

# MSc Thesis : Simulation of Brain Functional Connectivity on Empirical and Randomized Complex Networks, 1st Report

Şeyma Bayrak <sup>\*1</sup> and  
Advisers: Philipp Hövel, Vesna Vuksanović

<sup>1</sup>Integrative Neuroscience MSc, Otto von Guericke University, Magdeburg

14 Apr 2014

## 1 Randomization Methods and Measures of Randomized Networks

The empirical data matrices in our work are obtained from fMRI-BOLD and DW-MRI measures, one is the functional connectivity matrix (FCM) and the other is the anatomical connection probability matrix (ACM). fMRI-BOLD reveals the correlations of functional activity among anatomically labeled brain regions. DW-MRI technique illustrates the probability of anatomical connection by nerve fiber tracks among those regions, which are also called as nodes or vertices.

Our work begins with identifying the characteristic network topologies of FCM and ACM, such as network density, degree distribution, clustering coefficient etc.. In addition to the original graphs built on FCM and ACM, we construct further some random graphs and calculate their characteristic measures. The overall purpose is to understand the conditions that distinguish the original network topologies from that of randomly constructed networks.

Our random graphs depend on our empirical data. The empirical matrix is first converted to its adjacency matrix with a threshold value, which has only ones and zeros. At second, the corresponding real graph of the adjacency matrix is obtained. Thirdly, this graph is reconstructed with some randomization tools. The randomization methods built the new networks by preserving some of the original network measures, e.g. keeping the degree distribution the same as in the original graph, they are summarized in Table 1.

---

\*seyma.bayrak@st.ovgu.de

Abbreviation	Description	Reference
R0	original network	networkx
Ra	same number of nodes and edges in R0	networkx
Rd	swapping two edges in R0	networkx
Rg	same degree sequence in R0	networkx
Rh	same degree distribution in R0	BCT
Rk	swap edges, same degree distribution	BCT

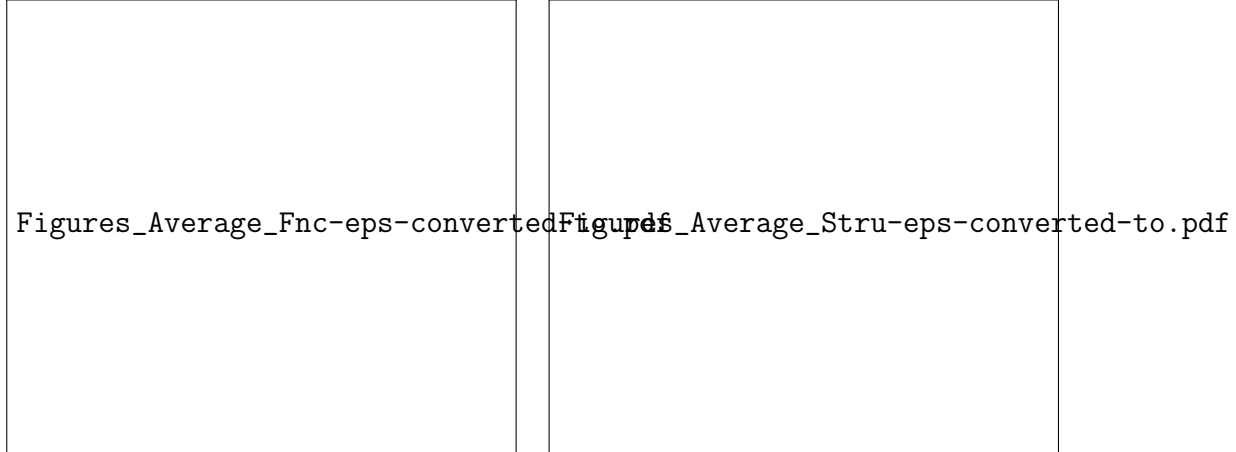
**Table 1:** Descriptions of randomization tools used, networkx is a Python package based on graph theory, BCT : Brain Connectivity Toolbox, Rubinov and Sporns, 2009

## 1.1 Average Degree

Degree (denoted by  $k_i$ ) is simply the number of edges connected to the node  $i$ . Average degree of a network (denoted by  $\langle k \rangle$ ) indicates the ratio of total number of edges,  $L$ , to total number of nodes,  $N$  in a graph.

$$\langle k \rangle = \frac{2L}{N} \quad (1)$$

In order not to count each link twice, the total number of edges is divided by  $\frac{N}{2}$  instead of  $N$ .



**Figure 1:** Average degrees of the original network and the randomized networks. Left:  $R0$  corresponds to the graph of FCM (source is *A\_aal.txt*), right:  $R0$  is that of ACM (source is *acp\_w.txt*). Successful  $r$  ranges for randomization methods of FCM :  $r_{Ra} = [0, 1]$  ,  $r_{Rd} = [0.25, 1.00]$  ,  $r_{Rg} = [0, 1.00]$  ,  $r_{Rh} = [0, 1.00]$  ,  $r_{Rk} = [0.08, 0.94]$ . Successful  $p$  ranges of ACM :  $p_{Ra} = [0, 0.99]$  ,  $p_{Rd} = [0.01, 0.99]$  ,  $p_{Rg} = [0, 0.99]$  ,  $p_{Rh} = [0.05, 0.98]$ .

Degree is one of the statistical tools to measure the centrality of network. The higher the average degree is, the more interaction the nodes in the graph have.

Increasing threshold and probability values diminishes number of edges inverse sigmoidally. As long as the total node numbers, total edge numbers and networks

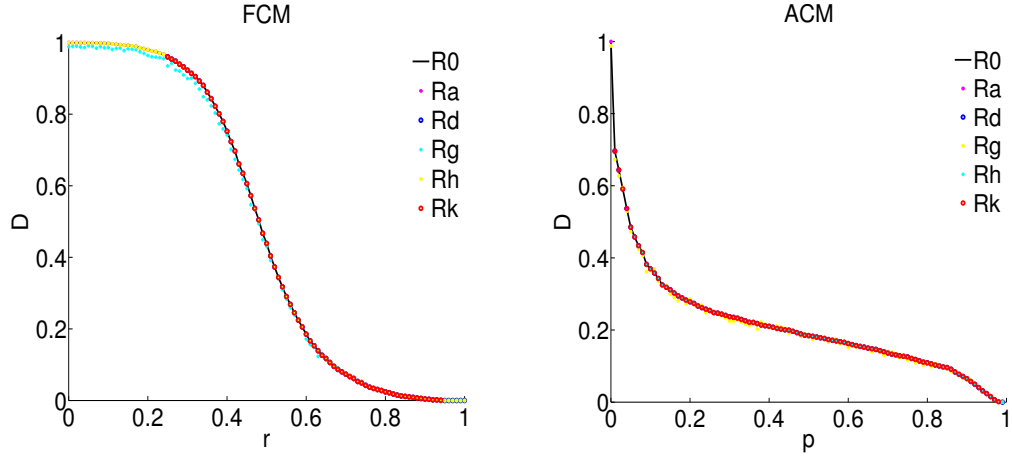
density are all preserved while constructing the random graphs, the average degree remains the same.

## 1.2 Network Density

The density of a network ( $D$ ) is given by the following equation.

$$D = \frac{2L}{N(N-1)} \quad (2)$$

The formula above describes the density as the ratio between the number of total edges and maximum number of possible edges,  $\binom{N}{2}$ . The density is a scaled version of average degree measure.



**Figure 2:** Network density measure of test and random networks are similar to the Figure 1, except that the network density represents a probability value over all possible nodes, therefore it is lying between 0 and 1.

