



Article

# Brain network modeling based on Mutual Information and Graph Theory for predicting the connection mechanism in the development of Alzheimer's disease

Shuaizong Si , Bin Wang \* , Xiao Liu \*, Chong Yu, Chao Ding, and Hai Zhao

School of computer science and engineering, Northeastern University, 110169 Shenyang, China;  
sishuaizong@gmail.com; bin.wang.brilliant@gmail.com; xiao\_6xiao@126.com; im.chong.yu@gmail.com;  
dingchao@stu.neu.edu.cn; zhaoh@neu.edu.cn

\* Correspondence: bin.wang.brilliant@gmail.com; xiao\_6xiao@126.com; Tel.: +86-188-4241-8067

**Abstract:** Abnormal connections in brain networks of healthy people always bring the problems of cognitive impairments and degeneration of specific brain circuits, which may finally result in Alzheimer's disease (AD). Exploring the development of the brain from normal controls (NC) to AD is an essential part of human research. Although connections changes have been found in the development, the connection mechanism that drives these changes remain incompletely understood. The purpose of this study is to explore the connection changes in brain networks in the process from NC to AD, and uncover the underlying connection mechanism that shapes the topologies of AD brain networks. In particular, we propose a model named MINM from the perspective of topology-based mutual information to achieve our aim. MINM concerns the question of estimating the connection probability between two cortical regions with the consideration of both the mutual information of their observed network topologies and their Euclidean distance in anatomical space. In addition, MINM considers establishing and deleting connections, simultaneously, during the networks modeling from the stage of NC to AD. Experiment results show that MINM is sufficient to capture an impressive range of topological properties of real brain networks such as characteristic path length, network efficiency, and transitivity, and it also provides an excellent fit to the real brain networks in degree distribution compared to experiential models. Thus, we anticipate that MINM may explain the connection mechanism for the formation of the brain network organization in AD patients.

**Keywords:** Alzheimer's disease; graph theory; mutual information; network model; connection mechanism; functional magnetic resonance imaging; topological structures; anatomical distance

## 1. Introduction

Alzheimer's disease (AD) is the primary form of dementia and the most common degenerative brain disease among older people [1]. AD patients show symptoms of a decline in memory, language, problem-solving and other cognitive functions that affects a person's ability to perform daily activities. This decline occurs because the functional connections between two brain regions involved in cognitive function have been damaged, which blocks the normally information transformation [2-4]. Although a large amount of research has been devoted to Alzheimer's disease, it is still a significant challenge to discover the underlying connection patterns that cause the alteration of functions in brain network of AD [5-7]. Addressing these problems have profound significance for a better understanding of how the abnormal interactions occur among brain regions, and it is crucial for the early detections and early treatments of AD.

More recently, both graph theory and mutual information (MI) have emerged as powerful and efficient mathematical techniques for investigating AD. Indeed, graph theory provides a mount of

topological properties for evaluating the characteristics of human brain networks [8-12]. Taking the form of a graph, people can learn the alterations of brain network properties in terms of network connectivity, transitivity, efficiency, degree distribution, modularity, and small-world-ness between normal controls (NC) and AD patients [13-15]. Particularly, graph theory also provides numerous methods of network modeling for simulating the evolution processes of real complex networks [16-19]. Through network modeling, one can surmise the fundamental causes that result in the existence of connections among nodes, and explain the underlying mechanisms of networks organization [20]. Previous studies have reported that network modelings could be effectively applied in various science fields to help explore the dynamic connection schemes in networks of real-world systems [21-25], i.e., friendships recommendation in social networks [21,22], spurious links identification in biological networks [23,24]. We can generate network topologies that incorporate desired properties by employing suitable network models. Network modeling is viewed as a promising way to help us understand how the inter-connection mechanism affects the topological structures in complex networks.

It is worth mentioning that network modeling has made some significant advances, particularly in the study of brain networks simulation, with the intersecting developments of graph theory and network neuroscience [26-31]. One of the critical contributions that should be mentioned, for instance, is the Economical Clustering Model (ECM) proposed in [26]. ECM achieved to construct the brain network by adopting the local topologies of common neighbors (CN) between two regions, and the authors tried to uncover the connection mechanism of the formation of brain networks. Simulation results show that networks modeled by ECM can dramatically capture a range of topological features of real functional brain networks. Similarly, Betzel et al. proposed a series of generative models of human brain networks according to different network structures. Their primary goal is to explore the wiring rules that shape the topologies of the brain connectome [27]. Their efforts play a vital role in understanding how the cognitive function changes across the lifespan. Previous studies indicate that network modeling has emerged as an essential way in brain network investigations. By studying the associated connection mechanisms of the network models, we can explain the workings of systems built upon those networks.

Besides, as a fundamental quantity of information theory, MI provides a measure of the statistical dependence between two random variables [32, 33]. At present, MI has been intensively investigated in the evaluation of the functional brain connectivity and the changes in interactions between mild cognitive impairment (MCI)<sup>1</sup> and AD [35]. Besides, MI is applied to estimate the probability of successfully information transmission over the brain connections between different cortical regions of patients suffering from AD [36]. Moreover, previous studies have found that MI could be used to quantify the effect of correlations between the MMSE<sup>2</sup> scores of AD and cognitive stage [38, 39]. By comparing the MMSE scores between MCI and the moderate stages of AD, people found that the cognitive ability of probable AD patients declines more severe than that of MCI [40]. Furthermore, MI is treated as an effective measure in evaluating the inter-relationships of gene expression networks for AD and other neurodegenerative diseases [41, 42]. As a consequence, previous studies indicate that MI has a significant influence in brain research of AD.

However, it is worthwhile mentioning that mutual information between different cortical regions in brain networks has not been concerned in previous brain network models. Additionally, although extensive research has been carried out on brain network modeling, far too little attention has been paid to the dynamic process of brain networks simulation from one existing state, e.g. NC to another, e.g. AD. Most of the previously proposed models start their simulation with the assumptions that nodes

---

<sup>1</sup> Mild cognitive impairment (MCI) causes a slight but noticeable and measurable decline in cognitive abilities, including memory and thinking skills. A person with MCI is at an increased risk of developing Alzheimer's or another dementia [34].

<sup>2</sup> Mini-mental state examination (MMSE) is a method to evaluate the cognitive state of AD patients and have been routinely used in clinical settings [37].

are isolated and no connections exist in the initial simulation networks. Apparently, these models are not suitable for simulating the brain networks from NC to AD. Moreover, these models generate the desired synthetic networks by continually adding connections into the initial simulation networks. They have ignored the fact that the real brain networks are combining with not only emerging of new connections but also disappeared of old connections, in the development process from NC to AD [43, 44]. Therefore, models that only take adding connections into account is not realistic. To realize a better brain networks construction of AD patients, the models should consider adding and deleting connections, simultaneously.

Taking facts as mentioned above into consideration, in this work, we proposed a network model named MINM for brain network modeling. Both MI and graph theory are taken into account to simulate the connection changes in the development of human brain networks of AD. The ultimate goal of the study is to uncover the connection mechanism that produces synthetic networks with properties similar to those of real observed brain networks topologies. The rest of this paper is organized as follows. In Section 2, we introduce the materials and methods, including the construction of real brain network, the presentation of our proposed models MINM in detail and the evaluation of synthetic networks. Then, we show the results of our experiments in Section 3. Section 4 is devoted to performance analysis and discussion of our proposed MINM. Finally, in Section 5, we draw the conclusion of this work.

## 2. Materials and Methods

### 2.1. Data acquisition and Participants Selection

Data used in this study were recruited from the public resting-state functional magnetic resonance imaging (rs-fMRI) datasets named Alzheimer's Disease Neuroimaging Initiative (ADNI) (<http://adni.loni.ucla.edu>) consisting of a total of 147 participants. They are divided into three groups: normal controls (NC) group, mild cognitive impairment (MCI) group, and Alzheimer's disease (AD) group. Table 1 shows the demographic and clinical characteristics of the three groups. The NC participants were non-depressed, non-demented, and had an average MMSE score of 28.72. The MCI group had an average MMSE score of 27.68. Patients with AD had an average MMSE score of 22.36. Each participant underwent a scan session using a 3.0T Philips MRI scanner. All the resting fMRI scans were collected axially by adopting an echo-planar imaging (EPI) sequence with the following parameters: repetition time (TR) = 3000 ms; echo time (TE) = 30 ms; axial slices = 48; slice thickness = 3.313 mm; slice acquisition order = sequential ascending; flip angle (FA) = 80.0°. Participants were informed to relax their minds and keep their eyes closed during the scanning so that we can get the resting state MRIs.

**Table 1.** Demographic and clinical characteristics of the participants in Normal controls (NC), Mild cognitive impairment (MCI) and Alzheimer's disease (AD) groups.

