Detecting Network Anomalies Using Forman-Ricci Curvature and A Case Study for Human Brain Networks

Tanima Chatterjee[1], Réka Albert[2], Stuti Thapliyal[1], Nazanin Azarhooshang[1], and Bhaskar DasGupta[1,*]

[1]University of Illinois at Chicago, Department of Computer Science, Chicago, 60607, USA
[2]Pennsylvania State University, Department of Physics, University Park, 16802, USA
[*]bdasgup@uic.edu

ABSTRACT

We analyze networks of functional correlations between brain regions to identify changes in their structure caused by Attention Deficit Hyperactivity Disorder (ADHD). We express the task for finding changes as a network anomaly detection problem on temporal networks. We propose the use of a curvature measure based on the Forman-Ricci curvature, which expresses higher-order correlations among two connected nodes. Our theoretical result on comparing this Forman-Ricci curvature with another well-known notion of network curvature, namely the Ollivier-Ricci curvature, lends further justification to the assertions that these two notions of network curvatures are not well correlated and therefore one of these curvature measures cannot be used as an universal substitute for the other measure. Our experimental results indicate nine critical edges whose curvature differs dramatically in brains of ADHD patients compared to healthy brains. The importance of these edges is supported by existing neuroscience evidence. We demonstrate that comparative analysis of curvature identifies changes that more traditional approaches, for example analysis of edge weights, would not be able to identify.

Introduction and motivation

It is by now a common research practice to study the properties of complex interconnected systems by representing them as heterogeneous networks and then using various network-theoretic tools for their analysis[1,2]. Such heterogeneous networks may vary in diversity from simple undirected networks to edge-labeled directed networks. One such class of network models are *temporal* networks[3] (networks whose edges vary over time) where *elementary components* of the network (such as nodes or edges) are added and/or removed as the network *evolves* over time. Examples of such networks include biological signal transduction networks with node dynamics, biochemical reaction networks, infectious disease contact networks, and time-evolving correlation networks[3]. Typically, such networks may have a set of *critical* elementary components (or simply "critical" components) whose presence or absence alters a significant *global* property of these networks between two time steps. Finding such a set of critical components in the context of temporal networks is more popularly called the *anomaly* detection or the *change-point* detection problem in statistics, computer science or data mining literature[4,5], and prior widely used application areas of these problems include medical condition monitoring[6,7], weather change detection[8,9] and speech recognition[10,11].

In this paper, we will use the two terms "graph" and "network" *interchangeably*. We provide a formal definition of the network anomaly detection problem following a mathematical framework similar to what is described in[12]. To identify critical components of a temporal network, one first needs to provide details for the following four specific items:

(*i*) the network *model* under consideration,

(*ii*) a definition of the *elementary components* of the network, and

(*iii*) how the network *changes* over time,

(*iv*) the *property* of the network that will be used to identify *critical* components.

Let $t \geq 0$ be the time variable. For this paper, the details of these four items are as follows:

(*i*)′ The temporal network model considered in this paper is a *complete undirected* graph $G(t) = (V, E, \omega(t))$ with a fixed set of n nodes V, a set of (all possible) edges $E = \{\{u,v\} \mid u,v \in V, u \neq v\}$, and an edge-weight function $\omega : E \times [0, \infty) \mapsto \mathbb{R}^+ \cup \{0\}$ that assigns a *non-negative real* number $\omega(e,t)$ to each edge $e \in E$ for every time t. We will

use the notations $w(u,v,t)$, $w(v,u,t)$ and $w(e,t)$ interchangeably to indicate the weight of an edge $e = \{u,v\}$ at time t. Note that alternatively we could also view $G(t)$ as a $n \times n$ connectivity matrix $A(t) = [a_{u,v}(t)]$ where the $(u,v)^{\text{th}}$ entry of the matrix is given by $a_{u,v}(t) = w(u,v,t)$.

$(ii)'$ The elementary components are the edges of the network, or equivalently following the convention as described above, they are every pair of nodes in the network.

$(iii)'$ The network changes over time by modification of these elementary components, *i.e.*, the network changes over time by changing the weights of edges while keeping the *same* set of nodes. Note that a change of the weight of an edge from zero to a non-zero value is interpreted as the *addition* of a *new* edge, whereas a change of the weight of an edge from a non-zero value to zero is interpreted as the *deletion* of an *existing* edge.

$(iv)'$ The property of the network studied in this paper is the curvature of G given by a suitable version of the *Forman's combinatorialization* of *Ricci curvature* for networks (henceforth will be referred to simply as the "*Forman-Ricci curvature*")[13–17]. For now, assume that the Forman-Ricci curvature for a complete graph $G(t) = (V,E,w(t))$ at time t is a real-valued scalar function $\mathfrak{C}_G : \left(V \times V \setminus \{(u,u)\,|\,u \in V\}\right) \times [0,\infty) \mapsto \mathbb{R}$ that maps every pair of nodes of G at time t to a real number with the assumption that

$\rhd$ $\mathfrak{C}_G(u,v,t) = \mathfrak{C}_G(v,u,t)$, *i.e.*, $\mathfrak{C}_G(u,v,t)$ is *symmetric* with respect to its first two arguments, and

$\rhd$ $\mathfrak{C}_G(u,v,t) = 0$ if $w(u,v,t) = 0$.

Then, our *network anomaly detection problem* can be defined as follows.

Given two complete graphs $G(t_1) = (V,E,w(t_1))$ and $G(t_2) = (V,E,w(t_2))$, find one or more pair(s) of nodes u,v such that the value of $|\mathfrak{C}_G(u,v,t_1) - \mathfrak{C}_G(u,v,t_2)|$ is sufficiently large.

In the above formulation of the network anomaly detection problem, we want $|\mathfrak{C}_G(u,v,t_1) - \mathfrak{C}_G(u,v,t_2)|$ to be sufficiently large since we want to identify large changes in the network to be more confident that the network is altered. This is equivalent to identifying connected components of the network with similar curvature values in the two time steps and focusing on the bridges among these components. Further technical details are provided later in the "Methods and Materials" section. Here we solve the network anomaly detection problem in the context of changes of the human brain network caused by *Attention Deficit Hyperactivity Disorder* (ADHD). ADHD is one of the most common neuro-developmental disorders of childhood impacting parts of the brain that help us plan, focus on, and execute tasks. ADHD impacts approximately 11% of children and 5% of adults in the US alone. It is usually first diagnosed in childhood and often lasts into adulthood. Children with ADHD may have trouble paying attention, exhibit controlling impulsive behaviors, or be overly active. Unfortunately, the causes and risk factors for ADHD are still unknown, and as of yet there is *no* single clinical test that helps diagnose ADHD before its onset. There are several published neuroimaging studies that link the behavioral symptoms of ADHD to altered connections between brain regions. In the last few years, network analysis methods have been extensively used for studying properties of human brain networks[18–21]. The human brain can be divided into different regions based on functional or anatomical properties[22]. One can consider these regions as nodes of a brain network and define the edges as functional correlations among brain regions. Two prior graph theoretical studies of ADHD[23, 24] reported changes at the global level of the entire brain but did not study any altered connection patterns between different regions in the brain. In this paper we use our curvature-based network anomaly detection algorithms to detect statistically significant altered connection patterns between different regions of the brain.

Another related work[25] characterized the brain networks affected by *Autism Spectrum Disorder* (ASD) using *Ollivier-Ricci* curvature. Individuals with ASD have altered white matter developmental patterns compared to individuals without ASD. White matter development can be altered by neuro-inflammation, which in turn is associated with abnormalities in cerebrospinal fluid circulation. Autologous cord blood infusion, a potential therapy, is believed to reduce neuro-inflammation and promote white matter development, thus triggering a reconfiguration of connectivity patterns in the brain. Simhal *et al.*[25] used the Ollivier-Ricci network curvature (based on the mass transportation distance) to quantify the changes in the brain network after administering ASD patients a single infusion of autologous umbilical cord blood. They calculated the Spearman correlation between changes in clinical behavioral scores and changes in curvature following treatment, and identified a relationship between clinical improvement and altered curvature in three white matter pathways that are implicated in social and communication abilities. We, on the other hand, use the combinatorial Forman-Ricci network curvature to determine and quantify the local and global changes in the brain network due to the onset of ADHD and highlight the target nodes and edges which undergo major alterations due to the disease. We believe that our work in localizing the changes in the brain network will aid biomedical science in targeting specific treatment of the disease before the behavioral symptoms manifest.

Notational simplifications

For simplified exposition and ensuring that notations are minimalistic, we will use the following conventions in suitable places through the article including supplementary documents.

- ▷ We may omit the time variable t from the argument of G, w, $\mathfrak{C}_G$, or any other notation that uses t in its argument if it is completely clear from the context.

- ▷ We may interpret an edge e with $w(e,t) = 0$ as an edge that is actually *not* present in the given graph at time t.

- ▷ If we use the above conventions, then we will further simplify exposition by renaming the two (complete) input graphs $G(t_1) = (V, E, w(t_1))$ and $G(t_2) = (V, E, w(t_2))$ of the network anomaly detection problem as $G_1 = (V, E_1, w_1)$ and $G_2 = (V, E_2, w_2)$, respectively, where

  - ▷ $E_1 = E \setminus \{e \,|\, w(e, t_1) = 0\}$,

  - ▷ $E_2 = E \setminus \{e \,|\, w(e, t_2) = 0\}$,

  - ▷ $w_1 : \{(u, v) \,|\, u, v \in V, u \neq v\} \mapsto \mathbb{R}^+ \cup \{0\}$ is given by $w_1(u, v) = w(u, v, t_1)$, and

  - ▷ $w_2 : \{(u, v) \,|\, u, v \in V, u \neq v\} \mapsto \mathbb{R}^+ \cup \{0\}$ is given by $w_2(u, v) = w(u, v, t_2)$.

Brief history of various notions of curvature for networks

Various notions of curvature are already widely used in disciplines such as physics and mathematics to study properties of high-dimensional objects of certain types [26, 27]. However, extensions of these curvature concepts to graphs and hyper-graphs are quite non-trivial for several reasons such as the discreteness and the lack of a preferred geometric embedding for combinatorial objects. There are several ways previous researchers have attempted to formulate such an extension in addition to what is used in this paper; we briefly review two such major approaches. We note that the references cited at various places in the paper also show that these kinds of network curvature measures can encode non-trivial topological properties that are not expressed by more established graph-theoretic measures such as degree distributions, clustering coefficients or betweenness centralities. Moreover, some of these references also show that these curvature measures can explain many phenomena one frequently encounters in real network-theoretic applications that are *not* easily explained by other measures.

The first kind of extension that gives rise to network curvatures involves some appropriate discretization of the Ricci curvature for a Riemannian manifold to capture metric properties of the manifold. Typically, these extensions defines a curvature value for every edge of a graph. In addition to the Forman-Ricci discretization used in this paper, another relevant and more direct discretization is the *Ollivier's* discretization of Ricci curvature (the "Ollivier-Ricci" curvature)[28–31]. Roughly speaking, the Ollivier-Ricci curvature is calculated by defining a probability distribution on the neighborhoods of two nodes of an edge and then calculating the difference between the weight of the edge and the L_1 Wasserstein distance between the two above distributions. This kind of curvature is very different compared to the Forman-Ricci curvature used in this paper.

In another direction and historically much before the discretizations of Ricci curvature were formulated, Gromov and others defined a notion of network curvature (the "Gromov-hyperbolic" curvature) via geodesic triangles that captures properties of the set of exact and approximate geodesics of the entire network[26, 32]. There is a large body of research works dealing with various aspects of this measure, e.g., see[12, 33–37]. The Gromov-hyperbolic curvature is a global measure in the sense that it assigns a scalar value to the entire graph, and therefore is not directly suitable for comparing components of two graphs.

There are also other notions of network curvatures explored in past literatures. For example, Chow and Luo[38] provided a notion of network curvature based on circle packings.

Methods and Materials

Formal definition of 2-complex based Forman-Ricci network curvature

This paper used the combinatorial complex-based Forman-Ricci network curvature as defined below. For other notions of network curvatures, we refer the reader to references[26, 28, 30]. Assume that $G = (V, E, w)$ is the given graph with n nodes and m edges. For a formal definition we require some basic concepts from topology as available in introductory textbooks such as[39, 40]; for convenience of the reader, we summarize these concepts in supplementary document S1. Conceptually, we define Forman-Ricci curvature of a network by "extrapolating" the network to higher-dimensional complexes via topological association[13]. To define such a topological association, the following definitions and assumptions are used for a simplicial complex:

▷ We define a partial order relation $\prec$ between faces of various dimensions of a simplex or a convex polytope in the usual manner: a ℓ-face $\mathfrak{f}^\ell$ is a parent of a ℓ'-face $\widehat{\mathfrak{f}}^{\ell'}$ (denoted by $\widehat{\mathfrak{f}}^{\ell'} \prec \mathfrak{f}^\ell$) if $\widehat{\mathfrak{f}}^{\ell'}$ is contained in $\mathfrak{f}^\ell$. Also, two ℓ-faces $\mathfrak{f}^\ell$ and $\widehat{\mathfrak{f}}^\ell$ are *parallel* (denoted by $\mathfrak{f}^\ell \parallel \widehat{\mathfrak{f}}^\ell$) if they have either at least one common *immediate predecessor* or at least one common *immediate successor* (in the partial order $\prec$) *but not both*.

