Volume entropy and information flow in a brain graph

Hyekyoung Lee\textsuperscript{a,b}, Eunkyung Kim\textsuperscript{a,c}, Hyejin Kang\textsuperscript{b,d}, Youngmin Huh\textsuperscript{b,e}, Youngjo Lee\textsuperscript{d,f}, Seonhee Lim\textsuperscript{b,***}, Dong Soo Lee\textsuperscript{b,c,g,h,i,***}

\textsuperscript{a}Biomedical Research Institute, Seoul National University Hospital
\textsuperscript{b}Department of Nuclear Medicine,
\textsuperscript{c}Department of Rehabilitation Medicine, Seoul National University College of Medicine,
\textsuperscript{d}Data Science for Knowledge Creation Research Center,
\textsuperscript{e}Interdisciplinary Program in Cognitive Science,
\textsuperscript{f}Department of Statistics, College of Natural Sciences,
\textsuperscript{g}Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, and College of Medicine or College of Pharmacy, Seoul National University
\textsuperscript{h}Department of Mathematical Sciences, Seoul National University, Seoul
\textsuperscript{i}Korea Brain Research Institute, Daegu, Republic of Korea

Abstract

Entropy is a classical measure to quantify the amount of information or complexity of a system. Various entropy-based measures such as functional and spectral entropy have been proposed in brain network analysis. However, they are less widely used than traditional graph theoretic measures such as global and local efficiencies because they are not well-defined on a graph or difficult to interpret its biological meaning. In this paper, we propose a new entropy-based graph invariant, called volume entropy. It measures the exponential growth rate of the number of graph paths, based on the assumption that information flows through a graph forever. We model the information propagation on a brain graph by the generalized Markov system associated to a new edge transition matrix. The volume entropy is estimated by the stationary equation of the generalized Markov system. Moreover, its stationary distribution shows the information capacity of edge and the direction of information flow on a brain graph. The simulation results show that the volume entropy distinguishes the underlying graph topology.

\*Corresponding author
\**Corresponding author
Email addresses: slim@snu.ac.kr (Seonhee Lim), dsl@plaza.snu.ac.kr (Dong Soo Lee)
and geometry better than the existing graph measures. In the clinical application, the volume entropy of brain graphs was significantly related to healthy normal aging from 20s to 60s. In addition, the stationary distribution of information propagation gives a new insight into the information flow of functional brain graph.

1. Introduction

Brain regions function not only on their own way but also are linked functionally to each other. Brain is considered as a network whose nodes are brain regions, which are connected to each other according to the intensity of their functional links [1]. The functional connection between brain regions is indirectly derived from interregional correlation of brain regions on brain imaging studies such as functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) [2, 3, 4, 5, 6, 7]. On the efficient brain network, the information is quickly transmitted over the whole brain because there are sufficient short paths between all the pairs of brain regions. Furthermore, when we say brain networks are locally efficient, it implies that sufficient alternative paths exist between brain regions, even when a connection is damaged and nonfunctioning. These global and local efficiencies of brain networks change along with aging [8, 9, 10]. The change of a network has been quantified by complex graph theoretic measures including global and local efficiencies, path length, and clustering coefficient [11, 12, 13]. They were proposed to be used as network-based biomarkers of normal aging.

Information entropy is the measure of information or randomness of a system in information theory [14]. It is clearly defined as the expected value of the negative logarithm of the probability distribution of the system. Various kinds of information entropy, especially functional and spectral entropies, have been proposed as a graph invariant of human brain [15, 16]. Functional entropy is the information entropy of edge weights of brain network [15]. It approximated the probability distribution by the relative frequency histogram of edge weights.
Thus, if two networks have the same histogram of edge weights, they have the same functional entropy even if they have different topologies. Spectral entropy is the information entropy defined on the eigenvalues of the adjacency matrix of binary network [16]. The eigenvalues of the adjacency matrix provide some information about the shape of a graph, especially that related to connectedness. However, the relationship between the eigenvalues and connectedness is not clearly defined. Thus, it is difficult to interpret its biological meaning. Like the functional entropy, the spectral entropy also needed the procedure to approximate the probability distribution of the eigenvalues using Gaussian kernel regression. The approximation procedure included the parameter selection such as the number of bins in the functional entropy and the width of Gaussian in the spectral entropy.

Topological entropy is the complexity measure of a topological dynamical system in ergodic theory and geometry [17]. It measures how much energy has flowed in the system or how widely it has spread out over the system [17]. The topological entropy on a graph is called volume entropy [18]. The volume entropy assumes that the information flows through the links on a brain graph. If the time goes to infinite, the network paths through which information flows will increase exponentially. The volume entropy is the exponential growth rate of the number of network paths through which information flows. The larger the volume entropy is, the more information flows on the network. In this sense, the volume entropy is the network invariant of efficiency in terms of information spread. In this study, we introduce the volume entropy as a new invariant of brain network. To compute the volume entropy, we model the information flow on a graph by the generalized Markov system associated to a new edge transition matrix [18]. The volume entropy is obtained by the stationary equation of the generalized Markov system. Furthermore, its stationary distribution shows the edge capacity of information as well as the direction of information flow on a network. Thus, we can derive a directed network that represents the information propagation on a graph at the stationary state.

In simulations, we compare the volume entropy of various artificial graphs
such as regular, small-world, random, scale-free, hyperbolic, and modular graphs.

The results show that the volume entropy distinguishes the underlying graph topology and geometry better than the existing network measures such as global and local efficiencies and entropy-based invariants. We also apply the volume entropy to the resting state fMRI and PET data obtained from 38 normal controls between 20s and 60s. The volume entropy reveals the change of information flows in a brain graph during healthy normal aging. The main contributions of our work are as follows:

- We introduce the new invariant of brain graph, called volume entropy. It quantifies the efficiency of brain graph in terms of information propagation.
- The information flow on a brain graph is modelled by the generalized Markov system associated to a newly defined edge transition matrix. We also derive information flow on a brain graph based on the stationary equation of the generalized Markov system.
- The proposed method is applied to the functional and metabolic graphs obtained from resting state fMRI and PET data, respectively. The results reveal the relationship between the information propagation on a brain graph and age.

2. Materials and methods

2.1. Clinical data set

PET and fMRI data were simultaneously acquired from 38 healthy normal subjects (M/F: 19/18, age: 43.9±13.9) from 20s to 60s using a Siemens Biograph mMR 3T scanner (Siemens Healthcare Sector, Germany). MR images had 116 volume of images a subject. The first 4 volumes were discarded among 116 volumes and 112 volume of images a subject were used for network analysis. After preprocessing using the AFNI [19] and the FSL [20], we parcellated the brain into 116 regions of interest (ROIs) according to automated anatomical labelling (AAL) [21]. Among the 116 regions, 90 brain regions were selected as
the nodes and 26 cerebellar regions were not included in a graph (the number of nodes, \( p = 90 \)). The measurement of each node was obtained by averaging blood-oxygen-level dependent (BOLD) signals in the ROI of fMRI data. Each node had \( n \) measurements, which were the number of time points a subject in fMRI data (\( n = 112 \)). The measurement vectors of 90 ROIs were written by \( x^j_1, \ldots, x^j_p \in \mathbb{R}^n \) of the \( j \)th subject (\( j = 1, \ldots, 38, p = 90, n = 112 \)).

PET images were preprocessed using the Statistical Parametric Mapping (SPM8, www.fil.ion.ucl.ac.uk/spm) and PVElab software [22]. The image intensity of gray matter was globally normalized to 50. The measurement of a node was obtained by averaging FDG uptakes in the corresponding ROI. We divided the data into two groups, young (age: 32.2 ± 6.9) and old (age: 55.6 ± 7.7) depending on whether a subject was more than 45 years old. The number of subjects in each group was 19. We had the measurement vectors of PET data, \( x^Y_1, \ldots, x^Y_p \in \mathbb{R}^n \) for young group, denoted by \( Y \), and \( x^O_1, \ldots, x^O_p \in \mathbb{R}^n \) for old group, denoted by \( O (p = 90, n = 19) \).