	NC	MCI	AD
Number	62	45	40
Gender(Male/Female)	27/35	20/25	21/19
Age	73.95 ± 4.83	74.38 ± 4.92	74.86 ± 5.52
MMSE score	28.72 ± 1.06	27.68 ± 1.86	22.36 ± 2.77
CDR score	0.00 ± 0.00	0.51 ± 0.17	0.93 ± 0.16

Values of Age, MMSE score and CDR score are expressed as the mean ± SD (standard deviation). MMSE: Mini-Mental State Examination; CDR: Clinical Dementia Rating. Significant differences were noted in MMSE scores between any two groups ( $p < 0.05$ , the  $p$ -value was obtained by two sample  $t$ -test).

### 2.2. Data preprocessing

A standard data preprocessing strategy was performed in our study using the Data Processing Assistant for Resting-State fMRI (DPARSF) software (<http://www.rfmri.org/DPARSF>)[45]

and the well-known Statistical Parametric Mapping software package (SPM8) (<http://www.fil.ion.ucl.ac.uk/spm>) [46]. The data preprocessing for each resting-state scan contains the following steps: (1) To guarantee the stabilization of the magnetic field, the first ten slice time points were discarded. (2) **Slice timing correction** was performed to ensure all remaining time points in the correct time domain. (3) **Realignment** was then executed to eliminate the movement artifact in the BOLD time series. Participants whose head translation exceeded 3.0 mm and participants whose head rotation more than  $3.0^\circ$  were discarded. (4) The functional volumes would subsequently be **Spatially normalized** to the standard EPI template and re-sliced to  $3 \times 3 \times 3\text{mm}^3$  resolution in Montreal Neurological Institute (MNI) space. (5) **Spatially smooth** was further performed on the normalized images using a Gaussian kernel of 4 mm full width at half-maximum (FWHM). (6) To reduce the influence of low-frequency drifts and high-frequency noise, temporal band-pass **filtering** in the frequency range from 0.06Hz to 0.11 Hz was achieved over each smoothed images. (7) Both linear and quadratic trends were removed. (8) Nuisance covariates such as six head motion parameters, whole-brain signal, cerebrospinal fluid, and white matter were regressed out from the preprocessed data.

### 2.3. Construction of Real Brain Network

The real brain network of each participant is represented by a binary graph in this work. First, we use the automated anatomical labeling (AAL) template [47], which functionally parcels the cerebrum into 90 regions (45 regions for each hemisphere) of interest (ROIs), to define the nodes in the graph. Second, we calculate the Pearson correlation coefficient between the whole-run BOLD time courses of any pairs of nodes, to obtain one  $90 \times 90$  inter-regional symmetric correlation matrix for each participant. Then, Fisher's  $r$ -to- $z$  transformation was performed to improve the normality of the correlation coefficients in the matrix. Third, each inter-regional correlation matrix was threshold to retain a fraction of the strongest connections for statistical significance. The element in the matrix is set to 1, with the condition that the corresponding correlation coefficient of the node pair is greater than a given threshold  $\theta$ . Alternatively, it is set to 0. Finally, we obtain a binary graph  $g$  to represent the real brain network of one participant. The element  $g(u, v) = 1$  means that there is a functional connection between node  $u$  and node  $v$ . Moreover, different connectivity densities of the real brain networks will be generated by defining different  $\theta$ .

### 2.4. Synthetic Brain network Modeling

In this study, we aim to simulate the development progress from NC to AD through network modeling. Here, we call the networks generated by our proposed model synthetic brain networks, and the networks constructed from the preprocessed images real brain networks. By comparing the properties of synthetic networks with real target brain networks, we can predict the connection mechanism that causes the topological alterations of AD. In the following subsection, we make a detailed description of the network modeling steps.

#### 2.4.1. Network modeling steps

The dynamic process of synthetic brain network modeling is executed step by step, starting from the **Initialization** step, followed by the **Connection probabilities calculation** step and the **Evolution** step. In the Connection probabilities calculation step, we compute the probability of the existence of a connection between two nodes according to the Connection probability models, i.e., ECM and MINM. In the Evolution step, connections are added into or deleted from the initial network. The modeling will not proceed to the end until a given number of connections are added, and another given number of connections are deleted successfully. The detailed network modeling steps are listed in the following:

1. Initialization: At the beginning of modeling, all participants in groups of NC, MCI and AD are preprocessed and the corresponding inter-regional correlation matrices  $G_n = (g_{n1}, g_{n2}, \dots, g_{nk_1})$ ,  $G_m = (g_{m1}, g_{m2}, \dots, g_{mk_2})$ ,  $G_a = (g_{a1}, g_{a2}, \dots, g_{ak_3})$  are obtained, where  $k_1$ ,  $k_2$  and  $k_3$  represent the number of participants in NC, MCI, and AD, respectively. We calculate the average correlation matrix of each group, and each average correlation matrix is threshold with the same  $\theta = 0.15$  to construct the corresponding real brain networks, i.e.,  $\bar{G}_n$  for NC,  $\bar{G}_m$  for MCI and  $\bar{G}_a$  for AD. The real brain network in each group consists of a constant number of nodes,  $|V| = 90$ . The connection number of  $\bar{G}_n$ ,  $\bar{G}_m$  and  $\bar{G}_a$  are represented by  $|E_n|$ ,  $|E_m|$  and  $|E_a|$ , respectively. In current work, we study the evolution process of AD networks from two stages, i.e., the stage from NC to MCI and the stage from NC to AD. By comparing the elements in the binary graphs  $\bar{G}_n$ ,  $\bar{G}_m$  and  $\bar{G}_a$ , we can get a constant number  $\alpha_1$  of connections that need to be added and another constant number  $\beta_1$  of connections that need to be delete in the stage from NC to MCI; and  $\alpha_2$  connections to be added and  $\beta_2$  connections to be deleted from NC to AD. Here,  $\alpha_1 \leq \beta_1$  and  $\alpha_2 \leq \beta_2$  because a declining number of connections is found when comparing  $|E_m|$  and  $|E_a|$  with  $|E_n|$  ( $|E_a| \leq |E_n| \leq |E_m|$ ). It is worthwhile mentioning that we call the real brain network  $\bar{G}_n$  of NC the initial network in our modeling, and call  $\bar{G}_m$  and  $\bar{G}_a$  the real target brain networks (TN).
2. Connection probabilities calculation: After initialization, we calculate the connection probabilities of any node pairs in  $\bar{G}_n$  according to the proposed connection probabilities models introduced in the following subsection. Then, we sort each node pair in line with its connection probability. The node pair with the largest connection probability and the node pair with the smallest connection probability are recorded, respectively.
3. Evolution : Our model starts to evolve from the initial graph  $\bar{G}_n$ . In each iteration, a random number is generated to decide whether to add or delete one connection in  $\bar{G}_n$ . The node pair with the largest connection probability will establish a link if its two nodes disconnected with each other. Meanwhile, the node pair with the smallest connection probability will cut off its link if there is a connection between its two nodes. It should be noted that each node must have a connection to ensure the connectivity of the synthetic network. Therefore, a new pair of nodes must be chosen according to the sorted connection probabilities, if either node's connection number in the node pair is equal to 1 when deleting the link between them. We upgrade connection set in  $\bar{G}_n$  at the end of this step.
4. End of the modeling: Our model runs the above steps of Connection probabilities calculation and Evolution round by round. The simulation will not proceed to the end, until  $\alpha_1$  connections are established and  $\beta_1$  connections are deleted successfully for  $\bar{G}_n$  in the stage from NC to MCI; or  $\alpha_2$  connections are established and  $\beta_2$  connections are deleted in the stage from NC to AD. Finally, we get two synthetic networks with the same connection size as  $\bar{G}_m$  and  $\bar{G}_a$ , respectively.

#### 2.4.2. Connection Probabilities Models

In this part, we give a detailed description of our proposed connection probability model based on mutual information for calculating the connection probability between two nodes in the network modeling. Before introducing our model, we introduce one previous model that has been intensively investigated over recent years.