▷ We assume that there exists a weight function $\omega(\mathfrak{f})$ that assigns a non-negative weight (real number) to every face $\mathfrak{f}$. We will provide precise details of our implementation of the weight function in the next sub-section.

Informally, the higher-dimensional complex in the topological association is obtained by "gluing" nodes, edges, cycles and other sub-graphs of the given graph. There are many alternate ways such topological associations can be performed[12–14, 41, 42]. Our topological association is similar to that used in[12] and is described as follows. For $q \in \{0, 1, 2\}$, we topologically associate a q-simplex with a $(q+1)$-clique $\mathcal{K}_{q+1}$, *i.e.*, 0-simplexes, 1-simplexes and 2-simplexes are associated with nodes, edges and 3-cycles (triangles), respectively. Next, we use the concept of an *order* of a 2-simplex for more non-trivial topological association. Consider a p-face f^p of a q-simplex for $q \in \{0, 1, 2\}$. An order d association of such a face, which we will denote by the notation f^p_d with the additional subscript d, is associated with a sub-graph of *at most d* nodes that is obtained by starting with $\mathcal{K}_{p+1}$ and then *optionally* replacing each edge by a path between the two nodes. In other words,

▷ f^0_d is a node of G for all $d \geq 1$.

▷ f^1_2 is an edge, and f^1_d for $d > 2$ is a path having at most d nodes between two nodes adjacent in G.

▷ f^2_3 is a triangle (cycle of 3 nodes or a 3-cycle), and f^2_d for $d > 3$ is obtained from 3 nodes in a 3-cycle by connecting every pair of nodes by a path such that the total number of nodes in the sub-graph is at most d.

For a node v, an edge e and a cycle $\mathscr{C}$, let the notations $v \sim e$ and $e \sim \mathscr{C}$ indicate that v is an end-point of e and e is an edge of $\mathscr{C}$, respectively. The basic formula of the 2-complex based order d Forman-Ricci curvature of an edge $e = \{u, v\} \in E$ is given by[13]:

$$\mathfrak{C}_G^{2,d}(e) \stackrel{\text{def}}{=} \mathfrak{C}_G^{2,d}(u,v) = \omega(e) \left[\left(\sum_{e \sim f_d^2} \frac{\omega(e)}{\omega(f_d^2)} + \sum_{v \sim e} \frac{\omega(v)}{\omega(e)} \right) - \sum_{e' \parallel e} \left| \sum_{e', e \sim f_d^2} \frac{\sqrt{\omega(e)\omega(e')}}{\omega(f_d^2)} - \sum_{v \sim e, v \sim e'} \frac{\omega(v)}{\sqrt{\omega(e)\omega(e')}} \right| \right] \quad (1)$$

Note that $\mathfrak{C}_G^{2,d}(u,v) = 0$ if $\omega(e) = 0$ as stated previously in the "Introduction" section.

Based on our need, we make one slight modification of the above formula. Finding all edges e' such that e' is parallel to an edge e would involve an worst-case running time of $O(n^d)$ which is quite prohibitive for us since we select d as 5. To make our calculations computationally tractable we therefore take only those edges parallel to e that belong to the same face, *i.e.*, instead of $\sum_{e' \parallel e}$ we use $\sum_{e' \parallel e, e' \sim f_d^2}$. Note that two parallel edges $e.e'$ belonging to the same face cannot have a common end-point. Thus, the simplified formula used in this paper is the following:

$$\mathfrak{C}_G^{2,d}(e) \stackrel{\text{def}}{=} \mathfrak{C}_G^{2,d}(u,v) = \omega(e) \left[\left(\sum_{e \sim f_d^2} \frac{\omega(e)}{\omega(f_d^2)} + \sum_{v \sim e} \frac{\omega(v)}{\omega(e)} \right) - \sum_{e' \parallel e, e \sim f_d^2} \frac{\sqrt{\omega(e)\omega(e')}}{\omega(f_d^2)} \right] \quad (2)$$

Equation (2) may look complicated at first glance to some readers. For the convenience of the reader, we illustrate the calculation of $\mathfrak{C}_G^{2,5}(e)$ for a small-size graph in supplementary document S2.

Let $\mathfrak{C}_G^{2,d}(e)^{(1)}$ and $\mathfrak{C}_G^{2,d}(e)^{(2)}$ be the values of $\mathfrak{C}_G^{2,d}(e)$ as computed by (1) and (2), respectively. What properties should a graph possess so that the two values $\mathfrak{C}_G^{2,d}(e)^{(1)}$ and $\mathfrak{C}_G^{2,d}(e)^{(2)}$ are identical? Call an edge e' adjacent to the edge e if e and e' share a node, *i.e.*, if e' is of the form $\{u, v'\}$ or $\{u', v\}$ for some $u' \neq u$ or $v' \neq v$. We define an edge e' to be an *order d hanging edge* with respect to an edge $e = \{u, v\}$ (or, simply a hanging edge if other parameters are clear from the context) provided it satisfies the following two conditions (see Figure 1 for an illustration):

▷ The edges e and e' are adjacent.

▷ G does not contain a (simple) cycle of length (number of edges) at most d containing the edges e and e'.

Proposition 1. *If G has no hanging edges with respect to e then $\mathfrak{C}_G^{2,d}(e)^{(1)} = \mathfrak{C}_G^{2,d}(e)^{(2)}$.*

A proof of Proposition 1 is provided in supplementary document S3; note that the claim in Proposition 1 does *not* depend on the weights of the faces. Some examples of graph classes that satisfy the condition in Proposition 1 are as follows (these are just some examples, and do *not* necessarily list *every* possible graph class that satisfy the condition):

▷ The class of *complete k*-partite graphs for any $k \geq 3$ satisfy the condition in Proposition 1.

▷ The class of *complete* bipartite graphs with every partition having at least two nodes satisfy the condition in Proposition 1 for any $d \geq 4$.

▷ Any biconnected graph of circumference at most d satisfy the condition in Proposition 1.

▷ The class of standard hypercube graphs[43] satisfy the condition in Proposition 1 for any $d \geq 4$. To see this, let $n = 2^k$ and let $G = (V, E)$ be the standard k-dimensional hypercube where $V = \{(b_1, \ldots, b_k) \mid b_1, \ldots, b_k \in \{0, 1\}\}$ and two nodes are connected by an edge if and only if they differ in exactly one coordinate. Let $u = (x_1, x_2, x_3, \ldots, x_k)$, $v = (1 - x_1, x_2, x_3, \ldots, x_k)$, and $u' = (1 - x_1, 1 - x_2, x_3, \ldots, x_k)$. Then, a cycle of length at most 4 containing the edges e and e' consists of the edges $\{u, v\}, \{v, u'\}, \{u', u''\}, \{u'', u\}$ where $u'' = (x_1, 1 - x_2, x_3, \ldots, x_k)$.

▷ Increasing d allows more graph classes to satisfy the condition in Proposition 1. For the largest possible value of d, namely when $d = n$, all biconnected graphs satisfy the condition in Proposition 1.

▷ Consider the classical Erdös-Rényi random graph model[44] $G(n, p)$, namely the class of random graphs, parameterized by p, in which each possible edge $\{u, v\}$ is selected *independently* for inclusion in G with a probability of p for some $0 < p < 1$, and for convenience let the notation $\text{diam}(H)$ denote the diameter of a graph H. The random graph $G(n, p) \setminus \{u, v\}$ obtained from $G(n, p)$ by deleting the two nodes u and v and all the edges incident on these two nodes is itself an Erdös-Rényi random graph on $n - 2$ nodes with the same p (since edges are selected independently). Note that $G(n, p)$ satisfies the condition in Proposition 1 if $\text{diam}(G(n, p) \setminus \{u, v\}) \leq d - 3$, and there are at most $\binom{n}{2} < n^2/2$ choices of the two nodes u and v. Thus, using known extremal results for $\text{diam}(G(n, p))$ and the union bound for probabilities we get the following bounds (the standard phrase "with high probability (*w.h.p.*)" in probabilistic methods refers to a probability of at least $1 - o(1)$, *i.e.*, a probability whose limit is 1 as n tends to infinity):

  ▷ If $p \geq \frac{2 \log_2 n}{n}$ then for all sufficiently large n $G(n, p)$ satisfies the condition in Proposition 1 *w.h.p.* for all $d \geq \frac{2 \log_2 n}{\log_2 \log_2 n}$. For this result we need to use the bounds in Theorem 7.1 and Theorem 7.2 in[45].

  ▷ If p is a fixed constant less than 1 then for all sufficiently large n $G(n, p)$ satisfies the condition in Proposition 1 *w.h.p.* for all d. For this result we need to use the bounds in Corollary 10.11 in[46].

Existence of a cycle $\mathscr{C}$ of at most d edges containing the edges e and e' implies the existence of a path $\mathscr{P}$ between u' and v of $\alpha \leq d - 2$ edges $\{u, u_1\}, \{u_1, u_2\}, \ldots, \{u_{\alpha - 1}, u'\}$ where the nodes $u_1, \ldots, u_{\alpha - 1}$ are distinct from the nodes u' and v. As mentioned in the "Introduction" section, our graph could be viewed as an $n \times n$ connectivity matrix $A = [a_{u,v}]$ where zero entries correspond to edges *not* present in the graph. Viewed in this context, satisfying the condition in Proposition 1 is *tantamount* to satisfying the following claim:

> for every ordered sequence of $\alpha - 1 \leq d - 3$ nodes from $V \setminus \{u, v, u'\}$, say the nodes $u_1, \ldots, u_{\alpha - 1}$, at least one of the α quantities (henceforth to be referred to as the "weights of the sequence") $w(u, u_1), w(u_1, u_2), \ldots, w(u_{\alpha - 1}, u')$ must be zero.

The total number Λ of ordered sequences of at most $d - 3$ nodes from $V \setminus \{u, v, u'\}$ is given by $\Lambda = \sum_{j=1}^{d-3} \binom{n-3}{j} j!$. Since Λ rapidly grows with n and d, it is unlikely that each of these Λ ordered sequences will have at least one zero weight, unless n and d are relatively small or unless the connectivity matrix A has a *large* number of zeroes. For the connectivity matrix considered in this paper, $n = 200$ and $d = 5$, giving $\Lambda = \binom{197}{2} \times 2 + \binom{197}{1} = 38809$, a very large number.

Details of the ADHD **brain networks analyzed in this study**

The data for our empirical analysis was collected from the UCLA Multimodal Connectivity Database[47] as available via the website `http://umcd.humanconnectomeproject.org`. This is a web based data repository with specific data analysis tools. The data here are in the form of connectivity matrices derived from neuroimaging data. The site is administered by MGH/UCLA Human Connectome Project. The source of the data %added[comment=newly added] available via various links in the website are of various imaging modalities such as fMRI, DTI, structural MRI and EEG, however the imaging modality for the specific study that we use is only fMRI. The data is also divided into various subject groups, ages, gender and spans different disorders such as Alzheimer's, Autism, and ADHD. Further details are available in supplementary document S5.

Algorithmic and other specific details of our analysis of the ADHD brain networks

Since computation of $\mathfrak{C}_G^{2,d}(e)$ via Equation (2) requires (implicit or explicit) enumeration of cycles of up to d edges, the running time *is* prohibitive for arbitrary correlation matrices and arbitrarily large d. Thus, to reduce the time complexity we made the following decisions so that curvatures and z-scores can be computed within reasonable time:

▷ We selected $d = 5$ so that using appropriate data structures we could still get the empirical results discussed in this paper within reasonable time.

▷ The original connectivity matrix had 200 nodes with a total of 34036 non-zero entries (edges). In order to reduce the computation time, we sparsified the given correlation matrix by zero-ing out all entries that are less than a threshold of 0.4. We observed that for a threshold of 0.3, the total number of non-zero entries was 27701 and our algorithm took a *very* long time to compute the edge curvatures. However, for a threshold of 0.4, the total number of non-zero entries was 25220 and the time taken was reduced to a *manageable* level. We also observed that for a higher threshold of 0.57 the graph corresponding to the non-zero entries in the thresholded connectivity matrix became disconnected, and the total number of non-zero edges was *significantly* reduced to 8800.

