}^+ + (-1) \\ &\cdot (1 - P_{i,j}^+) = 2P_{i,j}^+ - 1 = 1 - 2P_{i,j}^- \end{aligned}$$

Next, let us compute and simplify  $\overline{H(\rho)}$  using the above equations, coupled with mean-field approximation, by separately considering ferromagnetic versus antiferromagnetic interactions (i.e., with respect to the sign of  $J_{ij}$ ):

$$J^+ = \frac{\sum_{(i,j) \in E, J_{ij} > 0} J_{ij}}{|\{(i,j) \in E, J_{ij} > 0\}|} > 0; \quad J^- = -\frac{\sum_{(i,j) \in E, J_{ij} < 0} J_{ij}}{|\{(i,j) \in E, J_{ij} < 0\}|} > 0$$

Thus,

$$\begin{aligned} -\overline{H(\rho)} &= \sum_{(i,j) \in E} J_{ij} \overline{\sigma_i \sigma_j} = \sum_{(i,j) \in E, J_{ij} > 0} J_{ij} \overline{\sigma_i \sigma_j} + \sum_{(i,j) \in E, J_{ij} < 0} J_{ij} \overline{\sigma_i \sigma_j} \\ &\approx J^+ \sum_{(i,j) \in E, J_{ij} > 0} \overline{\sigma_i \sigma_j} - J^- \sum_{(i,j) \in E, J_{ij} < 0} \overline{\sigma_i \sigma_j} \\ &= J^+ \sum_{(i,j) \in E, J_{ij} > 0} (2P_{i,j}^+ - 1) + J^- \sum_{(i,j) \in E, J_{ij} < 0} (2P_{i,j}^- - 1) \end{aligned}$$

Note, since here “mean-field” is constructed by averaging over ferromagnetic/antiferromagnetic interaction terms, our formulation may instead be called a mean-interaction approach.

Last, realizing that maximizing the joint likelihood of the observed data with respect to unknown ensemble interactions  $J$  (and thus unknown mean-interaction approximations  $J^+$  and  $J^-$ ) is equivalent to maximizing the negative mean Hamiltonian  $-\overline{H(\rho)}$ , we examine the right-hand side of the above equation (and note that both  $J^+$  and  $J^-$  are nonnegative) and deduct that a general maximization strategy can be devised by:

1. assigning as much as possible any two nodes  $i, j$  that are highly likely to exhibit co-activation  $e_{i,j} > 0$  (and thus the term  $[2P^+_{i,j} - 1]$  more likely to be positive) to have ferromagnetic interactions (thus  $i, j$  more likely to be placed in the same community), and at the same time
2. assigning as much as possible nodes  $i, j$  that are highly likely to exhibit anti-activation  $e_{i,j} < 0$  (and thus the term  $[2P^-_{i,j} - 1]$  more likely to be positive) to have antiferromagnetic interactions (thus  $i$  and  $j$  more likely to be placed in different communities).

This is exactly the intuition of PACE, that is, we maximize *the difference between mean inter-community edge negativity & mean intracommunity edge negativity* (or equivalently *maximizing the difference between mean intra-community edge positivity & mean intercommunity edge positivity*).

### 2.3 | Relaxation of the powers-of-two constraint: constructing the PACE null model and testing the statistical significance of each bifurcation

As PACE attempts, for each branch at a specific PACE level, to further split nodes within that branch into two subsequent groups, it is thus natural to ask if a procedure can be constructed in order to determine the level of statistical significance for such a split. By stopping a branch from further splitting when there is evidence against it, PACE can in theory yield any number of communities (no longer restricted to powers of 2).

Here, we propose such a procedure by first constructing the null distribution based on the observed data. Indeed, we can sample the null distribution (i.e., there is no modular patterns of co-/antiactivation) of the PACE benefit function  $\Psi$  by first randomly permuting the pair  $P^+_{i,j}/P^-_{i,j}$  for all  $(i, j)$  (i.e., randomly exchanging edge positivity with negativity, or simply put a probability value is replaced by 1 minus this value), followed by re-running PACE with shuffled edge positivity/neg-ative. Then, at each split the actual  $\Psi$  achieved by the original data is compared to the  $\Psi$  values of the reshuffled data at the same PACE level; if the former lies within the top 5% of the latter, such a split is determined to be significant ( $p < .05$ ).

In sum, using this data-informed permutation procedure, we relax the powers-of-two constraint during PACE optimization, thus letting the observed data to inform us the statistically most meaningful number of modules. This number can now be any positive integer which is no longer constrained to be a power of two.

## 3. | RESULTS

### 3.1 | Data description

We tested our PACE framework on two publicly available connectome data sets (Biswal et al., 2010; Brown, Rudie, Bandrowski, Van Horn, & Bookheimer, 2012). The first one is a 986-subject resting state fMRI connectome data set from the 1,000 functional connectome project (17 subjects' connectomes were discarded due to corrupted files), downloaded from the USC multimodal connectivity database (<http://umcd.humanconnectomeproject.org>). The dimension of the network is  $177 \times 177$ . The second data set is 820 subjects' resting state



fMRI connectome from the HCP (released in December 2015, named as HCP900 Parcellation + Timeseries + Netmats, [https://db.humancon-nectome.org/data/projects/HCP\\_900](https://db.humancon-nectome.org/data/projects/HCP_900)). The dimension of the network is  $200 \times 200$ , derived using ICA. For details of these two data sets, please refer to their respective websites and references.

### 3.2 | Simulation study

Here, we created a  $100 \times 100$  edge-negativity probability map, which contains five modules (each module has 20 nodes; Figure 1a). The edge negativity values within each module are uniformly randomly generated between 0 and 0.5 (less likely antiactivation within module) and the values across modules are uniformly randomly assigned from 0.5 to 1. Then, three-level PACE was applied to generate eight modules (Figure 1b), followed by sampling the null distribution of  $\Psi$  with 1,000 permutations using the procedure described in Section 2.3.

Results indicated that PACE correctly recovered the five-module ground truth, and the null distribution procedure indeed rejected any further splitting beyond five modules (Figure 1c; blue lines indicate statistically meaningful bifurcations). Figure 1d further shows the performance of PACE across different levels of noise (the exact procedure of how noise is applied is discussed in the Supporting Information).

### 3.3 | Stability analysis

To better understand the stability of the PACE with respect to the number of subjects used in estimating edge negativity/positivity, we tested PACE on subsets of HCP and F1000 randomly generated with a bootstrapping procedure (Chung, Lu, & Henry, 2006) (sampling with replacement) to investigate the reproducibility of the extracted community structure as a function of the sample size (from  $N = 50$  to 900 for F1000 and 50 to 800 for HCP; in increments of 50). For each  $N$ , 1,000 bootstrap samples were generated and the resulting 1,000 community structures were compared to the community structure derived from the entire HCP/F1000 sample using the normalized mutual information (NMI; Alexander-Bloch et al., 2012). NMI values are between 0 and 1, with 1 indicating two modular structures are identical. Figure 2 illustrates the mean NMI ( $y$  axis) as a function of sample size  $N$  ( $x$  axis), for all levels of PACE. Careful examinations of these NMI values suggest that stable PACE-derived modularity can be obtained with as few as  $\sim 100$  subjects (note here we include all subjects, regardless of age and sex, during bootstrapping, it is likely that the NMI values would be even higher if we restrict the analysis to a narrower age range and/or one specific sex).

### 3.4 | Optimal number of PACE bifurcations informed by the null model in F1000 and HCP

To determine the significance of PACE-derived hierarchical modularity at each bifurcation, for HCP and F1000, we generated 1000 samples of  $\Psi$  under the null hypothesis using the procedure in Section 2.3, and tested the significance of each split up to the fourth level. For both data sets, all bifurcations up to the third level were significant. At the fourth level, in F1000 none of the eight possible bifurcations was significant (thus resulting in a total eight of modules) and in HCP only one of the eight bifurcations was ( $p = .002$ , which remained significant after Bonferroni correction with a cut-off of  $0.05/8$ ), thus yielding a total of nine

modules. Figure 3 illustrates the whole procedure in HCP. (Please refer to Supporting Information Figure S2 for the final community structures for F1000 and HCP.)

### 3.5 | Modular structures revealed using PACE versus weighted- $Q$ maximization based methods

In this section, we compared PACE-derived modularity with  $Q$ -based modularity computed from the mean F1000 or HCP functional connectome (mean connectome is computed by element-wise averaging). As the optimal number of PACE-derived communities is eight in F1000 and nine in HCP (with a relatively small fifth community, C5, shown in Figure 3e) while  $Q$ -based methods primarily yield three to five communities, we selected a comparable PACE level, up to level 3, for our analyses.

Table 1 lists six  $Q$ -based methods adopted in this study (five weighted and one binarized). We conducted 100 runs for each of the six methods as well as PACE, and quantified pairwise similarity between two modular structures using NMI. We report summary statistics of these pairwise NMI values in Table 2 (the total number of NMI values are  $4950 = 100 \times 99/2$ ). As shown in this table,  $Q$ -based methods produced substantially variable modular structures across runs (and the number of communities across runs is also variable). By contrast, PACE produced identical results up to the third level (i.e., eight communities) for HCP and F1000.

