Some Studies of Complex Networks in Multidisciplinary Fields

Thesis Submitted for the Degree of
Doctor Of Philosophy (Science)
of the University of Jadavpur

Abhijit Chakraborty
Satyendra Nath Bose National Centre for Basic Sciences
Block-JD, Sector-III, Salt Lake
Kolkata - 700098, India

July, 2014

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the thesis entitled "**Some Studies of Complex Networks in Multidisciplinary Fields**" submitted by **Sri Abhijit Chakraborty**, who got his name registered on **June 23, 2010**, for the award of **Ph.D. (Science) degree** of **Jadavpur University**, is absolutely based upon his own work under the supervision of **Dr. Subhrangshu Sekhar Manna** and that neither this thesis nor any part of it has been submitted for any degree / diploma or any other academic award anywhere before.

Date: 28/07/2014

Subhrangshu Sekhar Manna
Senior Professor and Dean (Faculty)
Department of Theoretical Sciences
Satyendra Nath Bose National Centre for Basic Sciences
Block-JD, Sector-III, Salt Lake, Kolkata - 700098

Dr. Subhrangshu Sekhar Manna
Professor
S.N. Bose National Centre for Basic Sciences
Block-JD, Sector-III, Salt Lake
Kolkata-700 098

Acknowledgements

First and foremost I want to thank my supervisor, Prof. Subhrangshu Sekhar Manna for his support and guidance to complete my thesis. He has always been available to advise in tough times. The joy and enthusiasm he has for research is motivational and inspirational to me. I am very grateful for his scientific advice, many insightful discussions and suggestions.

I am extremely thankful to Dr. Gautam Mukherjee who was a member of our group. It has always been exciting to collaborate with him.

I would like to thank our past group members, Dr. Kunal Bhattacharya and Dr. Anjan Nandy for their constant encouragement. Special thanks to our present group members Biplab Bhattacherjee, Chandreyee Roy and Sumanta Kundu.

I acknowledge Dr. Punyabrata Pradhan, Dr. Sakuntala Chatterjee and all the members of the Statistical Physics Group for fruitful discussions in journal club meetings.

I also acknowledge the inspiration that I have got from Prof. Jayanta Kumar Bhattacharjee .

I would also like to thank my M. Sc. Project supervisor, Prof. Sitangshu Bikas Santra of IIT-Guwahati for his encouragement to pursue research in Statistical Physics.

I also thank all my friends of SNBNCBS with whom I have spent a lot of time. Specially I want to mention the names of Ambika P Jena, Devraj Roy, Prashant Singh, Soumyajit Sarkar, Sandeep Singh, Sandeep Agarwal, Sujay Pal, Rajiv Nath, Manotosh Chakraborty, Debmalya Mukherjee and Sudipta Kanungo.

I thankfully acknowledge the funding and facilities at SNBNCBS.

I also take this opportunity to express my gratitude to the Monks of the Ramkrishna Mission Calcutta Students' Home, Belgharia for their constant support, inspiration and love throughout my academic career.

Finally, I acknowledge the unconditional support and love from my Father, Mother and Sister.

Abhijit Chakraborty
Satyendra Nath Bose National Centre for Basic Sciences
Block-JD, Sector-III, Salt Lake, Kolkata - 700098

List of Publications

1. **Weighted trade network in a model of preferential bipartite transactions**
   Abhijit Chakraborty and S. S. Manna
   Phys. Rev. E **81**, 016111 (2010).

2. **Conservative self-organized extremal model for wealth distribution**
   Abhijit Chakraborty, G. Mukherjee and S. S. Manna
   Fractals **20**, 163 (2012).

3. **Disease spreading model with partial isolation**
   Abhijit Chakraborty and S. S. Manna
   Fractals **21**, 1350015 (2013).

4. **Space-filling Percolation**
   Abhijit Chakraborty and S. S. Manna
   Phys. Rev. E **89**, 032103 (2014).

5. **Weighted network analysis of earthquake seismic data**
   Abhijit Chakraborty, G. Mukherjee and S. S. Manna
   Submitted to Physica A.

Synopsis of the Thesis

We often come across many networks in our daily lives, for example, the electronic communication network, surface and air transport networks etc. In addition, international trade among different countries defines the economic trade networks, biologists define the protein interaction network and gene regulatory network as the examples of biological networks etc. Therefore, quite expectedly, 'Study of Complex Networks' has been recognized as a multidisciplinary topic.

During the last decade, extensive research efforts have been devoted towards the study of structure, function and activities of different complex networks. Many new information and properties of these networks have been known. Though these networks are random in nature, yet it has become increasingly apparent that the well-known model of Graph Theory like 'Random Graphs (RG)' is no more appropriate to describe these complex networks. In particular, unlike RG, many networks have been observed to have power law degree distributions, (degree of a nodes is the number of links meeting at that node) and consequently few very large degree nodes, called as 'hubs'. Such networks are called 'Scale-free Networks' since they lack a characteristic value for the nodal degrees in the asymptotic limit of very large sizes. Further, it has also been realized that the efficiency of transport networks has been hidden in their 'Small-world' properties. In the present thesis, we report the study of complex networks and related phenomena from the point of view of Statistical Physics from different disciplines of science.

A. Econophysics:

i. Modeling the structure and properties of a Trade Network

The evolution of economic status of a society takes place in terms of mutual trades among its different members, they may be individuals or corporates. To understand the intricacies of the trade dynamics it is necessary to understand the underlying network of mutual trades among different trading members. When a pair of traders take part in a mutual business, a trade relationship is established between them. We have studied a model of trade network where each individual trader or corporate is a node of the network and when two such members take part in a mutual business, a link is established between the corresponding nodes. Using a model of wealth distribution, where traders are characterized by their individual quenched random saving propensities and trade among themselves by bipartite transactions, we mimic the enhanced rates of trading of the rich by introducing the preferential selection rule using a pair of continuously tunable parameters. The bipartite trading defines a growing trade network of traders linked by their mutual trade relationships. With the preferential selection rule this network appears to be highly heterogeneous characterized by the scale-free nodal degree and the link weight distributions and presents signatures of nontrivial strength-degree correlations **(Phys. Rev. E, 81, 016111 (2010))**.

ii. Modeling the self-organized critical evolution of the wealth distribution in a society

In a society, all the individual members tend to improve their economic status. However, poorer the member more is the social pressure felt to uplift its economic condition and consequently the poorest agent feels the strongest pressure. Using the framework of the Pianegonda *et. al.* model, we have studied a conservative

self-organized extremal model based on the above observation with a stochastic bipartite trading rule. More precisely, in a bipartite trade one agent must be the poorest one and the other one is selected randomly from the neighbors of the first agent. The two agents then randomly reshuffle their entire amount of wealth without saving. This model is one of the few examples of non-dissipative self-organized critical systems where the entire wealth of the society is strictly conserved. We estimate a number of critical exponents which indicate this model is likely to be in a new universality class, different from the well established models of Self-organized Criticality. How long a typical agent has to wait to get a chance for a mutual trade? The time interval between two successive updates of an agent is referred as the 'Persistence Time' and it has been observed for the first time that in the stationary state it follows a non-trivial power law distribution **(Fractals, 21, 163 (2012))**.

B. Spreading Phenomena in Networks:

iii. SIS and SIR type disease spreading models with partial isolation on networks

Spreading of an infectious disease from an infected person to other susceptible individuals depends on the existing number of people in the contact neighborhood of the infected person in a population. If it is possible to maintain that the infected individuals are completely isolated, the disease would not spread in the society. However, in all practical cases, this kind of isolation is not perfect but only partial. Here, we studied the effect of partial isolation in disease spreading processes using the well-known models of susceptible-infected-susceptible (SIS) and susceptible-infected-recovered (SIR) models where individuals are located at the nodes of several graphs representing the contact networks in a society. In this model we impose a restriction: each infected individual can probabilistically infect only up to a maximum number n of his susceptible neighbors. Numerical study of this model shows that the critical values of the spreading rates for endemic states are non-zero in both models and decreases as 1/n with n, on all graphs including scale-free graphs. In particular, the SIR model on square lattice with n=2 found to be special case, characterized by a new bond percolation threshold **(Fractals, 21, 1350015 (2013))**

C. Long Range Connectivity in a System of Growing Discs:

iv. Network of growing discs in a plane leading to the Space-filling Percolation

Nature of transition in Explosive Percolation (EP) is studied extensively in recent years. Though wide class of EP models show very sharp change in their order parameters for finite size systems and appear to exhibit discontinuous transition are actually turned out to be continuous transition in the asymptotic limit of the large system sizes. We propose and study a variant of the continuum percolation (CP) model to exhibit a similar discontinuous-like continuous transition. A pattern of circular discs inside a square is generated by filling a large number of growing circular discs one by one at random position with 'slight' overlapping. More elaborately, every disc grows from a nucleation centre that is selected at a random location within the uncovered region. The growth rate δ is a fixed parameter of the model, which is continuously tunable. When a growing disc overlaps with at least another disc, it stops growing and is called to be 'frozen'. Numerical simulation of the

model shows the signature of a discontinuous-like-continuous transition similar to Achlioptas process. Critical area coverage at the transition point approaches to unity at $\delta \to 0$, implying the limiting pattern is space-filling. Fractal dimension of the pore space is found to be 1.42(10) and the contact network of the discs is found to be a scale-free network **Phys. Rev. E, 89, 032103 (2014)**).

D. Properties of the Earthquake Network:

v. Weighted network analysis of earthquake seismic data

We have used the method of Abe *et. al.* to generate a weighted earthquake network associated with the time series of occurrences of the tremors over a long duration and the positions of their epicenters. Here, the entire earthquake region is digitized into a grid, where a cell represents a node if and only if at least one tremor occurs within this cell. In addition, a bond is drawn between every pair of successive events. In our analysis, the number of bonds between a pair of nodes is defined as the weight of the link connecting the nodes. Weighted network is useful to gain better insights about the structural properties and correlations present in the network. It is observed that different properties of the weighted network are quite different from those of their un-weighted counterparts **(Submitted to Physica A)**.

Contents

Chapter 1

Introduction

1.1 Complex Network

'Complex Network' is a research topic of current interest that involves multidisciplinary fields. Researchers from diverse fields such as Physics, Mathematics, Computer Science, Biology, Sociology, Economics etc., have participated in studying different aspects of Complex Networks. Expectedly, different problems attempted in the study of Complex Networks are widely different in nature. In the present thesis, we have investigated the problems of Complex Network in Econophysics, phenomenon of Percolation and Self-organized Criticality. Therefore, in this introductory chapter, we begin with a brief introduction of Complex Network, the Percolation phenomenon, systems displaying Self-organized Criticality and the ideas of Econophysics.

Networks are useful to study the systems as diverse as Internet, world-wide-web, power grid, food web, trade relationships among individuals or countries and social acquaintances etc. In Mathematics, networks are represented by graphs. A graph (network) is defined in terms of a pair of sets: a non-empty set $\mathbf{V}$ of N vertices (nodes) and another set $\mathbf{E}$ of E edges (links). Each element of the set $\mathbf{E}$ corresponds to a pair of nodes of the set $\mathbf{V}$. Many complex systems can be mapped as networks by this simple definition of network, which does not incorporate additional detailed features of the systems. Networks are mathematically represented using the points (nodes) and the lines (links) between them. For example, in the most detailed description of the Internet, computers represent the nodes and the connection between the computers are represented by the links; in a Food Web, various species form the nodes and the links are established by the predator-prey relationship; in the trade network of the individuals, traders are the nodes and business transactions among the traders are represented by the links.

An extremely interesting Mathematical Problem of seven bridges of Königsberg is regarded as the one of the early problems tackled in terms of the Graph Theory. The problem is to find a walk through the city that would cross each bridge only once, with the additional requirement that the trip ends in the same place it began (Fig. 1.1). In 1736, Leonhard Euler showed that such paths are impossible [1]. Originating from Mathematics, Graph Theory and its application spreads in the field of Sociology [2] in the mid-twentieth century. At the same time, mathematicians also did a vast work in the field of graph theory [3]. At the early stage, scientists were

mostly interested in pure mathematical part of the Graph Theory and the graph properties of the network of a small number of nodes. In recent age, the availability of modern computation facilities advanced this field; scientists have become interested in studying vast networks of millions of nodes. Modeling of real-world network, studying structural and dynamical properties of network and dynamical processes on network are the current trends in the complex network field.

However, the breakthrough in attention to the study of networks came after the publication of two seminal papers, one by Watts and Strogatz [4] introducing "Small-world Networks" and the other one by Barabási and Albert [5] introducing "Scale-free Networks".