To explore the formation mechanism of generating human brain network topologies, Vértés, et al. proposed the Economical Clustering Model (ECM) [26]. Their experiment results show that both the **topological similarity** of common neighbors (CN) and the **Euclidean distance similarity** between two brain regions are treated as essential impactors in brain networks modeling. The connection probability of ECM is defined in the following function:

- Connection probability of ECM: The more common neighbors (CN) that node  $u$  and node  $v$  have, the higher topological similarity to one another [48]. The number of common neighbors between node  $u$  and node  $v$  can be described by

$$S_{(u,v)}^{CN} = |\Gamma(u) \cap \Gamma(v)| \quad (1)$$

where,  $\Gamma(u)$  and  $\Gamma(v)$  represent the set of neighboring nodes of node  $u$  and node  $v$ , respectively. Scaling  $S_{(u,v)}^{CN}$  by the Euclidean distance similarity  $E(u, v)$  between two brain regions, the ECM connection probability between node  $u$  and node  $v$ , i.e., the probability that  $u$  prefers to build the connection with  $v$  is given by

$$P(u, v) = S_{(u,v)}^{CN \gamma} \cdot E(u, v)^{-\eta}. \quad (2)$$

In this expression,  $P(u, v)$  is the connection probability of ECM.  $S_{(u,v)}^{CN}$  is the contribution of the topological similarity computed by CN. The other term,  $E(u, v)$ , is the contribution of the Euclidean distance similarity which represents the anatomical distance between nodes  $u$  and  $v$ .  $\gamma$  and  $\eta$  represent the parameters of topological similarity and anatomical distance penalization, respectively. Although ECM can construct networks similar to the real target brain networks on several essential topological features, there are still two major problems that have been pointed out.

One problem is that ECM devotes to calculating the topological similarity between node  $u$  and node  $v$  only according to the number of their common neighbors, it does not consider the individual characteristics of node  $u$  and node  $v$ . In fact, the individual topological characteristics of node  $u$  and node  $v$ , are paramount for the formation of a connection between them. Additionally, the other problem is that ECM gives each common neighbor the same score to the topological similarity. However, various common neighbors may have different topological characteristics, i.e., degree, clustering coefficient. Therefore, they make a different contribution to establishing the connection. In this study, we propose a novel brain network model named MINM from the perspective of mutual information to solve above issues. Mutual information provides a measure of the statistical relationship between two random variables. More specifically, it is a measure of the reduction in uncertainty about one random variable given knowledge of another[32]. MINM adopts not only the individual features node  $u$  and node  $v$ , but also the topological-based mutual information of their common neighbors. The topological-based mutual information is used to make a distinguish for the different contributions of the common neighbors of the node pair  $(u, v)$  in calculating the existence probability of one connection. Moreover, higher mutual information indicates a substantial reduction in the uncertainty of the formation of a connection; otherwise, lower mutual information means a smaller probability of the existence of one connection. The question we investigate is whether the mutual information between node  $u$  and node  $v$  can be helpful for the modeling of brain networks during the evolution process from NC to AD. Next, we introduce the definitions of self-information and mutual information, and then, we give a precise definition of our proposed connection probability of MINM.

**Definition 1.** (Self-information) *The Self-information of a random variable is a function which concentrates on quantifying the information involved in the value of a random variable. Given a random variable  $X$  with a probability distribution  $P(X)$ , where  $X$  takes on values in a set  $X = \{x_1, x_2, \dots, x_n\}$ , the self-information of the random variable can be expressed by*

$$I(X) = - \sum_{x \in X} P(x) \log P(x). \quad (3)$$

**Definition 2.** (Mutual information) *Mutual information is a quantity that measures the amount of message which can be acquired about one random variable by observing another. Formally, the mutual information of two random variables  $X$  and  $Y$ , whose joint distribution is given by  $P(X, Y)$  can be defined as*

$$\begin{aligned}
I(X; Y) &= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(y)p(x|y)}{p(x)p(y)} \\
&= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x|y)}{p(x)} = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x|y) - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x) \\
&= - \sum_{x \in X} p(x) \log p(x) - [- \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x|y)] \\
&= I(X) - I(X|Y)
\end{aligned} \tag{4}$$

where  $P(X)$  and  $P(Y)$  represent the *marginal distributions* of  $X$  and  $Y$ , respectively.  $I(X; Y) = 0$  if and only if  $X$  and  $Y$  are independent.

- Connection probability of MINM: Given a pair of node  $(u, v)$ , whose common neighbors can be represented by  $\omega_{u,v} = \Gamma(u) \cap \Gamma(v)$ , the connection probability of MINM between them can be given by

$$P(u, v) = S_{(u,v)}^{MI}{}^\gamma \cdot E(u, v)^{-\eta}. \tag{5}$$

where  $S_{u,v}^{MI}$  is the topological similarity between node  $u$  and node  $v$ , which represents the topology-based mutual information between them;  $E(u, v)$  is the Euclidean distance similarity, which is similar with ECM.  $S_{u,v}^{MI}$  can be given by

$$S_{u,v}^{MI} = -I(L_{u,v}^1 | \omega_{u,v}) \tag{6}$$

where  $I(L_{u,v}^1 | \omega_{u,v})$  denotes the conditional self-information of an event that there is a connection between node  $u$  and node  $v$ , whose common neighbors are known as  $\omega_{u,v}$ . From the equation (6), we can know that the smaller  $I(L_{u,v}^1 | \omega_{u,v})$  is, the higher topological similarity  $S_{u,v}^{MI}$  is, and this indicates that a larger probability for node  $u$  and node  $v$  to establish one connection between them.

Next, we give a comprehensive illustration about how to calculate  $S_{u,v}^{MI}$ . According to the definition of Mutual Information in equation (4), the topological similarity of  $S_{u,v}^{MI}$  can be given by

$$S_{u,v}^{MI} = -I(L_{u,v}^1 | \omega_{u,v}) = I(L_{u,v}^1; \omega_{u,v}) - I(L_{u,v}^1) \tag{7}$$

where  $I(L_{u,v}^1; \omega_{u,v})$  denotes the mutual information between two events  $L_{u,v}^1$  and  $\omega_{u,v}$ .  $L_{u,v}^1$  describes the event that there is one link between node  $u$  and node  $v$ , and  $\omega_{u,v}$  represents the event that the common neighbors between node  $u$  and node  $v$  are observed.  $I(L_{u,v}^1; \omega_{u,v})$  evaluates the increment of the probability for the formation of one connection between node  $u$  and node  $v$  by knowing the information of their common neighbors.  $I(L_{u,v}^1)$  represents the self-information of the event that node  $u$  and node  $v$  are connected. According to the definition in equation (3),  $I(L_{u,v}^1)$  can be obtained by calculating the probability of  $p(L_{u,v}^1)$ . In this work, we assume that the common neighbors in  $\omega_{u,v}$  are independent with each other, and thus,

$$I(L_{u,v}^1; \omega_{u,v}) = \sum_{z \in \omega_{u,v}} I(L_{u,v}^1; z). \tag{8}$$

Given one common neighbor  $z$  in  $\omega_{u,v}$ ,  $I(L_{u,v}^1; z)$  can be estimated by calculating the average mutual information of all pairs of nodes that have node  $z$  as their neighbor in common.

$$\begin{aligned} I(L_{u,v}^1; z) &= \frac{1}{|\Gamma(z)|(|\Gamma(z)| - 1)} \sum_{m,n \in \Gamma(z)} I(L_{m,n}^1; z), \forall m \neq n, \\ &= \frac{1}{|\Gamma(z)|(|\Gamma(z)| - 1)} \sum_{m,n \in \Gamma(z)} \{I(L_{m,n}^1) - I(L_{m,n}^1|z)\}, \forall m \neq n \end{aligned} \quad (9)$$

where  $\Gamma(z)$  is the neighbor set of  $z$ ;  $I(L_{m,n}^1)$  represents the self-information of an event that node  $m$  and node  $n$  are connected;  $I(L_{m,n}^1|z)$  is used to describe the conditional self-information of the event that node  $m$  and node  $n$  are connected on the condition that node  $z$  is known as one of their common neighbors. Here, we can get  $I(L_{m,n}^1)$  through calculating  $p(L_{m,n}^1)$ .  $p(L_{m,n}^1)$  can be given by

$$p(L_{m,n}^1) = 1 - p(L_{m,n}^0) = 1 - \prod_{i=1}^{k_n} \frac{|E| - k_m - i + 1}{|E| - i + 1} = 1 - \frac{C_{|E|-k_m}^{k_n}}{C_{|E|}^{k_n}} \quad (10)$$

where  $L_{m,n}^0$  represents the event that node  $u$  and node  $v$  are disconnected with each other;  $|E|$  represents the existing number of connections in the network;  $k_m$  and  $k_n$  denote the connection number of node  $m$  and node  $n$ , respectively.