To visualize these modularity results, we show axial slices of representative modular structures, for the HCP data set, generated using different methods (Figure 4, also see Figure 3e for rearranged connectome matrices based on PACE).

As  $Q$ -based methods yielded variable results (with variable number of communities, see Table 2), for a fair comparison, we randomly select a 4-community modular structure to visualize each of the five  $Q$ -based methods. Visually, except for the  $Q$ -amplitude and  $Q$ -negative-only,  $Q$ -based results shared strong similarities with results generated using second-level PACE (variability among  $Q$ -based methods notwithstanding). Table 3 summarizes, for each  $Q$ -based method, the mean and standard deviation of NMI between the 100 runs and second-level PACE-derived modularity.

Last, to better visualize the effect of variable numbers of modules in  $Q$ -based methods, we randomly selected one 3-community and one 4-community  $Q$ -derived HCP modular structure and compared them (Figure 5), with the visualizations supporting the potential issue of reproducibility with  $Q$  (for comparison, the first-level 2-community and second-level 4-community PACE HCP results are also shown).

### 3.6 | Variability in the modular structure computed using $Q$ -based thresholding-binarizing method

For the sixth  $Q$ -based modularity method, which applies an arbitrary nonnegative threshold to the mean connectome followed by binarization (all edges below threshold set to zero, and above threshold to one), we again conducted 100 runs for each threshold (starting, as a fraction of the maximum value in the mean functional connectome, from 0 to 0.5 with increments of 0.02) using the unweighted Louvain method routine implemented in the BCT toolbox. Figure 6 plots the mean pairwise NMI  $\pm$  SD as a function of the threshold, between

each of the 100 runs and those generated using the *Q*-Comb-Sym or *Q*-Comb-Asym methods. Results again demonstrated the substantial variability in *Q* as we vary the threshold, especially in the case of HCP.

### 3.7 | Comparison between PACE modularity and spectral graph cut

Last, for completeness, we also evaluated a network clustering algorithm derived from spectral graph theory (the normalized spectral cut or Ncut) (Ng, Jordan, & Weiss, 2001). Since the Ncut algorithm only deals with positive edges and one needs to pre-specify the value of “*k*” (the number of clusters to be generated), we artificially set all negative edges to zero in the network and ran Ncut 100 times for  $k = 2, 4, \text{ and } 8$ . Our results revealed that clustering derived from Ncut was also variable, as evidenced by the mean/standard deviation of pairwise NMI between any two of the 100 runs (Table 4). Note this should not come as a surprise, since the Ncut algorithm requires a random initialization step during the *k*-means step (i.e., even after *k* is determined, results are still dependent on how one initializes the center locations of the *k* clusters).

Further, we also computed the NMI between each of the 100 Ncut-derived modularity and that from PACE and reported the results in Table 5, which suggests substantial differences between the two.

### 3.8 | Sex differences in RSNs using a PACE-based hierarchical permutation procedure

Because the HCP data set has a better spatial resolution ( $2 \text{ mm}^3$ ) and thus better suited for detecting modularity differences at a granular level (Van Essen et al., 2012, 2013), we demonstrate here that the stability of PACE makes it possible to pinpoint modularity differences between males and females in the HCP data set, while minimizing potential confounding influences of age. As PACE uses a hierarchical permutation procedure to create trees, controlling for multiple comparisons is straightforward. Here, if two modular structures exhibit significant differences at each of the *m* most-local levels of modular hierarchy (each of them controlled at 0.05), collectively it would yield a combined false positive rate of 0.05 to the power of *m*. For the actual permutation procedure, we first computed the NMI between the two PACE-derived modular structures generated from the 367 males and the 453 females in the HCP data set. Then, under the null hypothesis (no sex effect), we randomly shuffled subjects between male and female groups and recomputed the NMI between the permuted groups across all three levels of PACE-derived modularity. This shuffling procedure was repeated 10,000 times and the resampled NMI values were recorded.

By ranking our observed NMI among the re-sampled 10,000 NMI values, we detected significant sex differences in modularity starting at the first-level (*p* values:  $< 1\text{e-}04$ ,  $1\text{e-}04$ , and  $1.4\text{e-}03$  for hierarchical level 1–3, respectively; a combined *p* value would thus be in the scale of  $10^{-11}$ ). By contrast, a similar strategy to detect sex effect using any of the *Q*-based methods failed to identify significant differences in the two sex-specific modular structures. Figure 7 visualizes the PACE-identified modular structure sex differences (highlighted using blue arrows and rectangles) in HCP.

Figure 7 shows sex differences primarily in the bilateral temporal lobes, which was not detected using  $Q$ -based methods. These differences extended to the hippocampus and amygdala, which in females, were part of the green module, and in males formed part of the red module.

Last, to validate these modularity findings, we further replicated our hierarchical permutation procedure using a subset of F1000 in the age range of 20–30 (319 females at  $23.25 \pm 2.26$  years of age and 233 males at:  $23.19 \pm 2.35$ ), with results yielding not only visually highly similar PACE modularity (despite that F1000 and HCP are based on completely different brain parcellation techniques), but also similar sex differences (Figure 8; statistically significant at PACE level 1;  $p = .0378$ ) in the limbic system (including the hippocampus) and the frontotemporal junction (including the pars opercularis as part of the inferior frontal gyrus), here primarily lateralized to the right hemisphere.

## 4 | DISCUSSION

In this study, we proposed PACE, a novel way of understanding how anticorrelations help define modularity of the resting-state fMRI connectome. The benefit function to be optimized exploits the intuition that a higher probability of an edge being anticorrelated indicates a higher probability of it connecting regions in different modules. Importantly, PACE permits a symmetric equivalent dual form, such that it can be equally conceived as placing edges that are most consistently positive within modules. Thus, PACE is intrinsically symmetrized.

Conventional  $Q$ -maximization methods take a variable approach at negative edges. For example, many studies to date simply ignore anticorrelations by setting any values below a threshold (usually positive) to be zero (Sporns & Betzel, 2016), while others have proposed to down-weight negative edges in a somewhat heuristic fashion. The PACE method offers a novel and theoretically advantageous interpretation of left-tail fMRI networks, as traditionally, the left-tail network, that is, those formed by negative edges alone, has been at times considered to be weak correlations that may “compromise” network attributes. For example, Schwarz and McGonigle (2011) argued that the left tail networks may not be biologically meaningful, despite noting that some connections were consistently observed in the negative-most tail networks, both with and without global signal removal. Schwarz and McGonigle (2011) thus recommended a “soft thresholding” approach be taken by replacing the hard thresholding or binarization operation with a continuous mapping of all correlation values to edge weights, suppressing rather than removing weaker connections and avoiding issues related to network fragmentation.

Rather than taking an approach that interprets the magnitude of correlations as the strength of connectivity, PACE determines the probability of a correlation being positive or negative. Interestingly, PACE can also be thought of as a different way of binarizing, with a “two-way” thresholding at zero. Although thresholding at a different value is possible, it would compromise the equivalence of the PACE dual forms. Indeed, one could theoretically generalize PACE by setting  $p^{\alpha \pm}_{i,j}$  to compute the probability of  $e_{i,j}$  being larger or smaller

than an arbitrary threshold  $\alpha$ . However, in the case of a positive  $\alpha$ , the left tail is no longer strictly anticorrelations.

To validate PACE, we used full rather than partial correlations. We chose this method because recent literature has suggested that in general, partial correlation matrices need to be very sparse (Peng, Wang, Zhou, & Zhu, 2009), and partial correlations have a tendency to reduce more connections than necessary. In dense networks such as fMRI-based brain networks, partial correlations have not been shown to be necessarily better than the Pearson correlation. Thus, partial correlations are often used in small networks that (a) have small numbers of connections or (b) have been forced to be sparse by introducing a penalty term during whole-brain network generation (Lee, Lee, Kang, Kim, & Chung, 2011).

Following current practice in the literature, we compared PACE to  $Q$  maximization based modular structures in both the HCP and F1000 data sets using the default setting in the BCT toolbox (the Louvain method). Notably, for  $Q$ -maximization, we observed more variable modular structure, not only across different  $Q$ -based formulations (right tail, left tail, absolute value, and symmetric and asymmetric combined), but also across multiple runs within each formulation.

A secondary analysis further applied PACE to the investigation of potential sex differences in the resting functional connectome. Note, while sex differences have been reported in the structural connectome of the human brain (e.g., Szalkai, Varga, & Grolmusz, 2015), few studies have examined sex differences in the functional connectome in healthy individuals, and no studies to our knowledge have examined sex differences in higher-level connectome properties such as network modularity. Previously, one large study (Biswal et al., 2010) examined the functional connectome of the F1000 data set using three methods: seed-based connectivity, independent component analysis (ICA) and frequency domain analyses. Across the three analytic methods, they found consistent effects of sex, with evidence of greater connectivity in males than females in the temporal lobes, more so in the right hemisphere, and particularly when using ICA. Our results (Figure 8) are consistent with these reported findings.