1.2 Different Types of Networks

Considering additional details of the links, a network can be termed as directed network, undirected network and weighted network. In a directed network, a specific direction is associated with each link. The food web is an example of a directed network, where a directed link is present from prey to predator. On the other hand, an undirected network is the one in which links have no direction. In other words, here all links are bi-directional. For example, a road network is undirected, where links are represented by streets and nodes are represented by crossings; airport network is also undirected, where airports are the nodes and they are connected if there is a direct flight between a pair of airports. In many real-world networks, it is often seen that different links participate in network activities with different strengths. Such networks are represented by a weighted network. For example, in the airport network, passenger traffic of the direct flights between the airports, varies widely from one link to the other. Schematic diagram of the undirected, directed and weighted network is shown in Fig. 1.2.

1.3 Quantities Associated with Networks

Following is the discussion of some quantities that are associated with the description of networks.

1.3.1 Adjacency Matrix

An adjacency matrix is defined to represent the structure of a network. Different elements of the adjacency matrix are either 1 or 0, depending on if the connection exists between the nodes or not. More elaborately, a network of N nodes is presented by a N x N adjacency matrix $\mathbf{A}$. The elements of the matrix are a_{ij}; a_{ij}=1, if there exists a link between i-th and j-th node and a_{ij}= 0, otherwise. For a network with no self-loops, the diagonal elements of the adjacency matrix must be zero. For an undirected network, the adjacency matrix is symmetric. In a directed network a_{ij} = 1, if the link is directed from i to j, otherwise $a_{ij} = 0$. For example, the adjacency matrices $\mathbf{A_1}$ and $\mathbf{A_2}$ for the network shown in Fig. 1.2(a) and 1.2(b) can be written as follows:

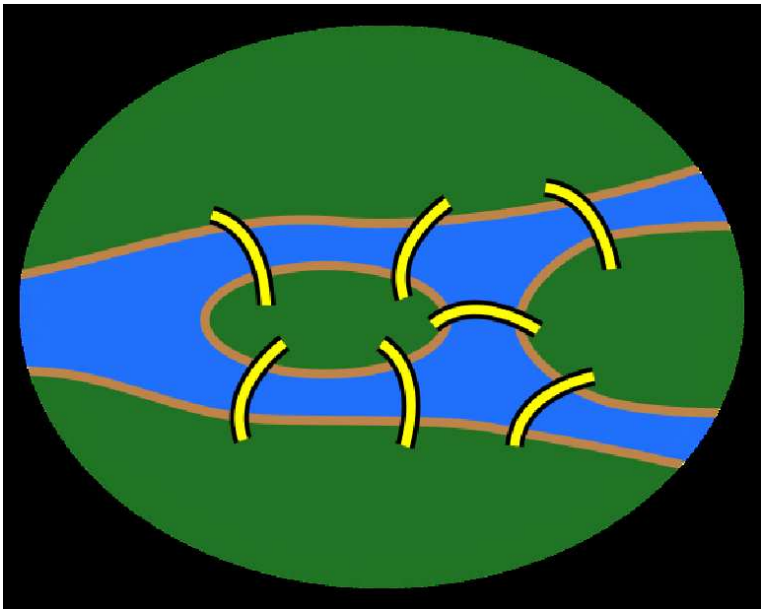


Figure 1.1: Königsberg bridge problem. Figure taken from:
http://en.wikipedia.org/wiki/Seven_Bridges_of_Königsberg

$$\mathbf{A_1} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \text{ and } \mathbf{A_2} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

1.3.2 Nodal Degree

The 'degree' k_i of a node i is defined as the number of links meeting at the node. This can be written in terms of the elements of the adjacency matrix:

$$k_i = \sum_{j=1}^{N} a_{ij}. \tag{1.1}$$

The mean degree of the network is evaluated from the expression

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i = \frac{1}{N} \sum_{i,j=1}^{N} a_{ij}. \tag{1.2}$$

In comparison, two different nodal degree values are defined for the directed networks. The number of links k_i^{in} pointing towards the node i is called the in-degree, where as the number k_i^{out} of links pointing outwards from the same node is called its out-degree. Total degree k is therefore $k = k_i^{in} + k_i^{out}$. For example, in Fig. 1.2(b) node 3 has $k_i^{in} = 2$ and $k_i^{out} = 1$.

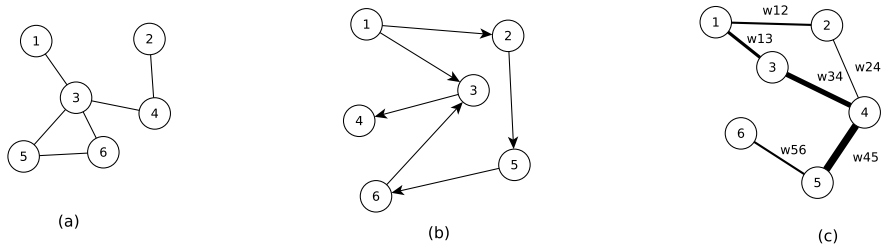


Figure 1.2: The depiction of (a) an undirected, (b) a directed and (c) a weighted network with $N = 6$ nodes and $L = 6$ links. The direction of the links is represented by arrowhead in (b). The thicknesses of the links represent the values of the weights associated with the links in (c).

1.3.3 Probability Distribution of Nodal Degrees

The nodal degree distribution is a very important and widely used quantity of the network structures. It is the basic tool for topological characterization of networks. The degree distribution $\mathcal{P}(k)$ is defined as the probability that an arbitrarily selected node has degree k. In other words, this is the fraction of nodes in the network that have degree k. The n-th moment of $\mathcal{P}(k)$ is defined as

$$\langle k^n \rangle = \sum_k k^n \mathcal{P}(k). \tag{1.3}$$

The first moment $\langle k \rangle$ is the mean degree of the network and second moment $\langle k^2 \rangle$ measures the fluctuation in the degree distribution. In general, networks are characterized by the functional form of their degree distributions. In Random Graphs, links are placed with uniform probabilities. It is known that Random Graphs have Poissonian degree distribution. On the other hand, scale-free networks are characterized by degree distributions having power law decaying tails.

1.3.4 The Clustering Coefficient

The clustering coefficient is a measure of the three point correlation among the local neighboring nodes A, B and C. It measures the probability that A is connected to C, when A is linked to B and B is linked to C. Quantitatively, if E_i is the number of links among the k_i neighbors of i then the clustering co-efficient $\mathcal{C}_i$ of the i-th node is

$$\mathcal{C}_i = \frac{2E_i}{k_i(k_i-1)} = \frac{\sum_{j,m=1}^{N} a_{ij}a_{im}a_{jm}}{k_i(k_i-1)}. \tag{1.4}$$

Therefore, the average clustering coefficient of the entire network is:

$$\mathcal{C} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{C}_i \tag{1.5}$$

According to the definition, $0 \leq \mathcal{C}_i \leq 1$ and $0 \leq \mathcal{C} \leq 1$. The minimum value of the average clustering coefficient of a network $\mathcal{C} = 0$ implies the absence of any triangle among any three nodes of the network. As an example, one can cite a loop-less tree graph. The maximum value of the average clustering coefficient of a network $\mathcal{C} = 1$ indicates all nodes of the network connected to one another and such a graph is called a 'complete graph' or an N-clique. In a more detailed analysis, the clustering co-efficient $\mathcal{C}(k)$ averaged over the subset of nodes of degree k is measured. For many real-world networks $\mathcal{C}(k)$ decreases with increasing k as $\mathcal{C}(k) \sim k^{-\beta_k}$ with $\beta_k \approx 1$. This implies that the neighbors of the large degree nodes are relatively less connected among themselves compared to the neighbors of the small degree nodes.

1.3.5 Degree Correlations and Assortativity

Correlation among nodes of different nodal degrees is measured in terms of conditional probability $\mathcal{P}(k_1|k)$. The conditional probability $\mathcal{P}(k_1|k)$ is the probability that a node of degree k is connected to a node with degree k_1. For correlated networks $\mathcal{P}(k_1|k)$ depends on k and for uncorrelated networks it is independent of k. Most of the real-world networks exhibit such degree correlations, but due to the finite size of the networks the correlation becomes extremely noisy and therefore difficult to estimate.

In an alternative way to measure degree correlation, one defines the average degree of the nearest neighbors of a node i as

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in k_i} k_j = \frac{1}{k_i} \sum_{j=1}^{N} a_{ij} k_j, \tag{1.6}$$

where the first sum runs over all the neighboring nodes k_i of node i. Similarly, $k_{nn}(k)$ is defined as the average degree of the neighbors of nodes of degree k. Therefore, $k_{nn}(k)$ can be expressed in terms of conditional probability by the following expression

$$k_{nn}(k) = \sum_{k_1} k_1 P(k_1|k). \tag{1.7}$$

For uncorrelated networks $k_{nn}(k)$ is independent of k. Correlated networks are called 'assortative' or 'disassortative' depending on $k_{nn}(k)$ increases or decreases with k respectively. In other words, the network is said to be assortative, if high degree nodes are preferentially connected with high degree nodes of the network. Similarly, the network is called disassortative if the high degree nodes are preferentially connected with the low degree nodes of the network.

1.3.6 Scale-free Networks

It has been observed that many real-world networks are highly heterogeneous in their structural connectivities. Few most common examples are the Internet [6], Air-traffic networks [7], Protein-protein interaction networks in Biology [8] etc. These networks have the following common characteristics: a large number of nodes have very small degrees, comparatively less frequently observed are the higher degree nodes and there are few nodes whose degrees are very high, called the hubs of the

network. It has also been observed that the probability distribution of their nodal degrees have power law tails. It is assumed that in the asymptotic limit of very large sizes, the degree distribution $\mathcal{P}(k)$ has the form

$$\mathcal{P}(k) \propto k^{-\gamma}, \tag{1.8}$$

where the constant γ is regarded as a critical exponent whose values are compared for different networks to determine the different universality classes.

For finite size N of the network, however, there exists an upper cut-off k_{max} in the degree distribution which is determined by the network size as $k_{max} \sim N^{\zeta}$. Therefore, the degree distribution data $\mathcal{P}(k, N)$ for different network sizes can suitably be scaled to obtain a good collapse of the data. This finite-size scaling has the following form:

$$\mathcal{P}(k, N) \propto N^{-\eta}\{k/N^{\zeta}\}. \tag{1.9}$$

One of the most well known scale-free network is known as the Barabási-Albert (BA) network [5]. This is a model of a growing network. Starting from a small cluster of few nodes, the network is grown by adding new nodes one by one. In addition, every new node is connected to the growing network by m distinct links. Each of these links selects one distinct node of the growing network using a preferential attachment rule. For BA network it is required that the probability of attachment is linearly proportional to the degree of the node $\pi(k) \propto k$. It has been shown that the exact values of the exponents $\gamma = 3$, $\eta = 3/2$ and $\zeta = 1/2$.

1.3.7 Shortest Paths and Diameter

The 'shortest path' between the nodes i and j is defined as the minimum number of edges required to cross for going to the j-th node, starting from the i-th node. In general, there can be multiple paths between an arbitrary pair of nodes. All possible shortest paths on a network are represented by a matrix $\mathbf{D_p}$ whose elements d_{ij} is the measure of the shortest path between the i-th and the j-th node. For a connected network, the average shortest path is given by

$$\langle d(N) \rangle = \frac{1}{N(N-1)} \sum_{i,j=1}^{N} d_{ij} \tag{1.10}$$

If for a network $\langle d(N) \rangle \propto ln(N)$, then the network is called Small-world Network. Many real world networks, for example biological, technological and social networks [4, 9], such as neural network, electric power grids network, world-wide web [10], railway network [11], co-authorship network [9, 12] and network of movie actors [4] show the small-world property. Diameter $D(N)$ of a network is defined as the largest shortest path between any two nodes of the network. The logarithmic variation of the diameter $\langle D(N) \rangle \propto ln(N)$ is also considered as the signature of the small-world property. The notion of path length plays an important role for transportation on networks.

1.3.8 Weighted Networks

In a real-world network, different links appear with different strengths. In fact, the link strengths are not at all uniform, actually often they vary over wide ranges. For example, the passenger traffic between a pair of airports in airport network [13], strength of the pair-interaction between two species in ecological network [14], the volume of trade between two countries in the international trade network [15,16] and the data traffic in a link of the Internet [17] etc. For these networks, a new variable called 'weight' is associated with every link of the network which represents the strengths of the links. Such networks are known as weighted networks. A weighted network is represented by a weight matrix $\mathbf{W}$, in analogy to the adjacency matrix of unweighted network. Element w_{ij} of the weight matrix $\mathbf{W}$ represents the weight of the link between the nodes i and j. A number of new properties of these networks have come to light when they are analyzed considering the link weights.