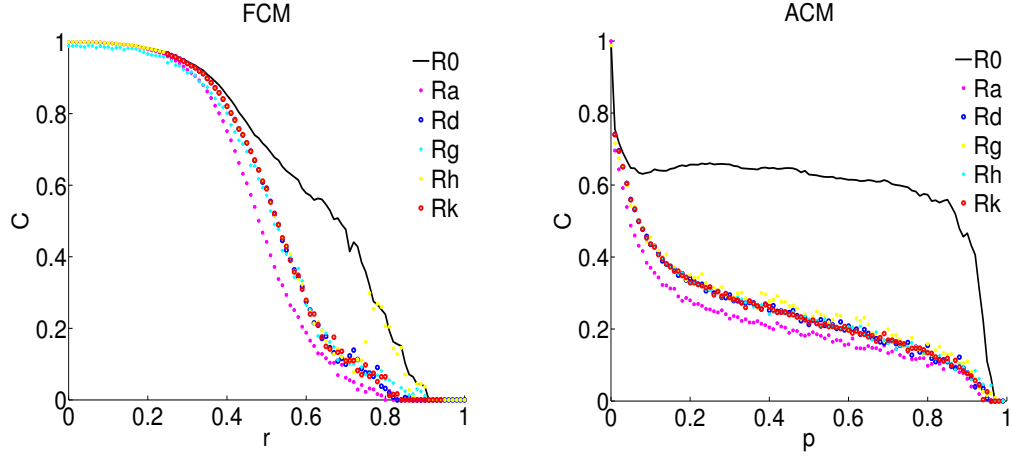
All the networks seem to be densely connected at lower  $r$  and  $p$  values. The FCM graphs reach zero network density faster than ACM graphs which stay almost constant for higher  $p$ .

## 1.3 Average Clustering Coefficient

The average clustering coefficient ( $C$ ) of network is calculated through the clustering coefficients of single nodes ( $C_i$ ):

$$C = \frac{1}{n} \sum_{i \in N} C_i = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)} \quad (3)$$

where  $t_i$  is the number of triangles (triplets) around node  $i$ ,  $k_i$  is the degree (number of links connected to the node) of node  $i$  (Watts and Strogatz, 1998). Clustering coefficient is a measure of segregation, it reveals how the single nodes in a graph cluster together.



**Figure 3:** Average clustering coefficient of the test networks and the randomized networks. All the randomization methods besides *Rk* lead random graphs having slightly lower average clustering coefficients (*cc*) than that of *R0*. However, the method *Rk* has almost the same *cc* as *R0* for both FCM and ACM.

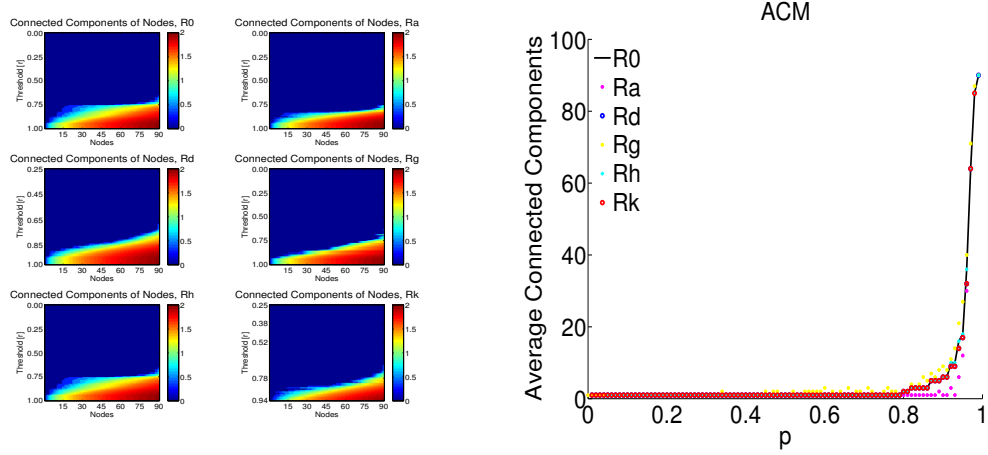
Clustering coefficient is formulated as the ratio of  $t_i$  over all possible edges of the node  $i$ ;  $\binom{k_i}{2}$ . Since this resembles the probability, all *cc* values are between 0 and 1. Figure 3 shows that at lower threshold, the nodes tend to cluster more due to higher number of existing edges. The empirically obtained original network has the highest *cc* compared to random graphs, and has a good agreement with *Rk* method. The real brain functional and structural networks seem to be higher clustered than our randomized networks in general.

## 1.4 Connected Components

The connected components of an undirected graph indicates the number of subgraphs in overall network. Subgraph can be imagined as a connected group of vertices which has globally no connection to any other subgraph. In order to visualize subgraphs algebraically, let us define number of edges  $L$  of graph  $G$  in terms of three subgraphs of  $G$ :  $L_G = L_{G_1} \cup L_{G_2} \cup L_{G_3}$ .

We can expect that the nodes are assumed to be well connected at lower  $r$  and  $p$ , it is always possible to visit any node in the network through the edges, when started from any other node. That means subgraphs should begin to be constructed after a high  $r$  and  $p$  values.

Figure 4 shows that, at higher  $r$  and  $p$  levels, the nodes become sparse as the network densities get lower in FCM and ACM graphs. When  $r > 0.95$  and  $p = 1.00$ , we can imagine each node as a single subgraph, since none of the nodes is connected.



**Figure 4:** Number of connected components of the test networks and the randomized networks. The number of subgraphs remains to be for the threshold values  $r < 0.65$  in the FCM graphs and for the probability values  $p < 0.80$  in the ACM graphs.

## 1.5 Transitivity

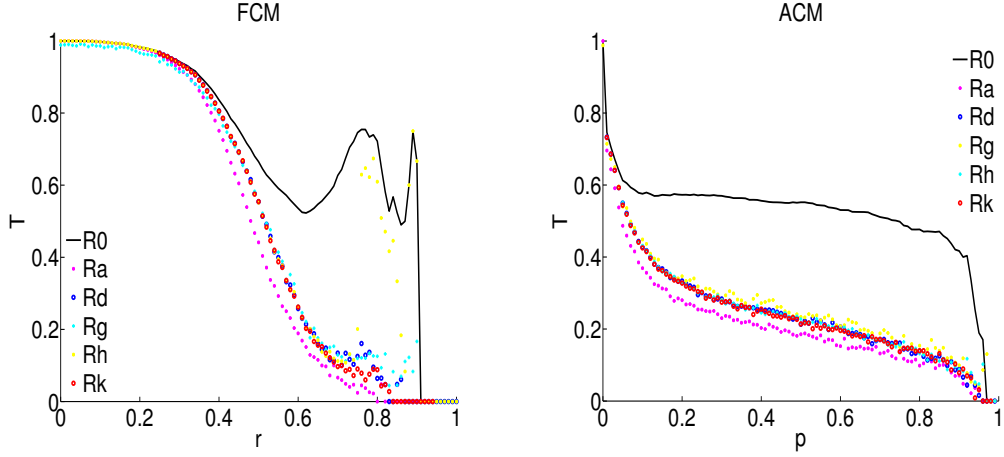
Transitivity is a similar measure to the clustering coefficient, it is also a measure for the segregation in the network. The mean clustering coefficient is normalized individually for each node [RUB10]. The corresponding equation represents the transitivity of a network (Newman, 2003):

$$T = \frac{\sum_{i \in N} 2t_i}{\sum_{i \in N} k_i(k_i - 1)} \quad (4)$$

If a node has links to two other nodes, transitivity inquires whether those two other nodes are also connected to each other. Transitivity is defined only for the whole network rather than single nodes.

In Figure 5, the transitivity of random graphs  $Ra$ ,  $Rd$ ,  $Rg$  and  $Rh$  seems to differ highly from that of the test network  $R0$  for both FCM and ACM. The transitivity of the random graph  $Rk$  tends to match the best with that of  $R0$ . The method  $Rk$  swaps the existing double edges in  $R0$ , as well as keeping degree distribution the same. We already expect that  $Rk$  graph is very similar to  $R0$ , it is expected that the network measures of  $Rk$  are in well agreement with that of  $R0$ .

Transitivity is one of severe shortcomings that real world networks and random networks strongly differ [NEW10]. Transitivity equation resembles the clustering coefficient, however transitivity becomes more sensitive with the changing number of subgraphs (Figure 3, 4, 5). The higher the number of average connected components in the graph is, the more the transitivity measure is distinguishable than clustering coefficient.