Using the HCP data set, our study also revealed higher-level sex-specific connectome modularity differences in the temporal lobes, including the middle temporal gyrus, amygdala, and hippocampus. The amygdala (Cahill, 2010) and hippocampus (Addis, Moscovitch, Crawley, & McAndrews, 2004) are important for emotional and autobiographic memory, while previous studies have reported sex differences in their activities in this context (Davis, 1999; Seidlitz & Diener, 1998; St. Jacques, Conway, & Cabeza, 2011; Young, Bellgowan, Bodurka, & Drevets, 2013). In line with these findings that likely reflect differential, sex-specific cognitive strategies for recalling memories related to self, we found that in female, the amygdala and hippocampus are within the module that also contains the default mode network, whereas in males they belong to the module largely consisting of the visual and somatomotor networks. It is also consistent with Damoiseaux, Viviano, Yuan, and Raz (2016), which found that females had greater connectivity between the hippocampus and medial PFC than males, and Kogler et al. (2016) which found that females had greater connectivity between the left amygdala and left middle temporal gyrus

than males. These medial prefrontal and lateral temporal regions form part of the brains default mode network (Fox et al., 2005). Thus, these recent and preliminary findings may reflect stronger coupling within the default mode network and between the amygdala and the default mode network in females than in males, supporting previous reports of greater regional homogeneity in the right hippocampus and amygdala in females than males (Lopez-Larson, Anderson, Ferguson, & Yurgelun-Todd, 2011).

## 5 | LIMITATIONS AND FUTURE DIRECTIONS

First, we note that the proposed PACE framework is based on the estimation of edge positivity/negativity frequency, which encodes details of functional co-activation/antiactivation. Unlike  $Q$ -based methods that encode such details using correlation magnitudes, PACE procedure discards edge weights, which may be part of the reason why it yields more stable results by discarding otherwise useful details and reducing accuracy. However, a counter argument can also be made in that a majority of noisy features tend to be close to zero with arbitrary signs; thus  $Q$ -based methods that employ thresholding can simply remove these noisy features and probably yield more stable results (although this is not supported by our thresholding-binarization experiments in Section 3.6).

Further, the idea that correlation magnitudes always encode meaningful details is also not universally accepted in the imaging community, as evidenced by studies that instead adopted a thresholding-binarization approach. For example, in (van den Heuvel et al., 2017), the authors extensively tested a proportional threshold (PT) approach that “includes the selection of the strongest PT% of connections in each individual network, setting all (in the binary case) surviving connections to 1 and other connections to 0” in order to “remove spurious connections and to obtain sparsely connected matrices, a prerequisite for the computation of many graph theoretical metrics.”

Second, with the PACE benefit function cast as a difference between inter-modular versus intra-modular mean edge negativity, the optimization problem is non-deterministic polynomial-time hard or NP-hard and thus the global solution is not computable in realistic terms. Thus, we instead used a top-down hierarchical bifurcating solver that was previously extensively tested in PLACE. Despite this limitation, we (a) outlined a theoretical connection between PACE and the Ising model, demonstrating that the PACE algorithm is a maximum likelihood estimation algorithm, (b) showed that PACE results were robust and insensitive to multiple runs while recovering known RSNs, and (c) showed that PACE-derived number of communities is not restricted to powers of 2, due to a permutation procedure that constructs the null distribution of the observed data allowing us to determine, at each branch, if further bifurcation is statistically meaningful.

Although the novel PACE-based symmetrized functional modularity is shown to be a powerful and mathematically elegant approach to understanding anticorrelations in fMRI connectomes, it cannot be computed without robust estimates of edge negativity/positivity frequencies, and thus there may be instances where  $Q$  yields more biologically meaningful results. Here, we tested PACE using large- $N$  cohorts of HCP and F1000 functional connectomes. Although theoretically feasible, individual-level PACE would require multiple

runs for each individual or alternatively a completely different mathematical formulation (e.g., noncorrelation based, see below) for estimating these frequencies.

Along this line, we note that recently several more sophisticated approaches for rigorous null modeling of correlation matrices and for multilayer multiscale  $Q$  maximization have been proposed (Bazzi et al., 2016; Betzel, Fukushima, He, Zuo, & Sporns, 2016; Betzel et al., 2015; MacMahon & Garlaschelli, 2015). For example, in MacMahon and Garlaschelli (2015), the authors used random matrix theory to identify nonrandom properties of empirical correlation matrices, leading to the decomposition of a correlation matrix into a “structured” and a “random” component. While beyond the scope of this study, we note that (a) such a decomposition requires strong assumptions that have been criticized and may not hold for the human brain, (b) PACE extracts modularity given *some estimation* of functional co-activation/antiactivation via edge positivity/negativity, which can be based on time series correlation (an approach we adopted here due to its conventional popularity), based on more advanced null modeling as in this cited study, or based on other information-theoretical approaches that we are currently exploring and are completely noncorrelation-based. Thus, individual subject-level PACE becomes possible.

Last, a new multi-scale modularity maximization approach has been recently investigated that seeks to generalize the  $Q$  modularity metric, and is thus likely to outperform the  $Q$  methods we studied here. However, in contrast to the simplicity of the PACE model (that does not require any parameter tuning) and in addition to the several caveats and nuances of the existing  $Q$  methods, this multi-scale approach introduces additional resolution parameter ( $\gamma$ ) that needs to be further tuned (a range from  $10^{-2.0}$  to  $10^0$  was studied in Betzel et al., 2015)

Notwithstanding the several limitations of correlation-based PACE noted above, we demonstrated that testing specific effects (e.g., sex) can be achieved with careful permutation testing while controlling for other variables (such as age), as in our secondary analyses showing significant sex effects in the temporal lobes. Last, a recent study has utilized the rich data set provided by the HCP to develop a new multimodal method for parcellating the human cerebral cortex into 180 areas per hemisphere (Glasser et al., 2016). This semi-automated method incorporates machine-learning classification to detect cortical areas. It would be interesting to apply PACE to this new parcellation once the classifier becomes publicly available.

## 6 | CONCLUSIONS

This methodological report outlines a novel PACE framework that complements the existing  $Q$ -based methods of defining modularity for brain networks in which negative edges naturally occur. When applied to the HCP and the F1000 data sets, we showed that PACE yielded stable reproducible results that are consistent with those derived from existing methods, providing evidence for convergent validity. Furthermore, given the high reliability of this new method, we have been able to demonstrate sex differences in resting state connectivity that are not detected with traditional methods.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

This work has been partially supported by NIH AG056782 to LZ and AL, NIH EB022856 to MK, NIH U54 EB020403 to PT, NSF IIS- 1213013 and IIP-1534138 to OW.

Funding information

This work has been partially supported by NIH AG056782 to LZ and AL, NIH EB022856 to MK, NIH U54 EB020403 to PT, NSF IIS-1213013 and IIP-1534138 to OW.