The computation of $\mathfrak{C}_G^{2,5}(e)$ requires a value for the weight $\omega(f_d^2)$ for $d \in \{4,5\}$. To satisfy this demand, we only considered those cycles that are "chordal", *i.e.*, they can be built by merging triangles along their common edge. We then specify the weight function ω for the given graph $G = (V, E, \omega)$ in the following manner:

▷ The weight $\omega(e)$ of an edge $e \in E$ is the correlation value of the edge, *i.e.*, $\omega(e) = \omega(e)$.

▷ The weight $\omega(v)$ of a node $v \in V$ is the average of the correlation values of the edges incident on v, *i.e.*, $\omega(v) = \frac{\sum_{v \sim e} \omega(e)}{\deg(v)}$ where $\deg(v)$ is the degree of node v.

▷ The weight of a triangle (3-cycle) f_3^2 consisting of edges e_1, e_2, e_3 is $\omega(f_3^2) = (\omega(e_1) + \omega(e_2) + \omega(e_3))/3$.

▷ The weight $\omega(f_d^2)$ for $d \in \{4,5\}$ were obtained by adding the weights of triangles from which f_d^2 was built.

Evaluation of statistical significance

As briefly mentioned in the "Introduction" section, given the two graphs $G_1 = (V_1, E_1, \omega_1)$ and $G_2 = (V_2, E_2, \omega_2)$ and a pair of nodes u, v such that the *absolute value* of $\Delta_{u,v} = \mathfrak{C}_{G_1}^{2,5}(u,v) - \mathfrak{C}_{G_2}^{2,5}(u,v)$ is sufficiently large, we need to evaluate the statistical significance of this difference. For this purpose, we need a subroutine that, given a brain network $G = (V, E, \omega)$ and a pair of nodes $u, v \in V$, generates a random network $\mathcal{G}^{u,v}$ similar to G in the sense that the random network keeps the relation between the nodes u and v (*i.e.*, if $\{u,v\} \in E$ then the edge is always kept but otherwise u and v are never connected by an edge), and preserves the first-order topological characteristics (such as the degree sequence) of G but randomizes higher-order statistics (such as distribution of paths) of G. Due to the absence of accurate generative null models for brain networks, we use the well-known and widely used Markov-chain algorithm for generating random networks[48] that starts with the given network G, and repeatedly swaps randomly chosen pairs of connections up to η times for a sufficiently large positive integer η (for our specific implementation we selected $\eta = 5|E|$). For the convenience of the reader, we provide the pseudo-code of the Markov-chain algorithm in supplementary document S4. This approach was used in papers such as[49,50]. Given such a subroutine for generating random networks, we use the following generic method which is by now standard in the network science literature dealing with various types of biological networks[49–51]. We repeat the following step 100 times: during the jth iteration we generate two independent random networks $\mathcal{G}_{1,j}^{u,v}$ and $\mathcal{G}_{2,j}^{u,v}$ from G_1 and G_2, respectively, and calculate the value $\widetilde{\Delta}_{u,v,j} = \mathfrak{C}_{\mathcal{G}_{1,j}^{u,v}}^{2,5}(u,v) - \mathfrak{C}_{\mathcal{G}_{2,j}^{u,v}}^{2,5}(u,v)$. We then calculate the standard Z-score[52] of the observed value $\Delta_{u,v}$ with respect to the 100 samples $\widetilde{\Delta}_{u,v,1}, \widetilde{\Delta}_{u,v,2}, \ldots, \widetilde{\Delta}_{u,v,100}$ as the desired statistical significance, and use a standard cut-off of 2 for labeling the Z-score as the criteria for being statistically significant. Using a Z-score to relate a property of an actual network to the distribution of the same property in an *ensemble* of randomized networks is a well established and popular method in network science, e.g., see[53].

Results

For our results, $G_1 = (V, E_1, \omega_1)$ and $G_2 = (V, E_2, \omega_2)$ correspond to the mean connectivity matrix of the 27 healthy individuals and the 24 diseased (ADHD) patients, respectively, and henceforth will be referred to as the "*control network*" and the "*disease network*", respectively. We implemented a program in PYTHON to efficiently compute $\mathfrak{C}^{2,5}$, and used it to compute the values of $\mathfrak{C}_{G_1}^{2,5}(u,v)$ and $\mathfrak{C}_{G_2}^{2,5}(u,v)$ for every pair of nodes u and v.

Theoretical results

Comparison of Ollivier-Ricci curvature and Forman-Ricci curvature

References such as[28,54] provide precise mathematical definitions of the Ollivier-Ricci curvature of a graph; for the convenience of the reader we do summarize these definitions in supplementary document S6. Let $\mathfrak{C}_G^{\text{O-R}}(e)$ denote the value of Ollivier-Ricci curvature of an edge e of a given graph $G = (V, E)$. It is not possible to directly compare the numerical values of $\mathfrak{C}_G^{2,d}(e)$ and $\mathfrak{C}_G^{\text{O-R}}(e)$ for a given graph G since $\mathfrak{C}_G^{\text{O-R}}(e)$ is a value in the range $[-2, 1]$[54] but $\mathfrak{C}_G^{2,d}(e)$ is a value in the range $[-\ell, \ell']$ for two numbers $\ell, \ell' > 0$ that depends on n and m via the topology of G. However, under the standard assumption that the zero value of a curvature represents the default flat curvature, one would expect the signs of the two curvatures to be *not* correlated if they indeed represent different discretization of the Ricci curvature for a Riemannian manifold. For further discussion, we recall the standard definition of the sign function sgn:

$$
\text{sgn}(x) = \begin{cases} 1, & \text{if } x > 0 \\ \text{undefined}, & \text{if } x = 0 \\ -1, & \text{otherwise} \end{cases}
$$

The following theorem shows the lack of correspondence of the signs of the two curvatures.

Theorem 1. *For every $s_1, s_2 \in \{-1, 1\}$ and for every $d \geq 5$ the following claim is true: there exists an infinite family $\mathscr{G}$ of graphs such that for every graph $G \in \mathscr{G}$ it holds that* $\text{sgn}(\mathfrak{C}_G^{2,d}(e)) = s_1$ *and* $\text{sgn}(\mathfrak{C}_G^{\text{O-R}}(e)) = s_2$ *for some edge e of G.*

The proof of Theorem 1 is lengthy, requires review of some definitions and notations of prior literature, and uses Theorem 13.3 and its corollary in[55]; we therefore provide the proof in supplementary document S6. To the best of our knowledge, the result in Theorem 1 has not been formally proved before.

On a related note, Samal *et al.*[17] *empirically* compared $\mathfrak{C}_G^{\text{O-R}}(e))$ to some version of the Forman-Ricci curvature, and found that those two measures were correlated for a few real networks (similar assertions were also very briefly mentioned in[56]). However, as Samal *et al.*[17] themselves cautioned, their results should not be construed as implying that one of these curvature measures can be used as a *universal substitute* for the other measure, but merely that for some real networks using one of those measures that allow faster implementation may suffice. Our results in Theorem 1 lends further theoretical justification to their caution. There is *no* contradiction between our theoretical results and the empirical observations in[17,56]. To provide an analogy, consider the standard simplex algorithm for the linear programming problem. It has been theoretically proved that there exists examples of inputs for which for which simplex takes exponential time[57]. However, it is also a well-known observation that simplex runs in polynomial time for many empirical data[58], and there is no contradiction between these two kinds of knowledge.

Empirical results

First-order statistics of edge curvatures of control and disease networks

To begin, we display the first-order statistics of the curvatures of edges for the two networks in Figure 2. These results shows that average curvature values of the two networks are *not* significantly different due to many edges of similar curvature values.

Edges with large statistically significant curvature differences

In accordance with the framework of our network anomaly detection problem as defined in the "Introduction" section, we next calculated the curvature differences $\Delta_{u,v} = \mathfrak{C}_{G_1}^{2,5}(u,v) - \mathfrak{C}_{G_2}^{2,5}(u,v)$ for every pair of nodes $u, v \in V$. The *average* and the *standard deviation* of the $\Delta_{u,v}$ values were found to be 0.384 and 6.9, respectively. Since the standard deviation was close to 7, we selected a value about twice the standard deviation (15) as a suitable cutoff for detecting drastic changes in curvature between corresponding edges. We found 32 pairs of nodes u, v which satisfied $|\Delta_{u,v}| > 15$, and each such pair of node u, v correspond to an edge $e = \{u, v\} \in E_1 \cap E_2$ (*i.e.*, node pairs with non-zero correlation values) in *both* the control network G_1 and the disease network G_2 (and henceforth referred to as the 32 "*outlier edges*"). These edges and their corresponding curvature differences in the disease and control networks are shown in Table 1. Following are a few observations about these 32 outlier edges:

1. The top 2 outlier edges (*i.e.*, the outlier edges with the largest and the second largest $|\Delta_e|$ values) belong to both the *occipital* region and to the *parietal lobe* region of the human brain.

2. 7 out of these 32 outlier edges lie in the *occipital* region of the brain.

3. 7 out of these 32 outlier edges belong to the *parietal lobe* region of the brain.

4. 12 out of these 32 outlier edges lie in the *left frontal pole* region of the brain.

The above findings show that most of the extreme curvature changes happen in the *occipital cortex* and the *frontal cortex* regions of the brain. This is consistent with previous neuroimaging studies of ADHD involving voxel-wise estimation of regional tissue volume changes by Wang *et al.*[59]. Their statistical results showed significant volume alteration in the brains of the patients with ADHD. While there were significant volume *reductions* in the *prefrontal*, *parietal* and *temporal* regions, volume *enlargements* were observed in the *occipital regions* and *posterial lateral ventricle*. Along the same line, more recently Sun *et al.* in[60] in 2017 conducted a comparative study by building a model using anatomic and diffusion-tensor MRI of different regions of the brains of children with ADHD with that of children without the disease via MRI. They found that there were differences in the cortical shape of the frontal lobe and areas in the occipital lobe along with central cortex in the brains of ADHD patients with those in (age and sex-matched) control groups.

We next focus our attention to the top 5 outlier edges (based on the $|\Delta_e|$ values) and 4 additional edges that connect brain regions spanned by a high number of outlier edges. In Table 2 we show these 9 outlier edges with their Δ_e values and, using the approach outlined in the "Methods and Materials" section, the corresponding *z*-scores of the Δ_e-values. Based on the standard assumption that a *z*-score of absolute value at least 2 is statistically significant, we note that 4 of the 9 outlier edges in Table 2 have statistically significant curvature differences. We make the following observations regarding these 4 edges with large and statistically significant curvature differences:

(*a*) The curvature values of all the four edges are higher in the disease network compared to the control network.

(*b*) One of the four edges, namely the topmost edge in Table 2, belong to the *occipital* region and the *superior parietal lobule* region of the brain.

   The *occipital* region of the brain is involved in *visual processing* activities such as object recognition, memory formation and distance perception, whereas the *superior parietal lobule* is important in planning movements, spatial reasoning and attention. As we have already mentioned before, the importance of these regions in ADHD development is consistent with previous studies by Wang *et al.*[59].

(*c*) The remaining three edges, namely the second, third and fourth edges from the top in Table 2, belong to the *occipital* region and the *frontal operculum* region of the brain.

   Generally speaking, the *frontal operculum* plays a role in thought, cognition, and planning behavior; for example, see the paper[61] regarding the role of *operculum* in cognition and the subsequent changes in ADHD patients. Studies such as[62] suggest that the *frontal operculum* is a critical node in a brain network controlling activities in other brain areas to perform a wide array of cognitive tasks. Activities in the *frontal operculum* are involved in predicting the efficacy of future tasks, and also in cessation of engagement in a task. Decreased functional connectivities between the left frontal inferior operculum and other areas of brain were studied in[63].

In addition, for illustrative purposes, we show in Figure 3 the histogram of the frequency of the difference of edge curvatures among corresponding randomly generated network pairs of disease and control networks in a given range for the edge between L_Precuneus_Ctx_Superior_Lateral_Occipital_Ctx_128 and R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneus_Ctx_143. As can be seen, the actual curvature difference of this edge *is* indeed an outlier with respect to the curvature differences of this edge in the random networks.

Edge curvature vs. edge weight for anomaly detection

A natural question that arises here is the following: "*why not simply use the edge-weights to find out significant differences between the disease and control network?*" Previous studies such as[64] have indicated that edge weight by itself is not a sufficiently informative measure, for example the set of nodes with the highest sum of edge weights do not capture all the *hubs* in the brain networks. It is thus of interest to see the differences in edge weights between the corresponding edges in the two networks, and check if a measure as simple as the edge weight difference could also encode the higher order correlations that the curvature method encodes. Let us denote this difference of edge weights between the control and disease networks by Λ_e. We need to be careful to ensure a *fair* comparison between the Δ_e and the Λ_e values since the Λ_e values lie in the same range of $[-1, 1]$, but this is *not* necessarily the case for the Δ_e values. Thus, for a fair comparison between the two sets of values, we "*normalize*" the Δ_e values to lie in the same range of $[-1, 1]$ using the linear map $\eta : \Delta_e \mapsto \frac{2\Delta_e - \text{high} - \text{low}}{\text{high} - \text{low}}$ where high and low are the maximum and minimum, respectively, of Δ_e values over all edges *e*.