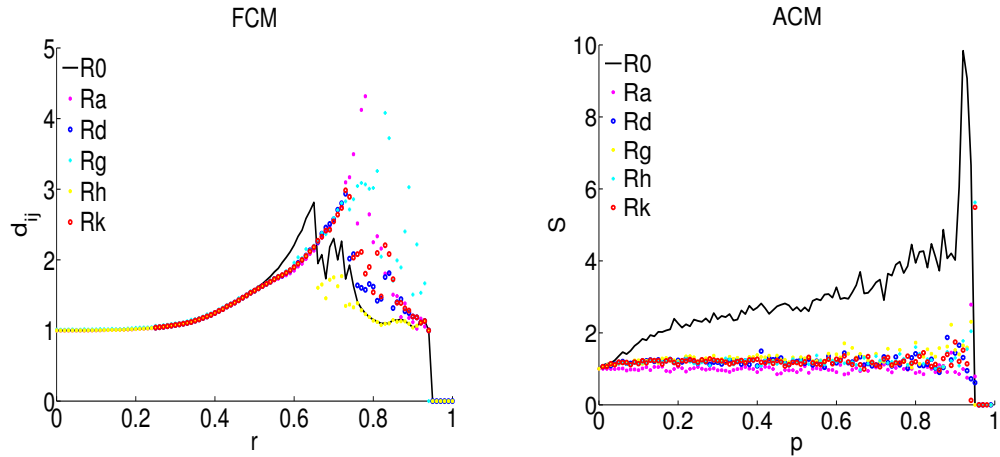
**Figure 5:** Transitivity of the test networks and the randomized networks. The transitivity evolves as the clustering coefficient up to a threshold value.

## 1.6 Shortest Pathway

Shortest pathway is a measure of integration in the network, opposite to the segregation measures. It corresponds to the shortest path length between two nodes in an unweighted graph.

$$d_{ij} = \sum_{a_{uv} \in g_{i \leftrightarrow j}} a_{uv} \quad (5)$$

where  $g_{i \leftrightarrow j}$  is the shortest path between nodes  $i$  and  $j$ .  $d_{ij}$  is assumed to be  $\infty$  for disconnected pairs.



**Figure 6:** Shortest pathway of the original networks and random graphs.

The  $R0$  network of FCM seems to be less segregated than the randomized networks while  $r$  lies between  $[0.65, 0.95]$ . This is the threshold value at which the  $R0$  network of FCM begins to get multiple components. The  $R0$  network of ACM tends

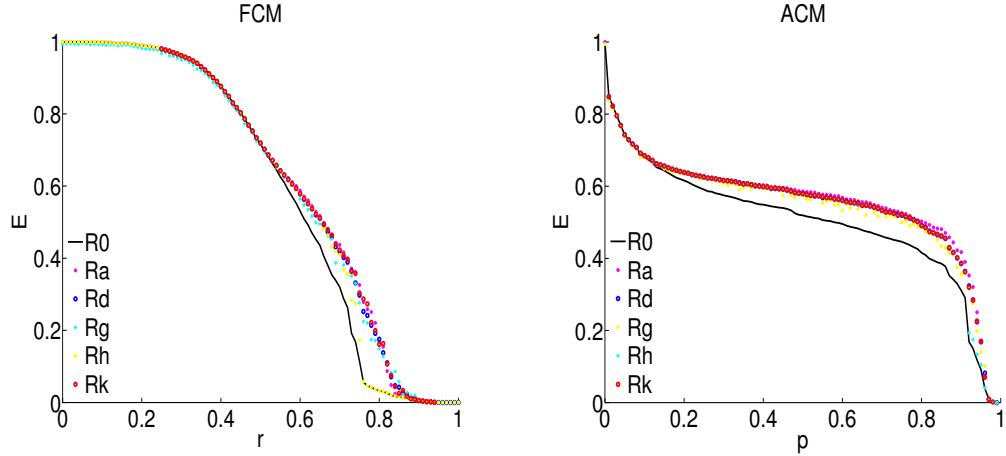
to be more segregated than random networks except for *Rk* method. The graph constructed with the method *Rk* seems to be the most segregated at all for ACM. Whenever all the nodes get sparse in both FCM and ACM networks, the shortest pathway is represented as 0.

## 1.7 Global Efficiency

The global efficiency is measured as the average of the inverse shortest pathway (Latora and Marchiori, 2001);

$$E = \frac{1}{n} \sum_{i \in N} E_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}^{-1}}{n-1} \quad (6)$$

where  $E_i$  is the global efficiency of node,  $d_{ij}$  is the shortest pathway between nodes  $i$  and  $j$ . As seen from the equation, global efficiency becomes larger with smaller shortest pathways between nodes. The global efficiency is a measure of the integration in the network. It reveals the strength of connections in a network. Global efficiency measures the ability of a network to transmit information at the global level (Latora and Marchiori, 2001, 2003).



**Figure 7:** Global efficiencies of the original networks and random graphs; FCM on left side, ACM on right side.

In figure 7, the random graph *Rk* seems to have the lowest global efficiency and best matching global efficiency to the *R0* for both FCM and ACM. Other random graphs have slightly larger global efficiency values than that of *R0*.

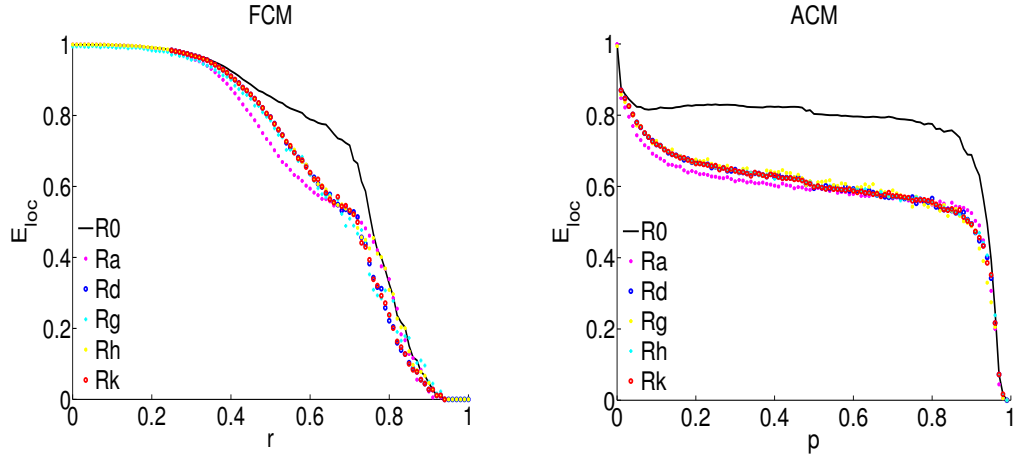
When Figure 6 and Figure 7 are compared, one can see that higher  $d_{ij}$  values reveals lower global efficiency in a network. If it is easier to visit a node starting from any other node in the graph, the information transmission is better.

## 1.8 Local Efficiency

The local efficiency is measured as the average of inverse shortest pathways between nodes in neighborhood of a specific node (Latora and Marchiori, 2001);

$$E_{loc} = \frac{1}{n} \sum_{i \in N} E_{loc,i} = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j,h \in N, j \neq i} a_{ij} a_{ih} [d_{jh}(N_i)]^{-1}}{k_i(k_i - 1)} \quad (7)$$

where  $E_{loc,i}$  is the local efficiency of node  $i$ ,  $d_{jh}(N_i)$  is the shortest pathway between nodes  $j$  and  $h$ , which are located in neighborhood of node  $i$ . Local efficiency measure the ability of a network to transmit information at the local level (Latora and Marchiori, 2001, 2003).



**Figure 8:** Local efficiencies of the test networks and random graphs

The random graph  $Rk$  and the actual graph  $R0$  have relatively higher local efficiency than that of other random graphs for both FCM and ACM. The anatomical connectivity network and its random graphs have in general higher local efficiency compared to the functional connectivity network. In general local information transmit is more efficient in ACM than in FCM. The graphs having higher global efficiencies in Figure 7 tends to have lower local efficiencies in Figure 8.