2.2. Distance of brain network

The edge weight between two nodes \( i \) and \( t \) is estimated by the Gaussian kernel based on Pearson correlation:

\[
    w_{it} = k(x_i, x_t) = \exp \left( -\frac{1 - \text{corr}(x_i, x_t)}{\sigma_i \sigma_t} \right),
\]

where \( \text{corr}(x_i, x_t) \) is the Pearson correlation between two measurement vectors \( x_i \) and \( x_t \) and \( \sigma_i \) is the width of Gaussian kernel. Because \( 1 - \text{corr}(x_i, x_t) = \frac{1}{2} \| x_i/\|x_i\| - x_t/\|x_t\| \|^2 \), \( \sqrt{1 - \text{corr}(x_i, x_t)} \) is conditionally negative semi-definite for \( x_i, x_t \in \mathbb{R}^n \) [23]. The Gaussian kernel based on correlation in (1) is positive definite for all \( \sigma_i > 0 \) and satisfies Mercer’s theorem [23]. Thus, it transforms the original data in a nonlinear manifold into a higher dimensional feature space where the transformed features have a linear representation. The distance of
the kernel $w_{it}$ is estimated by a kernel trick [24]:

$$d_{it} = \| \phi(x_i) - \phi(x_t) \|$$

$$= \langle \phi(x_i) - \phi(x_t), \phi(x_i) - \phi(x_t) \rangle^{1/2}$$

$$= [k(x_i, x_i) + k(x_t, x_t) - 2k(x_i, x_t)]^{1/2}$$

$$= \sqrt{2 - 2k(x_i, x_t)}. \quad (2)$$

If an edge $e$ connects two nodes $i$ and $t$, $d_{it}$ is also denoted by $l(e)$.

The kernel-based distance is a Euclidean distance between two nodes in a higher dimensional feature space. When the kernel width is small in (1), the local neighbors that are highly positively correlated in the original data space are more clearly separated in the feature space, while non-local neighbors are not. The kernel width $\sigma_i$ in (1) is determined by the tenth smallest one among all $1 - corr(x_i, x_t) \ (t = 1, \ldots, i-1, i+1, p)$ [25].

The 38 brain graphs of 38 subjects were constructed from fMRI data by the kernel-based distance in (2) and (1). Two brain graphs of two groups, Y and O were constructed from PET data. We call the brain graphs constructed by fMRI and PET data functional and metabolic graphs, respectively.

2.3. Volume entropy

Suppose that $\mathcal{N} = \mathcal{N}(V, E)$ is a connected finite graph with a node set $V$ and edge set $E$. We will assume that $\mathcal{N}$ does not have any terminal node. We will be given a length $l(e)$ for each edge $e$, which determines a distance on $\mathcal{N}$. Let $S$ be a subset of edges with multiplicities, where each edge can be counted several times. The length of $S$ is defined by $l(S) = \sum_{e \in S} l(e)$. For example, $S = \{e, e, f\}$ for $e, f \in E$ is allowed and the volume of $S$ is $2l(e) + l(f)$. Edge $e$ is assumed to have an orientation from the initial node $i(e) \in V$ to the terminal node $t(e) \in V$. For a given oriented edge $e$ from $i(e)$ to $t(e)$, we denote by $\overline{e}$ the oriented edge from $i(\overline{e}) = t(e)$ to $t(\overline{e}) = i(e)$. Note that for any $e \in E$, both $e$ and $\overline{e}$ exists in $\mathcal{N}$.

The sequence of $n$ consecutive edges without backtracking is denoted by a path $\mathcal{P} = e_1e_2 \cdots e_n \ (e_{j+1} \neq \overline{e_j}, e_j \in E)$. The set of all possible metric paths...
Figure 1: (a) Toy example of a weighted graph that consists of 7 nodes and 12 edges. The degree of all nodes is more than three. (b) Distance matrix of (a). (c) Tree $B(v_0, r)$ starting from $v_0 = v_1$ with the radius $r = 6$. All nodes in (c) have the same connections to nodes in (a).

of length $r$ starting from a node $v_0 \in V$ in $\mathcal{N}$ has a structure of a tree, which we denote by $B(v_0, r)$. Because $\mathcal{N}$ is assumed to have no terminal node, the number of possible paths $B(v_0, r)$ increases exponentially as $r \to \infty$. The limit of the ball $B(v_0, r)$ as $r \to \infty$ is called the universal covering tree of $\mathcal{N}$.

**Example 2.1.** Fig. 1 shows the universal covering tree of toy example. The weighted graph and its distance matrix are given in (a) and (b), respectively. The tree $B(v_1, 6)$ of (a) is illustrated in (c). All nodes in the tree (c) have the same edges to nodes of the graph (a).

The volume entropy $h_{vol}$ is defined as [18]

$$h_{vol} = \lim_{r \to \infty} \frac{\log |B(v_0, r)|}{r}.$$  \hspace{1cm} (3)

The volume entropy $h_{vol}$ does not depend on $v_0$. When $h_{vol} > 0$, it is easy to
see that \( l(B(v_0, r)) \) is concentrated on the outer shell, i.e.
\[
l(B(v_0, r) - B(v_0, r - r_0)) = e^{h_{\text{vol}}r} - e^{h_{\text{vol}}(r-r_0)} = e^{h_{\text{vol}}r}(1 - e^{-r_0}).
\]
It follows that
\[
h_{\text{vol}} = \lim_{r \to \infty} \frac{\log l(B(v_0, r) - B(v_0, r - r_0))}{r}, \text{ for any } r_0 > 0.
\]
Note also that
\[
h_{\text{vol}} = \lim_{r \to \infty} \frac{\log N_r(v_0)}{r},
\]
where \( N_r(v_0) \) is the number of metric paths of length \( r \) starting from \( v_0 \), since
\[
r_0N_{r-r_0}(v_0) \leq l(B(v_0, r) - B(v_0, r - r_0)) \leq r_0N_r(v_0).
\]
In other words, the volume entropy \( h_{\text{vol}} \) is the exponential growth rate of the number of graph paths \( N_r(v_0) \) as \( r \to \infty \), i.e. as the information flows through a graph forever.

2.4. The generalized Markov system

Recall that in any given graph \( \mathcal{N}(V, E) \), for every edge \( e \), the inverse edge \( \overline{e} \) is also in \( \mathcal{N} \) and \( l(e) = l(\overline{e}) \). \( q \) is denoted as the number of edges in \( E \) including the inverse edges.