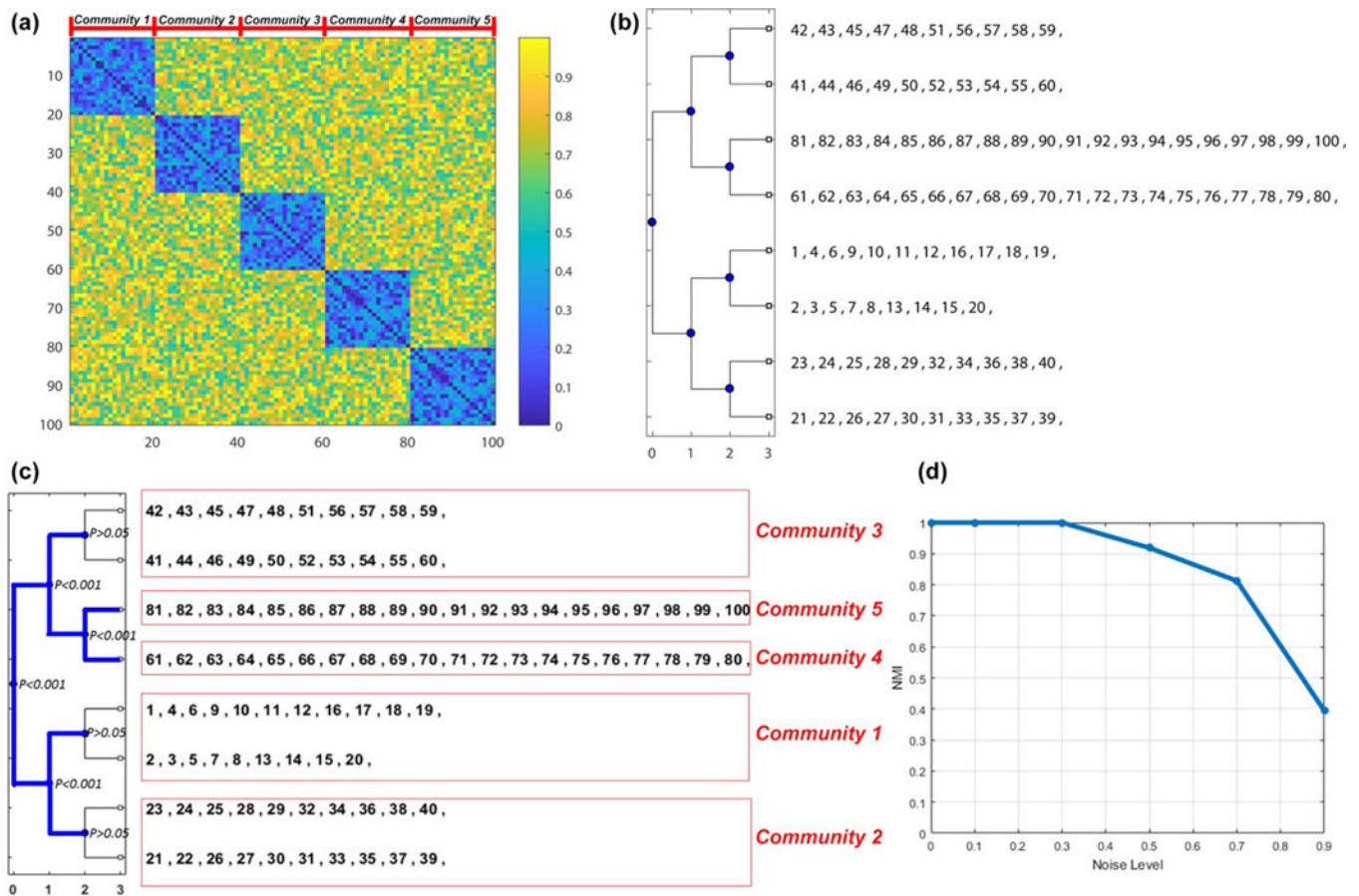
## REFERENCES

- Addis DR, Moscovitch M, Crawley AP, & McAndrews MP (2004). Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Hippocampus*, 14, 752–762. [PubMed: 15318333]
- Ajilore O, Zhan L, GadElkarim JJ, Zhang AF, Feusner J, Yang SL, ... Leow AD (2013). Constructing the resting state structural connectome. *Frontiers in Neuroinformatics*, 7, 30. doi:10.3389/fninf.2013.00030 [PubMed: 24409139]
- Alexander-Bloch A, Lambiotte R, Roberts B, Giedd J, Gogtay N, & Bullmore ET (2012). The discovery of population differences in network community structure: New methods and applications to brain functional networks in schizophrenia. *Neuroimage*, 59, 3889–3900. [PubMed: 22119652]
- Aynaud T, Blondel VD, Guillaume JL, & Lambiotte R (2013). Multilevel local optimization of modularity In Bichot C-E & Siarry P (Eds.), *Graph partitioning*. Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9781118601181.ch13
- Bazzi M, Porter M, Williams S, McDonald M, Fenn D, & Howison S (2016). Community detection in temporal multilayer networks, with an application to correlation networks. *Multiscale Modeling & Simulation*, 14, 1–41.
- Beckmann CF, DeLuca M, Devlin JT, & Smith SM (2005). Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 1001–1013.
- Betzel RF, Fukushima M, He Y, Zuo XN, & Sporns O (2016). Dynamic fluctuations coincide with periods of high and low modularity in resting-state functional brain networks. *Neuroimage*, 127, 287–297. [PubMed: 26687667]
- Betzel RF, Misisic B, He Y, Rumschlag J, Zuo X-N, & Sporns O (2015). Functional brain modules reconfigure at multiple scales across the human lifespan. arXiv:1510.08045.
- Biswal BB, Mennes M, Zuo XN, Gohel S, Kelly C, Smith SM, ... Milham MP (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 4734–4739. [PubMed: 20176931]
- Blondel VD, Guillaume JL, Lambiotte R, & Lefebvre E (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008, P10008. doi: 10.1088/1742-5468/2008/10/P10008
- Brown J, Rudie JD, Bandrowski A, Van Horn JD, & Bookheimer SY (2012). The UCLA multimodal connectivity database: A web-based platform for brain connectivity matrix sharing and analysis. *Frontiers in Neuroinformatics*, 6, 17. [PubMed: 22654753]
- Bullmore ET, & Sporns O (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10, 186–198. [PubMed: 19190637]
- Butts CT (2003). Network inference, error, and informant (in)accuracy: A Bayesian approach. *Social Networks*, 25, 103–140.
- Cahill L (2010). Sex influences on brain and emotional memory: The burden of proof has shifted In Savic I (Ed.), *Sex differences in the human brain, their underpinnings and implications* (Vol. 186, pp. 29–40). Amsterdam, The Netherlands: Elsevier Science Bv.