Then, we give an explanation of the calculation of  $I(L_{m,n}^1|z)$ , which is described in equation (9). According to the definition of self-information, to obtain  $I(L_{m,n}^1|z)$  we need to calculate the  $p(L_{m,n}^1|z)$ . In the current work,  $p(L_{m,n}^1|z)$  is equal to the clustering coefficient of node  $z$

$$p(L_{m,n}^1|z) = \frac{N_{\wedge z}}{N_{\wedge z} + N_{\vee z}}, \quad (11)$$

where  $N_{\wedge z}$  and  $N_{\vee z}$  represent the connected and disconnected link number in  $\Gamma(z)$ , respectively. Obviously,  $N_{\wedge z} + N_{\vee z} = |\Gamma(z)|(|\Gamma(z)| - 1)/2$ , which represents the total connections that could possibly exist within  $\Gamma(z)$ , then,  $p(L_{m,n}^1|z)$  can be described by

$$p(L_{m,n}^1|z) = \frac{2|E_z|}{|\Gamma(z)|(|\Gamma(z)| - 1)}, \quad (12)$$

where  $\Gamma(z)$  represents neighbors set of node  $z$ ,  $|E_z|$  denotes the existing link number in  $|\Gamma(z)|$ . Therefore,  $I(L_{u,v}^1; z)$  is given by

$$\begin{aligned} I(L_{u,v}^1; z) &= \frac{1}{|\Gamma(z)|(|\Gamma(z)| - 1)} \sum_{m,n \in \Gamma(z)} I(L_{m,n}^1 - I(L_{m,n}^1|z)), \forall m \neq n \\ &= \frac{1}{|\Gamma(z)|(|\Gamma(z)| - 1)} \sum_{m,n \in \Gamma(z)} \{-\log p(L_{m,n}^1) - (-\log p(L_{m,n}^1|z))\}, \forall m \neq n \\ &= \frac{1}{|\Gamma(z)|(|\Gamma(z)| - 1)} \sum_{m,n \in \Gamma(z)} \left\{ \log \frac{C_{|E|}^{k_n}}{C_{|E|-k_m}^{k_n}} + \log \frac{2|E_z|}{|\Gamma(z)|(|\Gamma(z)| - 1)} \right\}, \forall m \neq n. \end{aligned} \quad (13)$$



Finally, substituting equations (8), (10) and (13) back into equation (7), we can get the topological similarity of  $S_{u,v}^{MI}$ .

$$S_{u,v}^{MI} = \sum_{z \in \omega_{u,v}} \left\{ \frac{1}{|\Gamma(z)|(|\Gamma(z)| - 1)} \sum_{m,n \in \Gamma(z)} \left\{ \log \frac{C_{|E|}^{k_n}}{C_{|E|}^{k_n} - C_{|E|-k_m}^{k_n}} + \log \frac{2|E_z|}{|\Gamma(z)|(|\Gamma(z)| - 1)} \right\} \right\} - \log \frac{C_{|E|}^{k_n}}{C_{|E|}^{k_n} - C_{|E|-k_m}^{k_n}}, \forall m \neq n. \quad (14)$$

As we can see from the equations (7) and (14), both the mutual information  $\sum_{z \in \omega_{u,v}} I(L_{u,v}^1; z)$  and the self-information  $I(L_{u,v}^1)$  are used to define the topological similarity between node  $u$  and node  $v$ . Thus, two different node pairs  $(u, v)$  who have the same number of common neighbors may get various topological similarity.

### 2.5. Evaluation of synthetic networks

To evaluate the performance of our proposed model, we define a similarity index  $SI$  function as described in [26] and [27], by comparing different topological properties between synthetic networks and real target brain networks (TN), i.e., the real brain networks of MCI and AD. Moreover, to make it a more convincing and reasonable evaluation, the most important network properties including clustering coefficient, local efficiency, modularity, characteristic path length, global efficiency, and transitivity are chosen. The detailed description of these features is shown in Table 2, and they can reflect the performance of one network from different aspects. The definition of  $SI$  is given by

$$SI = \frac{1}{\zeta_C + \zeta_{Eloc} + \zeta_M + \zeta_L + \zeta_{Eglob} + \zeta_T}. \quad (15)$$

where  $\zeta_C$  is the relative error of clustering coefficient between the synthetic network and real target brain network (TN);  $\zeta_{Eloc}$  and  $\zeta_{Eglob}$  represent the relative errors of local efficiency and global efficiency, respectively. Similarly,  $\zeta_M$ ,  $\zeta_L$ , and  $\zeta_T$  are the relative errors of modularity, the characteristic path length, and transitivity, respectively. A smaller value of  $SI$  indicates that the synthetic networks are more similar to the real target brain networks of MCI or AD. Obviously, we can get different values of  $SI$  with various  $\gamma$  and  $\eta$ . To find the optimal synthetic networks that overall most closely approximate the real target group, simulated annealing (SA) on  $SI$  is used. Moreover, the optimal  $\gamma$  and  $\eta$  in the parameter space that minimizes  $SI$  are recorded.

**Table 2.** Description of topological properties in complex networks.

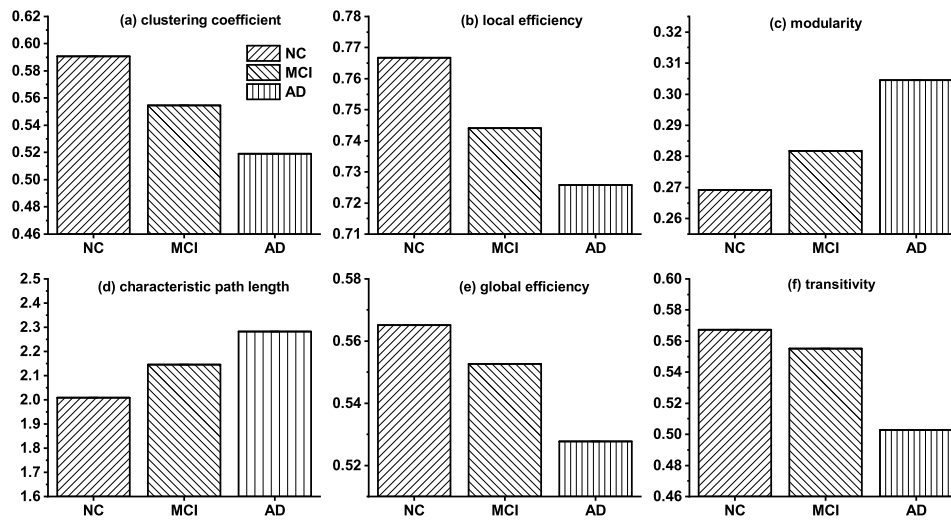
Property name	Symbol	Description
Clustering coefficient	C	It is a measure of the number of triangles in a graph.
Local efficiency	Eloc	It is a measure to quantify the efficiency of local information transmission.
Global efficiency	Eglob	It is a measure to quantify the efficiency of global information transmission.
Characteristic path length	L	L is the average shortest path length between all node pairs in the network.
Modularity	M	It is used to detect the strength of the division of a network into communities.
Transitivity	T	It measures the probability that the adjacent nodes of a node are connected.
Degree	k	It indicates the number of links connecting with a node.

## 3. Results

In this study, we proposed one novel brain network model named MINM for a better understanding of the connection mechanism of AD brain networks from two stages, i.e., the stage from NC to MCI, and the stage from NC to AD. Both topological-based mutual information and the anatomical distance are taken into account in MINM. We give detailed descriptions of the modeling results in the two stages, respectively, in the following subsections.

### 3.1. Topological differences in brain networks of NC, MCI and AD

Before introducing the modeling results, we analyze the topological differences among brain networks of NC, MCI, and AD. Figure 1 shows the changes of the six topological features among the real observed networks of NC, MCI, and AD groups. The results indicate that noticeable topological changes are found in the networks of these three groups. Specifically, clustering coefficient, local efficiency, global efficiency, and transitivity of brain networks decrease gradually from NC, via MCI to AD, whereas, characteristic path length and modularity increase progressively. The analysis of these topological differences can help us explain how connections change in brain networks, and it is valuable for proposing a more reasonable simulation model.



**Figure 1.** Topological differences among the real brain networks of NC, MCI and AD. NC represents the real brain network of normal control (NC) group; MCI and AD are the real brain networks of Mild cognitive impairment (MCI) and Alzheimer’s disease (AD) group, respectively.

#### 3.1.1. Network modeling of the stage from NC to MCI

In this subsection, we evaluate the performance of our proposed model MINM and other five well-known models in the stage from NC to MCI. The compared models are similar with MINM except for the definitions of the topological similarities. Table 3 makes a detailed description of the topological similarities in the compared models. In addition, we also perform a random model for comparison. In the random model, we select two disconnected nodes to establish one connection and two connected nodes to delete one connection, randomly.