1.3.9 Nodal Strength

Strength of a node is the weighted counterpart of degree, measured by the total amount of weight supported by the node:

$$s_i = \sum_{j=1}^{N} w_{ij}. \tag{1.11}$$

When the weights are independent from the topology, we obtain the strength of the nodes of degree k is $s(k) \simeq \langle w \rangle k$ where $\langle w \rangle$ is the average link weight. However, in presence of correlation one gets $s(k) \simeq Ak^{\beta}$ with $\beta \neq 1$ in general.

1.3.10 Weighted Clustering Coefficient

The weighted clustering coefficient of the i-th node of a network is defined as [13]

$$\mathcal{C}_i^w = \frac{1}{s_i(k_i-1)} \sum_{j,m=1}^{N} \frac{(w_{ij}+w_{im})}{2} a_{ij}a_{im}a_{jm}. \tag{1.12}$$

Since, $s_i = k_i(s_i/k_i) = k_i\langle w_i \rangle$ where $\langle w_i \rangle = (s_i/k_i) = \sum_{j=1}^{N} w_{ij}/k_i$ average weight of links associated with node i. $\mathcal{C}_i^w$ can also be written in the form [18]

$$\mathcal{C}_i^w = \frac{1}{k_i(k_i-1)} \sum_{j,m=1}^{N} \frac{(w_{ij}+w_{im})}{2 < w_i >} a_{ij}a_{im}a_{jm} \tag{1.13}$$

$\mathcal{C}_i^w$ shows the amount of strength of the i-th node is associated with the edges of the adjacent triangle. From the definition $0 \leq \mathcal{C}_i^w \leq 1$. In the limiting case, when w_{ij} is either 1 or 0, $\mathcal{C}_i^w$ will be the same as unweighted clustering coefficient $\mathcal{C}_i$. The $\mathcal{C}^w$ and $\mathcal{C}^w(k)$ are the weighted clustering coefficient averaged over all the nodes of the network and the subset of nodes having degree k, respectively. Weighted clustering coefficient reflects the correlation between topology and weights.

1.3.11 Weighted Average Nearest-neighbors Degree

The weighted average nearest-neighbors degree of the i-th node of a network is defined as [13]

$$k^w_{nn,i} = \frac{1}{s_i} \sum_{j=1}^{N} w_{ij} k_j. \tag{1.14}$$

One can define $k^w_{nn}(k)$ which is the average $k^w_{nn,i}$ over the nodes with degree k. It measures the weighted assortative or disassortative properties of the network.

1.3.12 Disparity

We have already mentioned that typically the link weights in a weighted network are highly heterogeneous and they vary over a wide range. The degree of heterogeneity is measured by a quantity, called the 'disparity'. For example, the disparity Y_i of the i-th node is evaluated by the following expression [19, 20]

$$Y_i = \sum_{j=1}^{N} [w_{ij}/s_i]^2. \tag{1.15}$$

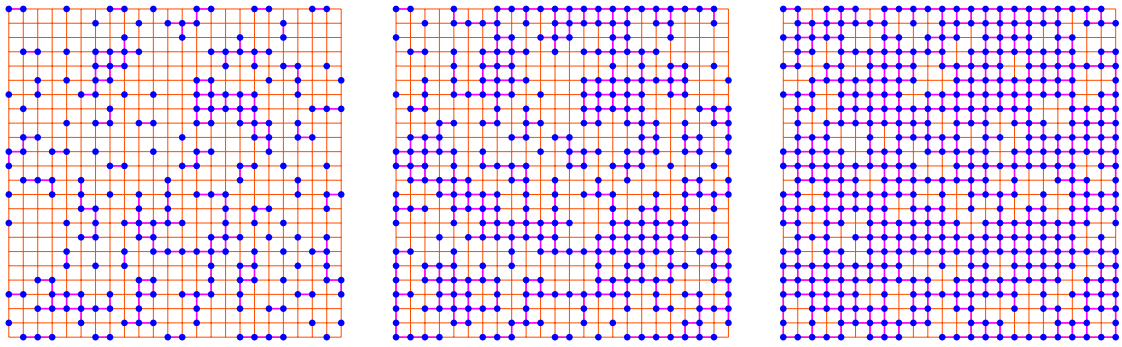
Similarly, $Y(k)$ is the average disparity of the nodes of degree k. If all the links have the weight of same order, then $Y(k) \sim 1/k$ whereas, if one link weight dominates over all other link weights, then $Y(k) \simeq 1$ [21].

1.3.13 Reciprocity

It has been observed that for a real-world directed network, some links are uni-directional and other links are bi-directional. A complete uni-directional network is the 'citation network' where a paper cites papers which are already published, the opposite cannot be true. A complete bi-directional network is actually undirected network. For example, the Internet is a bi-directional network where data is transferred in both directions along the links. Finally, the example of partially directed network, having both uni- and bi-directional links, is the World Trade Network. Here, a country i may export to another country j, but the opposite may or may not be true. The bi-directionality of a network is measured by defining a quantity called 'reciprocity' r [22]. The reciprocity of an unweighted directed network is defined as the ratio between the number of bi-directional links $\overleftrightarrow{L}$ and the total number of links L in the network.

$$r = \frac{\overleftrightarrow{L}}{L}. \tag{1.16}$$

Therefore, $r = 0$ represents a completely uni-directional network and $r = 1$ represents a completely bi-directional network. Most of the real-world networks have an intermediate value of r.



(a) (b) (c)

Figure 1.3: Site percolation on the square lattice of size of size 24 × 24. The site occupation probability p has been 0.4 in (a), 0.6 in (b) and 0.8 in (c). The spanning cluster exists only in (c).

1.3.14 Components

In general, a network may consist of multiple components. A connected component of a network is a sub-network where any two nodes of the sub-network have a path between them. The size of a connected component is measured by the number of nodes in the component. 'Giant component' of a network is the largest component in the network. If there exists a path between any two nodes of the network, the network is called a fully connected network.

1.4 Percolation

The concept of percolation phenomenon was introduced by Broadbent and Hammersley in 1957 [23]. On a regular lattice the edges or bonds are considered to be open with probability p $(0 \geq p \leq 1)$ and closed with probability $(1 - p)$. This process is called the 'bond percolation'. On the other way, if sites instead of bonds are considered to be occupied with probability p and empty otherwise, it is called site percolation. An occupied site / bond takes part in the flow of fluid or particles, but the unoccupied sites cannot. In percolation the stochastic nature of the medium affects the flow of the fluid that distinguishes it from the conventional diffusion process where the particle or fluid shows the stochastic behavior. Though the problem of percolation is simple to state, still a rigorous mathematics is involved with its solution. Percolation models are useful in explaining various phenomena, such as transport process in disorder media, disease spreading process in a population and electrical conduction problem etc.

The main interest in the percolation problem is the associated critical phenomena. Let us consider a square lattice whose sites are occupied randomly with probability p and are left vacant with probability $(1-p)$. As p is increased, at a specific value of p_c, called the percolation threshold, a global connectivity appears across the system. In other words, at p_c there will appear an infinite cluster of occupied sites in an infinite system for the first time (Fig. 1.3). A cluster is defined as a group of occupied sites

connected by the nearest neighbor distances. There is always an infinite cluster extending from one side of the system to another side of the system for $p \geq p_c$ and there is no infinite cluster for $p < p_c$. Therefore, across the percolation threshold p_c, the system changes from a locally connected phase to a globally connected phase. It is known that this phase transition is continuous and the size of the largest cluster is the order parameter for this problem. For square lattice, $p_c = 0.592746$ (site) and $1/2$ (bond) [24].

The problem of percolation has also been studied in continuous space, and it is referred as the continuum percolation (CP). There one finds the minimal density of Lilies, floating at random positions on the water surface of a pond, such that an ant will be able to cross the pond walking on the overlapping Lilies when the radii of the Lilies have a fixed value R [25]. This phenomenon can also be described as how the global connectivity is achieved in a Mobile ad hoc network where each node represents a mobile phone, located at a random position, with a range of transmission R [26]. Depending on R there exists a critical density of Lilies or phones where the long range correlation sets in. It is well-known that in both versions of the CP the transition is continuous and they belong to the same universality class of ordinary lattice percolation [27].

Apart from the studies of percolation on regular lattices, the phenomena of Percolation have been studied on networks as well [28-30]. The most important example is the Random Graph by Erdős and Rényi [3] where the phenomenon of percolation had been studied on the complete graphs. The Order Parameter, which is the size of the largest component, has a very small value of the order of $\log N$ when $p < p_c$. On the other hand, the giant component grows as N for $p > p_c$. Around the critical point, variations of different observables assume power law forms, as follows:
Order parameter near p_c varies as:

$$P(p) \sim |p - p_c|^{\beta} \tag{1.17}$$

Correlation length $\xi(p)$ diverges as $p \to p_c$:

$$\xi \sim |p - p_c|^{-\nu} \tag{1.18}$$

For infinite size systems, the cluster size distribution n_s at p_c has the form:

$$n_s \sim s^{-\tau} \tag{1.19}$$

At the sub-critical regime, only clusters of finite size exist with maximal size s_{max}. s_{max} diverges as $p \to p_c$:

$$s_{max} \sim |p - p_c|^{-1/\sigma} \tag{1.20}$$

The average size of the cluster $\langle s(p) \rangle$ diverges at $p \to p_c$:

$$\langle s(p) \rangle \sim |p - p_c|^{-\gamma} \tag{1.21}$$

Following scaling relations exist among the exponents:

$$\beta = \frac{(\tau - 2)}{\sigma} \tag{1.22}$$

	2d	3d
β	5/36	0.41
ν	4/3	0.875(1) [31]
τ	187/91	2.189(2) [31]
σ	36/91	0.445(10) [31]
γ	43/18	

Table 1.1: The values of the critical exponents in percolation model

$$\gamma = \frac{(3-\tau)}{\sigma} \tag{1.23}$$

$$\nu = \frac{(\tau-1)}{\sigma d} \tag{1.24}$$

The precise values of the critical exponents do not depend on the type of percolation, i.e., if it is site / bond percolation, or on the detailed structure of the lattice. These exponents only depend on the dimensionality d. For example, both site percolation and bond percolation on square or triangular lattices have the same critical exponents as the continuum percolation in two dimensions and they are said to belong to the same 'universality class'. The values of the critical exponents of standard percolation models are shown in Table 1.1.

1.5 Bak-Sneppen Model of Self-organized Criticality

An equilibrium system shows complex behavior characterized by power law at the critical point at the time of phase transition [32]. The system goes to an ordered state from a disordered state when the temperature is tuned. The system shows fluctuation of all length scales right at the critical point. Temperature is tuned properly to reach the critical point. But no one is present in nature to tune the temperature so that the system reaches a critical point and shows complex behavior.

In 1972, paleontologists Stephen Jay Gould and Niles Eldredge said that the apparent equilibrium is only a long period of quiescence between intermittent bursts of activity in which many species become extinct and new species emerge [33]. This phenomenon is called Punctuated equilibrium. The idea of punctuated equilibrium becomes the main mechanism of the dynamics of the complex system. Large intermittent bursts are impossible in an equilibrium system, but are present in biology, economics and history.

We know that complex behavior such as catastrophic events, fractal, $1/f$ noise, and Zipf's law ubiquitous in nature. Such phenomena can not be explained by equilibrium picture. In 1987, Per Bak, Chao Tang and Kurt Wiesenfeld adopted a non-equilibrium picture with the principle of Self-organized Criticality (SOC) [34]. The dynamics of the self-organized critical systems itself evolve to critical state without any fine tuning parameter from the outside. Fluctuations of such systems are measured in terms of avalanches. Avalanches of all length scales exist at the critical point. Sandpile is the canonical example of self-organized criticality.

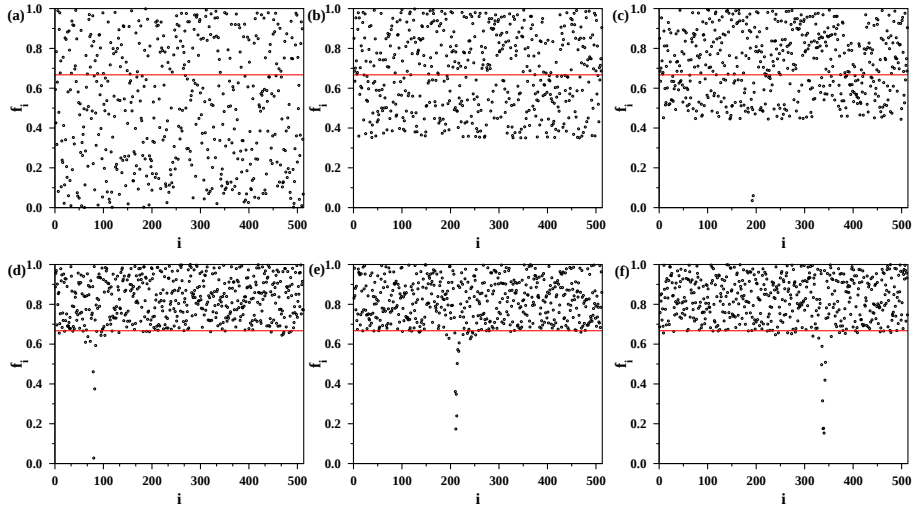


Figure 1.4: Time evolution of the fitness f_i at different sites of an 1d lattice of size $N = 512$. The series of snapshots are taken at time t = (a) 0, (b) 500, (c) 1000, (d) 500,000, (e) 5,000,000, (f) 15,000,000. The position of the threshold value of the fitness $f_c = 0.66702$ is shown by the red line. With time the fitness values gradually move up beyond f_c and stay there in the stationary state during further evolution.