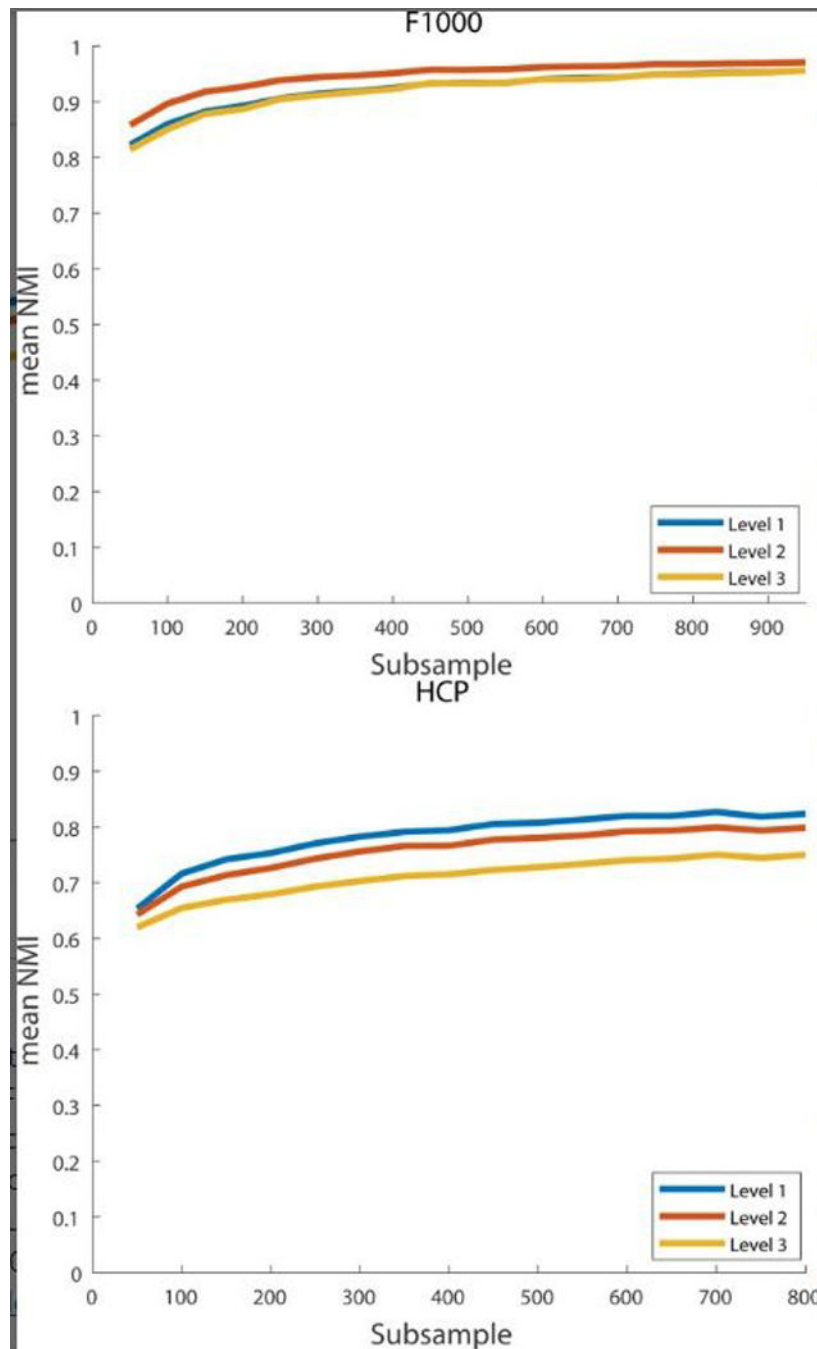


- Chung S, Lu Y, & Henry RG (2006). Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. *Neuroimage*, 33, 531–541. [PubMed: 16938472]
- Damoiseaux JS, Viviano RP, Yuan P, & Raz N (2016). Differential effect of age on posterior and anterior hippocampal functional connectivity. *Neuroimage*, 133, 468–476. [PubMed: 27034025]
- Danon L, Diaz-Guilera A, Duch J, & Arenas A (2005). Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, 2005, P09008. doi: 10.1088/1742-5468/2005/09/P09008
- Davis PJ (1999). Gender differences in autobiographical memory for childhood emotional experiences. *Journal of Personality and Social Psychology*, 76, 498–510. [PubMed: 10101879]
- Dijkstra EW (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1, 269–271.
- Fornito A, Zalesky A, & Breakspear M (2013). Graph analysis of the human connectome: Promise, progress, and pitfalls. *Neuroimage*, 80, 426–444. [PubMed: 23643999]
- Fortunato S, & Barthelemy M (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 36–41. [PubMed: 17190818]
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, & Raichle ME (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, 9673–9678. [PubMed: 15976020]
- GadElkarim JJ, Ajilore O, Schonfeld D, Zhan L, Thompson PM, Feusner JD, ... Leow AD (2014). Investigating brain community structure abnormalities in bipolar disorder using path length associated community estimation. *Human Brain Mapping*, 35, 2253–2264. [PubMed: 23798337]
- GadElkarim JJ, Schonfeld D, Ajilore O, Zhan L, Zhang AF, Feusner JD, ... Leow AD (2012). A framework for quantifying nodelevel community structure group differences in brain connectivity networks. *Medical Image Computing and Computer-Assisted Intervention*, 15(Pt 2), 196–203. [PubMed: 23286049]
- Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, ... Van Essen DC (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536, 171–178. [PubMed: 27437579]
- Guimera R, & Amaral LAN (2005). Functional cartography of complex metabolic networks. *Nature*, 433, 895–900. [PubMed: 15729348]
- Guimera R, & Sales-Pardo M (2009). Missing and spurious interactions and the reconstruction of complex networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 22073–22078. [PubMed: 20018705]
- Guimera R, Sales-Pardo M, & Amaral LAN (2004). Modularity from fluctuations in random graphs and complex networks. *Physical Review E*, 70, 4.
- Kogler L, Muller VI, Seidel EM, Boubela R, Kalcher K, Moser E, ... Derntl B (2016). Sex differences in the functional connectivity of the amygdalae in association with cortisol. *Neuroimage*, 134, 410–423. [PubMed: 27039701]
- Lamar M, Ajilore O, Leow A, Charlton R, Cohen J, GadElkarim J, ... Kumar A (2016). Cognitive and connectome properties detectable through individual differences in graphomotor organization. *Neuropsychologia*, 85, 301–309. [PubMed: 27037044]
- Lancichinetti A, & Fortunato S (2009). Community detection algorithms: A comparative analysis. *Physical Review E*, 80, 11.
- Lee H, Lee DS, Kang H, Kim BN, & Chung MK (2011). Sparse brain network recovery under compressed sensing. *IEEE Transactions on Medical Imaging*, 30, 1154–1165. [PubMed: 21478072]
- Lopez-Larson MP, Anderson JS, Ferguson MA, & Yurgelun-Todd D (2011). Local brain connectivity and associations with gender and age. *Developmental Cognitive Neuroscience*, 1, 187–197. [PubMed: 21516202]
- MacMahon M, & Garlaschelli D (2015). Community detection for correlation matrices. *Physical Review X*, 5, 021006.
- Meunier D, Lambiotte R, & Bullmore ET (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience*, 4, 200. [PubMed: 21151783]

- Newman MEJ (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 8577–8582. [PubMed: 16723398]
- Newman MEJ, & Girvan M (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69, 15.
- Ng AY, Jordan MI, & Weiss Y (2001). On spectral clustering: Analysis and an algorithm. *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, Vancouver, British Columbia, Canada 849–856.
- Peng J, Wang P, Zhou NF, & Zhu J (2009). Partial correlation estimation by joint sparse regression models. *Journal of the American Statistical Association*, 104, 735–746. [PubMed: 19881892]
- Reichardt J, & Bornholdt S (2006). Statistical mechanics of community detection. *Physical Review E*, 74, 14.
- Ronhovde P, & Nussinov Z (2009). Multiresolution community detection for megascale networks by information-based replica correlations. *Physical Review E*, 80, 18.
- Rubinov M, & Sporns O (2010). Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage*, 52, 1059–1069. [PubMed: 19819337]
- Rubinov M, & Sporns O (2011). Weight-conserving characterization of complex functional brain networks. *Neuroimage*, 56, 2068–2079. [PubMed: 21459148]
- Schwarz AJ, & McGonigle J (2011). Negative edges and soft thresholding in complex network analysis of resting state functional connectivity data. *Neuroimage*, 55, 1132–1146. [PubMed: 21194570]
- Seidlitz L, & Diener E (1998). Sex differences in the recall of affective experiences. *Journal of Personality and Social Psychology*, 74, 262–271. [PubMed: 9457787]
- Simon HA (2002). Near decomposability and the speed of evolution. *Industrial and Corporate Change*, 11, 587–599.
- Sporns O, & Betzel RF (2016). Modular brain networks. *Annual review of psychology* 67, 613–640.
- St. Jacques PL, Conway MA, & Cabeza R (2011). Gender differences in autobiographical memory for everyday events: Retrieval elicited by SenseCam images versus verbal cues. *Memory*, 19, 723–732. [PubMed: 20981611]
- Sun Y, Danila B, Josic K, & Bassler KE (2009). Improved community structure detection using a modified fine-tuning strategy. *EPL*, 86, 6.
- Szalkai B, Varga B, & Grolmusz V (2015). Graph theoretical analysis reveals: Women's brains are better connected than men's. *PLoS One*, 10, 30.
- van den Heuvel MP, de Lange SC, Zalesky A, Seguin C, Yeo BTT, & Schmidt R (2017). Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations. *Neuroimage*, 152, 437–449. [PubMed: 28167349]
- Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K, & WU-MH Consortium. (2013). The WU-Minn Human Connectome Project: An overview. *Neuroimage*, 80, 62–79. [PubMed: 23684880]
- Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, ... WU-Minn HCP Consortium. (2012). The Human Connectome Project: A data acquisition perspective. *Neuroimage*, 62, 2222–2231. [PubMed: 22366334]
- Ye AQ, Zhan L, Conrin S, GadElKarim J, Zhang AF, Yang SL, ... Leow A (2015). Measuring embeddedness: Hierarchical scale-dependent information exchange efficiency of the human brain connectome. *Human Brain Mapping*, 36, 3653–3665. [PubMed: 26096223]
- Young KD, Bellgowan PSF, Bodurka J, & Drevets WC (2013). Functional neuroimaging of sex differences in autobiographical memory recall. *Human Brain Mapping*, 34, 3320–3332. [PubMed: 22807028]
- Zhang A, Leow A, Zhan L, GadElKarim J, Moody T, Khalsa S, ... Feusner J (2016). Brain connectome modularity in weight-restored anorexia nervosa and body dysmorphic disorder. *Psychological Medicine*, 46, 2785–2797. [PubMed: 27429183]

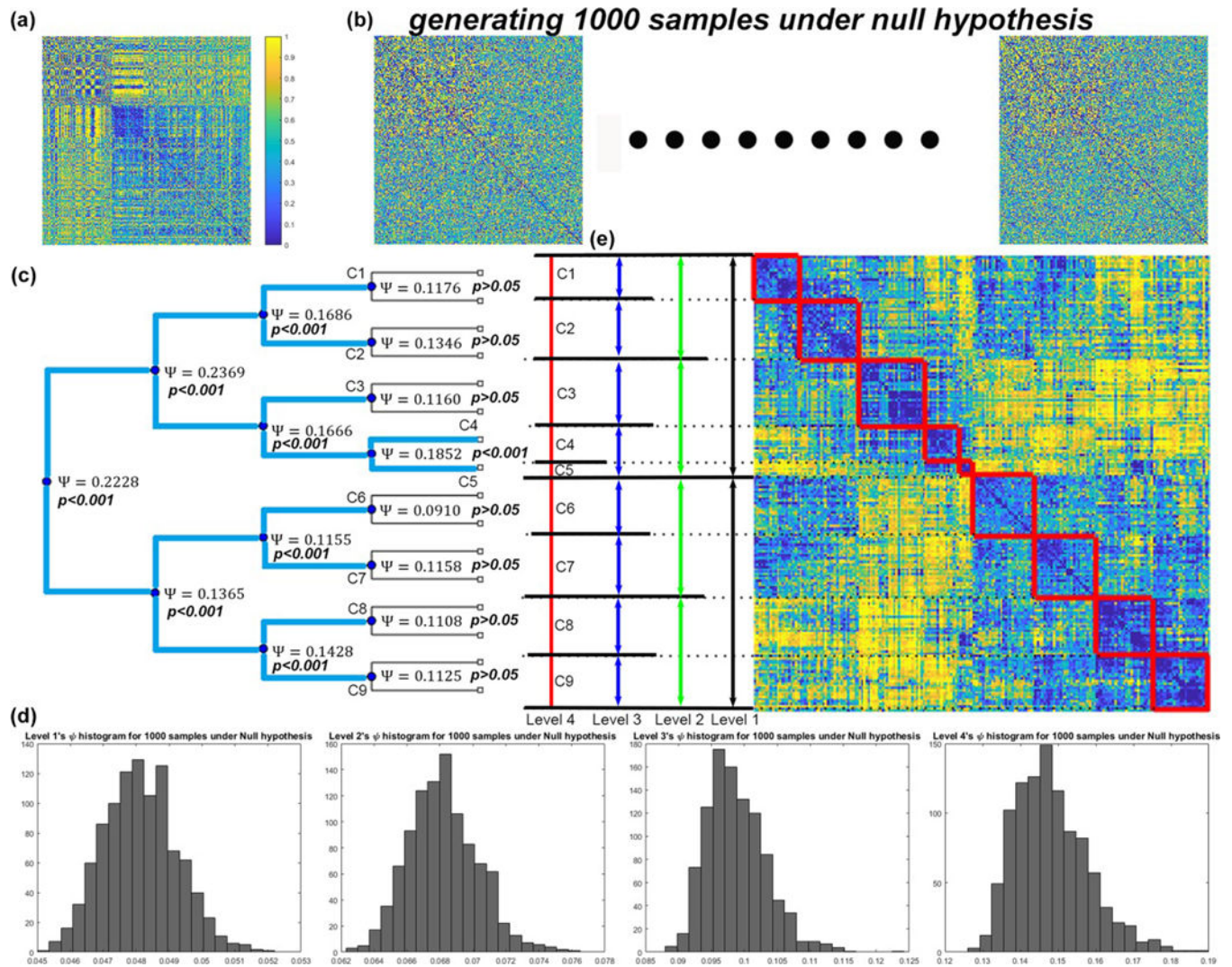
**FIGURE 1.**

Simulation study for PACE. (a) A  $100 \times 100$  5-community edge-negativity probability map was generated, where each module has 20 nodes. Within-community, the edge negativity value is uniformly randomly selected from 0 to 0.5 and between-community, the value from 0.5 to 1; (b) three-level PACE was computed and eight communities were generated; (c) using 1,000 permutations randomly exchanging edge positivity with negativity, we constructed the null distribution of  $\Psi$ , using which we tested the significance of each bifurcation. For the four possible bifurcations at the third-level PACE, only one was significant (significant bifurcations highlighted in blue), thus yielding a total of five modules, each of which matches the corresponding module in the ground truth (highlighted by the red square; from top to bottom, module correspondence is 3, 5, 4, 1, 2); (d) the performance of PACE in this toy example under different noise levels ( $L = 0.1, 0.3, 0.5, 0.7$ , and  $0.9$ ; the exact procedure of how  $L$  is applied is discussed in the Supporting Information) [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