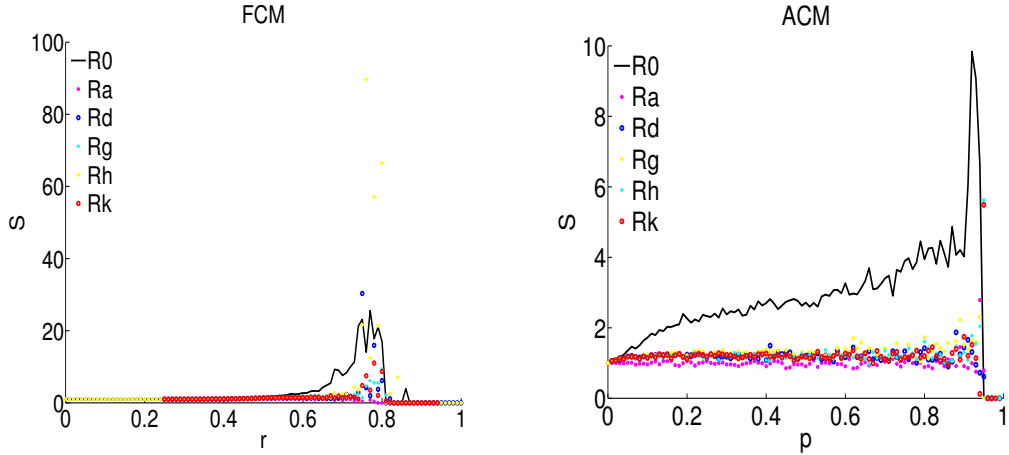
## 1.9 Small Worldness

A small world network is both highly segregated and integrated, a measure of small worldness was proposed to capture this effect in a single statistic (Humpries and Gurney,2008).

$$S = \frac{C/C_{rand}}{L/L_{rand}} \quad (8)$$

where  $C$  and  $C_{rand}$  are clustering coefficients,  $L$  and  $L_{rand}$  are characteristic path lengths of the original and random network respectively. The random network here is constructed with *Erdos-Renyi* method, which has the same number of nodes and links as the reference graph.

$$L = \frac{1}{n} \sum_{i \in N} L_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}}{n-1} \quad (9)$$



**Figure 9:** Small worldness of the test networks and random graphs.

In Figure 9,  $Rk$  graph has the highest small worldness among other random graphs for both FCM and ACM. The usual best matching pattern between  $R0$  and  $Rk$  is lost, when small worldness is measured.

Figure 3 representing cluster coefficients show that the  $R0$  and  $Rk$  networks have larger  $C$  than the other randomly constructed networks. The characteristic pathway  $L$  is truly proportional with the shortest pathway  $d_{ij}$ . Figure 6 gives a hint that  $R0$  network tend to have smaller  $d_{ij}$ , indicating smaller characteristic pathway for it. The division in equation 8 results in such an  $S$  value, much larger 1.

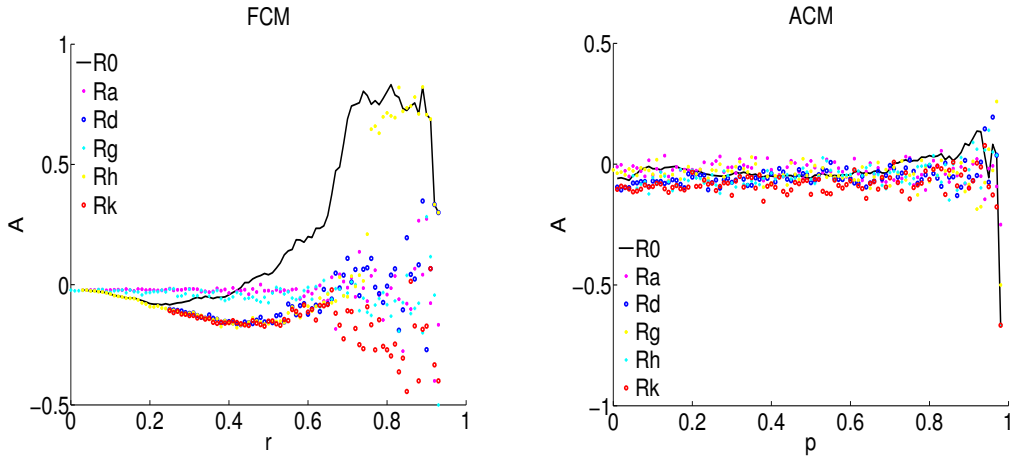
Figure 10 implies that the random networks except for  $Rk$  can not be equally segregated and integrated at the same time. The real world networks  $R0$  for both FCM and ACM tend to have higher small worldness, however  $Rk$  reaches the highest values only at some  $p$  and  $r$  values.

## 1.10 Assortativity

Assortativity measures the correlation coefficient between the degrees of all nodes on two opposite ends of a link [RUB10]. Assortativity coefficient of the network (Newman, 2002);

$$A = \frac{\frac{1}{l} \sum_{(i,j) \in L} k_i k_j - \left( \frac{1}{2L} \sum_{(i,j) \in L} k_i + k_j \right)^2}{\frac{1}{2L} \sum_{(i,j) \in L} (k_i^2 + k_j^2) - \left( \frac{1}{2L} \sum_{(i,j) \in L} k_i + k_j \right)^2} \quad (10)$$

where  $L$  is number of edges in,  $k_i$  is degree of node  $i$ .



**Figure 10:** Assortativity coefficients of the  $R0$  networks and random graphs

In Figure 10,  $R0$  and  $Rk$  graphs have the highest assortativity in general for FCM. However, assortativity values tend to be very similar for all the graphs in ACM.

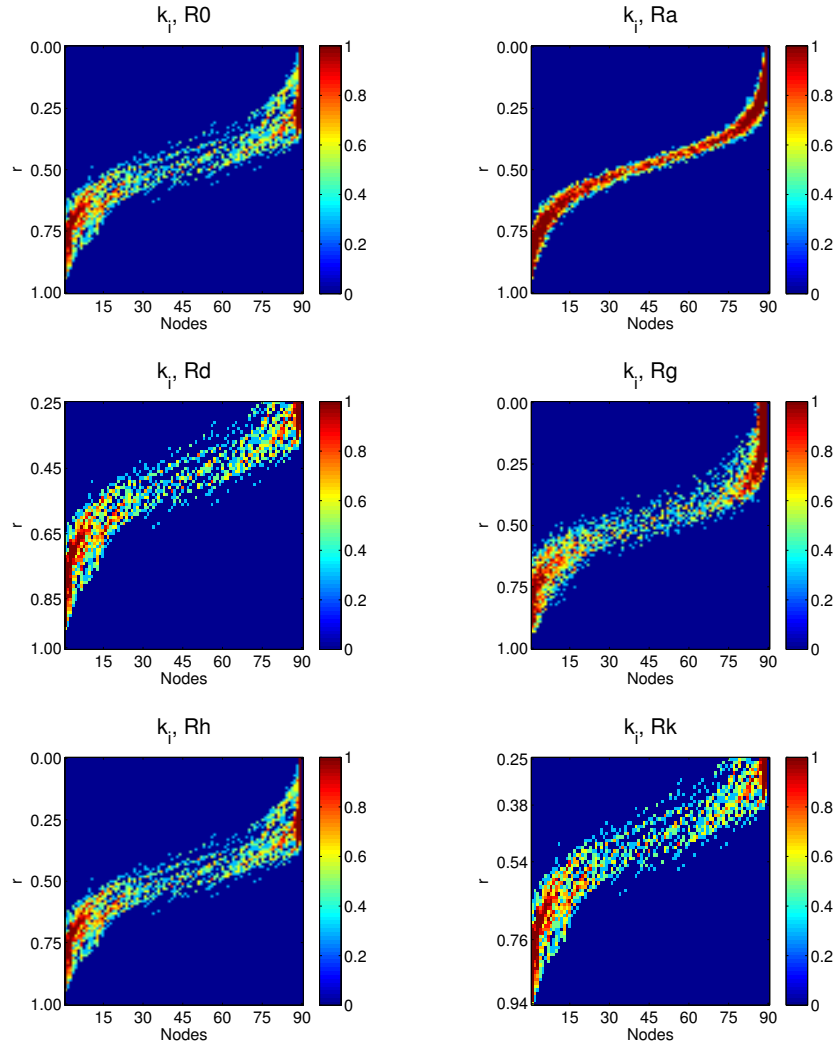
Negative assortativity presents a network having widely distributed high-degree hubs [RUB10]. On the other hand, assortativity coefficients close to 1 indicates a graph having fine correlated degree nodes.