Later in 1993, Bak and Sneppen introduced the SOC model for ecological evolution [35]. Using the idea of Darwinian principles [36] that the less fit species always gets mutated, they presented a stochastic extremal model of interacting species, which shows punctuated equilibrium behavior. They considered a one dimensional lattice as a food chain of L-species with a periodic boundary condition. Each lattice site represents a species. Each species is characterized by a single fitness variable f_i. Initially fitness values of the species are assigned from a distribution of uniformly distributed random numbers between 0 and 1. At an arbitrary time the system evolves according to following two steps: (i) A search is made to find the "active" site i_o which has the minimal fitness f_{min}. This "active" site i_o is get mutated .i.e. its fitness value is replaced by a new random number. (ii) Two nearest neighbors $(i_o + 1)$ and $(i_o - 1)$ of the "active" site i_o are mutated as well. Initially the system is uncorrelated in space, but as time passes the system becomes more and more correlated in space. In the stationary state the fitness distribution $P(f)$ assumes a time independent step like form: $P(f) = 0$, for $f < f_c$ and $P(f) =$ Constant, for $f > f_c$.

Initially f_{min} value is very small. The probability that f_{min} will be replaced by an even smaller fitness after evolution is also small. However, this probability gradually increases as f_{min} increases. As the sites with the minimal values of fitness are systematically replaced, very soon all sites with small f values are replaced by larger values of f resulting a vacancy in the small f region. This is explained pictorially in Fig. 1.4. We plot the lattice positions i along the abscissa and the corresponding fitness f_i along the ordinate on an 1d lattice with $N = 512$. As time increases a vacant region gradually forms for small values of f for all sites. Except for a localized zone where the avalanche occurs, the values of f_i are above the threshold value of the fitness $f_c = 0.66702$ in $1d$ in the stationary state. To

define an avalanche quantitatively, one has to take a reference fitness $f = f_o$. In the stationary state the fitness f_{min} of the active site i_o appears below and above f_o. A f_o-avalanche is defined as the sequence of successive evolution whose f_{min} values are smaller than f_o. The size s of the avalanche measures the duration of the avalanche, i.e., at times t and $(t+s+1)$ the $f_{min} > f_o$, whereas at every time step from $(t+1)$ to $(t+s)$ the $f_{min} < f_o$. When f_o is set equal to f_c it is called a critical avalanche. At the critical point the avalanche distribution has a power law behavior in limit of $N \to \infty$: $P(s) \sim s^{-\tau}$. For a finite system size L, $P(s, L)$ has a power law behavior up to a cut-off $s_{max}(L)$. $P(s, L)$ decreases very fast beyond $s_{max}(L)$. The dependence of $s_{max}(L)$ on L has the following power law behavior:

$$s_{max}(L) \sim L^{\nu} \tag{1.25}$$

The avalanche size distribution $P(s, L)$ obeys the standard finite size scaling ansatz:

$$P(s, L) \propto s^{-\tau} \mathcal{F}(sL^{-\nu}) \tag{1.26}$$

where the scaling function $\mathcal{F}(x) \sim constant$ in the limit of $x \to 0$ and $\mathcal{F}(x)$ approaches zero rapidly for $x >> 1$.

1.6 Econophysics of Wealth Distribution

One of the most important results in this field of Economics and Social Sciences was suggested by Vilfredo Pareto in 1897 [37]. Pareto analyzed the wealth distribution of the individual members in a society. From his empirical study in different countries, Pareto argued that the individual member's wealth distribution assumes a power law decay form

$$P(x) \sim x^{-(\nu+1)}, \tag{1.27}$$

where, $P(x)dx$ is the probability that a randomly selected individual has wealth between x and $x + dx$. From the empirical observations it has been found that the exponent $\nu \simeq 1.5$. However, Pareto suggested that for wealth distribution the exponent $\nu = 1$. Such a distribution is known as the Pareto distribution and the exponent ν is called Pareto exponent. Pareto exponent has been observed to be different for different countries. Some of these values have been listed in Table 1.2 [38]. However, recent studies have revealed that Pareto law is more applicable for the tail end of the wealth distribution of the rich people. The part of the wealth distribution corresponding to the low income group people are more appropriately described by an exponential [39, 40] or a Gaussian distribution [41] (Fig. 1.5).

Application of the concepts of Statistical Physics to the individual wealth or income distribution in a society goes back to 1931 when Saha and Srivastava had suggested that the form of the wealth distribution may be similar to the Maxwell-Boltzmann distribution of molecular speeds in an ideal gas [42, 43]. Similar interest along this line had been devoted when Majorana drew the analogy between Statistical Physics and the Social Sciences in 1936 [44]. Further, a large number of Physicists became highly active in this field in the latter part of the twentieth century, e.g., Stanley [45], Kadanoff [46], Montroll [47] etc. The term Econophysics was coined by H. E. Stanley in a Statistical Physics conference in Calcutta in 1995 [48]. Recently

Country	Year	ν
England	1843	1.50
	1879-80	1.35
	1893-94	1.50
Prussia	1852	1.89
	1876	1.72
	1881	1.73
	1886	1.68
	1890	1.60
	1894	1.60
Saxony	1880	1.58
	1886	1.51
Augsburg	1471	1.43
	1498	1.47
	1512	1.26
	1526	1.13

Table 1.2: Values of the Pareto exponent of different countries over a period of time [38].

Statistical Physics models such as Kinetic Exchange Models (KEM) [49, 50], Percolation models [51], Self-organized Critial models [52] etc. are applied extensively to simulate the wealth distribution of the individual. In a model of Econophysics a large number of individuals, also called agents, interact with one another. These types of models are known as Agent Based models.

1.7 Kinetic Exchange Models of Wealth Distribution

In this model, the society is considered as a collection of N agents, each possesses a certain amount of money equivalent to his wealth x_i, $\{i = 1, N\}$ which he uses for mutual trades with other agents. Typically at the initial stage, all agents start with an equal amount of money $P(x_i, t = 0) = \delta(x_i - a)$. The sequential time t is the number of bipartite trades. A trade consists of two parts, (i) a rule for the selection of a distinct pair of agents i and j, $(i \neq j)$ and (ii) a distribution rule for sharing their total money $(x_i + x_j)$ between them. Different KEMs differ from one another either in the selection rule, or in the distribution rule, or in both. Let, for the $(t + 1)$-th mutual trade, a pair of distinct agents i and j be selected. Every bipartite trade is conservative and therefore, if the i-th agent gains Δx amount of wealth then the j-th agent loses the same amount. The wealth of the agents i and j at time $(t + 1)$ can be written as

$$\begin{aligned} x_i(t+1) &= x_i(t) + \Delta x \\ x_j(t+1) &= x_j(t) - \Delta x \end{aligned} \tag{1.28}$$

By the rule of wealth conservation $x_i(t + 1) + x_j(t + 1) = x_i(t) + x_j(t)$ for all t. Consequently, the net wealth of the society remains conserved throughout its

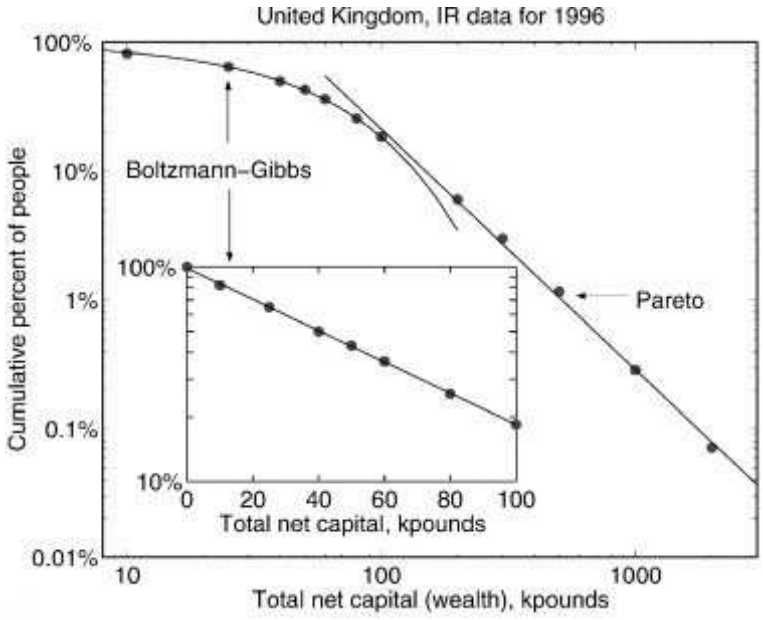


Figure 1.5: Cumulative probability distribution of total wealth in the UK shown on log-log scale. Points: the actual data from the Inland Revenue, solid lines: exponential fit (Boltzmann-Gibbs) and power(Pareto) law fit of the data. Same for the low wealth regime is shown in log-linear scale in the inset. Courtesy: Drăgulescu and Yakovenko [39].

dynamical evolution.

1.7.1 Drăgulescu and Yakovenko (DY) Model

In this model, at time $t+1$, a pair of agents i and j is selected randomly from the set of N agents and their total money $x_i(t) + x_j(t)$ is redistributed between them according to following expression:

$$\begin{aligned} x_i(t+1) &= \epsilon(t)(x_i(t) + x_j(t)) \\ x_j(t+1) &= \bar{\epsilon}(t)(x_i(t) + x_j(t)). \end{aligned} \tag{1.29}$$

Here, $\epsilon(t)$ is a freshly generated random number with uniform distribution within the range $[0:1]$ and $\bar{\epsilon} = 1 - \epsilon$. Starting from a uniform distribution of wealth, the system relaxes to a stationary state, where the wealth distribution assumes a time independent exponentially decaying form. Therefore, in the stationary state the wealth distribution can be written as:

$$P(x) = 1/\langle x \rangle \exp(-x/\langle x \rangle), \tag{1.30}$$

where, without the loss of generality the mean wealth is set at $\langle x \rangle = 1$ [53]. Simulation of the wealth distribution in DY model is shown in Fig. 1.6 for $N = 256$. A nice straight line fit in log-linear scale of the plot justifies the exponential decay of the wealth distribution. The exponential nature of the distribution arises from the fact that the net wealth is conserved and time-reversal symmetry is maintained in the dynamics of the model.

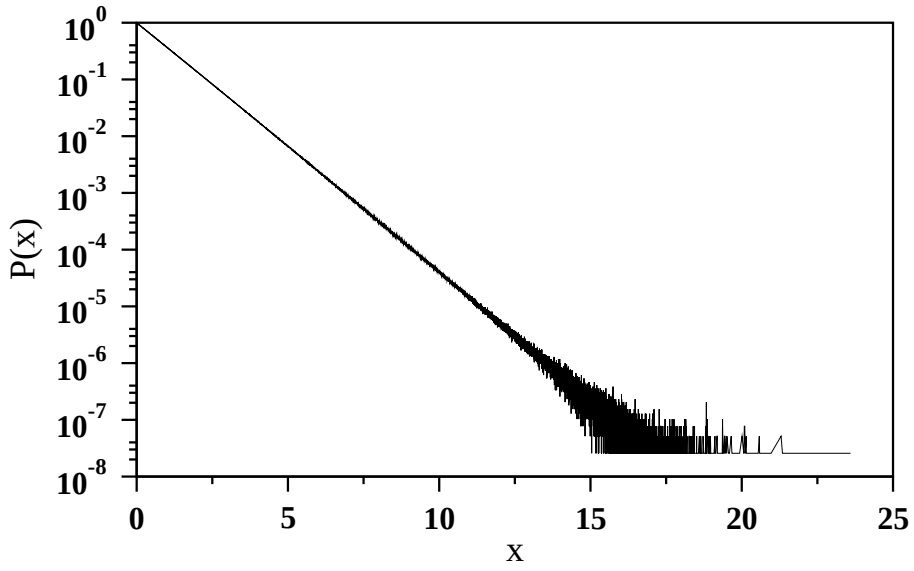


Figure 1.6: Probability density of wealth distribution for DY model has been shown for $N = 256$.