Figure 4 shows the first-order statistics of the normalized curvatures differences and the edge weight differences over all pairs of nodes in the disease and the control network. For both curvature differences and edge weight differences the standard deviations are very small, indicating that most individual values for them are concentrated within a small interval around their average values of 0.119 and 0, respectively. Table 3 shows the edge weight differences and the normalized curvature differences of the nine outlier edges reported in Table 2 between the diseased and control networks. As can be seen, the absolute values

of all the actual edge weight differences are *significantly* smaller than the absolute values of their corresponding normalized curvature differences, and therefore edge weight differences would not have detected the same critical edges as our curvature differences.

Discussion

Using the Forman-Ricci curvature measure as given by (2) in the framework of network anomaly detection, in this paper we have identified edges in brain networks of ADHD patients versus healthy humans that exhibit statistically significant changes of the curvature measure. Our usage of the Forman-Ricci curvature measure focused on 2-complex based associations of order 5 with a specific scheme for computing weights of faces. The empirical analysis can be further improved by using k-complex based associations of order d for higher values of k and d. However, a practical limitation of the extent to which such refinements can be carried out in practice stems from the computational complexity issues. Another improvement may come from using different weighting scheme for the faces; for example, for our case Heron's formula could not be used to provide weights of 3-cycles because it leads to complex numbers, but this obstacle may be avoided for other kinds of data.

Another direction for extension of our works in this paper would be to apply the Forman-Ricci curvature measure to analyze networks for other diseases such as schizophrenia, alzheimer's disease or epilepsy. A further but equally important extension in this direction would be to use curvature measures to analyze the progression of a disease when suitable temporal data are available. For example, suppose that we have the brain network data over k time steps spanning the progress of a disease in a patient. This provides a series of networks G^t for $t = 1, \ldots, k$, and applying our analysis methods over successive pairs of networks G^t, G^{t-1} for $t = 2, \ldots, k$ may provide insights into the progression of interactions of various parts of the brain as the onset of the disease progresses in a patient. This type of analysis has been done by prior researchers using other more standard network measures, e.g., authors in[65,66] showed that seizures evolve from a more random to regular and then back to random network structure before termination, but curvature analysis methods are likely to provide further non-trivial insights into the temporal structure of these networks.

A third direction of future research would be to compare another discretization of Ricci curvature, namely the Ollivier-Ricci curvature that was mentioned in Theorem 1, with the Forman-Ricci curvature in the context of our anomaly detection framework to compare the edges detected by the two measures. As we mentioned before, Samal *et al.*[17] did empirically compare the Ollivier-Ricci curvature to some version of the Forman-Ricci curvature for a few real networks, we believe a more comprehensive comparison may reveal salient distinguishing characteristics of these two curvatures for brain networks.

Finally, a further future research direction could be to investigate the biomedical significance of the curvature changes of the critical connections identified by curvature analysis methods such as ours.

Data availability

The data for our empirical analysis is publicly available from the UCLA Multimodal Connectivity Database[47] via the website http://umcd.humanconnectomeproject.org.

References

1. Newman, M. E. J. *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

2. Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A. & Boguná, M. Hyperbolic geometry of complex networks. *Phys. Rev. E* **82**, 036106, DOI: 10.1103/PhysRevE.82.036106 (2010).

3. Holme, P. Modern temporal network theory: a colloquium. *Eur. Phys. J. B* **88**, DOI: 10.1140/epjb/e2015-60657-4 (2015).

4. Aminikhanghahi, S. & Cook, D. J. A survey of methods for time series change point detection. *Knowl. Inf. Syst.* **8**, 339–367 (2017).

5. Kawahara, Y. & Sugiyama, M. Sequential change-point detection based on direct density-ratio estimation. In *SIAM International Conference on Data Mining*, 389–400 (Society for Industrial and Applied Mathematics, Philadelphia, 2009).

6. Yang, P., Dumont, G. & Ansermino, J. M. Adaptive change detection in heart rate trend monitoring in anesthetized children. *IEEE Transactions on Biomed. Eng.* **53**, 2211–2219 (2006).

7. Bosc, M. *et al.* Automatic change detection in multimodal serial mri: application to multiple sclerosis lesion evolution. *Neuroimage* **20**, 643–656 (2003).

8. Ducre-Robitaille, J. F., Vincent, L. A. & Boulet, G. Comparison of techniques for detection of discontinuities in temperature series. *Int. J. Climatol.* **23**, 1087–1101 (2003).

9. Reeves, J., Chen, J., Wang, X. L., Lund, R. & Lu, Q. Q. A review and comparison of changepoint detection techniques for climate data. *J. Appl. Meteorol. Climatol.* **46**, 900–915 (2007).

10. Chowdhury, M. F. R., Selouani, S. A. & O'Shaughnessy, D. Bayesian on-line spectral change point detection: a soft computing approach for on-line asr. *Int. J. Speech Technol.* **15**, 5–23 (2011).

11. Rybach, D., Gollan, C., Schluter, R. & Ney, H. Audio segmentation for speech recognition using segment features. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 4197–4200 (Institute of Electrical and Electronics Engineers, Piscataway, 2009).

12. DasGupta, B., Janardhanan, M. V. & Yahyanejad, F. How did the shape of your network change? (on detecting network anomalies via non-local curvatures). *Algorithmica* **82**, 1741–1783 (2020).

13. Forman, R. Bochner's method for cell complexes and combinatorial Ricci curvature. *Discret. Comput. Geom.* **29**, 323–374 (2003).

14. Weber, M., Jost, J. & Saucan, E. Forman-Ricci flow for change detection in large dynamic data sets. *Axioms* **5**, 26 (2016).

15. Sreejith, R. P., Mohanraj, K., Jost, J., Saucan, E. & Samal, A. Forman curvature for complex networks. *J. Stat. Mech. Theory Exp.* **2016**, 063206 (2016).

16. Sreejith, R. P., Jost, J., Saucan, E. & Samal, A. Systematic evaluation of a new combinatorial curvature for complex networks. *Chaos, Solitons Fractals* **101**, 59–67 (2017).

17. Samal, A. *et al.* Comparative analysis of two discretizations of Ricci curvature for complex networks. *Sci. Reports* **8**, 8650 (2018).

18. Sporns, O. *Networks of the Brain* (The MIT Press, Cambridge, MA, 2011).

19. Sporns, O. The human connectome: A complex network. *Annals New York Acad. Sci.* **1224**, 109–125 (2011).

20. Gallos, L. K., Makse, H. A. & Sigman, M. A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proc. Natl. Acad. Sci.* **109**, 2825–2830, DOI: 10.1073/pnas.1106612109 (2012). https://www.pnas.org/content/109/8/2825.full.pdf.

21. Meunier, D., Lambiotte, R. & Bullmore, E. Modular and hierarchically modular organization of brain networks. *Front. Neurosci.* **4**, 200, DOI: 10.3389/fnins.2010.00200 (2010).

22. Sporns, O. & Betzel, R. F. Modular brain networks. *Annu. Rev. Psychol.* **67**, 613–640, DOI: 10.1146/annurev-psych-122414-033634 (2016).

23. Lin, P. *et al.* Global and local brain network reorganization in attention-deficit/hyperactivity disorder. *Brain Imaging Behav.* **8**, 558–569 (2014).

24. Barttfeld, P. *et al.* Functional connectivity and temporal variability of brain connections in adults with attention deficit/hyperactivity disorder and bipolar disorder. *Neuropsychobiology* **69**, 65–75 (2014).

25. Simhal, A. K. *et al.* Measuring robustness of brain networks in autism spectrum disorder with Ricci curvature. *Sci. Reports* **10**, 10819, DOI: 10.1038/s41598-020-67474-9 (2020).

26. Bridson, M. R. & Häfliger, A. *Metric Spaces of Non-Positive Curvature* (Springer-Verlag, Berlin Heidelberg, 1999).

27. Berger, M. *A Panoramic View of Riemannian Geometry* (Springer-Verlag, Berlin Heidelberg, 2003), 1 edn.

28. Ollivier, Y. A visual introduction to Riemannian curvatures and some discrete generalizations. In Dafni, G., McCann, R. J. & Stancu, A. (eds.) *Analysis and Geometry of Metric Measure Spaces: Lecture Notes of the 50th Séminaire de Mathématiques Supérieures (SMS), Montréal, 2011*, vol. 56, 197–219, DOI: 10.1090/crmp/056/08 (American Mathematical Society, Providence, RI, USA, 2013).

29. Ollivier, Y. Ricci curvature of Markov chains on metric spaces. *J. Funct. Analysis* **256**, 810–864, DOI: 10.1016/j.jfa.2008.11.001 (2009).

30. Ollivier, Y. A survey of Ricci curvature for metric spaces and Markov chains. In Kotani, M., Hino, M. & Kumagai, T. (eds.) *Probabilistic approach to geometry*, 343–381, DOI: 10.2969/aspm/05710000 (Mathematical Society of Japan, Tokyo, Japan, 2010).

31. Ollivier, Y. Ricci curvature of metric spaces. *Comptes Rendus Math.* **345**, 643–646, DOI: 10.1016/j.crma.2007.10.041 (2007).

32. Gromov, M. Hyperbolic groups. In Gersten, S. M. (ed.) *Essays in Group Theory*, vol. 8, 75–263, DOI: 10.1007/978-1-4613-9586-7_3 (Springer, New York, NY, 1987).

33. DasGupta, B., Karpinski, M., Mobasheri, N. & Yahyanejad, F. Effect of Gromov-hyperbolicity parameter on cuts and expansions in graphs and some algorithmic implications. *Algorithmica* **80**, 772–800, DOI: 10.1007/s00453-017-0291-7 (2018).

34. Benjamini, I. Expanders are not hyperbolic. *Isr. J. Math.* **108**, 33–36, DOI: 10.1007/BF02783040 (1998).

35. Albert, R., DasGupta, B. & Mobasheri, N. Topological implications of negative curvature for biological and social networks. *Phys. Rev. E* **89**, 032811, DOI: 10.1103/PhysRevE.89.032811 (2014).

36. Chalopin, J. *et al.* Fast Approximation and Exact Computation of Negative Curvature Parameters of Graphs. In Speckmann, B. & Tóth, C. D. (eds.) *34th International Symposium on Computational Geometry (SoCG 2018)*, vol. 99 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 22:1–22:15, DOI: 10.4230/LIPIcs.SoCG.2018.22 (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2018).

37. Fournier, H., Ismail, A. & Vigneron, A. Computing the Gromov hyperbolicity of a discrete metric space. *Inf. Process. Lett.* **115**, 576–579, DOI: 10.1016/j.ipl.2015.02.002 (2015).

38. Chow, B. & Luo, F. Combinatorial Ricci flows on surfaces. *J. Differ. Geom.* **63**, 97–129, DOI: 10.4310/jdg/1080835659 (2003).

39. Henle, M. *A Combinatorial Introduction to Topology* (Dover publications, New York, 1994).

40. Gamelin, T. W. & Greene, R. E. *Introduction to Topology* (Dover publications, New York, 1999).

41. Bloch, E. Combinatorial Ricci curvature for polyhedral surfaces and posets. *arXiv:1406.4598v1* (2014).

42. M. Weber, E. S. & Jost, J. Can one see the shape of a network? *arXiv:1608.07838* (2016).

43. Harary, F., Hayes, J. P. & Wu, H.-J. A survey of the theory of hypercube graphs. *Comput. & Math. with Appl.* **15**, 277–289, DOI: 10.1016/0898-1221(88)90213-1 (1988).

44. Erdös, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hungarian Acad. Sci.* **5**, 17–61 (1960).

45. Frieze, A. & Karonski, M. *Introduction to Random Graphs* (Cambridge University Press, Cambridge, U.K., 2015).

46. Bollobás, B. *Random Graphs*. Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, U.K., 2001), 2 edn.

47. Brown, J., Jeffery, D., Bandrowski, A., Horn, J. & Bookheimer, S. The ucla multimodal connectivity database: A web-based platform for brain connectivity matrix sharing and analysis. *Front. Neuroinformatics* **6**, 28, DOI: 10.3389/fninf.2012.00028 (2012).

48. Kannan, R., Tetali, P. & Vempala, S. Markov-chain algorithms for generating bipartite graphs and tournaments. *Random Struct. Algorithms* **14**, 293–308 (1999).

49. Albert, R. *et al.* A new computationally efficient measure of topological redundancy of biological and social networks. *Phys. Rev. E* **84**, 036117 (2011).

50. Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *escherichia coli*. *Nat. Genet.* **31**, 64–68 (2002).