### 1.11 Degree Distribution

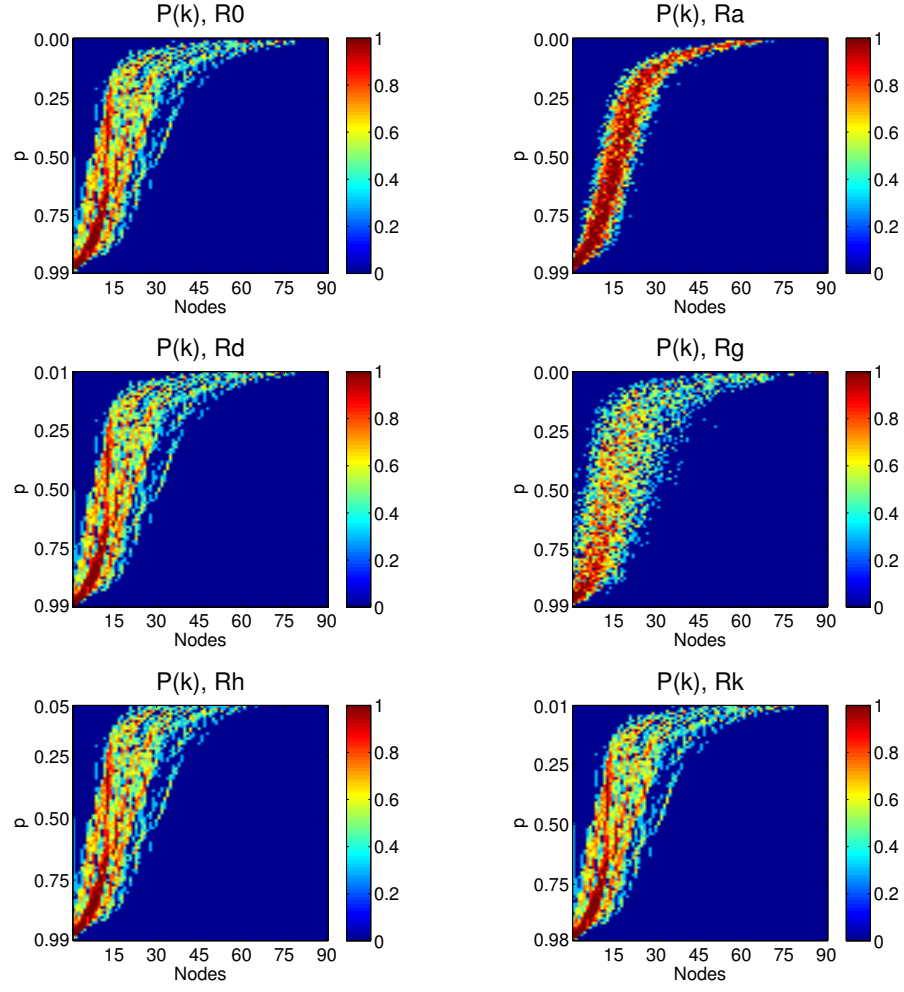
Degree distribution of a network reflects the probability ( $P$ ) of a node to have a given number of degree ( $k$ ). Degree distribution reveals the resilience of a graph.

$$P(k) = \sum_{k' \geq k} p(k') \quad (11)$$

where  $p(k')$  is the probability of a node having degree  $k'$  (Barabasi and Albert, 1999).



**Figure 11:** Heat maps of degree distributions of the FCM network. The limits of colorbar are  $[\log_{10}(10^0), \log_{10}(10^1)]$ .



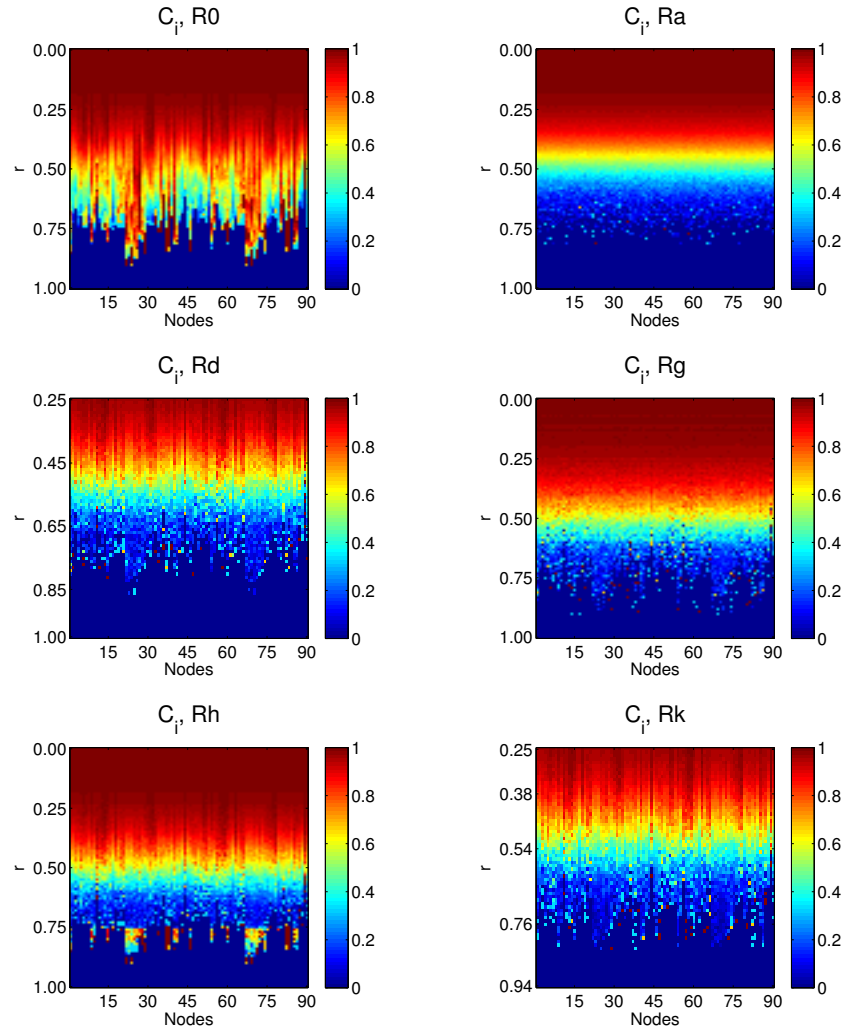
**Figure 12:** Heat maps of degree distributions of the ACM network and the randomized networks. The colors are coded as logarithmic scale. The limits of colorbar are  $[\log_{10}(10^0), \log_{10}(10^1)]$ .