1.7.2 Chakraborti and Chakrabarti (CC) Model

It is a natural tendency of an individual is to save money. Chakraborti and Chakrabarti have incorporated this saving mechanism and therefore modified the DY model. In this model, a parameter λ ($0 \geq \lambda \leq 1$) is introduced as the saving propensity, whose value has been maintained to be the same for all bipartite transactions [54]. The model evolves according to the following equation:

$$\begin{aligned} x_i(t+1) &= \lambda x_i(t) + \epsilon(t)\bar{\lambda}(x_i(t) + x_j(t)) \\ x_j(t+1) &= \lambda x_j(t) + \bar{\epsilon}(t)\bar{\lambda}(x_i(t) + x_j(t)), \end{aligned} \tag{1.31}$$

where, $\bar{\lambda} = 1 - \lambda$. It has been observed that in the stationary state the wealth distribution becomes a single peaked asymmetric function, which fits excellent with a Gamma distribution of the form [55]:

$$P_n(x) = \frac{1}{\Gamma(n)}(n/\langle x\rangle)^n x^{n-1} \exp(-nx/\langle x\rangle) \tag{1.32}$$

where $n = 1+3\lambda/(1-\lambda)$. The distribution has a peak at $x = 1-1/n = 3\lambda/(1+2\lambda)$. Probability density of wealth distribution in CC model is shown in Fig. 1.7 for $N = 8192$ and $\lambda = 1/2$.

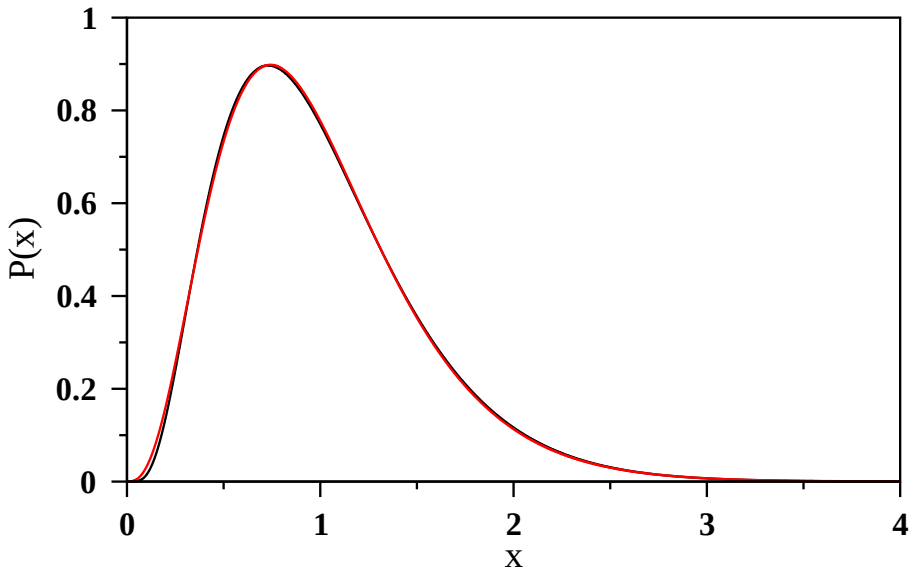


Figure 1.7: (Color online) Probability density of wealth distribution for CC model has been shown for $N = 8192$ and $\lambda = 1/2$. Simulated data have been plotted in black and the Gamma distribution has been fitted in red color.

Chapter 2

Weighted Trade Network in a Model of Preferential Bipartite Transactions

2.1 Introduction

In the last chapter we have seen that the economic status of the members of a society evolves in terms of mutual bipartite trades among different pairs of agents. This is the basic idea of Agent Based Models. It is also well known that the inequality in the economic strengths of the members in the society is an inherent phenomenon. Inequality arises through the economic interactions among the members.

It is also apparent that not all the agents take part in the trading process equally. The rich members of the society take part in the trade more frequently than the poor members. This suggests that there exists an underlying network and the economic evolution of the society takes place on this network. Therefore, to study the wealth distribution of the society, it is important to look into the detailed network structure of the trade dynamics.

In the trade network individual agents are the nodes. A connection or link is formed between a pair of agents when they take part in at least one transaction between them. Here we study the the growth and the structural properties of a trade network in the framework of a well-known model of wealth distribution, namely, the kinetic exchange model (KEM) with quenched random saving propensities [49, 50]. A similar study of the network structure of the International Trade has already been done [15,56]. In this network, different countries represent the nodes of the network, and a link between two nodes exists only when there is a non-zero volume of trade between them.

The trade network to be defined and studied in this chapter has its ingredients in the following KEM model. We first describe this model and then define our model.

2.2 Wealth Distribution Model with Quenched Saving Propensities

In this model Chatterjee, Chakrabarti and Manna (CCM) assigned a distribution of quenched saving propensities $\{\lambda_i, i = 1, N\}$ for the individual agents [57]. This implies that in a typical bipartite trade between the agents i and j, they save fractions λ_i and λ_j of their money and invest the remaining amounts to the mutual trade.

$$\begin{aligned} x_i(t+1) &= \lambda_i x_i(t) + \epsilon(t)(\bar{\lambda_i} x_i(t) + \bar{\lambda_j} x_j(t)) \\ x_j(t+1) &= \lambda_j x_j(t) + \bar{\epsilon}(t)(\bar{\lambda_i} x_i(t) + \bar{\lambda_j} x_j(t)) \end{aligned} \tag{2.1}$$

where, $\bar{\lambda_i} = 1 - \lambda_i$. The wealth distribution in the stationary state, after averaging over many different realizations of the quenched disorder $\{\lambda_i\}$, exhibits a power law decay with a value of the Pareto exponent $\nu \approx 1$ [57]. Subsequent detailed analyses have revealed that the CCM model has many interesting features [58]. For example, in contrast to the DY and CC models, CCM model is not ergodic. Therefore, the wealth distribution is not self-averaging and the single trader wealth distribution is totally different from the overall wealth distribution of the whole society. Consequently, the individual saving propensity factor λ_i plays the role of an identification label that determines the economic strength of a member in the society [58]. In fact, the wealth of a trader fluctuates around a mean value which depends very sensitively on the precise value of λ_i. Larger the value of λ higher is the mean wealth. Truly the wealth distribution averaged over many uncorrelated quenched $\{\lambda_i\}$ sets is the convolution of the individual members' wealth distributions [58]. This overall distribution for the whole system exhibits Pareto law but not the individual member distributions. The exponent $\nu = 1$ has been found to be exactly equal to unity in [59, 60].

2.3 Our Model

Quite generally, rich agents in the society take part in the mutual trade with other agents more frequently than poor agents. In contrast, in the CCM model all traders are selected with equal probabilities. Here, we modify the CCM model by incorporating a mechanism of selecting rich traders with higher probabilities [61]. This enhanced probability has been introduced by a pair of parameters α and β ($\alpha, \beta \geq 0$). A pair of traders i and j, ($i \neq j$) with money x_i and x_j are selected with probabilities

$$\pi_i(t) \propto x_i(t)^\alpha \qquad \text{and} \qquad \pi_j(t) \propto x_j(t)^\beta. \tag{2.2}$$

Clearly, α and β tune the preference of the traders in selection. For $\alpha = \beta = 0$ the model is same as the CCM model. For any non-zero values of α and β the rich traders are selected with larger probabilities. More rich a trader, higher is the probability that it will be selected for trading. After the selection of a pair of traders i and j, now they save λ_i and λ_j fractions of their money and invest the rest amounts to the mutual trade. The total invested amount by both the traders is therefore

$$\delta_{ij}(t) = \bar{\lambda}_i x_i(t) + \bar{\lambda}_j x_j(t), \tag{2.3}$$

where $\bar{\lambda} = 1 - \lambda$. This amount is then randomly divided into two parts and received by them randomly:

$$\begin{aligned} x_i(t+1) &= \lambda_i x_i(t) + \epsilon(t)\delta_{ij}(t) \\ x_j(t+1) &= \lambda_j x_j(t) + \bar{\epsilon}(t)\delta_{ij}(t). \end{aligned} \tag{2.4}$$

where $\epsilon(t)$ is a freshly generated random fraction and $\bar{\epsilon} = 1 - \epsilon$.

At first we estimate the relaxation time required for the system to reach the stationary state. When $(\alpha, \beta) > 0$, according to equation (2.7) richest and next rich are the most probable pair and poorest and the next poor are the least probable pair for selection. Relaxation time, which is the typical time required for the poorest pair to make a trade, can be estimated by considering the maximum wealth $x_{max} \sim N$ (with $\langle x \rangle$=1) and the minimal wealth $x_{min} \sim 1/N$. The probability that the poorest trader will be selected for a trade is $x_{min}^{\alpha}/\Sigma_i x_i^{\alpha}$. Approximating the denominator by its maximum value we get $(x_{min}/x_{max})^{\alpha} \sim N^{-2\alpha}$. Similarly, the selection probability for the next poorest trader is $N^{-2\beta}$ and for the poorest pair is $N^{-2(\alpha+\beta)}$. Therefore the time required for a trade between the poorest pair $T_2 \sim N^{2(\alpha+\beta)}$ and the relaxation time is several multiples of T_2. Thus for any $(\alpha, \beta) > 0$ the relaxation time grows very rapidly with N.

At the early stage rich traders at the top level quickly take part in the trading but gradually the inclusion of relatively poor traders becomes increasingly slower. As a result the number of distinct traders taking part in the trading process grows very slowly. Effectively this implies that the system passes through a very slow transient phase which is practically time independent. We call this state as the "quasi-stationary state (QSS)". It is to be noted that in the following sections we present our numerical results for large system sizes in the QSS only. To ensure that the system has indeed reached the QSS in our simulations we keep track of the quantity $\Sigma_i x_i^2$ and collect the data only after no appreciable change of its mean value is noticed. We mostly analyze the symmetric $\alpha = \beta$ cases except for a few asymmetric cases.

2.4 Wealth Distribution

The average money of a trader $\langle x(\lambda) \rangle$ with saving propensity λ diverges as $\lambda \to 1$ in CCM model [58]. It was also observed that the divergence is like $\langle x_i \rangle (1 - \lambda_i) =$ constant [59, 62]. The reasons behind it is that there is a trader with $\lambda = 1$, who will not invest any money in a trade but it will always get a share from the trading amount, which is invested by other traders. Therefore, the trader with $\lambda = 1$ finally will grab all the money from the society and this phenomena is similar to the condensation.

In Fig. 2.1(a) we plot the quantity $\langle x(\lambda) \rangle (1 - \lambda)$ vs. λ for five different values of $\alpha = \beta = 0$, 1/2, 1, 3/2 and 2. For $\alpha = \beta = 0$ we see the horizontal line as observed in [62]. However for other α, β values the variations of the same quantity are far from being uniform and are monotonically increasing with λ, their growth becoming increasingly faster with α. Therefore we try multiplying this function by $\lambda^{-\chi(\alpha)}$ where $\chi(\alpha)$ is a function of the parameter α. In Fig. 2.1(b) we plot $\langle x(\lambda) \rangle (1 - \lambda) \lambda^{-\chi(\alpha)}$ vs. λ using the same data of Fig. 2.1(a) using $\chi(\alpha) = 0.15$,

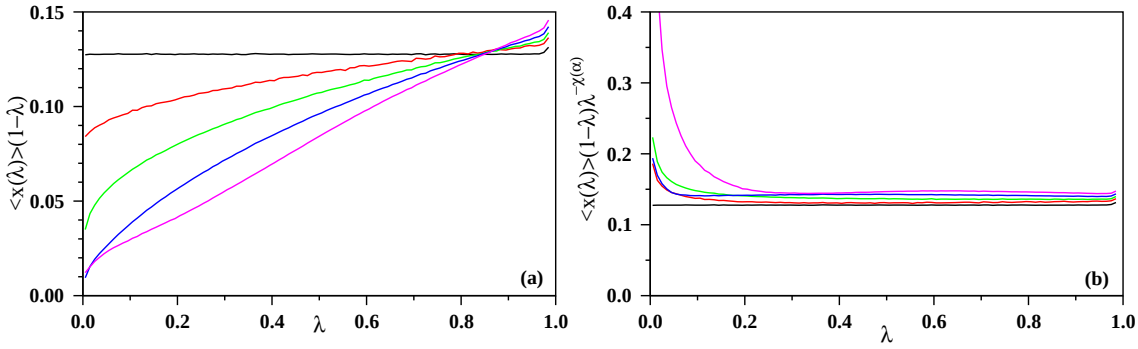


Figure 2.1: (Color online) (a) Plots of $\langle x(\lambda)\rangle(1-\lambda)$ vs. λ for $\alpha = \beta = 0$ (black), 1/2 (red), 1 (green), 3/2 (blue) and 2 (magenta) for $N = 1204$ (α increases from top to bottom). (b) The product function $\langle x(\lambda)\rangle(1-\lambda)\lambda^{-\chi(\alpha)}$ is plotted with λ using $\chi(\alpha)$ = 0.15, 0.35, 0.57 and 0.80 for α = 1/2, 1, 3/2 and 2 respectively using the same colors as in (a).