**Table 3.** The definitions of topological similarities in the compared models.

Models	Abbreviation	Mathematical
Preferential Attachment [49]	PA	$S_{u,v}^{PA} =  \Gamma(u)  \times  \Gamma(v) $
Jaccard [50]	JC	$S_{u,v}^{JC} = \frac{ \Gamma(u) \cap \Gamma(v) }{ \Gamma(u) \cup \Gamma(v) }$
Adamic-Adar [51]	AA	$S_{u,v}^{AA} = \sum_{\xi \in  \Gamma(u) \cap \Gamma(v) } \frac{1}{\log \Gamma(\xi) }$
Resource Allocation [52]	RA	$S_{u,v}^{RA} = \sum_{\xi \in  \Gamma(u) \cap \Gamma(v) } \frac{1}{ \Gamma(\xi) }$

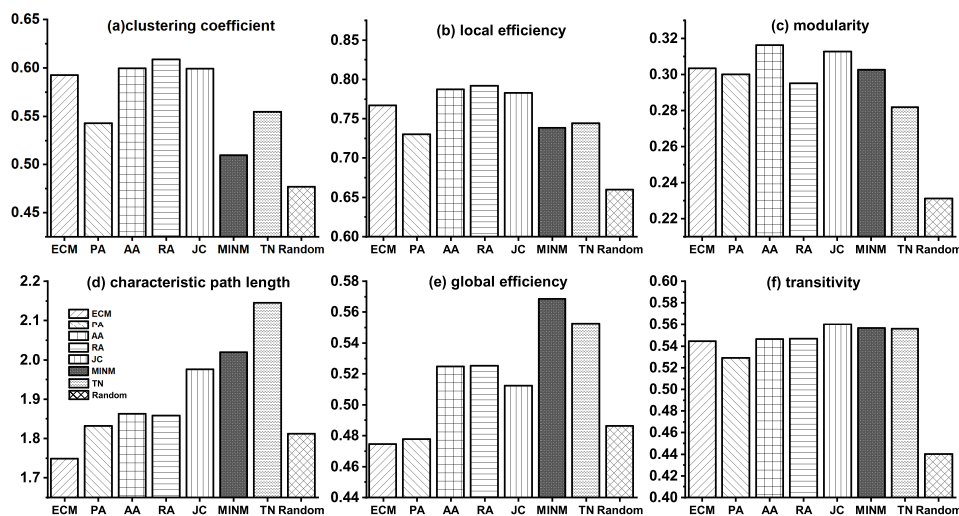
Table 4 shows the  $SI$  of different models with the best-fitting parameters  $\lambda$  and  $\eta$ . We find that MINM outperform other models with a largest  $SI = 3.9683$  and the optimal parameters  $\lambda = 0.4, \eta = 2.0$ , followed by RA ( $SI = 2.4358, \lambda = 0.4, \eta = 2.0$ ), PA ( $SI = 2.4010, \lambda = 0.4, \eta = 1.8$ ) and JC ( $SI = 2.3714, \lambda = 0.2, \eta = 0.2$ ). The results indicate that MINM considering both network topology-based mutual information and anatomical distance can generate synthetic networks capturing

all of the key topological features of real target brain networks. Specially, MINM minimizes the difference between the synthetic network and the real brain network on properties of local efficiency ( $\zeta_{Eloc} = 0.0077$ ), global efficiency ( $\zeta_{Eglob} = 0.0292$ ), the characteristic length ( $\zeta_L = 0.0598$ ), and transitivity ( $\zeta_T = 0.0010$ ). Therefore, MINM is the best model in simulating the underlying mechanism that causes the connection changes in brain networks in the stage from NC to MCI. Moreover, we find that PA generates synthetic networks most similar with the real brain network of MCI on properties of clustering coefficient, but it fails to match real MCI networks well on transitivity ( $\zeta_T = 0.0480$ ). RA generates synthetic networks most similar with the real brain network of MCI on properties of modularity, but mismatch real MCI networks on properties of clustering coefficient ( $\zeta_C = 0.0975$ ) and local efficiency ( $\zeta_{Eloc} = 0.0638$ ). In contrast, the results also indicate that the *SI* of ECM is the smallest among the six topology-based models. The performance of ECM was limited primarily by mismatches in characteristic path length ( $\zeta_L = 0.1851$ ) and global efficiency ( $\zeta_{Eglob} = 0.1413$ ). Additionally, we also compare the six topology-based models with the Random model. As shown in Table 4, we find that the *SI* of Random model has the lowest  $SI = 1.0912$ .

**Table 4.** The optimal *SI*-value of different brain network models for networks modeling in the stage from NC to MCI.  $\zeta_C$  is the relative error of clustering coefficient between synthetic networks and the real target brain network (TN);  $\zeta_{Eloc}$  and  $\zeta_{Eglob}$  represent the relative errors of local efficiency and global efficiency;  $\zeta_M$ ,  $\zeta_L$ , and  $\zeta_T$  are the relative errors of modularity, the characteristic path length, and transitivity, respectively. A larger value of *SI* indicates that the model can generate synthetic networks with properties more similar to the real target brain network of MCI.

Models	$\lambda$	$\eta$	$\zeta_C$	$\zeta_{Eloc}$	$\zeta_M$	$\zeta_L$	$\zeta_{Eglob}$	$\zeta_T$	<b>SI</b>
ECM	0.2	1.6	0.0687	0.0302	0.0713	0.1851	0.1413	0.0205	1.9339
PA	0.2	1.8	<b>0.0213</b>	0.0187	0.0639	0.1464	0.1356	0.0480	2.4010
AA	0.2	1.4	0.0811	0.0578	0.1210	0.1324	0.0514	0.0173	2.1692
RA	0.4	2.0	0.0975	0.0638	<b>0.0465</b>	0.1345	0.0502	0.0164	2.4358
JC	0.2	0.2	0.0844	0.0515	0.1082	0.0861	0.0789	0.0126	2.3714
MINM	0.4	2.0	0.0816	<b>0.0077</b>	0.0727	<b>0.0598</b>	<b>0.0292</b>	<b>0.0010</b>	<b>3.9683</b>
Random	-	-	0.1406	0.1132	0.1796	0.1558	0.1204	0.2068	<b>1.0912</b>

Furthermore, to make a detailed illustration of the above results we also present the values of the six topological properties in both synthetic networks generated by the topology-based models and the real target brain network (TN) of MCI, as shown in Figure 2.



**Figure 2.** Topological properties of the synthetic brain networks generated by various models and the real target brain network (TN) of MCI.

### 3.1.2. Network modeling of the stage from NC to AD

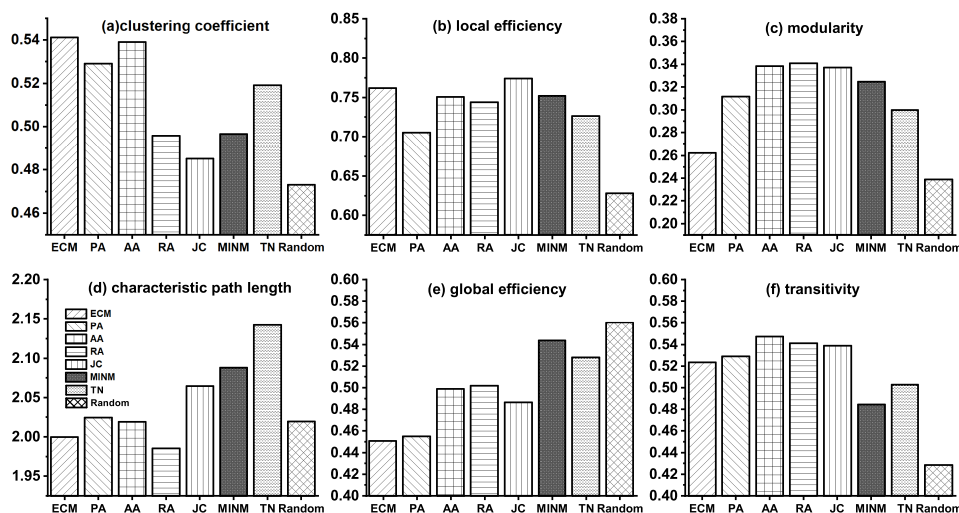
In this subsection, we study the evolution process of network connections in the stage from NC to MCI.  $SI$  is again used to evaluate the performance of our proposed model MINM and other five known topology-based models. A random model is also performed for comparison. By comparing  $SI$  and the relative errors of different topological properties, we want to estimate the connection mechanism that causes the changes of connections in AD patients.