51. DasGupta, B. & Liang, J. *Models and Algorithms for Biomolecules and Molecular Networks* (Wiley-IEEE Press, New Jersey, 2016).

52. Mendenhall, W. & Sincich, T. *Statistics for Engineering and the Sciences* (Pearson / Prentice Hall, New Jersey, 2007), 5 edn.

53. Milo, R. *et al.* Network motifs: Simple building blocks of complex networks. *Science* **298**, 824–827, DOI: 10.1126/science.298.5594.824 (2002).

54. Azarhooshang, N., Sengupta, P. & DasGupta, B. A review of and some results for Ollivier-Ricci network curvature. *Mathematics* **8**, 1416, DOI: 10.3390/math8091416 (2020).

55. Papadimitriou, C. H. & Steiglitz, K. *Combinatorial optimization: algorithms and complexity* (Prentice-Hall, Inc., NJ, USA, 1982).

56. Saucan, E., Samal, A., Weber, M. & Jost, J. Discrete curvatures and network analysis. *MATCH* **80**, 605–622 (2018).

57. Klee, V. & Minty, G. J. How good is the simplex algorithm? In Shisha, O. (ed.) *Inequalities III*, 159–175 (Academic Press, New York, 1972).

58. Shamir, R. The efficiency of the simplex method: A survey. *Manag. Sci.* **33**, 301–334, DOI: 10.1287/mnsc.33.3.301 (1987).

59. Wang, J., Jiang, T., Cao, Q. & Wang, Y. Characterizing anatomic differences in boys with attention-deficit/hyperactivity disorder with the use of deformation-based morphometry. *Am. J. Neuroradiol.* **28**, 543–547 (2007).

60. Sun, H. *et al.* Psychoradiologic utility of mr imaging for diagnosis of attention deficit hyperactivity disorder: A radiomics analysis. *Radiology* **287**, 620–630, DOI: 10.1148/radiol.2017170226 (2017).

61. Gordji, R. An investigation of abnormal brain connectivity associated with regions implicated in adhd. Honors Theses 802, University of Mississippi, Sally McDonnell Barksdale Honors College (2016).

62. Higo, T., Mars, R. B., Boorman, E. D., Buch, E. R. & Rushworth, M. F. S. Distributed and causal influence of frontal operculum in task control. *Proc. Natl. Acad. Sci.* **108**, 4230–4235, DOI: 10.1073/pnas.1013361108 (2011).

63. McCarthy, H. *et al.* Attention network hypoconnectivity with default and affective network hyperconnectivity in adults diagnosed with attention-deficit/hyperactivity disorder in childhood. *JAMA Psychiatry* **70**, 1329–1337, DOI: 10.1001/jamapsychiatry.2013.2174 (2013).

64. Farooq, H., Chen, Y., Georgiou, T. T., Tannenbaum, A. & Lenglet, C. Network curvature as a hallmark of brain structural connectivity. *Nat. Commun.* **10**, 4937, DOI: 10.1038/s41467-019-12915-x (2019).

65. Kramer, M. A. & Cash, S. S. Epilepsy as a disorder of cortical network organization. *Neuroscientist* **18**, 360–372, DOI: 10.1177/1073858411422754 (2012).

66. Schindler, K. A., Bialonski, S., Horstmann, M.-T., Elger, C. E. & Lehnertz, K. Evolving functional network properties and synchronizability during human epileptic seizures. *Chaos* **18**, 033119, DOI: 10.1063/1.2966112 (2008).

Acknowledgements

Author contributions statement

B.D. formalized the theoretical framework, T.C., S.T. and N.A. implemented the algorithms and conducted the simulations, T.C. and R.A. analyzed the results. All authors reviewed and edited the manuscript.

Competing interests

The authors declare *no* competing interests.

Software availability

The source code for our program to compute $\mathfrak{C}_G^{2,5}(e)$ is freely available at https://github.com/Tchatt2/Brain/blob/main/5cycle.py.
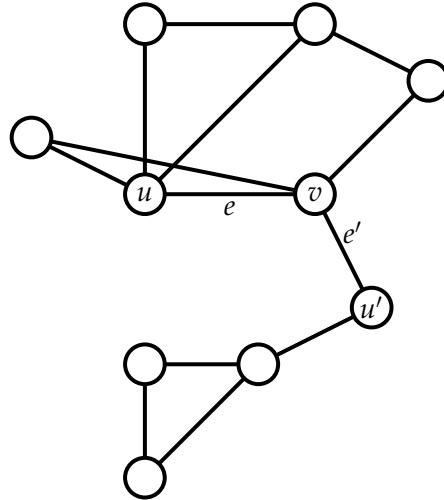
Figure 1. Illustration of a edge e' that is a hanging edge (of order d for any d) with respect to the edge e.


(a)

(b)
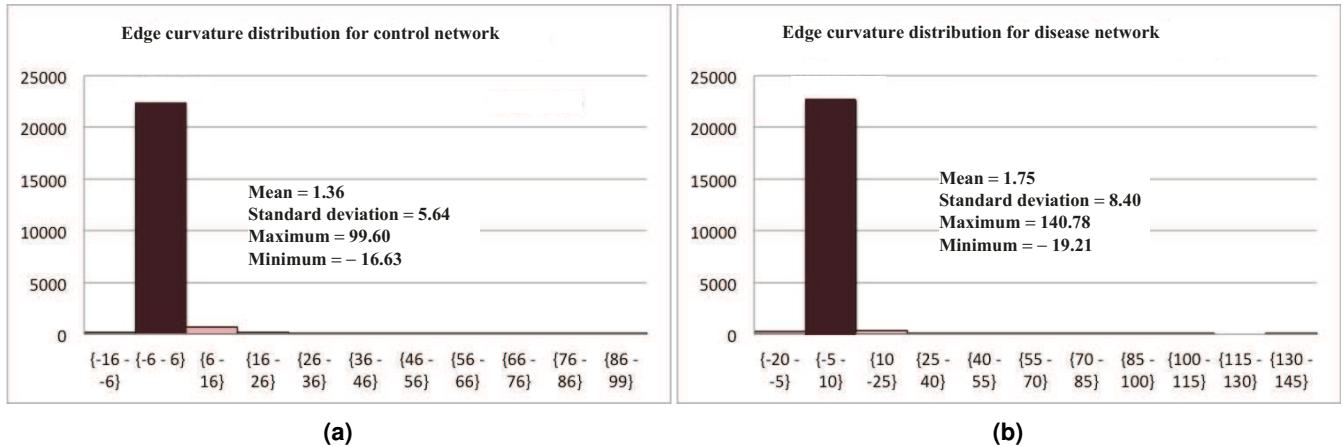
Figure 2. (a) Histogram showing the frequency of the edge curvatures in a given range for the control network. (b) Histogram showing the frequency of the edge curvatures in a given range for the disease network. For both (a) and (b), the heights of the bars along the y-axis indicate the number of samples belonging to the ranges given along the x-axis.
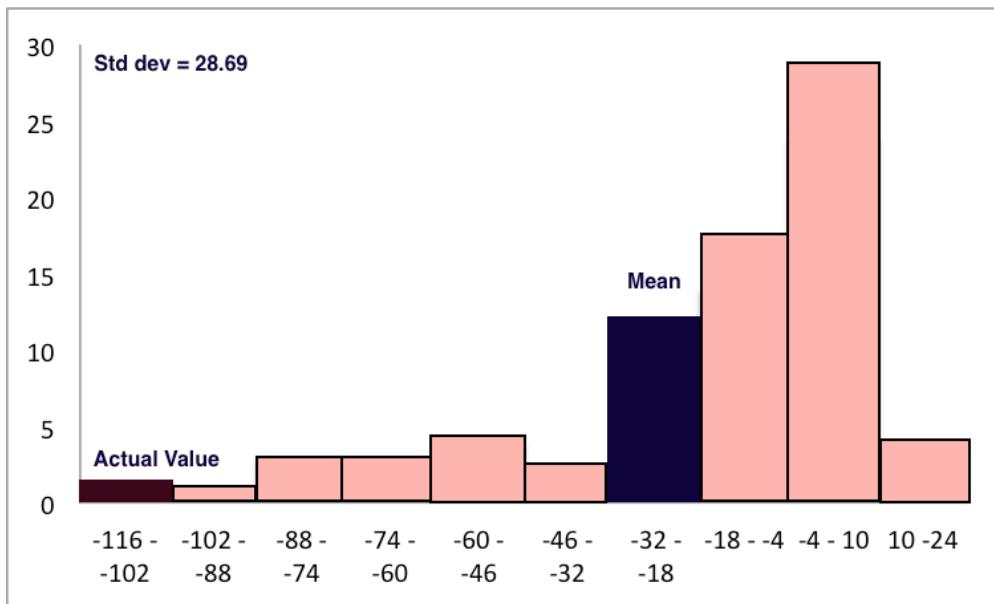
Figure 3. Histogram showing the frequency of the difference of edge curvatures among corresponding randomly generated network pairs of disease and control networks in a given range for the edge between L_Precuneous_Ctx_Superior_Lateral_Occipital_Ctx_128 and R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneous_Ctx_143 (the topmost edge in Table 2). The maroon bar indicates the range in which the curvature difference between the actual disease and control network belongs to while the blue bar indicates the population mean. The heights of the bars along the *y*-axis indicate the number of samples belonging to the ranges given along the *x*-axis.
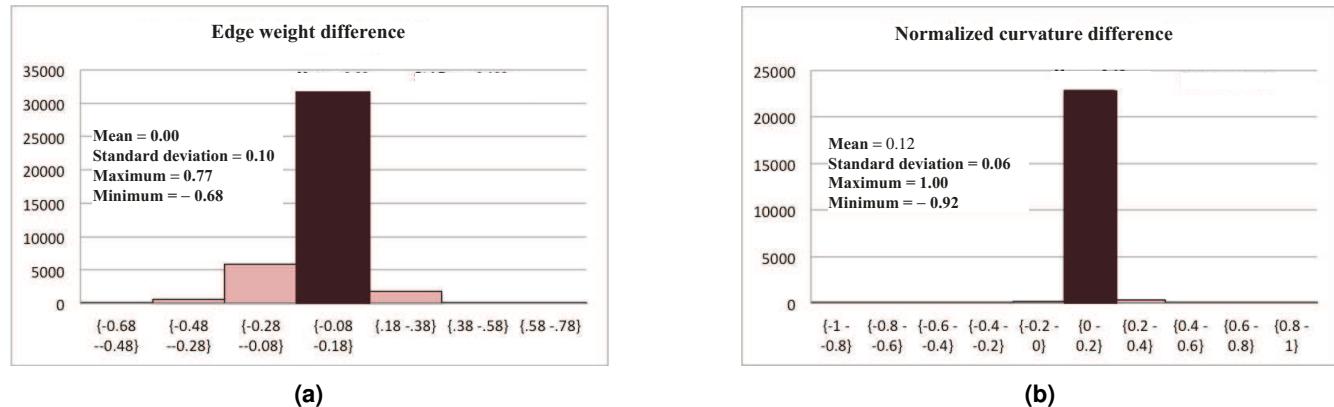
Figure 4. First-order statistics of the normalized curvatures differences and the edge weight differences over all pairs of nodes in the disease and the control network.