0.35, 0.57 and 0.80 for $\alpha = \beta$ = 1/2, 1, 3/2 and 2 respectively. We get nearly uniform variations between $\lambda = 0.3$ and 1. We assume that

$$\langle x(\lambda)\rangle(1-\lambda)\lambda^{-\chi(\alpha)} = constant. \tag{2.5}$$

If the distribution of λ values is denoted by $g(\lambda) = constant$, and since the wealth x and the saving propensity λ are the two mutually dependent variables associated with the same trader, their probability distributions follow the relation [59]

$$P(x)dx = g(\lambda)d\lambda. \tag{2.6}$$

Differentiating Eqn. (2.5) with λ one can find out the derivative $d\lambda/dx$ and substituting in Eqn. (2.6) one gets

$$P(x) = \frac{C}{x^2}[\lambda^{-\chi} + (1-\lambda)\chi\lambda^{-\chi-1}]^{-1}. \tag{2.7}$$

For this equation we see that for large λ the term within [..] is of the order of unity. Therefore in this range $P(x) \sim x^{-2}$ as in the Pareto law. This is an indication that even for $(\alpha, \beta) > 0$, Pareto law holds good and in the following we present numerical evidence in support of that.

The system is prepared by assigning uniformly distributed random fractions for the saving propensities λ_i to all N traders. Here λ_is are quenched variables and therefore they remain fixed during the subsequent time evolution of the trading system. Consequently all observable that we measured are averaged over different uncorrelated sets of the $\{\lambda_i\}$ values. While assigning the λ values we first draw N uniformly distributed random fractions, but then to avoid the situation when λ_{max} is very close to unity by chance we scale them proportionately so that $\lambda_{max} = 1 - 1/N$ in every $\{\lambda_i\}$ set. First a pair of values for (α, β) is selected. Two types of initial wealth distributions are used: (i) $x_i = 1$ for all i and (ii) x_is are uniformly distributed random numbers with $\langle x\rangle$ = constant. The sequence of bipartite trading begins by randomly selecting pairs of traders using Eqn. 2.2. Once a pair is selected, their

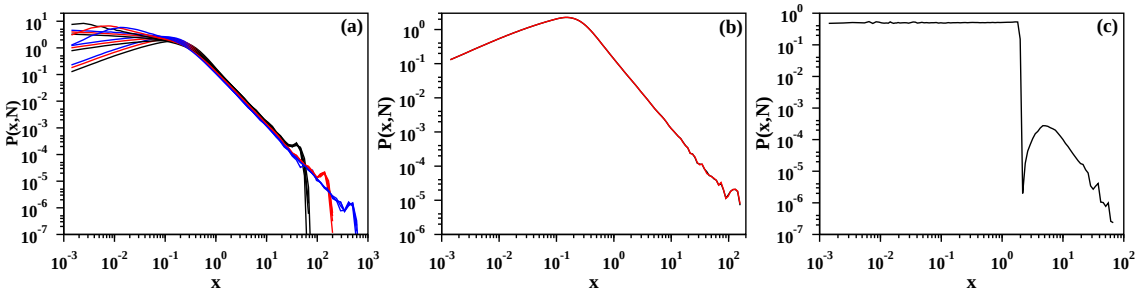


Figure 2.2: (Color online) (a) Wealth distribution $P(x, N)$ vs. x for $\alpha = \beta = 0$, 1/2, 1 and 2 and for $N = 256$ (black), 1024 (red) and 4096 (blue) (N increases from left to right). The slopes of these curves yield $\nu = 1.00(3)$ consistent with the Pareto law. (b) $P(x, N)$ for $(\alpha, \beta) = (\infty, 0)$ (black) and $(0, \infty)$ (red) for N= 1024 which almost overlapped. (c) $P(x, N)$ for $(\alpha, \beta) = (\infty, \infty)$, the distribution is uniform followed by a hump due to transactions between the richest and the next richest traders only.

total individual invested amount δ_{ij} is calculated using Eqn. 2.3 and this amount is shared again between them using Eqn. 2.4. This constitute a single bipartite trading and the dynamics is followed over a large number of such trading events.

The wealth distribution changes with time from the initial distribution to more and more flat distribution. After a certain time the system passes through the quasi stationary state when no appreciable change in the wealth distribution is observed. It is also observed that the distribution is robust with respect to the precise values of the parameters α and β used. In Fig. 2.2(a) the wealth distribution $P(x, N)$ has been plotted with x for four sets of parameter values namely, $\alpha = \beta = 0$, 1/2, 1 and 3/2 and for three system sizes $N = 256$, 1024 and 4096. Apart from slight fluctuations the four curves for a given system size nearly overlap on one another. On a double logarithmic scale the slopes of the curves give an average estimate for $\nu = 1.00(3)$ consistent with the Pareto law as observed in the CCM [57]. This indicates that the wealth distribution is robust with respect to the parameter values in this region. The non-zero values of α and β only controls the frequencies with which different traders are called for trading.

Next we consider the case when one of the two parameters (α, β) is infinity and the other one is zero. If $\alpha = \infty$ the richest trader is always selected as the first trader. The other trader is selected among the other $N - 1$ traders with uniform probability. As shown in Fig. 2.2(b) here also we see that the Pareto law holds good. For $\alpha = \infty$ and for finite β first the richest trader is selected and then the second trader is selected with probability $\propto x_j^{\beta}$. We observe numerically that here also Pareto law works very well.

However the situation is very different when both (α, β) take very large values. In this situation almost always only the rich traders are called for transactions. The system passes through an extremely long QSS and the number of traders taking part in trade does not increase at all. For example in the limiting case of $(\alpha, \beta) = (\infty, \infty)$ it implies that always only the richest and the next richest traders are selected for transactions with probability one but not any other trader. If their wealths are

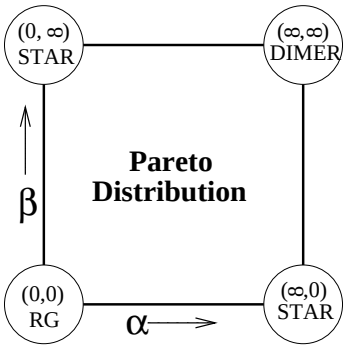


Figure 2.3: The phase diagram in the positive quadrant of the (α, β) plane. The Pareto law is valid in the entire region. The origin corresponds to the CCM model where the trade network is a random graph. At the corners $(\infty, 0)$ and $(0, \infty)$ the richest trader trade in every transaction, so that the network is a star-like graph but the wealth distribution still follows Pareto law as shown in Fig. 2.2(b). However at the corner (∞, ∞) the trade takes place between the top two richest traders and the network shrinks to a single dimer. The wealth distribution here is uniform followed by a hump as shown in Fig. 2.2(c).

very high then the trading will be limited only between them. Therefore the wealth distribution for the single λ_i set has two very high peaks and wealths of all other traders are small and uniformly distributed. Consequently the quench averaged wealth distribution is uniform throughout followed by a hump at the highest value of wealth (Fig. 2.2(c)). A systematic analysis with many different (α, β) pairs leads us to conclude that Pareto law holds good in the positive quadrant of the entire (α, β) plane.

In Fig. 2.3 we exhibit this behavior in the positive quadrant of the (α, β) plane where Pareto law is valid and the limiting points are marked by circles with their characteristics. The origin at $(\alpha = 0, \beta = 0)$ represents the CCM model where traders are selected randomly with uniform probabilities. As explained below, the trade network corresponding to this point is a random graph (RG). As explained before that at the two corners $(\infty, 0)$ and $(0, \infty)$ the richest trader always participates in every transaction. Therefore the corresponding trade networks have star-like structures. In the last corner of (∞, ∞) the trade takes place only between the richest and the next rich traders and therefore the graph reduced to a single dimer only.

2.5 The Trade Network

One can associate a network with this trading system. Each trader is a node of the network. Initially the network has only N nodes but no links. First the system is allowed to reach the QSS and then the network starts growing. Every time a pair of traders makes a trade for the first time, a link is introduced between their nodes. There after no further link is added between them irrespective of their subsequent trades and they remain connected with a single link. As the system evolves more and

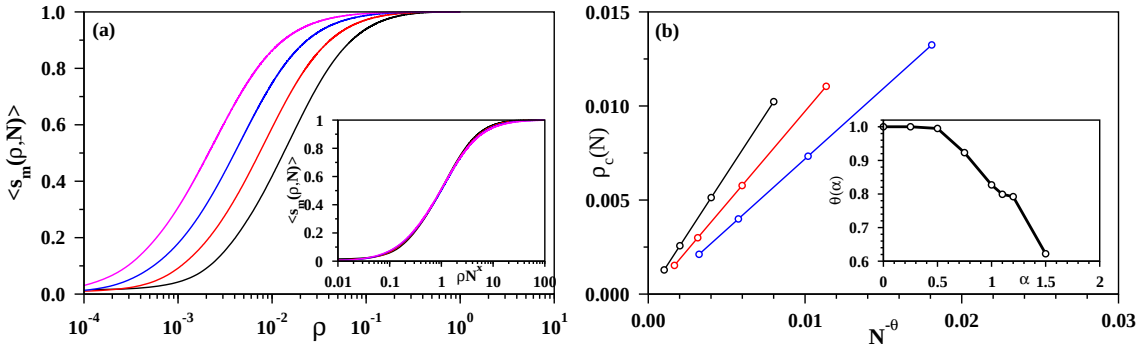


Figure 2.4: (Color online) Growth of the trade network. (a) Plot of the average size of the giant component $\langle s_m(\rho, N)\rangle$ with the link density ρ, for $\alpha = \beta = 1$ and for N = 128 (black), 256 (red), 512 (blue) and 1024 (magenta), (N increases from right to left). The inset shows a data collapse of the same plots with ρN^{θ}, and $\theta = 0.88$. (b) The percolation link density $\rho_c(N)$ is plotted with $N^{-\theta(\alpha)}$ where $\theta(\alpha)$ = 0.88, 0.92 and 1 for $\alpha = \beta$ =1 (black), 3/4 (red), and 1/2 (blue) respectively (α increases from left to right). The inset plots $\theta(\alpha)$ with α.

more new traders take part in the trading dynamics and consequently the number of links grow in the network. For $\alpha = \beta = 0$ the growth of the network is exactly the same as that of the random graph, however it is much different when $(\alpha, \beta) > 0$. Since the rich nodes are preferentially selected they acquire links at a faster rate than the poor nodes. The degree k_i of the node i is the number of distinct traders with whom the i-th trader has ever traded. The dynamics is continued for a certain time T till the average degree $\langle k\rangle$ of a node reaches a specific pre-assigned value.

In general there are two characteristic time scales involved. At the early stage the network grows with multiple components with different sizes. At time T_1 the growing network becomes a single component connected graph. A second time scale is T_2 when the whole network is a N-clique graph in which each node is linked to all others, which means each trader has traded at least once with all others. Unlike random graphs the growth of the network is highly heterogeneous and the rich traders have much larger degrees than the poor traders. Since poor traders are selected with low probabilities they take much longer times to be a part of the network. Consequently T_1 is found to be very large and of the same order as T_2. Numerically it is easier to calculate T_2, one keeps track of the number of distinct links and stops only when this number becomes just equal to $[N(N-1)]/2$. On the other hand to calculate T_1 one follows the growth of the giant component and stops when the giant component covers all N nodes. A Hoshen-Kopelman cluster counting algorithm [63] is used to estimate the size of the giant component. For the ordinary CCM with $\alpha = \beta = 0$ since both traders are chosen with uniform probability, the generated graph is a simple Erdős-Rényi random graph characterized by a Poissonian degree distribution [3].