Table 5 gives a detailed illustration of the largest  $SI$  for the six topology-based models and the Random model. The corresponding best-fitting parameters  $\lambda$  and  $\eta$  and the relative errors on the six topological properties are listed in the table. We find good correspondence between the synthetic network constructed by MINM and the real target brain network of AD on all of the key topological properties of global efficiency ( $\zeta_{Eglob} = 0.0303$ ), the characteristic length ( $\zeta_L = 0.0253$ ), and transitivity ( $\zeta_T = 0.0368$ ). Overall, the results confirm that our proposed model MINM achieves the largest  $SI = 3.9231$  with the optimal parameters  $\lambda = 0.2, \eta = 1.8$ , which means the corresponding synthetic network is significantly more AD-like than either of the other topological models previously considered. Therefore, MINM is the best model for anticipating the connection mechanism of the real brain network in the stage from NC to AD. In addition, the  $SI$  of PA is equal to 2.4358, with the optimal parameters  $\lambda = 0.4, \eta = 2.0$ , preceded only by MINM. The  $SI$  and the optimal parameters of AA and RA are  $SI = 2.4010, \lambda = 0.4, \eta = 1.8$  and  $SI = 2.3714, \lambda = 0.2, \eta = 0.2$ , respectively, followed by PA. Another finding is that the synthetic network generated by PA has the clustering coefficient and modularity that match exactly with that of the real brain network of AD.  $\zeta_{Eloc}$  of the synthetic network generated by RA is the closest to the real target brain network. However, the synthetic network fails to match the real network well regarding modularity ( $\zeta_M = 0.1370$ ) and the characteristic path length ( $\zeta_L = 0.0732$ ). Furthermore, the results of the current study also reveal that the performance of JC is not as clustering or local efficient as the real brain network of AD. ECM has the worst performance in simulating the global efficiency ( $\zeta_{Eglob} = 0.1413$ ), and it gets the smallest  $SI$  among the six topology-based models. Finally, the comparison results show that all the six topology-based models outperform the Random model, significantly.

**Table 5.** The optimal  $SI$ -value of different brain network models for networks modeling in the stage from NC to AD.  $\zeta_C$  represents the relative error in clustering coefficient between synthetic networks and the real target brain network (TN);  $\zeta_{Eloc}$  represents the relative error in local efficiency,  $\zeta_M$  is the relative error in modularity;  $\zeta_L$  is the relative error in the characteristic path length;  $\zeta_{Eglob}$  is the relative error in global efficiency;  $\zeta_T$  is the relative error in transitivity.

Models	$\lambda$	$\eta$	$\zeta_C$	$\zeta_{Eloc}$	$\zeta_M$	$\zeta_L$	$\zeta_{Eglob}$	$\zeta_T$	<b>SI</b>
CN	0.4	1.2	0.0426	0.0491	0.1263	0.0665	0.1460	0.0419	2.1169
PA	0.2	1.6	<b>0.0194</b>	0.0289	<b>0.0395</b>	0.0551	0.1377	0.0515	3.0111
AA	0.2	1.6	0.0386	0.0339	0.1291	0.0576	0.0551	0.0881	2.4851
RA	0.2	2.0	0.0448	<b>0.0245</b>	0.1370	0.0732	0.0496	0.0758	2.4697
JC	0.4	0.4	0.0653	0.0661	0.1249	0.0364	0.0786	0.0709	2.2614
MI	0.2	1.8	0.0432	0.0358	0.0835	<b>0.0253</b>	<b>0.0303</b>	<b>0.0368</b>	<b>3.9231</b>
Random	–	–	0.0888	0.1355	0.2304	0.0534	0.0611	0.1476	<b>1.3951</b>

Detailed information about the six topological properties of the synthetic networks and the real target brain network (TN) of AD are shown in Figure 3.

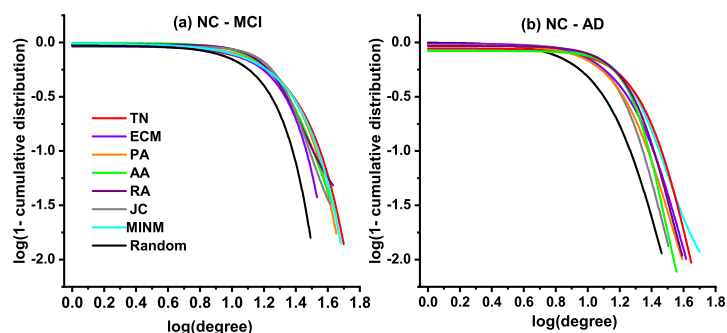


**Figure 3.** Topological properties of the synthetic brain networks generated by various models and the real target brain network (TN) of AD.

### 3.1.3. Degree distribution

In this subsection, we explore the degree distribution of the synthetic networks. In the study of graph theory, the degree of a node indicates the connections number it has to other nodes. The node degree describes the characteristic of the network from the view of local structures. The degree distribution is the probability distribution of these degrees over the whole network. It makes an evaluation of the structural and dynamical properties in the network from a global view. Therefore, it is necessary to make a detailed illustration of the degree distribution of the synthetic networks in the current study. Figure 4 shows the fitting results of the cumulative degree distribution of both the real target brain networks (TN) of MCI and AD, and synthetic networks generated by the topology-based models in the simulation of two specified stages, i.e., the stage from NC to MCI and the stage from NC to AD.

Figure 4(a) demonstrates the fitting results of the cumulative degree distribution for synthetic networks generated in the stage from NC to MCI, and the real MCI network in a log-log plot. Figure 4(b) shows the fitting results of the cumulative degree distribution for synthetic networks generated in the stage from NC to AD, and the real AD network. Both of the results show that the generated topologies of MINM are approximately the same as the real target brain networks (TN) of MCI and AD. Additionally, the results indicate that a significant difference exists between the synthetic network generated by the Random model and the real target brain networks (TN).

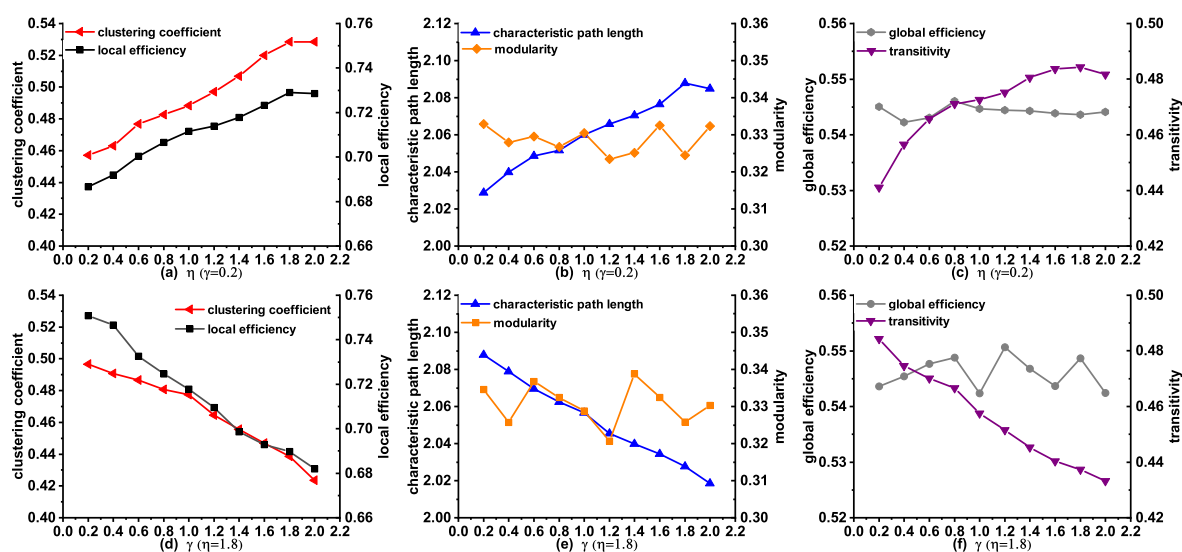


**Figure 4.** Degree distributions of the real target brain networks (TN) and the synthetic brain networks.

### 3.1.4. Topological properties of the synthetic networks generated by MINM with different $\lambda$ and $\eta$

We subsequently investigate the influence of the variety of parameters  $\lambda$  and  $\eta$  on the topological properties in the modeling stage from NC to AD. Figure 5(a),(b),(c) describe the changes of the

topological properties with  $\eta$ , when  $\lambda = 0.2$  is a constant. An interesting finding is that the clustering coefficient, local efficiency, characteristic path length, and transitivity are increasing with a gradual increase of  $\eta$ . These properties get the largest values when  $\eta = 1.8$ , which is consistent with the observation from the results in Table 5. It is worthwhile mentioning that modularity and the global efficiency do not change a lot with the increase of  $\eta$ , which confirms that the anatomical distance makes little influence on these two features. Additionally, we also investigate the topological properties of the networks by varying  $\lambda$  with  $\eta = 1.8$ . As can be seen from Figure 6(d),(e),(f), we observe a decreasing trend in clustering coefficient, local efficiency, characteristic path length, and transitivity with the increase of  $\lambda$  from 0.2 to 2.0. Especially, these properties get the largest values when  $\lambda = 0.2$ , which indicates a good agreement with the results in Table 5. It is interesting to note that both the modularity and global efficiency have little change with the increase of  $\lambda$ , which confirms that the topological similarity makes little influence on these two features.