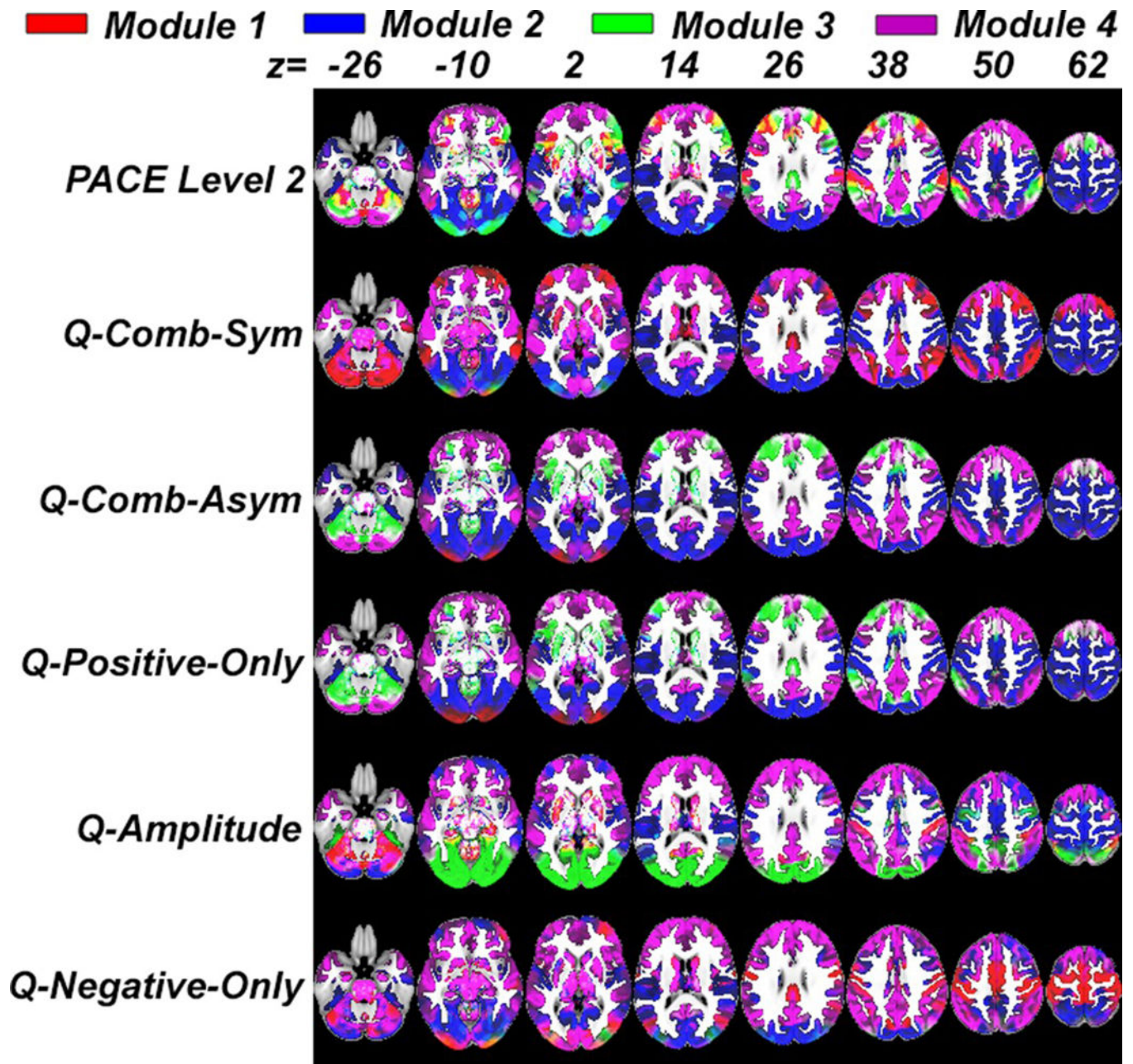
**FIGURE 2.**

Stability of PACE as a function of the number of subjects used in estimating edge negativity ( $x$  axis;  $N = 50$ – $900$  for F1000 and  $N = 50$ – $800$  for HCP) using a bootstrapping procedure. For each  $N$ , we generated 1,000 copies (random sampling with replacement) and computed the NMI between each of the 1,000 corresponding PACE modular structures and that derived from the full sample.  $y$  axis plots the mean of these 1,000 NMI values, for each  $N$  and each of PACE levels [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



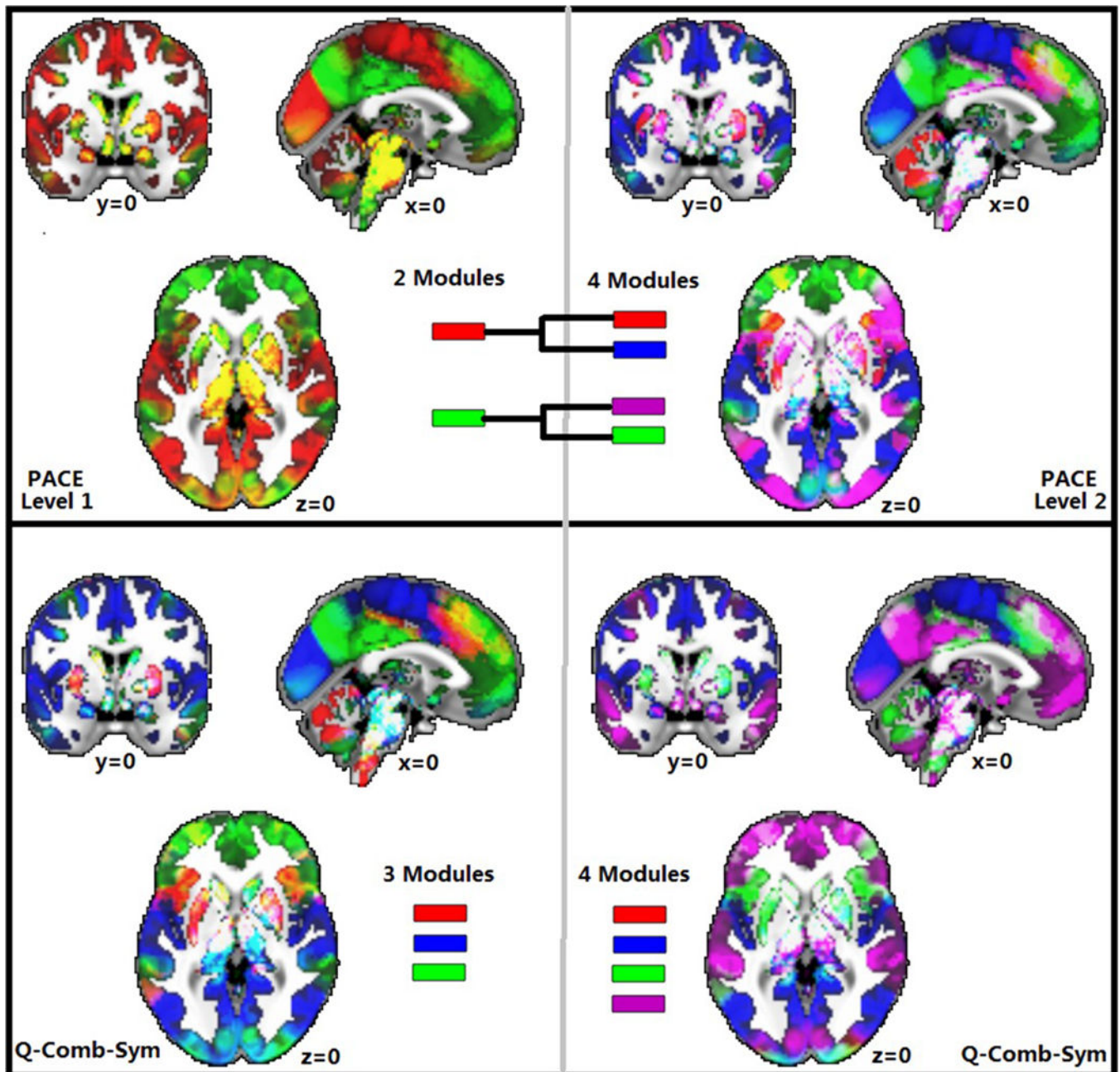
**FIGURE 3.**

Constructing the PACE null model in HCP. (a) Original HCP edge-negativity frequency map; (b) generating 1,000 samples of the same map under the null hypothesis, by randomly permuting the edge positivity/negativity pair (i.e., for each element of the matrix, its edge-negativity value  $p$  is randomly reassigned to  $1 - p$  with a probability of 50%); (c) Testing the significance of each PACE bifurcation in HCP up to the fourth-level tree structure. The  $\psi$  values achieved by the original data are shown at each bifurcation point, along with their statistical significance. (d) Histograms of the 1,000 PACE benefit function  $\psi$  values generated under the null model for each level. Only bifurcations with observed  $\psi$  values ranked in the top 50 (with respect to the 1,000 null-model values from the same PACE level) is considered “significant” (blue lines in c). At the fourth level, only one out of eight possible bifurcations was meaningful, resulting in nine final communities (C1-C9). (e) Rearranged matrix to show how these communities (C1-C9) are formed from level 1 to level 4 [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