The growth of the giant component is studied with increasing number of links n in the in the trade network. The average fraction of nodes in the giant component is denoted by $\langle s_m(\rho, N)\rangle$ which is the order parameter in this percolation problem. This has been plotted in Fig. 2.4(a) using a semi-log scale with link density ρ =

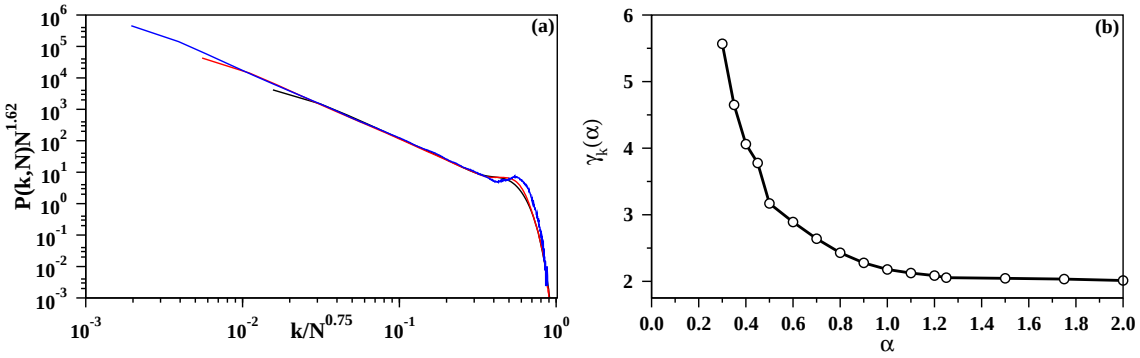


Figure 2.5: (Color online) (a) The finite-size scaling of the degree distributions $P(k, N)$ for $\alpha = \beta = 1$, for $N = 256$ (black), 1024 (red) and 4096 (blue), and for $\langle k \rangle = 1$. Direct measurement of slopes give $\gamma_k(1) = 2.18(3)$. The best data collapse corresponds to $\eta_k(1) = 1.62$ and $\zeta_k(1) = 0.75$ giving $\gamma_k(1) = 2.16(3)$. (b) The plot of $\gamma_k(\alpha)$ vs. α.

$n/[N(N-1)]/2$ in the network. Four curves shown in this figure correspond to $N = 128$, 256, 512 and 1024 for $\alpha = \beta = 1$, the system size increasing from right to left. The inset shows that a data collapse can be obtained by scaling the ρ axis by a factor N^{θ} with $\theta = 0.88$. The critical density of percolation transition $\rho_c(N)$ is defined as that particular value of ρ for which the average size of the giant component $\langle s_m(\rho, N) \rangle = 1/2$. In Fig. 2.4(b) we show that how the critical percolation threshold $\rho_c(N)$ depends on N by plotting it with $N^{-\theta}$ for $\alpha = \beta = 1/2$, 3/4 and 1. It has been observed that the exponent $\theta(\alpha)$ is dependent on α in general and in the inset of this figure we plot $\theta(\alpha)$ vs. α. We see that for $\alpha \leq 1/2$, $\theta(\alpha) = 1$ but for $\alpha > 1/2$, $\theta(\alpha)$ gradually decreases. For Erdős-Rényi random graphs it is known that $\theta = 1$ and therefore this result gives an indication that the trade network seems to be different from random graphs for $\alpha = \beta > 1/2$.

2.6 Degree Distribution

The degree distribution has been studied similar to random graphs. We keep track of the average degree $\langle k \rangle$ of the network which is related to the number of links n of the network as $n = \langle k \rangle \frac{N}{2}$. First the degree distribution has been studied for $\langle k \rangle = 1$ and for different system sizes. For an assigned set of values of (α, β), for a given set of values for the saving propensities $\{\lambda_i\}$ and for a specific value of N once $\langle k \rangle = 1$ is reached we calculate the degree distribution considering all components of the network on the same footing. The network is then refreshed by removing all links and a second network is constructed and so on. The dynamics is continued for the same values of the parameters and the same set of $\{\lambda_i\}$s till a large number of networks are generated and their mean degree distribution is calculated. The entire dynamical process is then repeated with another uncorrelated set of $\{\lambda_i\}$s and the degree distribution has been averaged over many such sets.

In Fig. 2.5(a) we show the finite-size scaling plot of the average degree distribution $P(k, N)$ vs. k for $\alpha = \beta = 1$ and for $N = 256$, 1024 and 4096. On a

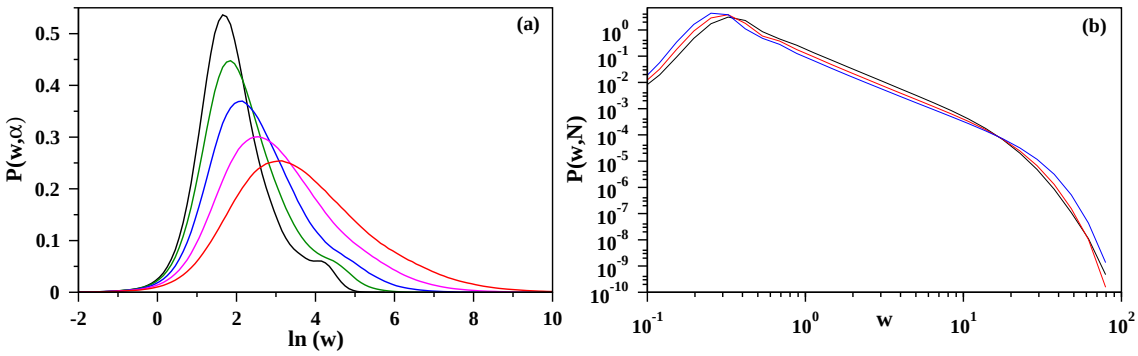


Figure 2.6: (Color online) Probability distribution $P(w, \alpha)$ of the link weights. (a) Plots for the N-clique graphs with N =64 and for α= 0 (black), 1/4 (green), 1/2 (blue), 3/4 (magenta) and 1 (red) and β =1 always. (b) Plots for $\langle k \rangle = 5$ with $\alpha = \beta = 1$ and for the system sizes 128 (black), 256 (red) and 512 (blue). Direct measurement of slopes gives 2.52, 2.53 and 2.51 respectively.

double logarithmic scale all three curves show quite long scaling regions followed by humps before the cut-off sizes of the degree distributions. The cut-off sizes of the distributions shifts to the larger values of k approximately by equal amounts on the double-log scale when the system size has been enhanced by the same factor. From the direct measurement of slope in the scaling region we estimate $\gamma(1) = 2.18(3)$. Almost the entire degree distribution obeys nicely the usual finite-size scaling analysis and an excellent collapse of the data is observed confirming the validity of the following scaling form:

$$P(k, N) \propto N^{-\eta(\alpha)} \mathcal{G}[k/N^{\zeta(\alpha)}] \tag{2.8}$$

where the scaling function $\mathcal{G}(y)$ has its usual forms like, $\mathcal{G}(y) \sim y^{-\gamma(\alpha)}$ as $y \to 0$ and $\mathcal{G}(y)$ approaches zero very fast for $y >> 1$. This is satisfied only when $\gamma(\alpha) = \zeta(\alpha)/\eta(\alpha)$. The exponents $\eta(\alpha)$ and $\zeta(\alpha)$ fully characterize the scaling of $P(k, N)$ in this case. To check the validity of the equation we attempted a data collapse by plotting $P(k, N)N^{\eta(1)}$ vs. $k/N^{\zeta(1)}$ by tuning the values of $\eta(1)$ and $\zeta(1)$. The values obtained for best data collapse are $\eta_k(1) = 1.62$ and $\zeta_k(1) = 0.75$ implying that in the infinite size limit $P(k, \infty) \sim k^{-\gamma(1)}$ with $\gamma(1) = 2.16(3)$. Tuning α and β to other values it is observed that the degree distribution exponent γ does depend on these two parameters. In Fig. 2.5(b) we show a plot of $\gamma_k(\alpha)$ with α which decreases to ≈ 2 at $\alpha = 2$.

2.7 The Weighted Network

Within a certain time T a large number of bipartite trades take place between any arbitrary pair of traders. The total sum of the amounts δ_{ij} invested in all trades between the traders i and j in time T is defined as the total volume of trade $w_{ij} = \Sigma^T \delta_{ij}$. Therefore w_{ij} is regarded as the weight of the link (ij). The magnitudes of weights associated with the links of the trade network are again found to be highly heterogeneous. This is primarily because within a certain time T a rich pair of traders trade many more times than a rich-poor or a poor-poor pair.

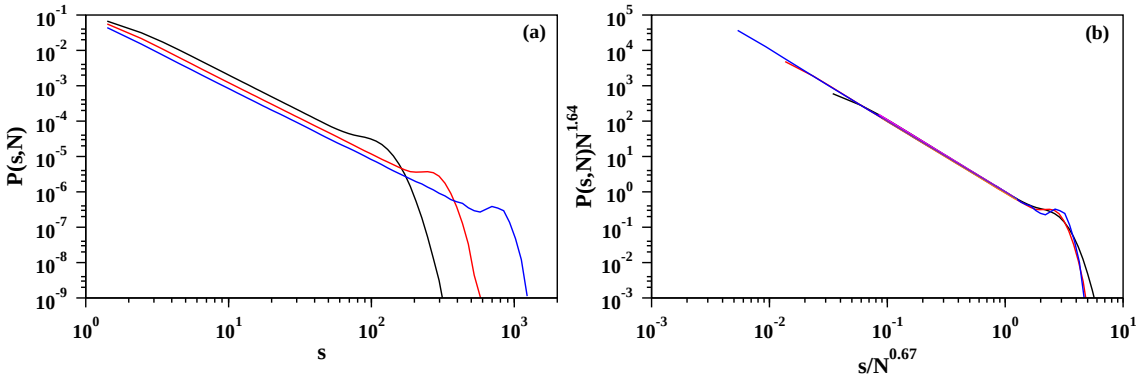


Figure 2.7: (Color online) (a) Nodal strength distributions $P(s, N)$ with strength s with $\alpha = \beta = 1$, for $N = 256$ (black), 1024 (red) and 4096 (blue) (N increases from left to right) and for $\langle k \rangle = 5$. Direct measurement of the slopes of these curves gives $\gamma_s = 2.50(5)$. (b) The Finite-size scaling analysis of the data (using same colors) in (a) gives $\eta_s = 1.64$ and $\zeta_s = 0.67$ estimating the value of $\gamma_s = \eta_s/\zeta_s = 2.45(5)$.

In addition the invested amounts depend on the mean wealths $\langle x_i \rangle$ of the traders involved as well as their saving propensity factors λ_i. The probability distribution $P(w, N)$ of the link weights are calculated when the average degree $\langle k \rangle$ reaches a specific pre-assigned value. As before, this distribution has also been averaged over many weighted networks for one $\{\lambda_i\}$ set and then further averaged over many uncorrelated $\{\lambda_i\}$ sets.

First we studied the case when the trade networks is a N-clique graph, i.e., when each trader has traded with all other traders at least once. Here each node has same degree i.e., $P(k) = \delta(k - (N - 1))$ and $\langle k \rangle = N - 1$. The required time T_2 increases rapidly with N as described in section 2.3 and we could study small system size $N = 64$ only. The distribution has a very long tail and therefore we used a lin-log scale for plotting. In Fig. 2.6(a) we show the plots of $P(w, \alpha)$ with $\ln(w)$ for different values of $\alpha = 0$, 1/4, 1/2, 3/4 and 1 and $\beta = 1$. Each curve is asymmetric and has a single maximum. The position of the peak shifts towards larger values of $\ln(w)$ as α increases. If Fig. 2.6(b) a similar plot has been shown for $\langle k \rangle = 5$ for three network sizes $N = 128$, 256 and 512 and for $\alpha = \beta = 1$. On a double-logarithmic scale each curve has a considerable straight portion. This indicates a power law decay like $P(w, N) \propto w^{-\gamma_w}$. The corresponding slopes give estimates for the exponent γ_w as 2.52, 2.53 and 2.51 for the three system sizes respectively, so that on the average $\gamma_w = 2.52(3)$.