## 1.12 Clustering Coefficient of Nodes

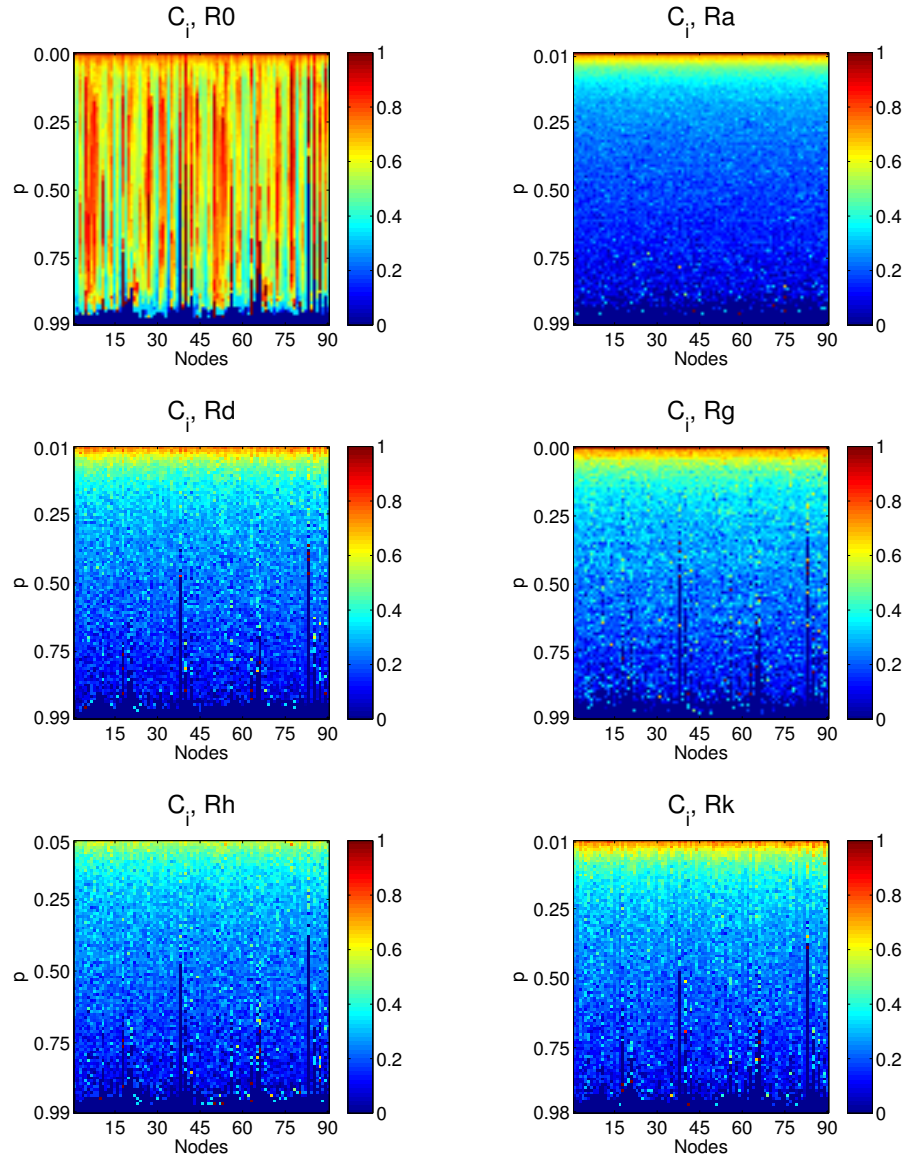
The clustering coefficient of each node is measured as ratio between number of triangles around a node and all possible edge connections of that node  $\binom{k_i}{2}$  (Watts and Stogatz, 1998);

$$C_i = \frac{2t_i}{k_i(k_i - 1)} \quad (12)$$

As the number of triangles around a node increased,  $C_i$  becomes larger indicating more segregated nodes in the network.



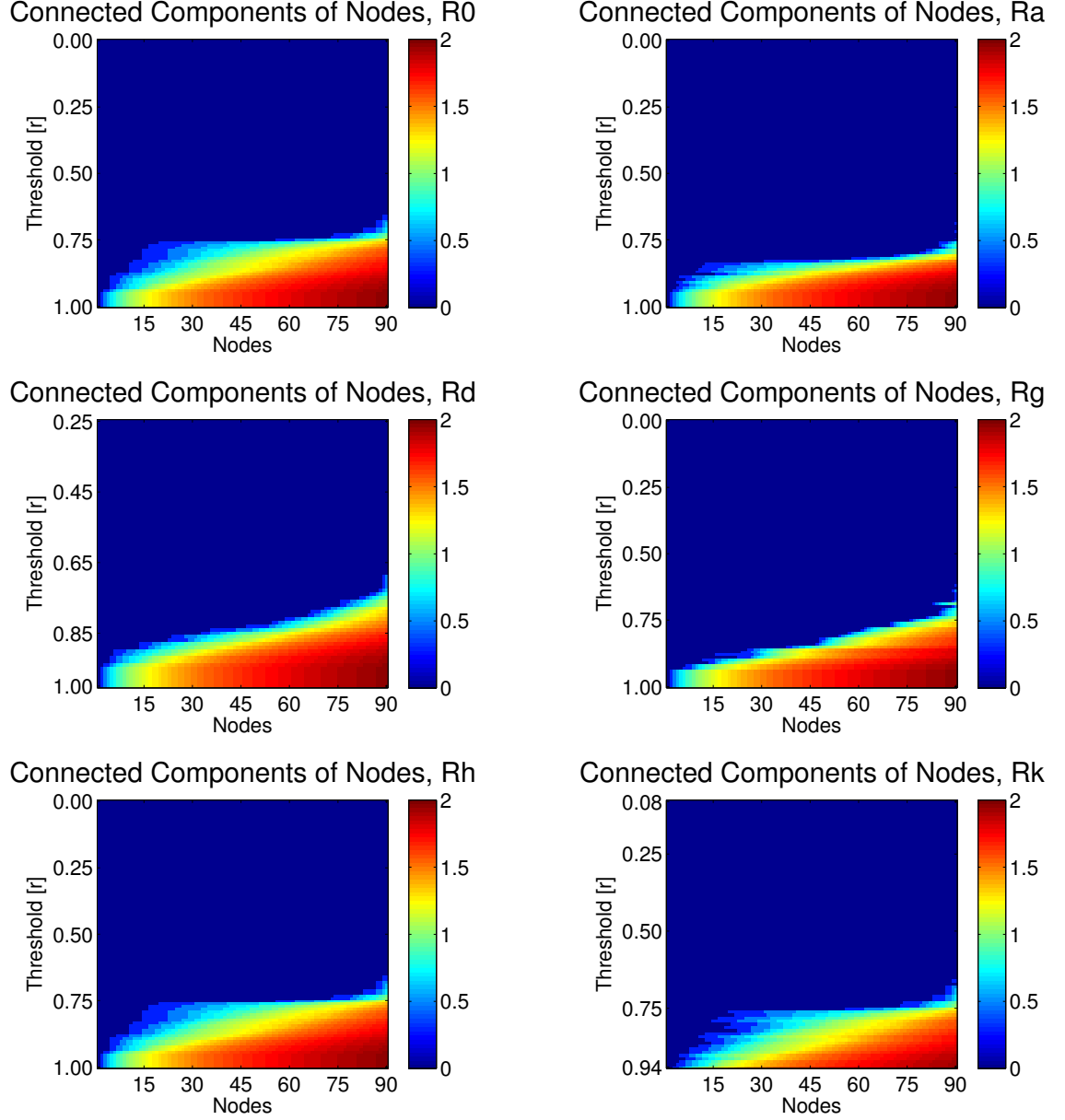
**Figure 13:** Heat maps of clustering coefficient distributions of FCM network.



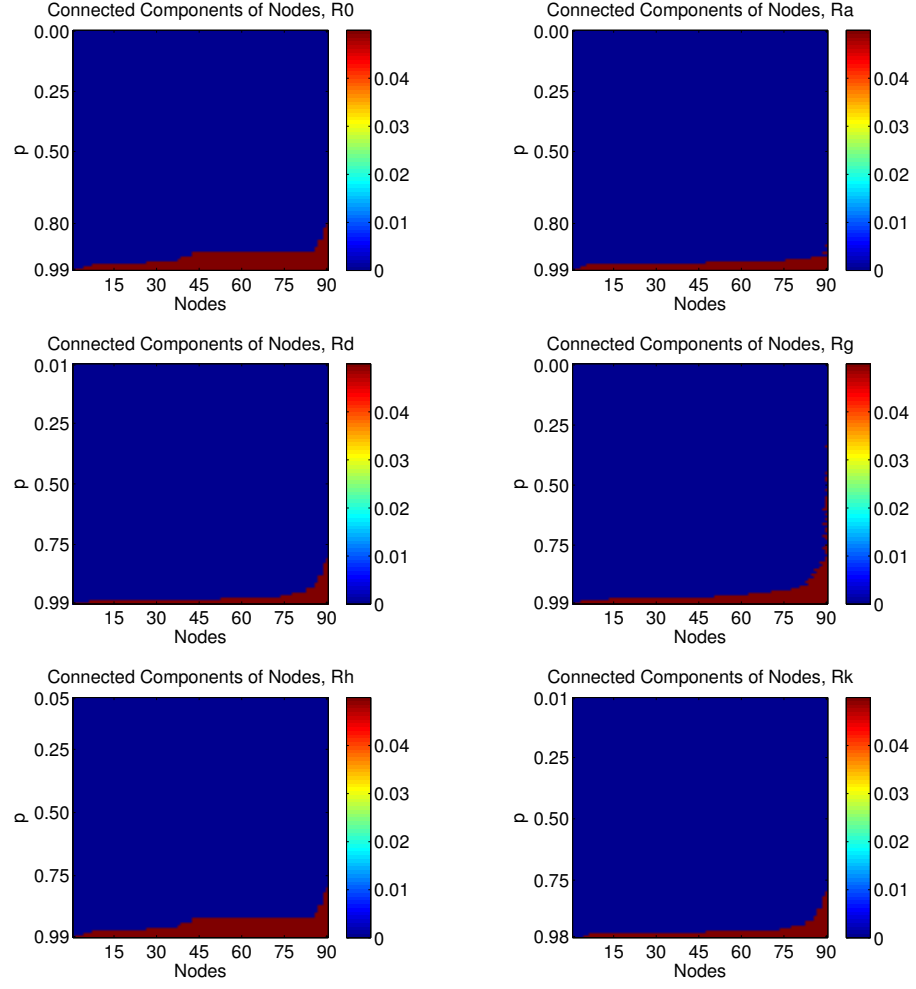
**Figure 14:** Heat maps of clustering coefficient distributions of ACM network.