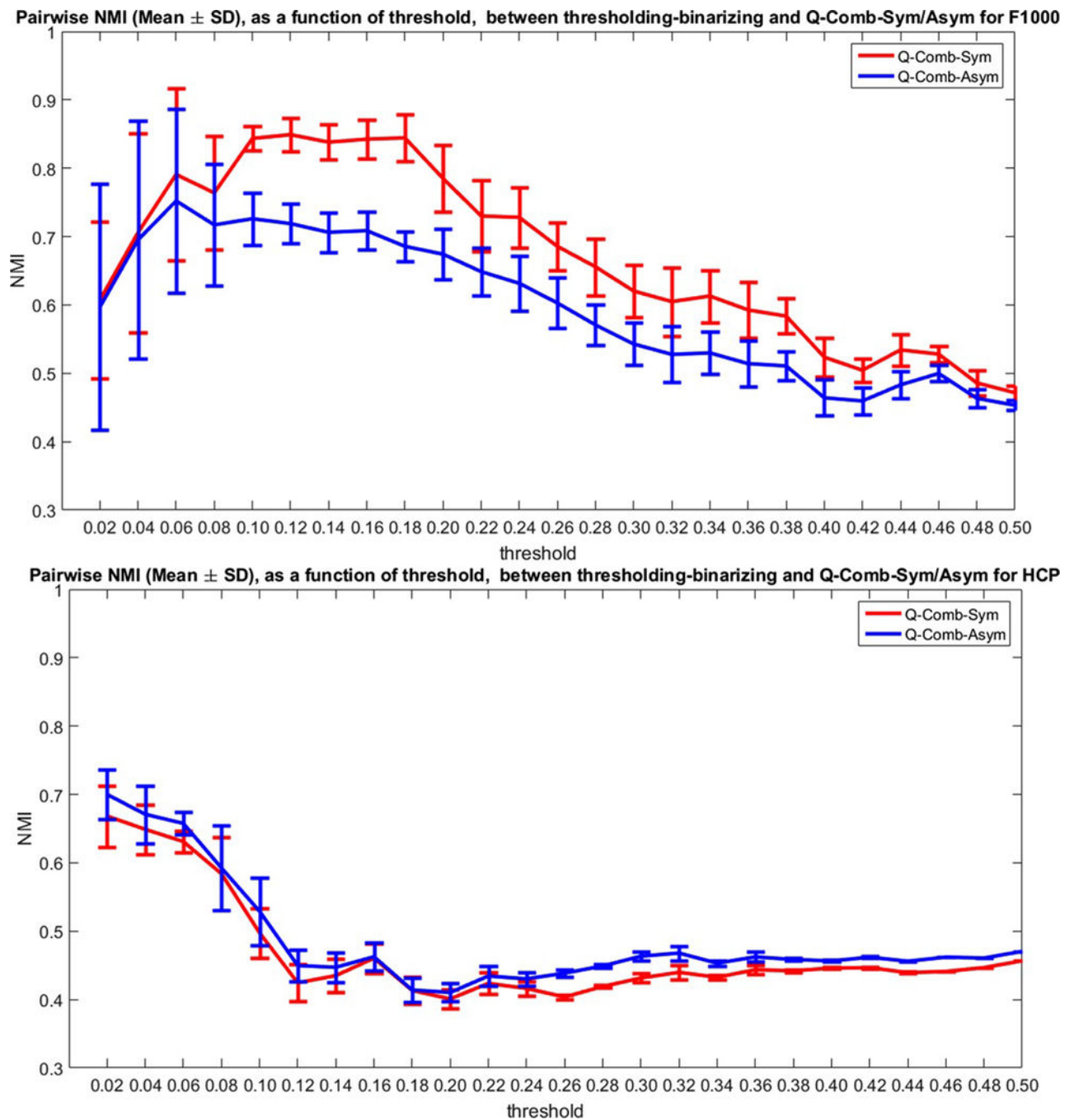
**FIGURE 4.**

Representative modular structures generated using different methods for the HCP data set. Regions coded in the same color (out of four: green, blue, red, and violet) form a distinct community or module. Note that unlike F1000, which uses structure parcellation to partition networks into nonoverlapping communities, HCP utilizes an ICA-based parcellation, which allows components (modules) to overlap (Sporns & Betzel, 2016), resulting in regions with mixed colors (e.g., yellow) [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



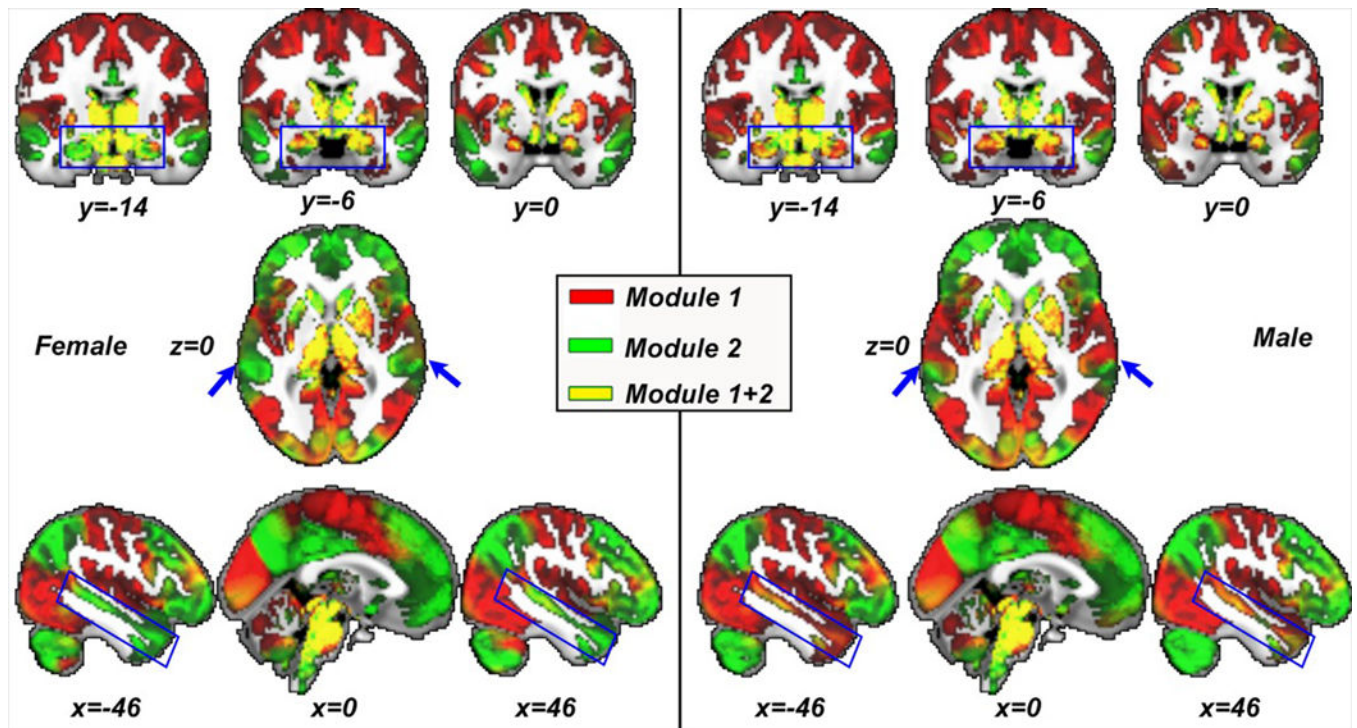
**FIGURE 5.**

Visualization of randomly selected 3-community and 4-community  $Q$  maximization-derived modular structures in HCP, demonstrating the suboptimal reproducibility with  $Q$  maximization. For comparison, the first-level 2-community and second-level 4-community PACE results are also shown [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**FIGURE 6.**