An edge transition matrix \( \mathbf{L}(h) = [L_{ef}(h)] \in \mathbb{R}^{q \times q} \) is defined by
\[
L_{ef}(h) = a_{ef} e^{-h_l(f)},
\]
where \( l(f) \) is the distance of an edge \( f \) and
\[
a_{ef} = \begin{cases} 
1 & \text{if } t(e) = i(f), i(e) \neq t(f) \\
0 & \text{otherwise}
\end{cases}.
\]
The generalized Markov system of \( \mathcal{N} \) associated to \( \mathbf{L}(h) \) is defined by
\[
z_t = \mathbf{L}(h)z_{t+1}, \quad (4)
\]
for \( h > 0, t > 0 \), and \( z_t \in \mathbb{R}^q \). When \( z = z_{t+1} = z_t \), (4) is called the stationary equation of the generalized Markov system.
Theorem 2.1 (Theorem 4 in [18]). Given a graph $N = (V, E)$, the volume entropy of $N$ is $h_{\text{vol}} = h$ satisfying that the generalized Markov system in (4) is stationary, i.e.,

$$z = L(h)z.$$  \hspace{1cm} (5)

$z$ in (5) is the eigenvector of $L(h_{\text{vol}})$ with the largest eigenvalue 1. If $z = [z_e]$ in (4), the sum of the squares of $z_e$ is equal to one, i.e., $\sum_e z_e^2 = 1$. Moreover, $z$ has all entries of the same sign according to the Perron-Frobenius theorem. Thus, we call $z_e^2 = [z_e^2]$ the stationary distribution of the generalized Markov system associated to the edge transition matrix $L(h)$. The generalized Markov system in (5) can be rewritten in a scalar form by

$$z_e = \sum_f a_{ef} e^{-h_{\text{vol}}l(f)} z_f.$$  \hspace{1cm} (6)

The equation (6) implies that when the information flows from an edge $e$ to $f$, the amount of information increases exponentially with the growth rate $h_{\text{vol}}l(f)$. For convenience, we reshape the vector $z_e^2$ into a matrix, $\Pi = [\pi_{it} = z_e^2] \in \mathbb{R}^{p \times p}$, where an edge $e$ has the initial node $i(e) = i$ and the terminal node $t(e) = t$ and $\pi_{it} = 0$ ($i = 1, \ldots, p$). Since $\Pi$ is an asymmetric matrix, $\Pi$ implies a directed graph that represents an information flow on a graph $N$. $\pi_{it}$ shows the information capacity of an edge from a node $i$ to $t$. We call it an edge capacity. In the stationary equation (6), $z_e(= \sqrt{\pi_{it}})$ is affected by the distance $l(f)$ for all edges $f$s connected with the terminal node $t$. Thus, $\pi_{it}$ is different from $\pi_{ti}$, and their difference is related to the imbalance of the connectivities of two nodes $i$ and $t$. We define a node capacity by the difference between the inward and outward information of a node, i.e., $\pi_i = \sum_t \pi_{ti} - \sum_t \pi_{it}$. If a node capacity is negative/positive, the amount of outgoing information is larger/smaller than that of incoming information.

Example 2.2. Given the weighted graph in the toy example 2.1, Fig. 2 shows (b) its edge transition matrix $L(h_{\text{vol}})$, (c) eigenvector $z$, (d) edge capacity matrix $\Pi$, and (e) the induced directed network. The example of the graph in Fig. 1
(a) has 7 nodes and 12 edges. To estimate volume entropy, each undirected edge is assumed to consist of bidirectional edges with the same distance. The number of oriented edges is \( q = 12 \cdot 2 = 24 \). Thus, \( L(h_{vol}) \) in Fig. 2 (b) is a \( 24 \times 24 \) dimensional sparse matrix, and its eigenvector is \( z \in \mathbb{R}^{24} \) in (c). In the induced directed network in Fig. 2 (e), the line width of edge is proportional to the edge capacity in the edge capacity matrix in (d). The size of node is proportional to the absolute of node capacity. In the original graph in Fig. 1 (a), the node sets \{1, 2, 3, 4\} and \{5, 6, 7\} form a clique, respectively called A and B for convenience. The module size, i.e., the number of the nodes of A is larger than that of B. Thus, more information flows within A than within B, and more information flows from B to A through three edges directed from the node 5 to 3, from 7 to 4, and from 6 to 4. In the induced directed network in Fig. 2 (e), while the bidirectional edges within A or B have similar capacity, the edges from A to B are much different from that from B to A.

3. Results

3.1. Simulations

3.1.1. Artificial unweighted graphs with different topology

We compared the performance of five global graph invariants,

- volume entropy \( h_{vol} \) (the proposed method),
- spectral entropy \( h_{spe} \) [16],
- functional entropy \( h_{fun} \) [15],
- global efficiency \( e_{glo} \), and
- average local efficiency \( e_{loc} \),

that distinguish five artificial unweighted graphs with different topologies,

- regular graph (RE),
- small-world graph (SW),
- random graph (RA),
- scale-free graph (SF), and
- hyperbolic graph (HY).
Figure 2: (a) Distance matrix of the graph in Fig. 1. (b) Edge transition matrix $L(h_{vol}) \in \mathbb{R}^{24 \times 24}$. If edges $e$ and $f$ are consecutive, the $(e,f)$th entry is $e^{-h_{vol}(f)}$. (c) Eigenvector $z = [z_i] \in \mathbb{R}^{24}$. (e) Edge capacity matrix $\Pi = [\pi_{ij} = z_i^2]$, where an edge $e$ connects between a node $i$ and $j$. (d) Directed information flow network. The line width is proportional to the entry of $\Pi = [\pi_{ij}]$. The size of the node $i$ is proportional to the absolute of node capacity, $|\pi_i = \sum_j \pi_{ji} - \sum_j \pi_{ij}|$.

RE is an unweighted graph where all nodes have the same degree. SW is a globally and locally efficient graph with short characteristic path length and large average clustering coefficient [26]. SF has heterogeneous degree distribution with a few number of heavily linked nodes, termed hubs, but many nodes with few connections [27]. Hubs make a great contribution to spreading information quickly throughout a network. On the other hand, it is vulnerable to targeted attacks on hubs. Thus, SF is known to be globally efficient and locally inefficient. HY is known to have both strong heterogeneity and high clustering coefficient [28]. Therefore, it may be a maximally efficient unweighted graph.
The five types of unweighted graphs were generated by CNM matlab toolbox [29]. The number of nodes was fixed by $p = 90$. The sparsity, which was the ratio of the number of edges to the number of maximally possible edges, was varied from 0.04 to 0.90. All nodes in a graph should have more than three edges for the estimation of volume entropy [18]. If there were nodes with degree less than three in the generated graph, we randomly took an edge connecting nodes with degree more than four and rewired it to a node with degree less than three. In this way, we generated 150 artificial graphs for each sparsity and each graph type. After five invariants were estimated in each graph, Wilcoxon rank sum test was performed to assess the statistical difference between graph types at each sparsity. We used brain connectivity toolbox for the estimation of global and local efficiencies [13].

Fig. 3 showed the results of (a) $h_{vol}$, (b) $h_{fun}$, (c) $h_{spe}$, (d) $e_{loc}$, and (e) $e_{glo}$. In each figure, large panel showed the change of graph invariant with respect to sparsity. The line color represented the type of graph: blue for RE, green for SW, red for RA, cyan for SF, and magenta for HY. When we sorted the graph types in ascending order of graph invariant, the order of the graph types was changed according to sparsity. Whenever the order of the graph types was changed over sparsity, we drew the box plot of graph types in ascending order of graph invariant in a small panel. Thus, the number of small panels in (a-e) was the same as the number of changes in the order of graph types over sparsity. If there was no change in the order of graph types over sparsity, the corresponding graph invariant would consistently distinguish network topology.

When the property of graph were measured by $e_{loc}$ in Fig. 3 (d), the order of graph types was changed four times between the sparsity 0.04 and 0.90. It was (1) $\text{RA} < \text{SF} < \text{HY} < \text{SW} < \text{RE}$, (2) $\text{RA} < \text{SF} < \text{SW} \leq \text{HY} < \text{RE}$, (3) $\text{RA} < \text{SF} < \text{HY} \leq \text{SW} < \text{RA}$, and (4) $\text{RA} < \text{SW} < \text{SF} < \text{HY} < \text{RE}$ ($p < .001$). The order of SW and HY was not consistent over the sparsity. Thus, it seemed that the local efficiency did not distinguish between SW and HY well. In Fig. 3 (e), the order of graph types in $e_{glo}$ was changed four times: (1) $\text{RE} < \text{SW} < \text{RA} < \text{SF} < \text{HY}$, (2) $\text{RE} < \text{SW} < \text{SF} < \text{RA} < \text{HY}$, (3)
Both global and local efficiencies distinguished RE, SW, and RA well regardless of the sparsity. However, they did not distinguish SF, HY, and SW.