### 1.13 Connected Components of Nodes

Connected components of nodes tells us how many connections a single node in any subgraph in its isolated neighborhood has. It implies how many other nodes we can reach through that single node.



**Figure 15:** Heat maps of connected component distributions of FCM network. The colors are coded in logarithmic scale, the limits are  $[\log_{10}(10^0), \log_{10}(10^2)]$ .



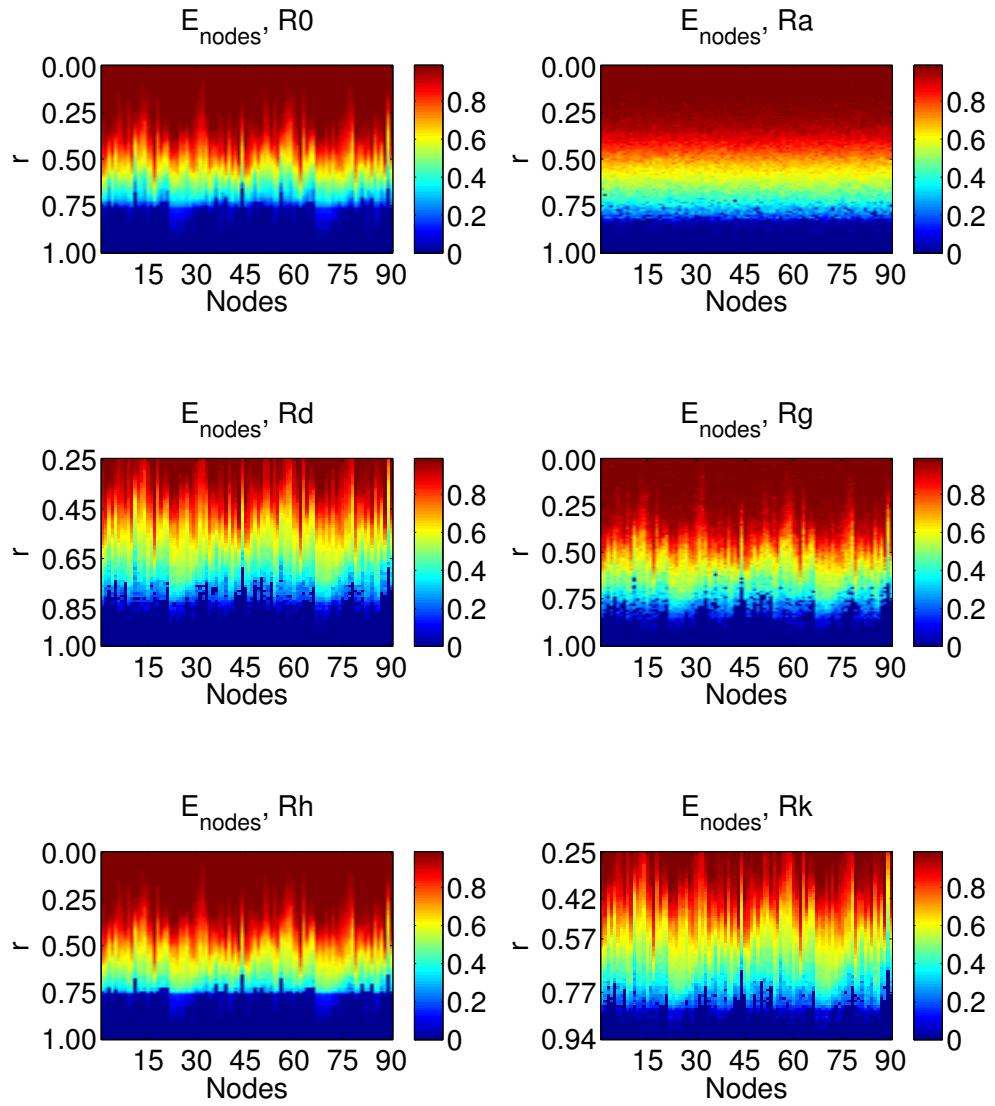
**Figure 16:** Heat maps of connected component distributions of ACM network. The colors are coded in logarithmic scale, the limits are  $[\log_{10}(10^0), \log_{10}(10^2)]$ . The color codes are again in logarithmic scale but connected components are mostly 1, corresponding to blue color  $10^0$ .



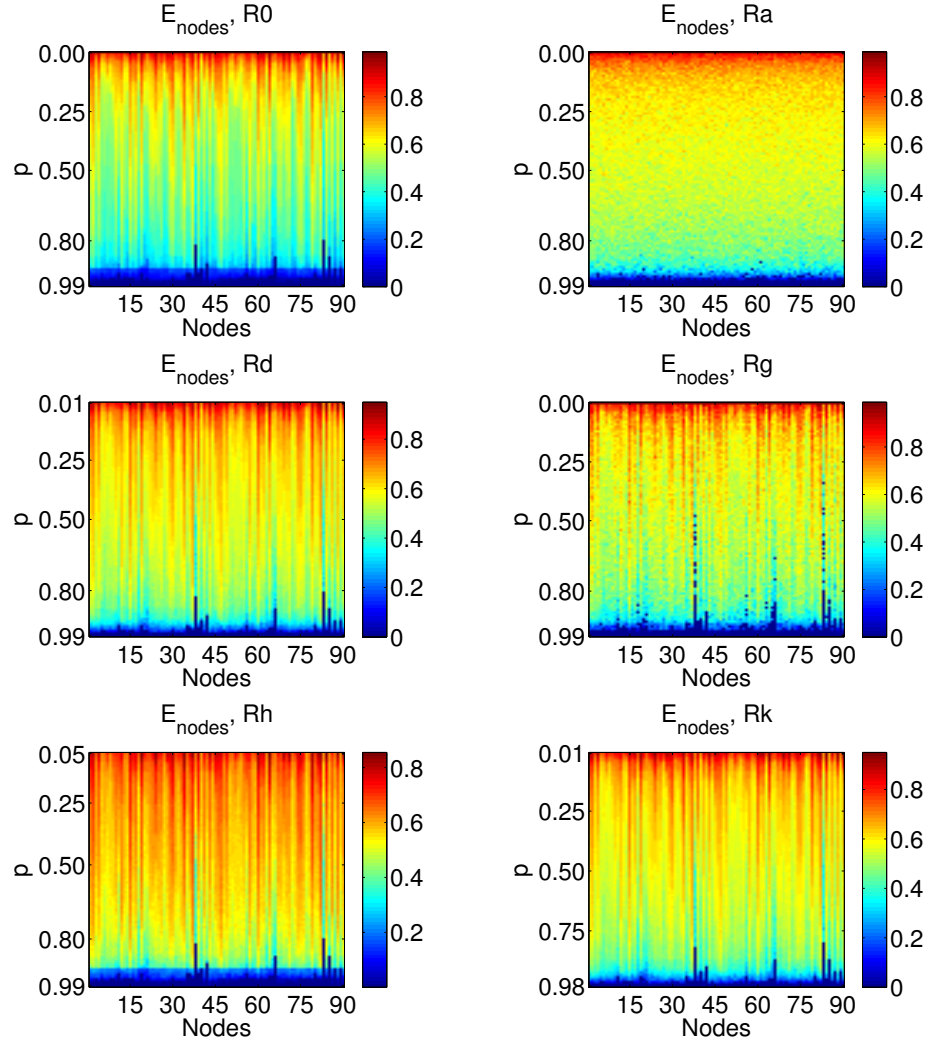
### 1.14 Global Efficiencies of Nodes

Global efficiency of single node can be expressed with the following equation as described previously;