**Figure 5.** The changes of topological properties of the synthetic networks generated by MINM with different  $\lambda$  and  $\eta$  in the stage from NC to AD.

#### 4. Discussion

In this study, we have investigated several brain network models to explore the fundamental reason that results in the alteration of connections in real AD brain networks. We have shown that the previous models, which consider either node feature, i.e., node degree, or shared nearest neighbors in common cannot satisfactorily account for all the topological properties of real target brain networks. We demonstrate that the addition of the topological-based mutual information to the model, favoring additional formation of connections between two nodes, can markedly improve the simulation of realistic brain network properties. Notably, our proposed model MINM provides a good account of network properties regarding local efficiency, global efficiency, the characteristic path length, and transitivity. MINM also provides an excellent fit to the degree distribution in both simulation stages of MCI and AD. Therefore, we confirm that MINM is the promising method for understanding the underlying mechanism that shapes the organization of brain networks during the development of AD.

We have intensively studied the topology differences between functional brain networks of NC and AD, and significant alterations of topological properties have been found, i.e., the increase of characteristic path length, the decrease of network efficiency. Specifically, the increasing of the characteristic path length indicates a long average time in dealing with the information transmission among brain areas; and the dropping down of both local efficiency and global efficiency implies the reduction of the ability in data processing. The alterations of these properties are related to the functional decline of cognition in AD patients, which is consistent with the results in literature [43] and

[44]. These previous works have shown that the abnormal alterations of networks topologies play vital roles in the influence of cognitive functions such as attention, working memory and executive control of human brains during the neuro-development. The disorder of network organizations will lead to the deficit of cognition and result in the emergence of AD. Therefore, the better understanding of the topological characteristics the more helpful for us to know the relationship between the functions and the architectures of AD brain networks.

In the current work, we study the changes of network properties for a further investigation of the fundamental rule that causes the formation of AD network topologies. Prior models such as ECM, AA, and PA, estimated the connection mechanism of the brain networks formation by emphasizing the importance of increasing connections among brain regions [26], [27]. Although the authors have confirmed that synthetic networks generated by these models can well match several properties of the real target brain networks, such an effort is accompanied by much debate in recent years in the simulation of the development process of AD. One obvious limitation of previous models is that they ignore the fact that the alterations in AD brain networks are caused by both decreasing and increasing connections between brain regions [43], [44]. As a consequence, previous models can not successfully identify the potential mechanism that results in the variation of network topologies in the development from NC and MCI, or NC and AD. It is worthwhile mentioning that our proposed model MINM takes both decrease and increase of connections among brain regions into account, and this consideration is much suitable for the actual situation of AD. Thus, it gets good performance in the modeling of AD brain networks.

It is also clear that mutual information has been extensively investigated in human brain research. The previous study has demonstrated that mutual information is associated with the cognitive functions of human brain networks [38-40]. Also, much work has suggested that mutual information makes a crucial influence on information transmission in brain networks [36]. Therefore, topological-based mutual information is considered an essential factor in MINM to make the modeling more efficiently. MINM can make an appropriate distinguish of the connection probabilities between any two pairs of nodes by considering the topological-based mutual information. This consideration may make it much reasonable for capturing the potential connection mechanism that shapes the topologies of AD brain networks. Experimental results are shown in both Table 4 and Table 5. On the basis of our findings, it can be concluded that our proposed topological-based mutual information model MINM can make a complete understanding of the connection mechanism that causes the formation of AD networks. Furthermore, we know that our proposed model MINM can generate networks with a better simulation in network properties of global efficiency and transitivity in both modeling stage from NC to MCI and from NC to AD. It means that the synthetic network has the same ability as the real ones in dealing with the global information transformation.

We have discussed several competitive models which take both topological similarity and distance penalization into consideration in this work. We make a further analysis of the influence of different topological structures on the reconstruction of AD brain networks. In the ECM model, two nodes get higher connection probability if they share more nearest neighbors in common. The PA model considers node degree of one node when calculating the connection probability, and the node with a large degree, i.e., the hub node, has more opportunity to establish connections with others. While, AA, RA, and JC are similar with ECM, they all take the number of common neighbors into consideration when calculating the connection probability between two nodes. Especially, both AA and RA take the individual characteristic of the node, i.e., node degree into account besides the common neighbors. Whereas JC defines the topological similarity for connection establishment with the consideration of not only the number of the common neighbors of two nodes, but also their union neighbor set. As can be seen from Table 4 and Table 5, all the topological-based models can generate synthetic networks much more similar to the real target brain networks than the Random model. Moreover, we can also know that models taking different topological structures into account generate various network topologies which present their characteristics. Additionally, through comparing  $SI$  of MINM to those

of other topological-based models, we confirm that MINM achieves the best performance in generating synthetic networks with properties most resemblance to those of the real AD brain network, with the best fitting parameters  $\lambda$  and  $\eta$  estimated by simulated annealing (SA). We devote our effort to propose an explicit simulation model for a complete understanding of the connection mechanism that results in the network topologies of AD. Also, our results reveal that the formation of the AD topologies is not randomly, it follows an appropriate rule. Our findings draw attention to the critical requirement for a comprehensive understanding of the relationship between the connection mechanism and the functions of the brain networks.

## 5. Conclusions

In conclusion, we have investigated how topological-based mutual information and Euclidean distance are adopted in the simulation of brain network topologies with AD. We concentrate on how connections are established or deleted among different brain regions. Our ambition is to uncover the fundamental connection mechanism that facilitates the alterations of brain networks in the development from NC to AD. We have demonstrated that by adding the mutual information into our model can promote the modeling performance in this development. Successful models of AD brain networks have been instrumental in understanding how structural brain organizations affect the ability of cognition. Our work has opened new avenues toward the diagnosis and treatment of AD.

**Author Contributions:** Si S.Z. designed the study, analyzed the data, and wrote the manuscript. Zhao H. proposed the study conception and interpreted the results. Ding C. and Yu C. took part in the collection of data and images preprocessing. Wang B. conducted experiments and analyzed the results. Liu X. revised the paper. All authors have read and approved the final manuscript.

**Funding:** This work is based upon work supported in part by China. Fundamental Research Funds for the Central Universities under Grant No. N161608001 and No. N171903002.

**Acknowledgments:** We would like to thank the Alzheimer's Disease Neuroimaging Initiative (ADNI) for allowing us to access the database used in this work. ADNI is a longitudinal multi-center study, data collection and sharing for this project was funded by the National Institutes of Health Grant U19 AG024904. ADNI aims at achieving the goal of early detection, tracking, and treatment of Alzheimer's disease (AD).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Alzheimer's Association. 2018 Alzheimer's disease facts and figures. *Alzheimers. Dement.* **2018**, *14*, 367-429.
2. Zhang Z.Q.; Liu Y.; Zhou B.; Zheng J.L.; Yao H.X.; An N.Y.; Wang P.; Guo Y.E.; Dai H.T.; Wang L.N.; et al. Altered functional connectivity of the marginal division in Alzheimer's disease. *Curr. Alzheimer Res.* **2014**, *11*, 145-155.
3. Jessen F.; Amariglio R.E.; van Boxtel M.; Breteler M.; Ceccaldi M.; Ch  telat G.; Dubois B.; Dufouil C.; Ellis K.A.; van der Flier W.M.; et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers. Dement.* **2014**, *10*, 844-852.
4. Cai S.P.; Chong T.; Peng Y.L.; Shen W.Y.; Li J.; von Deneen K.M.; Huang L.Y. Altered functional brain networks in amnesic mild cognitive impairment: a resting-state fMRI study. *Brain Imaging Behav.* **2017**, *11*, 619-631.
5. He Y.; Chen Z.; Evans A. Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. *J. Neurosci.* **2008**, *28*, 4756-4766.
6. Li W.; Wang M.; Zhu W.Z.; Qin Y.Y.; Huang Y.; Chen X. Simulating the evolution of functional brain networks in Alzheimer's disease: exploring disease dynamics from the perspective of global activity. *Sci. Rep.* **2016**, *6*, Art. No. 34156.
7. Li Y.P.; Qin Y.; Chen X.; Li W. Exploring the functional brain network of Alzheimer's disease: based on the computational experiment. *PLoS One* **2013**, *8*, Art. No. e73186.
8. Bullmore E.; Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **2009**, *10*, 186-198.