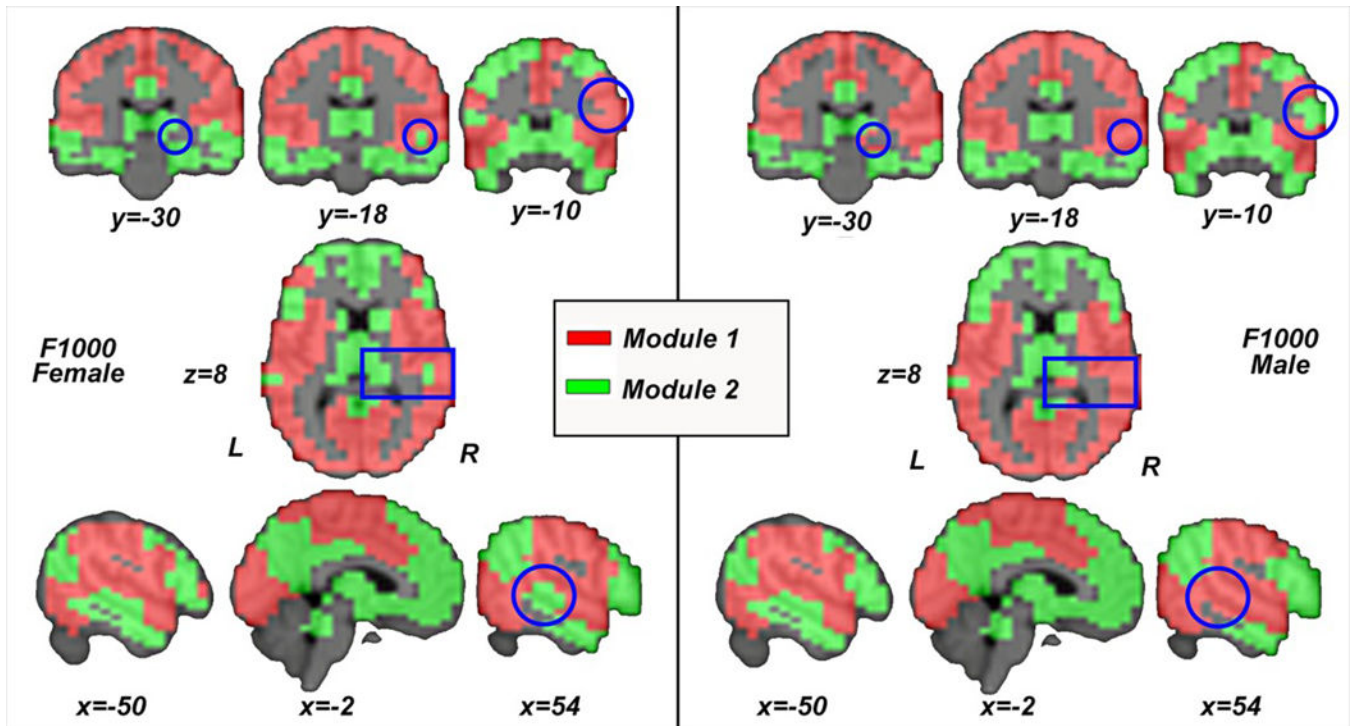
Mean and standard deviation of pair-wise similarity metric NMI, as a function of the threshold ( $x$  axis, as a fraction of the maximal value in the mean group connectome), between the modularity extracted using  $Q$ -based thresholding-binarizing and the weighted  $Q$ -Comb-Sym method or the  $Q$ -Comb-Asym method for F1000 (top) and HCP (bottom) [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]





**FIGURE 7.**

Visualization of PACE-identified sex-specific resting-state network (RSN) modularity in females (left) and males (right) from the HCP data set. Using permutation testing, sex-specific modularity differences are confirmed to be statistically significant throughout the entire PACE modular hierarchy starting at the first level. Here, the results are visualized at first-level PACE, yielding two modules coded in red (module 1) and green (module 2). As HCP utilizes an ICA-based parcellation, modules thus overlap, in this case resulting in some regions colored in yellow [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 8.**

Visualization of PACE-identified RSN modularity and its sex differences between females (left) and males (right) from the F1000 data set. The difference in spatial resolutions notwithstanding, note the visually highly similar PACE-derived modularity in HCP (Figure 7) and F1000, even though these two data sets utilize completely different brain parcellation techniques [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 1**

Summarizes the six  $Q$ -based methods, as implemented in the BCT toolbox, tested and compared in this study (Betzel et al., 2016; Rubinov & Sporns, 2011; Schwarz & McGonigle, 2011)

Method		Equation
Weighted version	Q-Comb-Sym	$Q = C^+ Q^+ - C^- Q^- \quad C^+ = C^-$ $Q^+ = f(W_{ij}^+) W_{ij}^+ = \begin{cases} W_{ij} & \text{if } W_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$ $Q^- = f(W_{ij}^-) W_{ij}^- = \begin{cases} -W_{ij} & \text{if } W_{ij} < 0 \\ 0 & \text{otherwise} \end{cases}$
	Q-Comb-Asym	$Q = C^+ Q^+ - C^- Q^- \quad C^+ \neq C^-$
	Q-Positive-only	$Q = f(W_{ij}^+) W_{ij}^+ = \begin{cases} W_{ij} & \text{if } W_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$
	Q-Amplitude	$Q = f( W_{ij} )  W_{ij}  = \begin{cases} W_{ij} & \text{if } W_{ij} > 0 \\ -W_{ij} & \text{otherwise} \end{cases}$
	Q-Negative-only	$Q = f(W_{ij}^-) W_{ij}^- = \begin{cases} -W_{ij} & \text{if } W_{ij} < 0 \\ 0 & \text{otherwise} \end{cases}$
Binarizing	Thresholding	$Q = f(B_{ij}) B_{ij} = \begin{cases} 1 & \text{if } W_{ij} > \text{thres} \\ 0 & \text{otherwise} \end{cases}$

TABLE 2

Mean and standard deviation of pair-wise normalized mutual information (NMI) across 100 repeated runs within each method

Method	F1000		HCP	
	NMI	Number of modules (number of runs)	NMI	Number of modules (number of runs)
PACE level 1	1.0 ± 0.0	2 (100)	1.0 ± 0.0	2 (100)
PACE level 2	1.0 ± 0.0	4 (100)	1.0 ± 0.0	4 (100)
PACE level 3	1.0 ± 0.0	8 (100)	1.0 ± 0.0	8 (100)
Q-Comb-Sym	<b>0.896 ± 0.093</b>	3 (97),4 (3)	0.731 ± 0.160	3 (38), 4 (62)
Q-Comb-Asym	0.835 ± 0.091	3 (63),4 (37)	0.772 ± 0.134	3 (31),4 (69)
Q-Positive-only	0.844 ± 0.103	3 (1),4 (99)	<b>0.834 ± 0.079</b>	3 (41),4 (59)
Q-Amplitude	0.819 ± 0.108	4 (18),5 (74),6 (8)	0.614 ± 0.135	3 (1),4 (49),5 (36),6 (14)
Q-Negative-only	0.617 ± 0.158	3 (66),4 (34)	0.460 ± 0.129	3 (3),4 (61),5 (36)

The first three rows are from PACE and the rest from Q. For Q-based methods, the most reproducible methods are highlighted in bold (for F1000 it was the Q-Comb-Sym, and for HCP the Q-Positive-only).

**TABLE 3**

For each  $Q$ -based method, this table summarizes the pair-wise NMI's mean and standard deviation between any of the repeated 100 runs and the second-level PACE-derived modularity

<b>Method</b>	<b>F1000</b>	<b>HCP</b>
Q-Comb-Sym	$0.725 \pm 0.026$	$0.576 \pm 0.051$
Q-Comb-Asym	$0.705 \pm 0.043$	$0.603 \pm 0.047$
Q-Positive-only	$0.740 \pm 0.061$	$0.539 \pm 0.035$
Q-Amplitude	$0.606 \pm 0.064$	$0.235 \pm 0.027$
Q-Negative-only	$0.170 \pm 0.013$	$0.113 \pm 0.017$

The stability/variability of the Ncut algorithm, which was run 100 times for a  $k$  value of 2, 4, and 8 (each run is different due to the random initialization during the  $k$ -means step)

**TABLE 4**

	HCP					
	F1000		4 modules		8 modules	
Ncut	0.9308 ± 0.0612	0.9374 ± 0.0814	0.6963 ± 0.0703	0.9469 ± 0.0520	0.6830 ± 0.1614	0.7184 ± 0.0735

We reported the pairwise NMI's mean and standard deviation between any two of the 100 runs for each  $k$  value.

**TABLE 5**

Comparing modularity derived from the Ncut algorithm and from PACE

	F1000			HCP		
	2 modules	4 modules	8 modules	2 modules	4 modules	8 modules
NMI between Ncut and PACE	$0.4369 \pm 0.0065$	$0.6635 \pm 0.0322$	$0.6388 \pm 0.0580$	$0.3745 \pm 0.0133$	$0.5284 \pm 0.0395$	$0.4529 \pm 0.0199$

Ncut was run 100 times with random initialization for a  $k$  value of 2, 4, and 8. We reported the pairwise NMI (mean and standard deviation) between any of the 100 runs and the corresponding PACE output (level 1–3, corresponding to 2, 4, and 8 modules).