$$E_i = \frac{\sum_{j \in N, j \neq i} d_{ij}^{-1}}{n - 1} \quad (13)$$



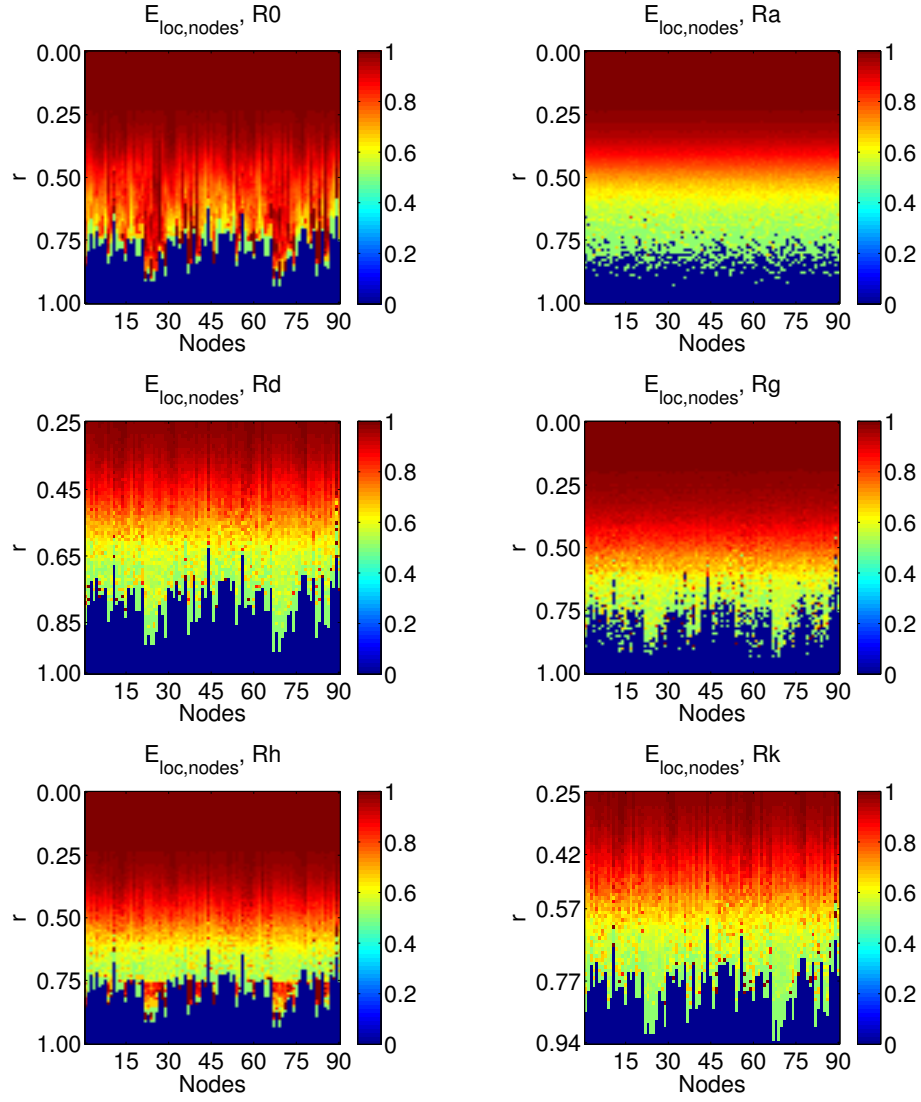
**Figure 17:** Heat maps of global efficiency distributions of FCM network.



**Figure 18:** Heat maps of global efficiency distributions of ACM network.

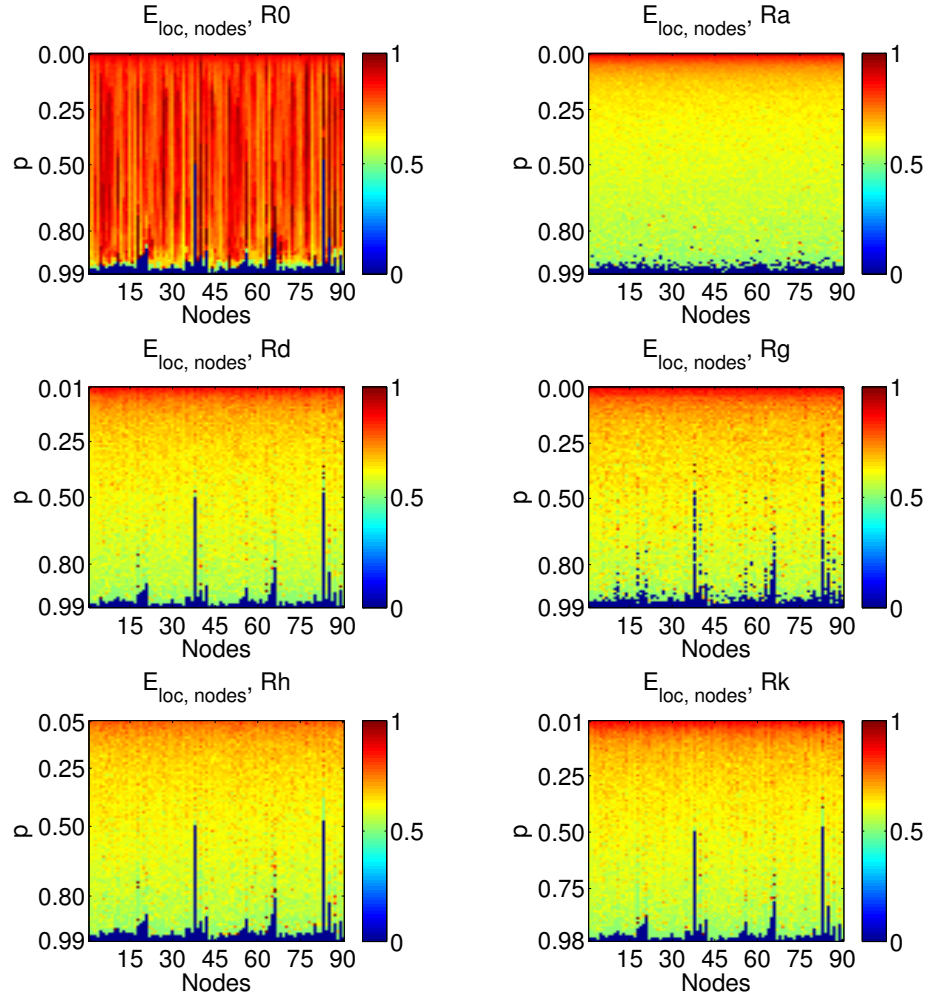
### 1.15 Local Efficiency of Nodes

$$E_{loc,i} = \frac{\sum_{j,h \in N, j \neq i} a_{ij} a_{ih} [d_{jh}(N_i)]^{-1}}{k_i(k_i - 1)} \quad (14)$$



**Figure 19:** Heat maps of local efficiency distributions of FCM network.

asdfasdfasdf[1]asdfasdfasdf



**Figure 20:** Heat maps of local efficiency distributions of ACM network.

## References

- [1] M. Rubinov, O. Sporns, C. Van Leeuwen, and M. Breakspear: *Symbiotic relationship between brain structure and dynamics*, BMC Neuroscience **10**, 55 (2009).