9. Bullmore E.; Sporns O. The economy of brain network organization. *Nat. Rev. Neurosci.* **2012**, *13*, 336-349.
10. Avena-Koenigsberger A.; Misić B.; Sporns O. Communication dynamics in complex brain networks. *Nat. Rev. Neurosci.* **2018**, *19*, 17-33.
11. Bassett D.S.; Sporns O. Network neuroscience. *Nat. Neurosci.* **2017**, *20*, 353-364.
12. Zuo X.N.; He Y.; Betzel R.F.; Colcombe S.; Sporns O.; Milham M.P. Human connectomics across the life span. *Trends Cogn. Sci.* **2017**, *21*, 32-45.
13. He Y.; Evans A. Graph theoretical modeling of brain connectivity. *Curr. Opin. Neurol.* **2010**, *23*, 341-350.
14. Barban F.; Mancini M.; Cercignani M.; Adriano F.; Perri R.; Annicchiarico R.; Carlesimo G.A.; Ricci C.; Lombardi M.G.; Teodono V. A pilot study on brain plasticity of functional connectivity modulated by cognitive training in mild Alzheimer's disease and mild cognitive impairment. *Brain sci.* **2017**, *7*, 50-61.
15. Vecchio F.; Miraglia F.; Piludu F.; Granata G.; Romanello R.; Caulo M.; Onofri V.; Bramanti P.; Colosimo C.; Rossini P.M. "Small World" architecture in brain connectivity and hippocampal volume in Alzheimer's disease: a study via graph theory from EEG data. *Brain Imaging Behav.* **2017**, *11*, 473-485.
16. Si S.Z.; Wang J.F.; Yu C.; Zhao H. Energy-efficient and Fault-tolerant Evolution Models based on Link Prediction for Large-scale Wireless Sensor Networks. *IEEE Access* **2018**, *6*, 73341-73356.
17. Liu X.; Wang J.F.; Jing W.; de Jong M.; Tummens J.S.; Zhao H. Evolution of the Internet AS-level topology: From nodes and edges to components. *Chin. Phys. B* **2018**, *27*, Art. No. 120501.
18. Jalili M.; Orouskhani Y.; Asgari M.; Alipourfard N.; Perc M. Link prediction in multiplex online social networks. *R. Soc. open sci.* **2017**, *4*, Art. No. 160863.
19. Cannistraci C.V.; Alanis-Lobato G.; Ravasi T. From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. *Sci. Rep.* **2013**, *3*, Art. No. 1613.
20. Newman M.; Barabasi A.L.; Watts D.J. In *The structure and dynamics of networks*; Princeton University Press: Princeton, UK, 2011; ISBN-13: 9780691113579.
21. Ma C.; Zhou T.; Zhang H.F. Playing the role of weak clique property in link prediction: A friend recommendation model. *Sci. Rep.* **2016**, *6*, Art. No. 30098.
22. Wang P.; Xu B.W.; Wu Y.R.; Zhou X.Y. Link prediction in social networks: the state-of-the-art. *Sci. China-Inf. Sci.* **2015**, *58*, 1-38.
23. Zhu L.; Deng S.P.; You Z.H.; Huang D.S. Identifying spurious interactions in the protein-protein interaction networks using local similarity preserving embedding. *IEEE-ACM Trans. Comput. Biol. Bioinform. (TCBB)* **2017**, *14*, 345-352.
24. Pan L.; Zhou T.; Lü L, Hu C.K. Predicting missing links and identifying spurious links via likelihood analysis. *Sci. Rep.* **2016**, *6*, Art. No. 22955.
25. Guimerà R.; Sales-Pardo M. Missing and spurious interactions and the reconstruction of complex networks. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 22073-22078.
26. Vértés P.E.; Alexander-Bloch A.F.; Gogtay N.; Giedd J.N.; Rapoport J.L.; Bullmore E.T. Simple models of human brain functional networks. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5868-5873.
27. Betzel R.F.; Avena-Koenigsberger A.; Goñi J.; He Y.; de Reus M.A.; Griffa A.; Vértés P.E.; Mišić B.; Thiran J.P.; Hagmann P. Generative models of the human connectome. *Neuroimage* **2016**, *124*, 1054-1064.
28. Bassett D.S.; Zurn P.; Gold J.I. On the nature and use of models in network neuroscience. *Nat. Rev. Neurosci.* **2018**, *19*, 566-578.
29. Cabral J.; Kringelbach M.L.; Deco G. Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* **2014**, *114*, 102-131.
30. Zhao T.D.; Xu Y.H.; He Y. Graph theoretical modeling of baby brain networks. *NeuroImage* **2018**, *185*, 711-727.
31. Sporns O. Contributions and challenges for network models in cognitive neuroscience. *Nat. Neurosci.* **2014**, *17*, 652-660.
32. Shannon C.E.; C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379-423.
33. Gibson J. Entropy Power, Autoregressive Models, and Mutual Information. *Entropy* **2018**, *20*, 750-770.
34. Ruiz-Gómez S.J.; Gómez C.; Poza J.; Gutiérrez-Tobal G.C.; Tola-Arribas M.A.; Cano M.; Hornero R. Automated multiclass classification of spontaneous EEG activity in Alzheimer's Disease and mild cognitive impairment. *Entropy* **2018**, *20*, 35-49.
35. Wang Z.; Alahmadi A.; Zhu D.; Li T.T. Brain functional connectivity analysis using mutual information[C]. In Proceedings of the 2015 IEEE Global Conference of Signal and Information Processing (GlobalSIP), Orlando, FL, USA, 14-16 December 2015; pp. 542-546.

36. Sayood K. Information Theory and Cognition: A Review. *Entropy* **2018**, *20*, 706-724.
37. Arevalo-Rodriguez I.; Smailagic N.; i Figuls M.R.; Ciapponi A.; Sanchez-Perez E.; Giannakou A.; Pedraza O.L.; Cosp X.B.; Cullum S. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev.* **2015**, Issue 3, Art. No. CD010783.
38. Jeong J.; Gore J.C.; Peterson B.S. Mutual information analysis of the EEG in patients with Alzheimer's disease. *Clin. Neurophysiol.* **2001**, *112*, 827-835.
39. Garn H.; Waser M.; Deistler M.; Benke T.; Dal-Bianco P.; Ransmayr G.; Schmidt H.; Sanin G.; Santer P.; Caravias G.; et al. Electroencephalographic Complexity Markers Explain Neuropsychological Test Scores in Alzheimer's Disease. In Proceedings of the 2014 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Valencia, Spain, 1-4 June 2014; pp. 496-499
40. Coronel C.; Garn H.; Waser M.; Deistler M.; Benke T.; Dal-Bianco P.; Ransmayr G.; Seiler S.; Grossegger D.; Schmidt R. Quantitative EEG markers of entropy and auto mutual information in relation to MMSE scores of probable Alzheimer's disease patients. *Entropy* **2017**, *19*, 130-143.
41. Hsu W.C.; Denq C.; Chen S.S. A diagnostic methodology for Alzheimer's disease. *J. Clin. Bioinf.* **2013**, *3*, 9-21.
42. Blokh D.; Stambler I. The application of information theory for the research of aging and aging-related diseases. *Prog. Neurobiol.* **2017**, *157*, 158-173.
43. Sulaimany S.; Khansari M.; Zarrineh P.; Daiyanu M.; Jahanshad N.; Thompsonc P.M.; Masoudi-Nejad A. Predicting brain network changes in Alzheimer's disease with link prediction algorithms. *Mol. Biosyst.* **2017**, *13*, 725-735.
44. Supekar K.; Menon V.; Rubin D.; Musen M.; Greicius M.D. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Comput. Biol.* **2008**, *4*, Art. No. e1000100.
45. Yan C.G.; Zang Y.F. DPARSF: a MATLAB toolbox for " pipeline" data analysis of resting-state fMRI. *Front. Syst. Neurosci.* **2010**, *4*, 13-19.
46. Penny W.D.; Friston K.J.; Ashburner J.T.; Kiebel S.J.; Nichols T.E. In *Statistical parametric mapping: the analysis of functional brain images*; Academic Press: London, UK, 2011; ISBN-13: 9780123725608.
47. Tzourio-Mazoyer N.; Landeau B.; Papathanassiou D.; Crivello F.; Etard O.; Delcroix N.; Mazoyer B.; Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **2002**, *15*, 273-289.
48. Liben-Nowell D.; Kleinberg J. (2007) The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci. Technol.* **2007**, *58*, 1019-1031.
49. Newman M.E.J. Clustering and preferential attachment in growing networks. *Phys. Rev. E* **2001**, *64*, Art. No. 025102.
50. Lü L.Y.; Zhou T. Link prediction in complex networks: A survey. *Physica A* **2011**, *390*, 1150-1170.
51. Adamic L.A.; Adar E. Friends and neighbors on the web. *Soc. Networks* **2003**, *25*, 211-230.
52. Zhou T.; Lü L.Y.; Zhang Y.C. Predicting missing links via local information. *Eur. Phys. J. B* **2009**, *71*, 623-630.