Edge name	Curvature difference Δ_e
R_Superior_Parietal_Lobule_Lateral_Occipital_Ctx_12, R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneous_Ctx_143	97.478
L_Precuneous_Ctx_Superior_Lateral_Occipital_Ctx_128, R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneous_Ctx_143	−116.717
R_Insular_Ctx_Frontal_Oper_Ctx_3, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	−87.363
L_Frontal_Pole_1, L_Posterior_MTG_32	−24.669
L_Frontal_Pole_1, L_Frontal_Pole_47	−62.764
L_Frontal_Pole_1, L_Frontal_Pole_57	−28.782
L_Frontal_Pole_1, L_Superior_Lateral_Occipital_Ctx_Angular_Gyrus_68	−18.273
L_Frontal_Pole_1, L_Superior_Lateral_Occipital_Ctx_Agular_Gyrus_90	−15.575
L_Frontal_Pole_1, L_Frontal_Pole_FOC_92	−29.940
L_Frontal_Pole_1, R_SFG_120	−16.636
L_Frontal_Pole_1, L_Frontal_Pole_122	−22.667
L_Frontal_Pole_1, M_ACC_127	−30.620
L_Frontal_Pole_1, R_Frontal_pole_133	−44.584
L_Frontal_Pole_1, R_PCC_148	−15.524
L_Frontal_Pole_1, R_Frontal_Pole_197	42.766
R_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_20, L_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_98	45.556
R_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_20, L_Precuneous_Ctx_Superior_Lateral_Occipital_Ctx_128	−48.615
R_Occipital_Pole_Occipital_Fusiform_Gyrus_25, R_Inferior_Lateral_Occipital_Ctx_101	−30.764
L_Precentral_60, R_Postcentral_Superior_Parietal_Lobule_Precuneous_Ctx_158	−54.354
M_Juxtapositional_Lobule_Ctx_75, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	−105.182
R_Postcentral_Precentral_76, R_Postcentral_Superior_Parietal_Lobule_Precuneous_Ctx_158	−28.090
L_Postcentral_81, R_Postcentral_Superior_Parietal_Lobule_Precuneous_Ctx_158	−70.806
R_Insular_Ctx_Central_Opercular_Ctx_Putamen_89, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	−18.388
R_Planum_Polare_Anterior_STG_Temporal_Pole_Insular_Cortex_105, R_Parietal_Operculum_Ctx_Planum_Temporale_Central_Operculum_Ctx_Heschls_193	17.111
R_Planum_Polare_Anterior_STG_Temporal_Pole_Insular_Cortex_105, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	−17.616
L_Insular_Ctx_Central_Opercular_Ctx_Heschls_110, R_Parietal_Operculum_Ctx_Planum_Temporale_Central_Operculum_Ctx_Heschls_193	−23.146
R_SFG_120, R_Frontal_Pole_197	−67.154
L_Lingual_Gyrus_TOFFC_129, R_TOFFC_168	21.590
R_Posterior_STG_Planum_Temporale_Central_Opercular_Ctx_152, R_Parietal_Operculum_Ctx_Planum_Temporale_Central_Operculum_Ctx_Heschls_193	55.522
M_ACC_156, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	−113.466
L_Temporal_Pole_Anterior_ITG_163, R_Anterior_ITG_171	−18.316
R_TOFFC_168, R_Lingual_Gyrus_Occipital_Fusiform_Gyrus_179	22.628

Table 1. The 32 outlier edges and their corresponding corresponding differences. The edge names are given in the format "u,v" where u and v are the names of the nodes in the network.

Edge name	Forman-Ricci curvature			
	Control $\mathfrak{C}^{2,5}_{G_1}(e)$	Diseased $\mathfrak{C}^{2,5}_{G_2}(e)$	Difference Δ_e	z-score of Δ_e
L_Precuneous_Ctx_Superior_Lateral_Occipital_Ctx_128, R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneous_Ctx_143	**24.066**	**140.784**	**−116.717**	**3.3**
M_ACC_156, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	**1.050**	**114.517**	**−113.466**	**2.3**
M_Juxtapositional_Lobule_Ctx_75, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	**1.338**	**106.521**	**−105.182**	**2.3**
R_Insular_Ctx_Frontal_Oper_Ctx_3, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	**1.421**	**88.785**	**−87.363**	**2.5**
L_Postcentral_81, R_Postcentral_Superior_Parietal_Lobule_Precuneous_Ctx_158	38.621	109.428	−70.806	1.8
R_Frontal_Pole_1, L_Frontal_Pole_122	10.066	32.734	−22.667	1.4
R_Superior_Parietal_Lobule_Lateral_Occipital_Ctx_12, R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneous_Ctx_143	99.605	2.126	97.478	−1.3
R_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_20, L_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_98	47.386	1.829	45.556	−1.8
R_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_20, L_Precuneous_Ctx_Superior_Lateral_Occipital_Ctx_128	0.698	49.314	−48.615	1.5

Table 2. 9 outlier edges with their curvature values and z-scores. The top 5 edges in the table are the top 5 of the 32 outlier edges, and the remaining 4 edges are selected from connecting those brain regions spanned by higher number of outlier edges. The edge names are given in the format "u,v" where u and v are the names of the nodes in the network. A z-score of absolute value at least 2 is generally considered to be statistically significant. Statistically significant curvature differences are shown in **bold blue**.

Edge Name	Edge weight difference Λ_e	Normalized curvature difference $\eta(\Delta_e)$
L_Precuneous_Ctx_Superior_Lateral_Occipital_Ctx_128, R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneous_Ctx_143	-0.059	-0.927
M_ACC_156, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	-0.303	-0.898
M_Juxtapositional_Lobule_Ctx_75, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	-0.142	-0.823
R_Insular_Ctx_Frontal_Oper_Ctx_3, L_Central_Opercular_Ctx_Frontal_Operculum_Ctx_Insular_Ctx_Putamen_198	-0.194	-0.663
L_Postcentral_81, R_Postcentral_Superior_Parietal_Lobule_Precuneous_Ctx_158	-0.106	-0.514
R_Frontal_Pole_1, L_Frontal_Pole_122	-0.017	-0.081
R_Superior_Parietal_Lobule_Lateral_Occipital_Ctx_12, R_Superior_Lateral_Occipital_Ctx_Superior_Parietal_Lobule_Precuneous_Ctx_143	-0.081	1.000
R_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_20, L_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_98	-0.031	0.533
R_Superior_Lateral_Occipital_Ctx_Precuneous_Ctx_20, L_Precuneous_Ctx_Superior_Lateral_Occipital_Ctx_128	-0.107	-0.314

Table 3. The edge weight differences and the normalized curvature differences of the 9 outlier edges reported in in Table 2 between the disease and the control network. The edge names are given in the format "u,v" where u and v are the names of the nodes in the network.

Supplementary documents

for

Detecting Network Anomalies Using Forman-Ricci Curvature and A Case Study for Human Brain Networks

Tanima Chatterjee, Réka Albert, Stuti Thapliyal, Nazanin Azarhooshang and Bhaskar DasGupta

Supplementary document S1

Summary of basic topological concepts

Here we review some basic concepts from topology; see introductory textbooks such as[39,40] for further information. For concreteness of exposition, let the underlying metric space be the r-dimensional real space $\mathbb{R}^r$ be for some integer $r > 1$.

- ▷ A subset $S \subseteq \mathbb{R}^r$ is *convex* if and only if for any pair $x, y \in S$, the *convex combination* of x and y is also in S (*i.e.*, $\lambda x + (1 - \lambda)y \in S$ for any real $0 \leq \lambda \leq 1$).

- ▷ A set of $k + 1$ points $x_0, \ldots, x_k \in \mathbb{R}^r$ are called *affinely independent* if and only if for all $\alpha_0, \ldots, \alpha_k \in \mathbb{R}$ $\sum_{j=0}^{k} \alpha_j x_j = 0$ and $\sum_{j=0}^{k} \alpha_j = 0$ implies $\alpha_0 = \cdots = \alpha_k = 0$.

- ▷ The *k-simplex* generated by a set of $k + 1$ affinely independent points $x_0, \ldots, x_k \in \mathbb{R}^r$ is the subset of $\mathbb{R}^r$ $\mathcal{S}(x_0, \ldots, x_k) = \left\{ \sum_{j=0}^{k} \alpha_j x_j \,|\, \forall j : \alpha_j \geq 0 \text{ and } \sum_{j=0}^{k} \alpha_j = 1 \right\}$ generated by *all* convex combinations of $x_0, \ldots, x_k$. For example, the equation of a k-simplex with *unit intercepts* is given by $\sum_{j=0}^{k} x_j = 1$ with $x_j \geq 0$ for all $0 \leq j \leq k$.

  - ▷ Each $(\ell + 1)$-subset $\left\{ x_{i_0}, \ldots, x_{i_\ell} \right\} \subseteq \left\{ x_0, \ldots, x_k \right\}$ defines the ℓ-simplex $\mathcal{S}(x_{i_0}, \ldots, x_{i_\ell})$ that is called a *face* of dimension ℓ (or a ℓ-*face*) of $\mathcal{S}(x_0, \ldots, x_k)$. A $(k - 1)$-face, 1-face and 0-face is called a *facet*, an *edge* and a *node*, respectively.

- ▷ A (closed) *halfspace* is a set of points satisfying $\sum_{j=1}^{r} a_j x_j \leq b$ for some $a_1, \ldots, a_r, b \in \mathbb{R}$. The convex set obtained by a bounded non-empty intersection of a finite number of halfspaces is called a *convex polytope* (called a *convex polygon* in two dimensions).

  - ▷ If the intersection of a halfspace and a convex polytope is a subset of the halfspace then it is called a *face* of the polytope. Of particular interests are faces of dimensions $r - 1$, 1 and 0, which are called *facets*, *edges* and *nodes* of the polytope, respectively.

- ▷ A *simplicial complex* (or just a complex) is a topological space constructed by the union of simplexes.

Supplementary document S2

Illustration of calculation of $\mathfrak{C}_G^{2,5}(e)$ for a simple six-node network

Consider the 6-node graph shown in Figure S1, and assume that all edge weights are $1/2$. We label various components of Equation (2) as follows:

$$\mathfrak{C}_G^{2,d}(e) = \omega(e)\left[\left(\underbrace{\sum_{e \sim f_d^2} \frac{\omega(e)}{\omega(f_d^2)}}_{\text{Term 1}} + \underbrace{\sum_{v \sim e} \frac{\omega(v)}{\omega(e)}}_{\text{Term 2}}\right) - \underbrace{\sum_{e'||e,\, e' \sim f_d^2} \frac{\sqrt{\omega(e)\omega(e')}}{\omega(f_d^2)}}_{\text{Term 3}}\right]$$

The various steps in the calculations for our graph $G = (V, E, \omega)$ are as follows (see Figure S1 for visual illustrations):

▷ The only f_3^2 containing the edge e is the 3-cycle $C_{2,3,4}$, the two f_4^2s containing the edge e are the 4-cycles $C_{2,3,4,5}$ and $C_{3,2,4,6}$, and the only f_5^2 containing the edge e is the 5-cycle $C_{3,2,5,4,6}$. The calculations of the relevant weight functions $\omega(\cdot)$ are as follows:

$\omega(u_2) = \frac{\omega(\{u_2,u_1\}) + \omega(\{u_2,u_3\}) + \omega(\{u_2,u_4\}) + \omega(\{u_2,u_5\})}{\deg(u_2)} = 1/2, \quad \omega(u_3) = \frac{\omega(\{u_3,u_2\}) + \omega(\{u_3,u_4\}) + \omega(\{u_3,u_6\})}{\deg(u_3)} = 1/2$

$\omega(e') = \omega(e') = 1/2$ for every edge $e' \in E$

$\omega(C_{2,3,4}) = \frac{\omega(e) + \omega(e_3) + \omega(e_4)}{3} = 1/2, \quad \omega(C_{2,4,5}) = \frac{\omega(e_1) + \omega(e_2) + \omega(e_4)}{3} = 1/2, \quad \omega(C_{3,4,6}) = \frac{\omega(e_3) + \omega(e_5) + \omega(e_6)}{3} = 1/2$

$\omega(C_{2,4,5}) = \frac{\omega(e_1) + \omega(e_2) + \omega(e_4)}{3} = 1/2$

$\omega(C_{2,3,4,5}) = \omega(C_{2,3,4}) + \omega(C_{2,4,5}) = 1, \quad \omega(C_{3,2,4,6}) = \omega(C_{2,3,4}) + \omega(C_{3,4,6}) = 1$

$\omega(C_{3,2,5,4,6}) = \omega(C_{2,3,4}) + \omega(C_{2,4,5}) + \omega(C_{3,4,6}) = 3/2$

▷ The total contributions of edge e to Term 1 and Term 2 are as follows:

Contribution to Term 1 is $\alpha_1 = \frac{\omega(e)}{\omega(C_{2,3,4})} + \frac{\omega(e)}{\omega(C_{2,3,4,5})} + \frac{\omega(e)}{\omega(C_{3,2,4,6})} + \frac{\omega(e)}{\omega(C_{3,2,5,4,6})} = 7/3$

Contribution to Term 2 is $\alpha_2 = \frac{\omega(u_2)}{\omega(e)} + \frac{\omega(u_3)}{\omega(e)} = 2$

▷ The calculations for Term 3 for various parallel edges are as follows:

▷ The edge e_2 is parallel to e in $C_{2,3,4,5}$ and $C_{3,2,5,4,6}$. The corresponding total contribution to Term 3 is as follows:

$$\gamma_1 = \frac{\sqrt{\omega(e)\omega(e_2)}}{\omega(C_{2,3,4,5})} + \frac{\sqrt{\omega(e)\omega(e_2)}}{\omega(C_{3,2,5,4,6})} = 5/6$$

▷ The edge e_5 is parallel to e in $C_{3,2,4,6}$ and $C_{3,2,5,4,6}$. The corresponding total contribution to Term 3 is as follows:

$$\gamma_2 = \frac{\sqrt{\omega(e)\omega(e_2)}}{\omega(C_{3,2,4,6})} + \frac{\sqrt{\omega(e)\omega(e_2)}}{\omega(C_{3,2,5,4,6})} = 5/6$$

▷ The final curvature value is then given by $\mathfrak{C}_G^{2,5}(e) = \omega(e) \times [\alpha_1 + \alpha_2 - \gamma_1 - \gamma_2] = 4/3$.