Fig. 3 (c) showed that the order of graph types in $h_{spe}$ was changed six times. It was (1) HY < RE < SF < RA < SW, (2) RE < HY < SF < RA < SW, (3) RE < HY < SF < SW < RA, (4) RE < HY < SW < SF < RA, (5) RE < SW < HY < SF < RA, and (6) RE < SW < SF < HY < RA ($p < .001$).

The order of network types in $h_{spe}$ highly depended on the sparsity. The spectral entropy $h_{spe}$ measured the connectedness of a graph. Because all nodes in RE had the same degree, RE did not have a modular structure and it always had the smallest $h_{spe}$ among all five graph types.

If the sparsity was the same, the functional entropy of five network types was always the same as shown in Fig. 3 (b). It was because the functional entropy consider only the distribution of edge weights, not the network topology. The functional entropy did not monotonically increase or decrease over sparsity. Thus, it could also not distinguish the difference in the sparsity of graph. When the volume entropy was applied to artificial unweighted graphs, the order of graph types was consistent for all sparsity. The order was RE < SW < RA < SF < HY ($p < .001$, FDR-corrected). The volume entropy of SF was similar to that of HY at large sparsity. It was because that as the number of edges increased in a graph, SF lost its sparse property. Unlike $e_{loc}$ and $e_{glo}$, the volume entropy also distinguished well between SW, SF, and HY.

### 3.1.2. Artificial weighted graphs with different geometry

In this section, we compared the graph invariants in discriminating three types of artificial weighted graphs. We generated 150 hyperbolic unweighted graphs using CNM toolbox, and defined the edge distance of the graph in three different ways,

- **U**: all edges had the same distance,
- **W**: the edge distance was proportional to the degree of its initial and
terminal nodes, \( i(e) \) and \( t(e) \), determined by

\[
l(e) = \frac{\log(k_{i(e)} - 1) + \log(k_{t(e)} - 1)}{\sum_{v \in V} k_v \log(k_v - 1)},
\]

where \( k_v \) is the number of edges connecting with a node \( v \in V \) [18].

- \( S \): the edge distance was inversely proportional to the degree of two connected nodes, determined by the inverse of \( l(e) \) in (7).

These three networks had the same topology, but different geometries. The edge connected to a node with a higher degree was longer in \( W \), but shorter in \( S \). Thus, it could be assumed that the information propagation was the fastest in \( S \), followed by \( U \) and \( W \). Before estimating the graph invariants, we normalized the volume of weighted graph to two, i.e., \( \sum_e l(e) = 2 \).

The results of network invariants were shown in Fig. 4. Three networks, \( S \), \( U \), and \( W \) were represented by red, blue, and green, respectively. The order of graph types, \( S \), \( U \), and \( W \) was consistent for all sparsity in \( h_{vol} \), \( h_{spe} \), and \( e_{loc} \).

The order was \( W < U < S \) in \( h_{vol} \) in Fig. 4 (a), \( S < U < W \) in \( h_{spe} \) in (c), and \( U < S < W \) in \( e_{loc} \) in (d) \((p < .001, \text{FDR-corrected})\). The result of \( h_{vol} \) was as we expected, but the result of spectral entropy was the reverse order. The results of \( h_{fun} \) and \( e_{glo} \) in Fig. 4 (b) and (e) were not consistent. The order of graph types in \( h_{fun} \) was changed twice: (1) \( U < S < W \), and (2) \( U < W < S \) \((p < .001)\). The functional entropy could not find the difference between \( S \) and \( W \). The order of graph types in \( e_{glo} \) was changed three times: (1) \( W < U < S \), (2) \( U < W < S \), and (3) \( U < S < W \) \((p < .001)\).

3.1.3. Artificial modular graphs and the stationary distribution of information flow

In this simulation, we constructed artificial modular graphs with two modules, and observed their directed networks induced by the stationary distribution of the generalized Markov system. The nodes of two modules were generated by two bivariate Gaussian distributions with mean \([-5, 0] \) and \([5, 0]\). The variance of the distributions was varied by 0.01, 0.1, 1, 10, and 100. The larger the variance was, the closer the two modules were. The total number of nodes
was $p = 100$. The ratio of node numbers in two modules was changed by $5:5, 4:6, 3:7, 2:8, 1:9$, and $0:10$. There was no module in a graph at the ratio $0:10$. After generating the 100 nodes of modular network at each variance and each ratio, the edge distance was estimated by Euclidean distance between any two nodes. Fig. 5 showed the example of generated modular networks. The ratio of node numbers was varied by $5:5, 4:6, 3:7, 2:8, 1:9$, and $0:10$ from left to right columns. The variance of Gaussian distribution was varied by $0.01, 0.1, 1, 10$, and $100$ from top to bottom rows. In each panel, two modules had two different colors: blue on the left and red on the right.

After constructing 100 artificial modular graphs for each variance and each ratio, we obtained the directed information flow network by using the stationary distribution of the generalized Markov system. We examined the direction of edges with the top 5 percent of edge capacities, i.e., $p \cdot (p - 1) \cdot 0.05 = 495$ edges with the largest edge capacity in each directed information flow network. Table 1 showed the mean and standard deviation of the ratio of edges from the left module to the right module among the top 5 percent of edge capacities. When the two modules had similar size at the ratio $5:5$, the direction of edges between two modules was not consistent in the first column of Table 1. However, when the number of nodes in two modules was 40 and 60, about 83% of the 495 edges had the orientation from the small-sized module to the large-sized module in the second column of the table. This percentage increased when the ratio was varied from $3:7$ to $1:9$ in Table 1. The modular network with the variance $100$ had the smallest percentage of edges from the small-sized module to the large-sized module at the fixed ratio between the sizes of two modules. It was because the two modules were too close to be discriminated.

Each panel in Fig. 5 showed the edges with the top 5 percent of edge capacities in the directed information flow network. The direction of edge was represented by the color. If the color of edge was blue, it was directed from the red nodes on the right to the blue nodes on the left. If the color of edge was red, the direction was opposite. The color saturation depended on the value of edge capacity as shown in the right colorbar. When the number of blue nodes on
the left was smaller than that of red nodes on the right, the capacity of edges from left to right modules was larger than that in the opposite direction. In this simulation, the nodes in the center of the larger module had large capacity. It meant that the edges with large capacity were directed to the nodes in the center of large module. The results of volume entropy, modularity, and global and local efficiencies over the variance and ratio between module sizes were shown in the supplementary material.

Table 1: Mean and standard deviation of the percentage of edges from the small-sized module to the large-sized module among the edges with the top 5 percent of edge capacities. The ratio of edges from left to right modules was obtained from 100 modular graphs at each ratio and variance. The row and column represent the variance of Gaussian distribution and the ratio between module sizes, respectively.

| var. | 5 : 5 | 4 : 6 | 3 : 7 | 2 : 8 | 1 : 9 |
|------|-------|-------|-------|-------|-------|
| 0.01 | 0.40 ± 0.49 | 0.82 ± 0.39 | 0.94 ± 0.24 | 1.00 ± 0 | 1.00 ± 0 |
| 0.1  | 0.48 ± 0.50 | 0.86 ± 0.35 | 0.94 ± 0.24 | 1.00 ± 0 | 1.00 ± 0 |
| 1    | 0.52 ± 0.50 | 0.86 ± 0.35 | 1.00 ± 0    | 1.00 ± 0 | 1.00 ± 0 |
| 10   | 0.66 ± 0.48 | 0.82 ± 0.39 | 1.00 ± 0    | 1.00 ± 0 | 1.00 ± 0 |
| 100  | 0.50 ± 0.51 | 0.58 ± 0.50 | 0.70 ± 0.46 | 0.88 ± 0.33 | 0.92 ± 0.27 |

3.2. Clinical dataset: resting state fMRI and PET

Before estimating the volume entropy, we normalized the volume of graph to two, i.e., \( \sum_{e} l(e) = 2 \). The volume entropy of 38 functional graphs was plotted with respect to age in Fig. 6 (a). The volume entropy and age were negatively correlated (\( p < .005 \)). The volume entropy of Y and O in metabolic graphs was shown in Fig. 6 (b) by green marker ‘X’. In the metabolic graph analysis, we performed 5000 permutations of Y and O to enable the assessment of statistical differences between two groups. If we called a graph constructed by permutation a null graph, the box plot in Fig. 6 (b) showed the volume entropy of 5000 null graphs. The volume entropy of O was significantly smaller than that of null
graphs, but the volume entropy of Y was not \((p < .05)\). The difference between Y and O was not significant, but showed the tendency that the volume entropy of Y was larger than that of O \((p < .13)\). The results of both functional and metabolic graphs showed that the volume entropy decreased with normal aging.

3.3. Information flow on a metabolic graph

Fig. 7 showed the edge capacity matrix and directed information flow network in the metabolic graph of Y and O. The edge capacity matrices of Y and O were shown in Fig. 7 (a) and (c), respectively. The obtained directed information flow networks of Y and O were shown in (b) and (d), respectively. In the edge capacity matrix, the first 45 rows and columns were the nodes in right hemisphere, and the last 45 rows were in left hemisphere. The nodes were sorted in the order of the frontal (F), limbic (L), parietal (P), temporal (T), basal ganglia (B), limbic (L), and occipital (O) lobes (more details in the supplementary material). The \(p_{i,t}q_{th}\) entry of the edge capacity matrix was the edge capacity directed from the node \(i\) to \(t\). As the edge capacity decreased, the color of entry was changed from dark red through yellow to white as shown in the right colorbar. The column of the edge capacity matrix had similar color. It meant that the edges with the same terminal node had similar information capacity.

In Fig. 7 (b) and (d), we plotted only edges with the top 5 percent of edge capacity in the directed information flow graphs of Y and O. The left and right panels showed the brain graph in the left and right views, respectively. In the information flow graph of Y in (b), the edges were mainly directed to the medial orbital part of superior frontal gyrus (SFGmorb) in the right hemisphere, bilateral putamen (PUT), left dorsolateral superior frontal gyrus (SFG), and left gyrus rectus (REG). In the information flow graph of O in (d), the edges were mainly directed to bilateral SFGmorb, right thalamus (THA), right posterior cingulate cortex (PCC), and left middle occipital gyrus (MOG). The color of node represents the location of node: red and orange in F, green in P, blue in T, purple in O, yellow in L, and yellow-green in B (more details in the supplementary material).
supplementary material). The size of node was proportional to the absolute of node capacity. We determined the color of edge by the color of its terminal node.

We performed 5000 permutations and Wilcoxon rank sum test for finding the difference between the edge capacities of Y and O. There was no significant edges that had larger capacity in Y than in O. On the other hand, the information capacity of Y< O were found in the connections from the most of brain regions to left angular gyrus (ANG) ($p < .05$, FDR-corrected). The node capacity of left ANG was also larger in O than in Y ($p < .05$, FDR-corrected).

3.4. Information flow on a functional graph

The directed information flow graphs of 38 subjects in the resting state fMRI were shown in the supplementary material. We estimated edge capacities that were significantly correlated with age. The negative correlation with age were found in the edges directed from the most of brain regions to right PUT and pallidum (PAL), and left THA ($p < .05$, FDR-corrected). The node capacity of right PUT and PAL, and left THA also decreased with age as shown in Fig. 8 ($p < .05$, FDR-corrected). The edge capacities to left PUT and PAL, and right THA and their node capacities also tended to be negatively correlated with age ($p < .05$, uncorrected).

The positive correlation with age were found in the bidirectional edges between left and right median cingulate cortex (MCC) and the edge from left superior temporal gyrus (STG) to right STG as shown in Fig. 8 ($p < .05$, FDR-corrected). We also estimated a quadratic relationship between the edge capacity and age. The capacity of the most of edges directed to right anterior cingulate cortex (ACC) had a U-shaped curve with respect to age ($p < .05$, FDR-corrected). It decreased to around 45 years of age and increased at older age. The node capacity of right ACC also had a U-shaped curve with respect to age. The minimum node capacity of right ACC was also found at around 45 years of age as shown in Fig. 8 ($p < .05$, FDR-corrected).
3.5. Global and local efficiencies, modularity, and age

To see the relationship between the volume entropy and the existing complex graph invariants, we also estimated the global and local efficiencies, and modularity. The global and average local efficiency highly depended on the volume of graphs. When we estimated the global and average local efficiency without the normalization, both of them significantly increased with age in the resting state fMRI ($p < .05$). However, when we estimated them after the normalization of graph volume, the global efficiency tended to decrease with age ($p = 0.085$), but the local efficiency tended to increase with age ($p = 0.070$). The volume of 38 functional graphs decreased with age ($p < .05$). The modularity decreased with age in the resting state fMRI regardless of the normalization of graph volume ($p < .05$). There was no node that was significantly related to age in node strength and local efficiency in functional graphs. In the metabolic graph of Y and O, there was no difference in the volume, global and local efficiencies, and modularity.

4. Discussion

4.1. Relationship between volume entropy and complex graph measures

To better understand the volume entropy, we discuss the relationship between the volume entropy and the existing complex graph invariants such as modularity, global and local efficiencies, and hubs. Firstly, the volume entropy was large when there were many edges in a graph. The simulation in Sec. 3.1.1 and 3.1.2 showed that the change of sparsity affected the volume entropy more than the change of network topology and geometry. These results were found not only in the volume entropy but also in the other graph invariants such as global and local efficiencies and spectral and functional entropies. Because we used fully connected weighted graphs in the clinical applications, there was no effect on the volume entropy from the difference of sparsity.

Secondly, the volume entropy was more related to global efficiency than local efficiency. When the volume entropy was applied to the graphs with different
topology in Sec. 3.1.1, the order of graphs was RE < SW < RA for all sparsities. The global efficiency was proportional to characteristic path length, while the local efficiency was inversely proportional to clustering coefficient [13]. According to the Watts-Strogatz model of the small world, the characteristic path length and the average clustering coefficient were the smallest in RE, followed by SW and RA [30]. Our results in Sec. 3.1.1 also showed that the global efficiency was the smallest in RE, followed by SW and RA, while the local efficiency was the opposite. If a graph had high clustering coefficient, but short characteristic path length, the information would not be spread throughout the graph because the information would turn around only in nodes with strong clustering coefficients. That might be the reason why the volume entropy of SW was smaller than that of RA.

Thirdly, the volume entropy was large when a graph had hubs. SF and HY had larger volume entropy than RE, SW, and RA in Sec. 3.1.1. SF and HY were a graph with hubs that played a decisive role in the exponential growth of the network path through which information was delivered [27, 28]. Fourthly, the volume entropy was not completely independent of the clustering coefficient, i.e., local efficiency of a graph. The volume entropy of HY was larger than that of SF in Sec. 3.1.1. HY was known as a network with high clustering coefficient and heterogeneous degree distribution, while SF had only heterogeneous degree distribution [28]. The high clustering coefficient might play a role in holding the information to highly clustered nodes, however, but also in creating many number of graph paths within them. Thus, if the paths out of the clustered nodes were appropriately created as in HY, the high local efficiency could also contribute to fast information propagation. In this sense, the volume entropy may be the first global invariant to measure the efficiency of hyperbolic graph.