Supplementary document S3

Proof of Proposition 1

Note that $\left| \sum_{(e'||e) \wedge (e', e \sim f_d^2)} \frac{\sqrt{\omega(e)\omega(e')}}{\omega(f_d^2)} \right|$ is the same as $\sum_{(e'||e) \wedge (e', e \sim f_d^2)} \frac{\sqrt{\omega(e)\omega(e')}}{\omega(f_d^2)}$ since all weights are positive numbers. Thus, it follows from a direct comparison of (1) and (2) that $\mathfrak{C}_G^{2,d}(e)^{(1)} = \mathfrak{C}_G^{2,d}(e)^{(2)}$ if the quantity $\Upsilon = \sum_{e'||e, v \sim e, v \sim e'} \frac{\omega(v)}{\sqrt{\omega(e)\omega(e')}}$ is zero. Obviously, Υ is zero if there is *no* such edge e' that satisfies the condition "$\Phi \equiv (e'||e) \wedge (v \sim e) \wedge (v \sim e')$" of the summation in the calculation of Υ. Φ is satisfied if and only if both the following conditions are satisfied:

(**a**) $(v \sim e) \wedge (v \sim e')$ is true, *i.e.*, edges e and e' are adjacent. Assume, without loss of generality, that $e' = \{u', v\}$ for the next condition.

(**b**) Edges e and e' are parallel edges. Since the edges e and e' have a common child (node) in the partial order $\prec$, the definition of parallel edges implies that e and e' must *not* have a common parent f_d^2 in the partial order $\prec$. Since f_d^2 in our topological association is a cycle of d edges, it follows that there is no cycle of length (number of edges) at most d in G containing the edges e and e'.

The proof is completed by noting that if G has no hanging edge that Φ is not satisfied since either (**a**) or (**b**) is not satisfied.

Supplementary document S4

Pseudo-code of the Markov-chain algorithm for generating random networks

repeat

    select a pair of distinct edges $e_1 = \{u_1, v_1\}, e_2 = \{u_2, v_2\} \in E \setminus \{u, v\}$ randomly uniformly

    if $\{u_1, v_1\} \cap \{u_2, v_2\} = \emptyset$ **then**

        if $\{u_1, v_2\} \notin E$ and $\{v_1, u_2\} \notin E$ **then**

            add the pair of edges $e'_1 = \{u_1, v_2\}$ and $e'_2 = \{v_1, u_2\}$ to E; set $w(e'_1) \leftarrow w(e_1)$ and $w(e'_2) \leftarrow w(e_2)$

            remove the pair of edges e_1 and e_2 from E

        else

            if $\{u_1, u_2\} \notin E$ and $\{v_1, v_2\} \notin E$ **then**

                add the pair of edges $e'_1 = \{u_1, u_2\}$ and $e'_2 = \{v_1, v_2\}$ to E; set $w(e'_1) \leftarrow w(e_1)$ and $w(e'_2) \leftarrow w(e_2)$

                remove the pair of edges e_1 and e_2 from E

    endif

until η pairs of edges have been swapped

Supplementary document S5

Details of data format collected from UCLA Multimodal Connectivity Database

The data that we used was in the form of two sets of mean/consensus matrices that are essentially the mean of the connectivity matrices of 27 healthy and 24 diseased patients respectively. The brain was categorized into 200 regions, hence each matrix had 200 nodes. Both the disease as well as control matrix had the *same* set of nodes. Each matrix comes in the form of the following text files:

▷ ADHD_grp_mean_24_region_names_abbrev_file.txt : This file essentially represents the abbreviations of the brain region names for the ADHD patients.

▷ ADHD_grp_mean_24_region_names_full_file.txt : This file contains the full names of the brain regions of the ADHD patients.

▷ ADHD_grp_mean_24_region_xyz_centers_file.txt : This file contains the (x, y, z) coordinates of the brain regions of the ADHD patients.

▷ ADHD_grp_mean_24_connectivity_matrix_file.txt : This file contains the actual mean connectivity matrix for the ADHD patients.

▷ CON_group_mean_27_region_names_abbrev_file.txt : This file contains the abbreviations of the brain region names for the control group.

▷ CON_group_mean_27_region_names_full_file.txt : This file contains the full names of the brain regions of the control group.

▷ CON_grp_mean_27_region_xyz_centers_file.txt : This file contains the (x, y, z) coordinates of the brain regions of the control group.

▷ CON_group_mean_27_connectivity_matrix_file.txt : This file contains the actual mean connectivity matrix for the control group.

Supplementary document S6

Proof of Theorem 1

We label various components of Equation (2) as follows:

$$\mathfrak{C}_G^{2,d}(e) = \omega(e)\left[\left(\underbrace{\sum_{e\sim f_d^2}\frac{\omega(e)}{\omega(f_d^2)} + \underbrace{\sum_{v\sim e}\frac{\omega(v)}{\omega(e)}}_{\text{Term 2}}}_{\text{Term 1}}\right) - \underbrace{\sum_{e'||e,e',e\sim f_d^2}\frac{\sqrt{\omega(e)\omega(e')}}{\omega(f_d^2)}}_{\text{Term 3}}\right] \tag{3}$$

All of our constructions will satisfy the following claim: edge e' is parallel to e if and only if both e and e' belongs to some f_d^2. Thus, our proof will hold irrespective of whether we use (1) or (2) for the calculation of $\mathfrak{C}_G^{2,d}(e)$. For a fair comparison of the two curvatures using the definitions in[54] for the Ollivier-Ricci curvature, we will ensure that all edge weights in our constructions are 1, implying $\omega(v) = 1$ for all $v \in V$. For this case, we simplify (3) in the following manner. Let $\mathscr{F}_d^2 = \{f_d^2 \mid e \in f_d^2\}$ be the set of 2-faces of order d containing the edge e, and let $\eta_d^2 = |\mathscr{F}_d^2|$. Our constructions will ensure that all 2-faces of order d containing the edge e are triangulated for any $d > 3$ so that our weighting scheme described in the "Methods and Materials" section is applicable to these faces; in particular this implies that $\omega(f_d^2) = d-2$ for all $d \geq 3$. Term 2 in (3) evaluates to precisely 2 for our case. For any $d \geq 3$, a d-cycle $f_d^2 \in \mathscr{F}_d^2$ contributes $\frac{1}{\omega(f_d^2)} = \frac{1}{d-2} =$ to Term 1. Moreover, since such a f_d^2 has exactly $d-3$ edges parallel to e, it contributes $d-3/\omega(f_d^2)\frac{1}{d-2} = \frac{d-3}{d-2}$ to Term 3. Thus the total contribution over all d-cycles for any specific $d \geq 3$ is $\left(\frac{1}{d-2} - \frac{d-3}{d-2}\right)\eta_d^2 = \left(\frac{4-d}{d-2}\right)\eta_d^2$. Thus the curvature equation for our constructions simplify to the following:

$$\mathfrak{C}_G^{2,d}(e)) = 2 + \sum_{d\geq 3}\left(\frac{4-d}{d-2}\right)\eta_d^2 = 2 + \eta_3^2 - \sum_{d\geq 5}\left(\frac{d-4}{d-2}\right)\eta_d^2 \tag{4}$$

Next we briefly review key definitions and notations from[54] as needed in our proof of Theorem 1; for further details see the original paper[54]. Let $G = (V,E)$ be a given undirected unweighted graph. The following two notations are used subsequently:

▷ $\mathsf{Nbr}(v) = \{u \mid \{v,u\} \in E\}$ and $\deg(v) = |\mathsf{Nbr}(v)|$ are the set of neighbors and the degree, respectively, of a node $v \in V$.

▷ $\mathrm{dist}_G(u,v)$ is the *distance* (*i.e.*, number of edges in a shortest path) between the nodes u and v in G.

Fix an edge $\{u,v\} \in E$. Let $V_{u,v} = \{u,v\} \cup \mathsf{Nbr}(u) \cup \mathsf{Nbr}(v)$, and define two probability distributions $\mathbb{P}_u$ and $\mathbb{P}_v$ over the nodes in $V_{u,v} = \{u,v\}$ as follows:

$$\mathbb{P}_u(x) = \begin{cases} \frac{1}{1+\deg(u)}, & \text{if } x \in \{u\} \cup \mathsf{Nbr}(u) \\ 0, & \text{otherwise} \end{cases} \qquad \mathbb{P}_v(x) = \begin{cases} \frac{1}{1+\deg(v)}, & \text{if } x \in \{v\} \cup \mathsf{Nbr}(v) \\ 0, & \text{otherwise} \end{cases}$$

The *Earth Mover's Distance* corresponding to these two probability distributions, denoted by $\mathrm{EMD}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v)$, is the value of the objective function of an optimal solution of the following linear program with a variable $z_{u,v}$ for every pair of nodes $u,v \in V_{u,v}$:

$$\begin{aligned} \textit{minimize} \quad & \sum_{u\in V_{u,v}}\sum_{v\in V_{u,v}}\mathrm{dist}_G(u,v)\,z_{u,v} \\ \textit{subject to} \quad & \\ & \sum_{v\in V_{u,v}}z_{u',v} = \mathbb{P}_u(u'), \text{ for each } u' \in V_{u,v} \\ & \sum_{u\in V_{u,v}}z_{u,v'} = \mathbb{P}_v(v'), \text{ for each } v' \in V_{u,v} \\ & z_{u,v} \geq 0, \text{ for all } u,v \in V_{u,v} \end{aligned}$$

The Ollivier-Ricci curvature is then defined as $\mathfrak{C}_G^{\text{O-R}}(e) = 1 - \mathrm{EMD}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v)$. For calculation of $\mathfrak{C}_G^{\text{O-R}}(e)$ for our constructions, it will be useful to prove the following result.

Lemma 1. *Fix $\gamma \in \{1,2,3\}$. Consider an edge $e = \{u,v\}$ of the graph $G = (V,E)$ such that $\deg(u) = \deg(v) = \alpha > 1$ and $\mathrm{dist}_G(u',v') = \gamma$ for every $u' \in \mathsf{Nbr}(u)\setminus(\mathsf{Nbr}(v)\cup\{v\})$ and $v' \in \mathsf{Nbr}(v)\setminus(\mathsf{Nbr}(u)\cup\{u\})$. Then, letting $\beta = |\mathsf{Nbr}(u)\setminus(\mathsf{Nbr}(v)\cup\{v\})| = |v' \in \mathsf{Nbr}(v)\setminus(\mathsf{Nbr}(u)\cup\{u\})|$ it follows that*

$$\mathrm{EMD}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v) = \frac{\gamma\beta}{\alpha+1} \equiv \mathfrak{C}_G^{\text{O-R}}(e) = 1 - \mathrm{EMD}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v) = 1 - \frac{\gamma\beta}{\alpha+1}$$

Proof. Consider the edge-weighted complete bipartite graph $H = (A, B, F, w)$, where there is a node $a_{u'} \in A$ for every node $u' \in \mathrm{Nbr}(u) \cup \{u\}$, there is a node $b_{v'} \in A$ for every node $v' \in \mathrm{Nbr}(v) \cup \{v\}$, and the weight function $w : A \times B \mapsto \{0, 1, 2, 3\}$ is given by $w(u', v') = \mathrm{dist}_G(u', v')$. Build a directed single-source single-sink flow network[55] $H_f =$ from H in the following manner: add a new source node s and a new sink node t, add an arc (directed edge) from s to every node of A of weight zero and capacity 1, add an arc from every node of B to t of weight zero and capacity 1, orient every edge $\{a_{u'}, b_{v'}\} \in F$ from $a_{u'}$ to $b_{v'}$ and set its capacity to 1. Since $|A| = |B| = \alpha + 1$, we have $\mathbb{P}_u(u') = \mathbb{P}_v(v') = \frac{1}{\alpha+1}$ for all $u' \in \mathrm{Nbr}(u) \cup \{u\}$ and $v' \in \mathrm{Nbr}(v) \cup \{v\}$. Thus, since H is a complete bipartite graph, by a simple scaling it follows that $\mathrm{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v) = \frac{\mathcal{M}}{\alpha+1}$ where $\mathcal{M}$ is the total weight of a minimum-weight maximum s-t flow on H_f. Since the node-arc incidence matrix of a directed graph is totally unimodular, the flow value of every arc of any extreme-point optimal solution of the minimum-weight maximum s-t flow on H_f is integral and therefore 0 or 1 (see Theorem 13.3 and its corollary in[55]). This integrality of flow values and the fact that H is a complete bipartite graph imply $\mathcal{M}$ is also the total weight of a minimum-weight *perfect* matching of H.