Finally, the volume entropy was also related to the modular structure of network as shown in Sec. 3.1.3. There were high clustering coefficient and short path length within a module, but not between modules in a modular graph. Since the global invariants of clustering coefficient and shortest path length were estimated in an average manner, they were not proper to represent
the heterogeneous clustering coefficient and shortest path length of modular graph. However, the volume entropy was calculated by the fastest growth rate of network paths, and not affected by the heterogeneous property of graph. The smaller the difference between module sizes of two modules in artificial modular graph was in Sec. 3.1.3, the larger the volume entropy and modularity were, but the smaller the global and local efficiencies were.

4.2. Normalization of graph volume

Like complex graph measures in Sec. 3.5, the volume entropy is influenced by the volume of graph. If we denote the sum of all edges in $\mathcal{N}(V, E)$ as $\text{vol}(E)$ and the unnormalized edge distance of $f$ as $\tilde{l}(f)$, the normalized distance $l(f)$ in (6) is obtained by $l(f) = \frac{2}{\text{vol}(E)} \tilde{l}(f)$. The stationary equation (6) is rewritten by

$$z_e = \sum_f a_{ef} e^{-\frac{2}{\text{vol}(E)} \tilde{h}_{\text{vol}} l(f)} z_f.$$  

Then, the volume entropy of unnormalized graph is obtained by

$$\tilde{h}_{\text{vol}} = \frac{2}{\text{vol}(E)} h_{\text{vol}}.$$  

The stationary distribution $z_e^2$ does not depend on the normalization of graph.

After the normalization of volume, the volume entropy significantly decreased with age. However, the volume entropy of unnormalized functional graphs had no relationship with age because the volume of functional graphs decreased with age in Sec. 3.5. The decline of brain graph volume with age might mean that the connection between any brain regions was generally short, and the decrease in the volume entropy of normalized brain graph might mean that the inherent topological structure of the brain graph became increasingly inefficient. Since the volume entropy of unnormalized functional graph had no relationship with age, it could be interpreted that the connections in the functional brain graph became shorter, i.e., the correlations between brain regions became stronger with age in order to compensate the inefficient topological change of brain graph across the lifespan.
4.3. Comparison of the results with previous studies

Previous studies on resting-state functional connectivity have shown somewhat inconsistent change of global and local efficiencies across the lifespan [31, 9, 32, 33]. The human brain has known to have a modular architecture [34, 35, 36]. As discussed in Sec. 4.1, the modular network had heterogeneous shortest path lengths and clustering coefficients. Thus, it was possible that the global and average local efficiency of modular brain network has not been calculated exactly. That might result in the inconsistent change of global and local efficiencies.

On the other hand, there were consistent reports of the age-related reorganization in the modular structure of functional connectivity [31, 37, 38]. Especially, they have consistently shown that the modularity decreased after 40 years of age [31, 37, 38]. The results of our resting state fMRI data also showed the age-related decline of modularity in Sec. 3.5. The volume entropy also significantly decreased with age. The age-related change in modularity might be related to the age-related inefficient topological change, which was well quantified by the volume entropy.

4.4. Information flow on a metabolic graph

The stationary distribution of the generalized Markov system in (5) is the edge capacity of information flow when the total amount of information capacity in a brain graph is fixed to one. Therefore, the increase or decrease of the edge capacity with age should be interpreted as the change of the relative proportion of edge capacity in the whole brain, not the change of absolute value.

The result in Sec. 3.3 showed that the role of left ANG became more important in the information propagation with age in a metabolic graph. The larger the size of module was, the larger the information capacity was in the proposed generalized Markov system. The information flowed from the small-sized module to the large-sized module in Sec. 3.1.3. Especially, the edges directed to nodes at the center of the large-sized module had larger information capacity. Thus, it could be assumed that the size of the module including left ANG was
larger in O than in Y, and the amount of information coming into the module of left ANG would also increased. In addition, more information would flow into ANG which was known as the functional hub of DMN [39]. The reason why only the left ANG had large information capacity might be because the left hemisphere had less age-related decline than the right hemisphere [40].

4.5. Information flow on a functional graph

The functional graph had a topological structure where the information propagation slowed down along with age. At the same time, the contributions of PUT, PAL, and THA to information propagation decreased with age. Previous studies consistently indicated that the circuit linking PUT, PAL, THA, and cortical areas played a key role in motor ability across the human lifespan [41, 42]. The functional and structural alterations in the basal ganglia-thalamocortical circuits have been found in the progression of Alzheimer’s disease and Parkinson’s disease as well as normal aging [43, 44, 45, 42].

ACC has been known as a key area involved in cognitive and emotional processing [46, 47]. Previous study on resting-state fMRI showed that the decreased functional connectivity between ACC and default mode network would be associated with the deficit of cognitive processing in aging, while the increased functional connectivity between ACC and the emotion-related brain regions such as STG, inferior frontal gyrus (IFG), PUT, and amygdala (AMYG) would be associated with the well-maintained emotional well-being in aging [47]. On the other hand, our result showed that the role of right ACC in information propagation decreased until around 45 years of age, but increased at the older age. This result was somewhat different from the previous studies, and its biological meaning needs further discussion in the future.

The information capacities of bidirectional edges between right and left MCCs had a linear relationship with age. In our results, the edge capacity between bilateral brain regions tended to be slightly smaller than the other edge capacities. This might be because the bilateral brain regions were highly correlated and likely to be in the same module. If the node capacity of the
bilateral MCCs had significantly increased with age, it could be interpreted that the role of the bilateral MCCs became increasingly important with age. However, since only the edge capacity between bilateral MCCs increased with age, we assumed that the bilateral MCCs consistently belonged to a module, and that the role of the module itself became increasingly important. MCC has been known to be related to environmental monitoring and response selection[48, 49]. Therefore, it could be speculated that there was the age-related change in social decision-making of human [50].

The edge capacity from left to right STGs also increased with age. In the clinical applications, the node capacity of only right STG tended to increase with age, while that of left STG was not changed (p < .05, uncorrected). From the result, we inferred that while the contribution of right STG slightly increased, but left STG did not. STG has been known to be involved in language processing, multisensory integration, and social perception [51, 52]. Especially, the dysfunction of right STG has been found to be related with the social cognition deficit in normal aging [53].

4.6. Limitations and conclusions

In our study, we introduced a new network invariant, called a volume entropy. It measured the fastest growth rate of network paths through which the information was propagated over a brain. The larger the volume entropy was, the more information was propagated in a graph. Thus, it could be regarded as a new graph invariant of efficiency in terms of the information propagation. The simulation results showed that the volume entropy was proper to measure the efficiency of graph with heterogeneous property such as modular and hyperbolic graphs. The information flow in a graph was modelled by the generalized Markov system associated with a newly defined edge transition matrix. The volume entropy was estimated by the stationary equation of the generalized Markov system. At the same time, we could obtain the stationary distribution of information flow in a graph. It provided a new insight of how much and in what direction the information flowed on a brain.
However, the information capacity and direction of edge highly depended on the terminal node of edge. Thus, the node capacity, which was the difference between the nodes’ incoming and outgoing information, was sometimes enough to represent the directed information flow network induced by the stationary distribution. Another disadvantage was that the direction and capacity of edge were difficult to interpret their biological meaning. If we mathematically prove the relationship between the proposed method and the existing complex network measures, it can be easier to interpret the biological meaning of the stationary distribution. In the clinical applications, the significance in the difference between Y and O in metabolic graphs was rarely found due to the small number of subjects. In addition, we could not exploit the advantage of simultaneously acquired PET and fMRI data. The tendency of the volume entropy to decrease with age was similar for both two modalities, however, the local changes in the directed information flow network were quite different. Thus, we need to improve the proposed method to enable multi-modal network analysis in the future. The proposed method can be applied not only to the brain imaging data of normal control but also to that of various disease groups. It can also be applied to effective functional connectivity of which connection represents the causal relationship between brain regions.

5. Acknowledgement

This work is supported by Basic Science Research Program through the National Research Foundation (NRF) (No.2013R1A1A2064593 and No.2016R1D1A1B03935463), NRF Grant funded by MSIP of Korea (No.2015M3C7A1028926 and No.2017M3C7A1048079), and NRF grant funded by the Korean Government (No. 2016R1D1A1A02937497, No.2017R1A5A1015626, and No.2011-0030815).

References

[1] O. Sporns, Networks of the Brain, The MIT Press, 2010.
[2] M. R. Brier, J. B. Thomas, A. Fagan, J. Hassenstab, D. Holtzman, T. L. Benzinger, J. C. Morris, B. M. Ances, Functional connectivity and graph theory in preclinical Alzheimer’s disease, Neurobiology of Aging 35 (2013) 757–768.

[3] E. J. Sanz-Arigita, M. M. Schoonheim, J. S. Damoiseaux, S. A. R. B. Rombouts, E. Maris, F. Barkhof, P. Scheltens, C. J. Stam, Loss of ‘small-world’ networks in Alzheimer’s disease: Graph analysis of fMRI resting-state functional connectivity, PLoS ONE 5 (2010) e13788.

[4] P. J. Toussaint, V. Perlberg, P. Bellec, S. Desarnaud, L. Lacomblez, J. Doyon, M. O. Habert, H. Benali, for the ADNI, Resting state FDG-PET functional connectivity as an early biomarker of Alzheimer’s disease using conjoint univariate and independent component analyses, NeuroImage 63 (2012) 936–946.

[5] M. P. Van den Heuvel, C. J. Stam, M. Boersma, H. E. Hulshoff Pol, Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain, NeuroImage 43 (2008) 528–539.

[6] K. Wang, M. Liang, L. Wang, L. Tian, X. Zhang, K. Li, T. Jiang, Altered functional connectivity in early Alzheimer’s disease: a resting-state fMRI study, Human Brain Mapping 28 (2007) 967–978.

[7] C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, J. Kurths, Hierarchical organization unveiled by functional connectivity in complex brain networks, Phys. Rev. Lett. 97 (2006) 238103.

[8] M. Daianu, N. Jahaneshad, T. M. Nir, A. W. Toga, C. R. Jack Jr., M. W. Weiner, P. M. Thompson, Breakdown of brain connectivity between normal aging and Alzheimer’s disease: A structural k-core network analysis, Brain Connectivity 3 (4) (2013) 407–422.

[9] L. K. Ferreira, G. F. Busatto, Resting-state functional connectivity in nor-
MAL brain aging, Neuroscience & Biobehavioral Reviews 37 (3) (2013) 384 – 400.

[10] M. D. Greicius, G. Srivastava, A. L. Reiss, V. Menon, Default-mode network activity distinguishes Alzheimer’s disease from healthy aging: Evidence from functional MRI, Proceedings of the National Academy of Sciences of the United States of America 101 (13) (2004) 4637–4642.

[11] E. Bullmore, O. Sporns, The economy of brain network organization, Nature Reviews.Neuroscience 13 (5) (2012) 336–49.

[12] M. P. Van den Heuvel, C. J. Stam, R. S. Kahn, H. E. Hulshoff Pol, Efficiency of functional brain networks and intellectual performance, J. Neurosci. 29 (2009) 7619–7624.

[13] M. Rubinov, O. Sporns, Complex network measures of brain connectivity: Uses and interpretations, NeuroImage 52 (2010) 1059–1069. URL http://www.brain-connectivity-toolbox.net

[14] T. M. Cover, J. A. Thomas, Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing), Wiley-Interscience, 2006.

[15] Y. Yao, W. L. Lu, B. Xu, C. B. Li, C. P. Lin, D. Waxman, J. F. Feng, The increase of the functional entropy of the human brain with age, Scientific Reports 3 (2013) 2853.

[16] J. R. Sato, D. Y. Takahashi, M. Q. Hoexter, K. B. Massirer, A. Fujita, Measuring network’s entropy in ADHD: A new approach to investigate neuropsychiatric disorders, NeuroImage 77 (2013) 44–51.

[17] R. Frigg, C. Werndl, Entropy-A Guide for the Perplexed, Oxford University Press, 2011.

[18] S. Lim, Minimal volume entropy on graphs, Trans. Amer. Math. Soc. 360 (2008) 5089–5100.
[19] R. W. Cox, AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages, Computers and Biomedical Research 29 (1996) 162–173.

[20] S. M. Smith, M. Jenkinson, M. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobajak, D. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. M. Brady, P. M. Matthews, Advances in functional and structural MR image analysis and implementation as FSL, NeuroImage 23 (2004) 208–219.

[21] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, NeuroImage 15 (2002) 273–289.

[22] M. Quarantelli, K. Berkouk, A. Prinster, B. Landeau, C. Svarer, L. Balkay, B. Alfano, A. Brunetti, J. C. Baron, M. Salvatore, Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction, J Nucl Med 45 (2004) 192–201.

[23] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, M. Harandi, Kernel methods on riemannian manifolds with gaussian RBF kernels, IEEE Trans. Pattern Anal. Mach. Intell. 37 (12) (2015) 2464–2477.

[24] B. Schölkopf, The kernel trick for distances, in: Advances in Neural Information Processing Systems 13, MIT Press, 2001, pp. 301–307.

[25] L. Zelnik-manor, P. Perona, Self-tuning spectral clustering, in: Advances in Neural Information Processing Systems 17, MIT Press, 2004, pp. 1601–1608.

[26] D. S. Bassett, Small-world brain networks, Neuroscientist 12 (2006) 512–523.

[27] V. M. Eguiluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, A. V. Apkarian, Scale-free brain functional networks, Phys. Rev. Lett. 94 (2005) 018102.
[28] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, M. Boguñá, Hyperbolic geometry of complex networks, Phys. Rev. E 82 (2010) 036106.

[29] G. Alanis-Lobato, CNM: A a matlab toolbox for the construction of artificial complex networks, https://kr.mathworks.com/matlabcentral/fileexchange/45734-cnm/, [Online; accessed 03-Mar-2014] (2014).

[30] D. J. Watts, S. H. Strogatz, Collective dynamics of small-world networks, Nature 393 (1998) 440–442.

[31] M. Cao, J.-H. Wang, Z.-J. Dai, X.-Y. Cao, L.-L. Jiang, F.-M. Fan, X.-W. Song, M.-R. Xia, N. Shu, Q. Dong, M. P. Milham, F. X. Castellanos, X.-N. Zuo, Y. He, Topological organization of the human brain functional connectome across the lifespan, Developmental Cognitive Neuroscience 7 (Supplement C) (2014) 76 – 93. doi:https://doi.org/10.1016/j.dcn.2013.11.004.
URL http://www.sciencedirect.com/science/article/pii/S1878929313000960

[32] R. Sala-Llonch, D. Bartrés-Faz, C. Junqué, Reorganization of brain networks in aging: a review of functional connectivity studies, Frontiers in Psychology 6 (2015) 663.

[33] X.-N. Zuo, Y. He, R. F. Betzel, S. Colcombe, O. Sporns, M. P. Milham, Human connectomics across the life span, Trends in Cognitive Sciences 21 (1) (2017) 32–45.

[34] Z. J. Chen, Y. He, P. Rosa-Neto, J. Germann, A. C. Evans, Revealing modular architecture of human brain structural networks by using cortical thickness from MRI, Cereb. Cortex 18 (2008) 2374–2381.

[35] O. Sporns, R. F. Betzel, Modular brain networks, Annual Review of Psychology 67 (2016) 19.1–19.28.
[36] M. Valencia, M. A. Pastor, M. A. Fernandez-Seara, J. Artieda, J. Martinerie1, M. Chavez, Complex modular structure of large-scale brain networks, Chaos 19 (2009) 023119.