We now estimate the total weight $\mathcal{M}$ of a minimum-weight *perfect* matching of H in the following manner. Let $A' = \{a_{u'} \mid u' \in \{u, v\} \cup (\mathrm{Nbr}(u) \cap \mathrm{Nbr}(v))\}$ and $B' = \{b_{v'} \mid v' \in \{u, v\} \cup (\mathrm{Nbr}(u) \cap \mathrm{Nbr}(v))\}$. Note that the edge-weights of H satisfy the following based on our assumptions: $\mathrm{dist}_G(a_{u'}, b_{v'}) = 0$ if $a_{u'} \in A'$, $b_{v'} \in B'$, $u' = v'$. We claim that there exists a minimum-weight perfect matching of H that uses the edges $\{a_{u'}, b_{u'}\}$ for all $u' \in \{u, v\} \cup (\mathrm{Nbr}(u) \cap \mathrm{Nbr}(v))$. For a contradiction, suppose that the edge $\{a_{u'}, b_{u'}\}$ is not used for some $u' \in \{u, v\} \cup (\mathrm{Nbr}(u) \cap \mathrm{Nbr}(v))$. Since our solution is a perfect matching, the nodes $a_{u'}$ and $b_{u'}$ must be matched to some other nodes, say to nodes $b_{v''} \in B$ and $a_{u''} \in A$, respectively. Then, if we instead use the edges $\{a_{u'}, b_{u'}\}$ and $\{a_{u''}, b_{v''}\}$ then using the total weight of this modified perfect matching is no more than that of the original perfect matching since

$$w(a_{u'}, b_{u'}) + w(a_{u''}, b_{v''}) = w(a_{u''}, b_{v''}) \leq w(a_{u''}, b_{u'}) + w(a_{u'}, b_{u'}) + w(a_{u'}, b_{v''}) = w(a_{u''}, b_{u'}) + w(a_{u'}, b_{v''})$$

Thus, there exists a perfect matching that $\{a_{u'}, b_{u'}\}$ for all $u' \in \{u, v\} \cup (\mathrm{Nbr}(u) \cap \mathrm{Nbr}(v))$ of total weight 0, and some perfect matching of the nodes in $A \setminus A'$ and $B \setminus B'$. Since $\mathrm{dist}_G(a_{u'}, b_{v'}) = \gamma$ for all $a_{u'} \in A \setminus A'$ and $b_{v'} \in B \setminus B'$, we get $\mathcal{M} = \gamma \times |A \setminus A'| = \gamma\beta$, and the desired bound follows. $\qquad\square$

We now continue with the proof of Theorem 1.

Proof for $s_1 = 1$, $s_2 = 1$

Consider the n-node graph as shown in Figure S2a where $n > 2$. Since $\eta_3^2 = 1$ and $\eta_d^2 = 0$ for all $d \geq 5$, using (4) we get $\mathfrak{C}_G^{2,d}(e)) = 2 + \eta_3^2 = 3 > 0$. For the Ollivier-Ricci curvature, note that (in the notations of Lemma 1 and[54]) $\beta = 0$ giving $\mathfrak{C}_G^{\text{O-R}}(e) = 1 > 0$.

Proof for $s_1 = 1$, $s_2 = -1$

Consider the n-node graph as shown in Figure S2b where $n > 25$. By straightforward calculation $\eta_3^2 = 6$, $\eta_5^2 = 15$, $\eta_d^2 = 0$ for all $d > 5$, and thus using (4) we get $\mathfrak{C}_G^{2,d}(e)) = 2 + \eta_3^2 - \frac{\eta_5^2}{3} = 3 > 0$. For the Ollivier-Ricci curvature, note that (in the notations of Lemma 1 and[54]) $\alpha = 16$, $\beta = 9$ and $\gamma = 2$, giving $\mathfrak{C}_G^{\text{O-R}}(e) = 1 - \frac{2 \times 9}{17} = -\frac{1}{17} < 0$.

Proof for $s_1 = -1$, $s_2 = 1$

Consider the n-node graph as shown in Figure S2c where $n > 5$. By straightforward calculation $\eta_3^2 = 4$, $\eta_5^2 = \eta_6^2 = 24$, $\eta_d^2 = 0$ for all $d > 6$, and thus using (4) we get $\mathfrak{C}_G^{2,d}(e)) = 2 + \eta_3^2 - \frac{\eta_5^2}{3} - \frac{\eta_6^2}{2} = -14 < 0$. For the Ollivier-Ricci curvature, note that (in the notations of Lemma 1 and[54]) $\beta = 0$ giving $\mathfrak{C}_G^{\text{O-R}}(e) = 1 > 0$.

Proof for $s_1 = -1$, $s_2 = -1$

Consider the n-node graph as shown in Figure S2d where $n > 10$. By straightforward calculation $\eta_3^2 = 1$, $\eta_5^2 = 16$, $\eta_d^2 = 0$ for all $d > 5$, and thus using (4) we get $\mathfrak{C}_G^{2,d}(e)) = 2 + \eta_3^2 - \frac{\eta_5^2}{3} = -\frac{7}{3} < 0$. For the Ollivier-Ricci curvature, note that (in the notations of[54]) $\alpha = 6$, $\beta = 4$ and $\gamma = 2$, giving $\mathfrak{C}_G^{\text{O-R}}(e) = 1 - \frac{2 \times 4}{7} = -\frac{1}{7} < 0$.

Supplementary figures

for

Detecting Network Anomalies Using Forman-Ricci Curvature and A Case Study for Human Brain Networks

Tanima Chatterjee, Réka Albert, Stuti Thapliyal, Nazanin Azarhooshang and Bhaskar DasGupta
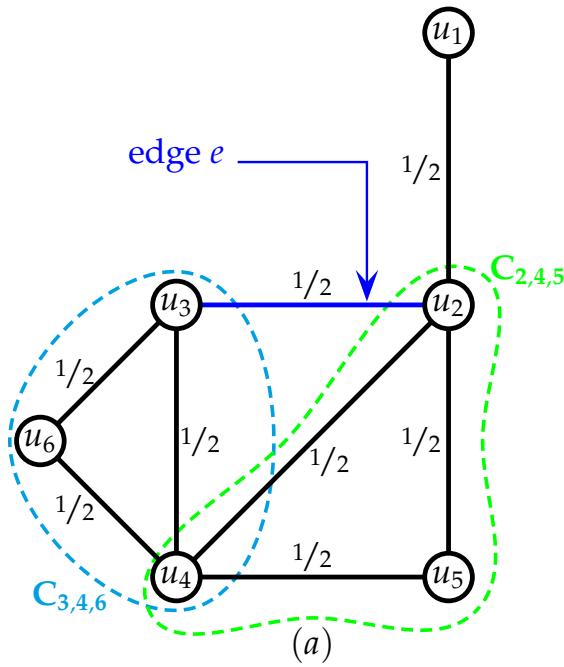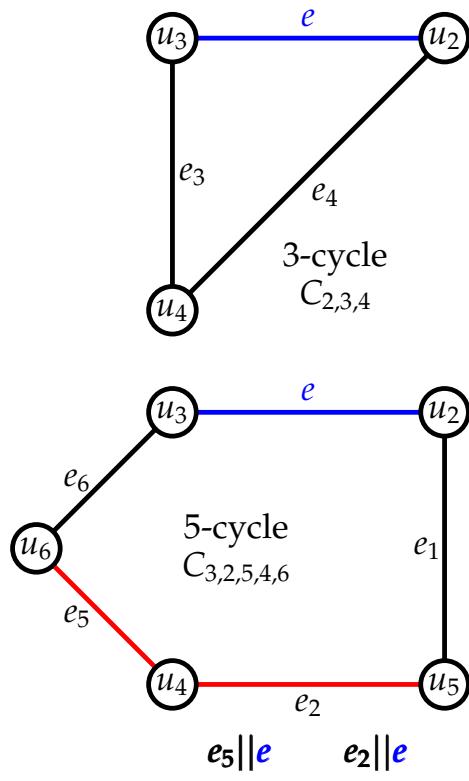
Figure S1. A simple example for illustration of calculation of $\mathfrak{C}^{2,5}(e)$. (a) The given graph and the edge e (all edge weights are $1/2$). (b) The connectivity matrix for the graph. (c) Listing of all f_d^2's (for $d \in \{3,4,5\}$) along with the relevant parallel edges used in the calculation of $\mathfrak{C}^{2,5}(e)$.
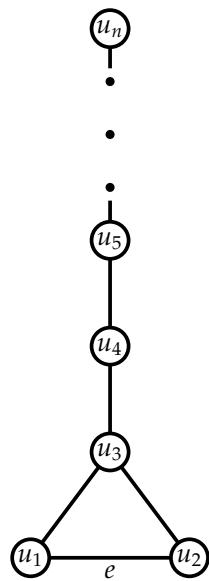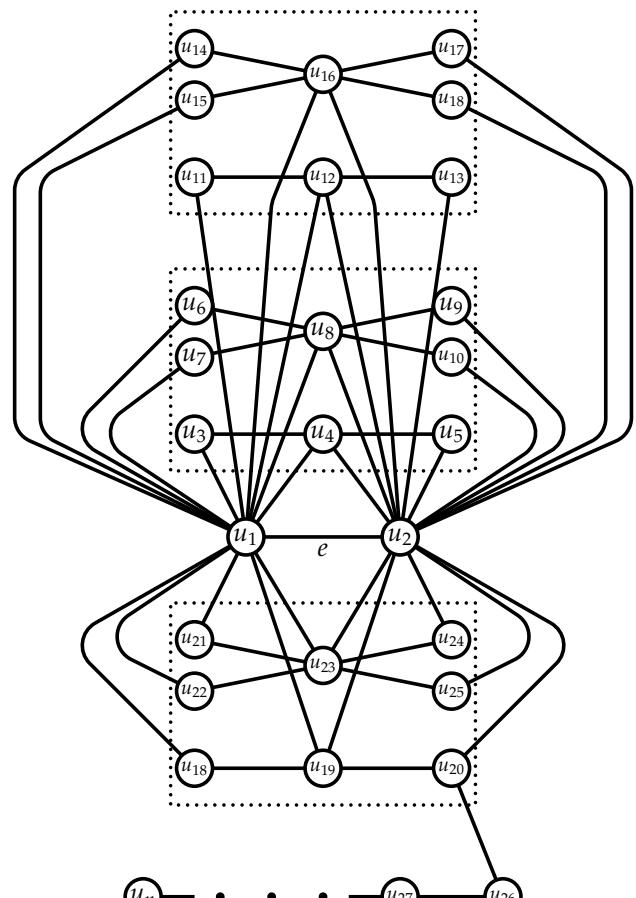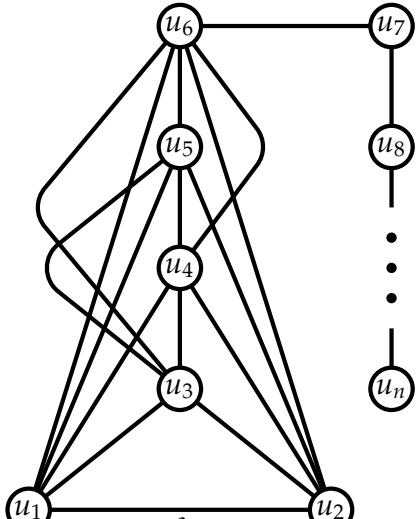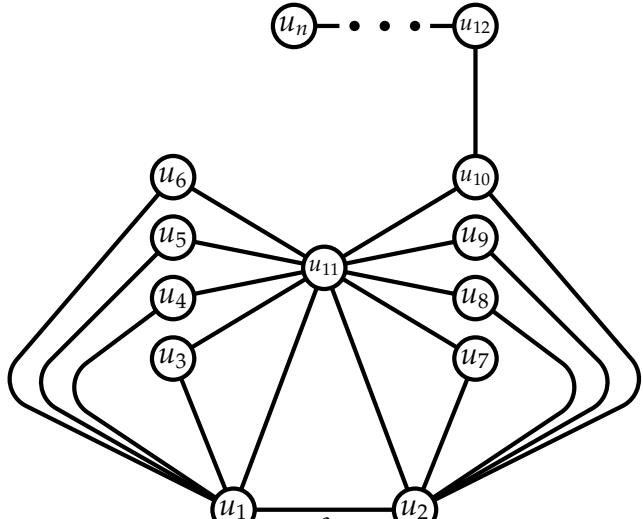
Figure S2. **a.** A n-node graph for which $\text{sgn}(\mathfrak{C}_G^{2,d}(e)) = \text{sgn}(\mathfrak{C}_G^{\text{O-R}}(e)) = 1$. **b.** A n-node graph for which $\text{sgn}(\mathfrak{C}_G^{2,d}(e)) = 1$ and $\text{sgn}(\mathfrak{C}_G^{\text{O-R}}(e)) = -1$. **c.** A n-node graph for which $\text{sgn}(\mathfrak{C}_G^{2,d}(e)) = -1$ and $\text{sgn}(\mathfrak{C}_G^{\text{O-R}}(e)) = 1$. **d.** A n-node graph for which $\text{sgn}(\mathfrak{C}_G^{2,d}(e)) = \text{sgn}(\mathfrak{C}_G^{\text{O-R}}(e)) = -1$.