[37] L. Geerligs, R. J. Renken, E. Saliasi, N. M. Maurits, M. M. Lorist, A brain-wide study of age-related changes in functional connectivity, Cerebral Cortex 25 (7) (2015) 1987–1999.

[38] J. Song, R. M. Birn, M. Boly, T. B. Meier, V. A. Nair, M. E. Meyerand, V. Prabhakaran, Age-related reorganizational changes in modularity and functional connectivity of human brain networks, Brain connectivity 4 (9) (2014) 662–76.

[39] J. R. Andrews-Hanna, J. Smallwood, R. N. Spreng, The default network and self-generated thought: component processes, dynamic control, and clinical relevance, Annals of the New York Academy of Sciences 1316 (2014) 29–52.

[40] F. Dolcos, H. J. Rice, R. Cabeza, Hemispheric asymmetry and aging: right hemisphere decline or asymmetry reduction, Neuroscience & Biobehavioral Reviews 26 (7) (2002) 819–825.

[41] G. E. Alexander, M. R. DeLong, P. L. Strick, Parallel organization of functionally segregated circuits linking basal ganglia and cortex, Ann. Rev. Neurosci. 9 (1986) 357–81.

[42] P. Manza, S. Zhang, S. Hu, H. H. Chao, H.-C. Leung, C. R. Li, The effects of age on resting state functional connectivity of the basal ganglia from young to middle adulthood, NeuroImage 107 (2015) 3117–322.

[43] J. P. Coxon, D. J. Goble, A. Van Impe, J. De Vos, N. Wenderoth, S. P. Swinnen, Reduced basal ganglia function when elderly switch between coordinated movement patterns, Cerebral Cortex 20 (10) (2010) 2368–2379.

[44] de Jong L. W., K. van der Hiele, I. M. Veer, J. J. Houwing, R. G. J. Westendorp, E. L. E. M. Bollen, P. W. de Bruin, H. A. M. Middelkoop,
M. A. van Buchem, J. van der Grond, Strongly reduced volumes of putamen and thalamus in Alzheimer’s disease: an MRI study, Brain 131 (12) (2008) 3277–3285.

[45] A. Garg, S. Appel-Cresswell, K. Popuri, M. J. McKeown, M. F. Beg, Morphological alterations in the caudate, putamen, pallidum, and thalamus in Parkinson’s disease, Front Neurosci. 9 (2015) 101.

[46] B. G., L. P., P. M. I., Cognitive and emotional influences in anterior cingulate cortex, Trends Cogn. Sci. 4 (2010) 215–222.

[47] W. Cao, C. Luo, B. Zhu, D. Zhang, L. Dong, J. Gong, D. Gong, H. He, S. Tu, W. Yin, J. Li, H. Chen, D. Yao, Resting-state functional connectivity in anterior cingulate cortex in normal aging, Frontiers in Aging Neuroscience 6 (2014) 280.

[48] A. H. Gutchess, E. A. Kensinger, D. L. Schacter, Aging, self-referencing, and medial prefrontal cortex, Social Neuroscience 2 (2) (2007) 117–133.

[49] M. A. J. Apps, P. L. Lockwood, J. H. Balsters, The role of the midcingulate cortex in monitoring others’ decisions, Frontiers in Neuroscience 7 (2013) 251.

[50] K. T. K. Lim, R. Yu, Aging and wisdom: Age-related changes in economic and social decision making, Frontiers in Aging Neuroscience 7 (2015) 120.

[51] G. Hein, R. T. Knight, Superior temporal sulcus - it’s my area: or is it?, Journal of Cognitive Neuroscience 20 (12) (2008) 2125–2136.

[52] R. J. Jou, N. J. Minshew, M. S. Keshavan, M. P. Vitale, A. Y. Hardan, Enlarged right superior temporal gyrus in children and adolescents with autism, Brain Research 1360 (2010) 205–212.

[53] J. M. Moran, E. Jolly, J. P. Mitchell, Social-cognitive deficits in normal aging, Journal of Neuroscience 32 (16) (2012) 5553–5561. arXiv:http:
Figure 3: Comparison of graph invariants for distinguishing graph topology. (a) Volume entropy ($h_{vol}$), (b) functional entropy ($h_{fun}$), (c) spectral entropy ($h_{spe}$), (d) average local efficiency ($e_{loc}$), and (e) global efficiency ($e_{glo}$). The line color represents RE (blue), SW (green), RA (red), SF (cyan), and magenta (HY). The horizontal axis on the large panel represents a sparsity. Whenever the order of five graph types is changed by varying the sparsity, the box plot in a small panel shows the changed order of RE, SW, RA, SF, and HY. The order of five graph types measured by the volume entropy in (a) was consistent for the sparsity, while the order of graph types in the other invariants in (c-e) was changed more than three times. The functional entropy in (b) was exactly the same for all graph types at the fixed sparsity.
Figure 4: Comparison of graph invariants for distinguishing graph geometry. (a) Volume entropy, (b) functional entropy, (c) spectral entropy, (d) average local efficiency, and (e) global efficiency with respect to sparsity. Blue, green, and red represented three types of weighted graph, U, W, and S, respectively. U, S, and W had the same topology, but different geometry. The box plots in a small panel showed the order of U, S, and W in ascending order of graph invariants. The number of small panels was the same as the change of the order of graph types. The volume entropy, spectral entropy, and average local efficiency in (a), (c), and (d) showed consistent result for all sparsity, while the functional entropy and global efficiency in (b) and (e) were changed more than twice.
Figure 5: Directed information flow network of modular graph. Each graph consists of two modules generated by two Gaussian distributions with mean ($-5, 0$) and $(5, 0)$. The variance is varied from 0.01 to 100 from top to bottom. The number of nodes is 100. The ratio between module size is varied from 5:5 to 0:10 from left to right. In each graph, the nodes on the left are blue, and the nodes on the right are red. After the edges with the top 5 percent of edge capacities are obtained by the stationary distribution of the generalized Markov system, they are plotted in each graph. The color of edge represents the direction of edge. It is the same as the color of its terminal node. The color saturation represents the edge capacity as shown in the right colorbar.
Figure 6: (a) Volume entropy of 38 functional graphs of resting state fMRI with respect to age. The volume entropy was significantly correlated with age (negative correlation, \( p < .005 \)). (b) Volume entropy of Y and O in the metabolic graphs of PET. Box plots showed the volume entropy of 5000 null graphs constructed by permuted PET data set. The green marker 'X' represented the volume entropy of true metabolic graphs, Y and O. The volume entropy of O was significantly different from that of null graphs, but the volume entropy of Y was not (\( p < .05 \)). The difference of volume entropy between Y and O was not significant, but showed the tendency of Y > O (\( p < .13 \)).
Figure 7: (a,c) Edge capacity matrices of Y and O in PET. In the edge capacity matrix, the first 45 rows and columns corresponds to right hemisphere and the last 45 rows and columns corresponds to left hemisphere. F, L, P, T, B, L, and O represents frontal, limbic (cingulate cortex), parietal, temporal, basal ganglia, limbic (hippocampus and parahippocampal gyrus), and occipital lobes. (b,d) Directed information flow graphs of Y and O. Only the top 5 percent of edge capacity was plotted in the directed network. The color of node represents the location of node (more details in the supplementary material). The size of node is proportional to the absolute of node capacity. The color of edge depends on the color of its terminal node.
Figure 8: (Upper) Node capacity of right PUT and PAL, left THA, and right ACC with respect to age in 38 functional graphs. The node capacity of right PUT and PAL, and left THA significantly decreased with age, and that of right ACC was a U-shaped curve with age ($p < .05$, FDR-corrected). (Lower) Edge capacity directed from right MCC to left MCC, in the opposite direction, and from left STG to right STG with respect to age. The edge capacity increased with age ($p < .05$, FDR-corrected).