The strength of a node $s_i = \Sigma_j w_{ij}$ where j runs over all neighbors k_i of i, is a measure of the total volume of trade handled by the i-th node. Nodal strengths varies over different nodes over a wide range. We first study the probability distribution of nodal strengths. In Fig. 2.7(a) the strength distribution $P(s, N)$ has been plotted for the average degree $\langle k \rangle = 5$, for $\alpha = \beta = 1$ and for the network sizes 256, 1024 and 4096. Extended scaling regions at the intermediate regions of the curves indicate that $P(s, N)$ also follows a power-law decay function $P(s, N) \sim s^{-\gamma_s}$ in the limit of $N \to \infty$. Direct measurements gives an estimate of $\gamma_s(1) \approx 2.5$. In Fig. 2.7(b) we try a similar finite size scaling of the same data giving $\eta_s(1) = 1.64$ and $\zeta_s(1) = 0.67$

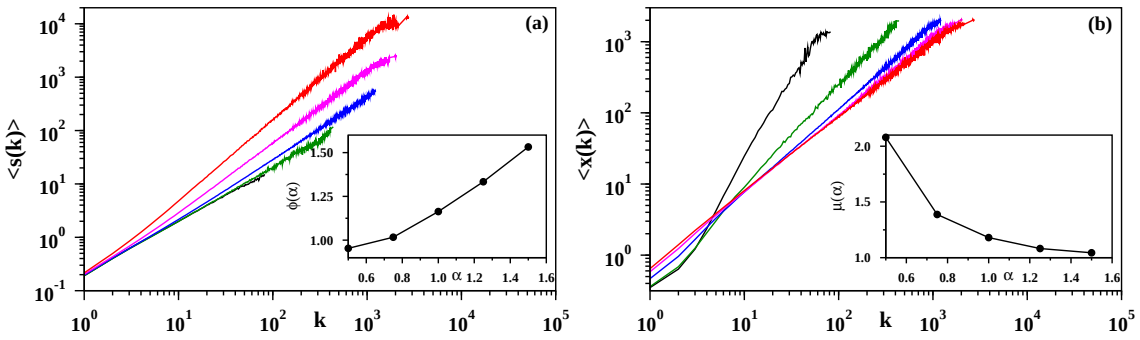


Figure 2.8: (Color online) (a) Plot of nodal strength $\langle s(k)\rangle$ with degree k for $N = 16834$ and for $\alpha = \beta = 1/2$ (black), 3/4 (green), 1 (blue), 5/4 (magenta) and 3/2 (red) (α decreases from left to right). The slopes $\phi(\alpha)$ of these curves are plotted in the inset. (b) Plot of average wealth $\langle x(k)\rangle$ with degree k for $N = 16384$ and for $\alpha = \beta = 1/2$ (black), 3/4 (green), 1 (blue), 5/4 (magenta) and 3/2 (red) (α increases from left to right). The slopes $\mu(\alpha)$ of these curves are plotted in the inset.

giving $\gamma_s(1) = 2.45(5)$.

Quite often weighted networks have non-linear strength-degree relations reflecting the presence of non-trivial correlations, example of such networks are the airport networks and the international trade network. For a network where the link weights are randomly distributed, the $\langle s(k)\rangle$ grows linearly with k. However a non-linear growth like $\langle s(k)\rangle \sim k^{\phi}$ with $\phi > 1$, exhibits the presence of non-trivial correlations. In Fig. 2.8(a) we plot the variation of $\langle s(k)\rangle$ vs. k for a system size $N = 16384$ and for different values of $\alpha = \beta = 1/2$ (black), 3/4 (green), 1 (blue), 5/4 (magenta) and 3/2 (red) (from bottom to top). The slopes of these plots give estimates for the exponent $\phi(\alpha)$ which gradually increased with α and the variation has been plotted in the inset. In the same context we also studied how the mean wealth of a trader depends on its degree. The mean wealth of a trader $\langle x(k)\rangle$ has been plotted in Fig. 2.8(b) with its degree k for the same system sizes as in Fig. 2.8(a) and for the same values of parameters. A power law growth has been observed for all values of α: $\langle x(k)\rangle \sim k^{\mu(\alpha)}$. The slopes of these plots give estimates for the exponent $\mu(\alpha)$ which has been plotted in the inset of Fig. 2.8(b).

2.8 Conclusion

To summarize we have studied the different structural properties of a trade network associated with the dynamical evolution of a model of wealth distribution with quenched saving propensities. In this model distinguishable traders make preferential bipartite trades among themselves and in this way create links. They are selected for trade preferentially using a pair of continuously tunable parameters, where the rich traders are picked up more frequently for trade than poor traders. This creates huge heterogeneity in the system which has been reflected in the power-law distributions of the nodal degree and the link weight distributions measuring the volumes of trade. We present numerical evidence that the associated individual wealth distribution follows the well known Pareto law robustly for all positive values

of the parameters.

Chapter 3

Conservative Self-organized Extremal Model of Wealth Distribution

3.1 Introduction

Pareto distribution shows the inherent inequality presence in the wealth distribution of the society. Most of the people are poor and few are rich in the society. It is observed from the income data of various countries that Pareto law is more appropriate for rich people, i.e., the tail end of the wealth distribution. Wealth distribution of the poor people is different from Pareto distribution. For example, tax return amount distribution in US and Japan shows log-normal distribution in the middle range followed by a power law for high income people [64], UK data of income shows an exponential decay which is followed by a power law in the high income range [39], income data in Brazil for 2004 shows an almost Gaussian law for the low and middle income groups where as high salary groups are described approximately by a power law [41] and there is evidence for power law tail of the wealth distribution in India as well [65].

Application of Statistical Physics with the presence of Complex Network in Econophysics has been discussed in Chapter 2. Here we will show another application of Statistical Physics in Econophysics and the topological effect on it. An individual member of the society is known as an agent in this description. The wealth of the people evolves due to the economic interaction between themselves, called trades. The microstate specified by the precise description of the wealth of every agent changes after each transaction. Their wealth change due to interaction among themselves. This interaction is the mutual bipartite trade among different pairs of agents. Thus the wealth distribution evolves due to such repeated interactions and finally assumes a time independent form. A review of all these models has been published in [50]

3.2 An Extremal Model of Wealth Distribution

Pianegonda, Iglesias, Abramson and Vega (PIAV) introduced the idea of extremal dynamics for the evolution of wealth in one dimension [52, 66]. They assumed that though all agents try to improve their economic status, it is the poor agents who feel the stronger pressure to uplift their economic status. In PIAV model, the extreme situation is modeled when the poorest agent initiates the trade. In PIAV model all agents are arranged on an one dimensional lattice with periodic boundary condition. Each site represents an agent. Initially, all the agents have been assigned random values of wealth between 0 and 1 and are drawn from a uniform distribution. An extremal dynamics is applied to uplift the status of the poorest agent. The dynamics consists of the following two steps: (i) A search is carried out to find the poorest agent in the system and its wealth is replaced with a new random number. (ii) If the poorest agent gains or loses ΔW amount of wealth, $\Delta W/2$ amount of wealth will be subtracted from or added to the wealth of the two nearest neighbors to keep the total wealth conserved in the system. After a large number of such recursions of the two steps the system evolves to a time independent stationary state. In the stationary state wealth distribution $P(w)$ jumps from zero to a maximal value at a critical value of w_c and then it decays as w increases as the Boltzmann-Gibbs exponential function. The PIAV model is a Self-organized Critical (SOC) [34, 67] model where starting from an arbitrary distribution of wealth the dynamical evolution of the model takes the system to the stationary state marked by a fixed critical value of w_c in the absence of any tuning parameter. The stationary state is critical since fluctuations of all length and time scales emerge spontaneously in such a system. This model is very similar to the non-conservative self-organized extremal model for the ecological evolution of interacting species introduced by Bak and Sneppen (BS) [35]. However we will see in the following that in spite of this similarity our modified version of the PIAV model is likely to have a different critical behavior than the BS model. In general SOC models are often compared with the deterministic [34] and the stochastic [68] versions of the sandpile model.

3.3 Continuous Phase Transition in a Wealth Distribution Model

Ghosh, Basu, Chakraborti and Chakrabarti studied a related model with the bipartite transaction rule [69]. They Considered N particles and the state of the particles are characterized by the wealth or energy $\{w_i\}$, where $i = 1, 2....N$. At any arbitrary time t a particle i of wealth x_i below a threshold energy level w_c is chosen and it will do a bipartite trade with a another agent j randomly chosen from the remaining $(N-1)$ agents. The bipartite trading rule is defined as:

$$\begin{aligned} w_i(t+1) &= \epsilon(t)(w_i(t) + w_j(t))) \\ w_j(t+1) &= \bar{\epsilon}(t)(w_i(t) + w_j(t))). \end{aligned} \tag{3.1}$$

Where $\epsilon(t)$ is a freshly generated random number from a uniformly distributed random number within range $[0:1]$ and $\bar{\epsilon} = 1 - \epsilon$. The fraction of agents having

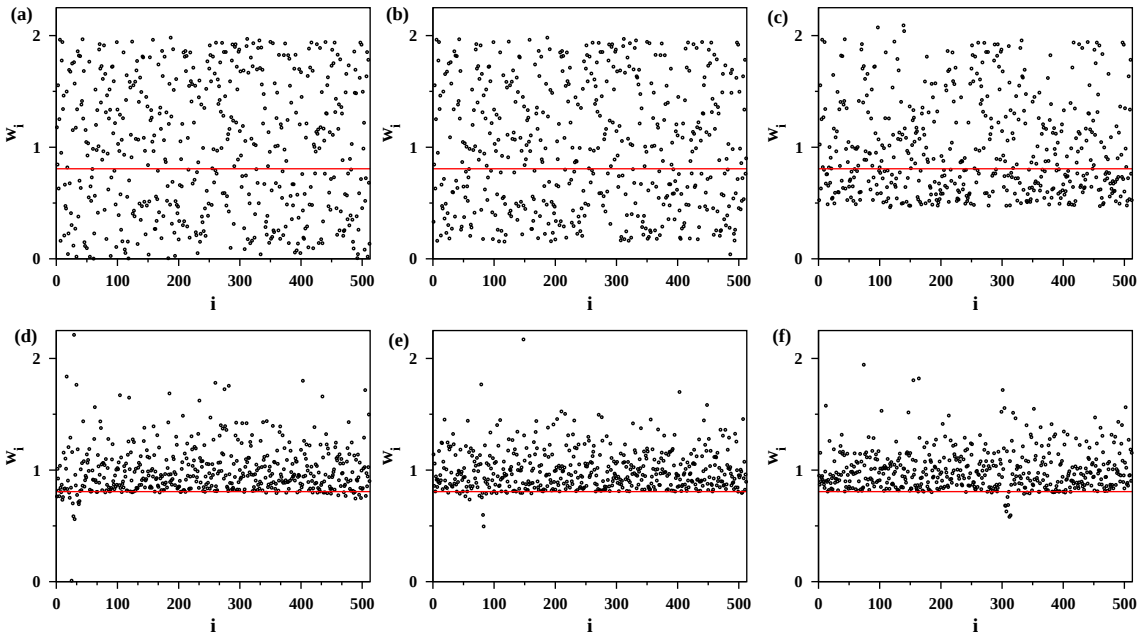


Figure 3.1: (Color online) Time evolution of the wealth w_i at different sites of an $1d$ lattice of size $N = 512$. The series of snapshots are taken at times $t =$ (a) 0, (b) 500, (c) 1000, (d) 500000, (e) 5000000 and (f) 15000000. The position of the poverty line at $w_c^{half}(512) = 0.8070$ is shown by the red line. With time the wealth values gradually move up beyond the poverty line and stays there in the stationary state during further evolution.

wealth less than w is considered as order parameter. They claimed that the system exhibits a continuous phase transition at a critical value of the threshold wealth w_c. A number of critical exponents have been measured to characterize the transition and some of them are found to be close to the exponents in the Manna model of Self-Organized Criticality [68].

Based on our publication [70] we present an extensive numerical study of the modified version of a conservative self-organized extremal model whose motivation has been drawn from the wealth distribution of the people in a society. We consider this model as one of the few examples of non-dissipative SOC systems where the entire wealth of the society is strictly conserved, for example the fixed energy sandpile [71]. Estimation of a number of critical exponents of our model suggests that this model belongs to a new universality class perhaps because of strict conservation of wealth is maintained in its dynamical rules.

3.4 The Minimal Wealth Model

We have considered a model with a conservative extremal dynamics for studying the evolution of wealth distribution in a society. In a bipartite transaction one agent is necessarily selected as the one with the globally minimal value of wealth w_{min}. The second agent is chosen randomly with uniform probability from neighbors of the first agent. This neighbor list has been defined in different ways for different graphs. We

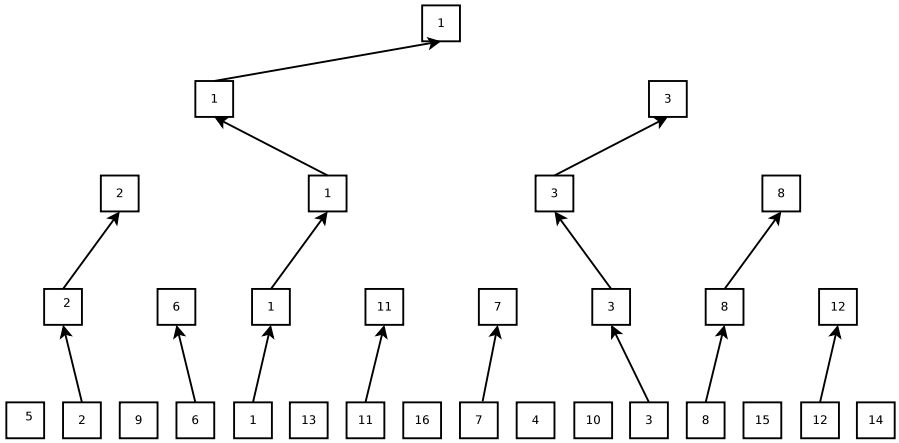


Figure 3.2: (Color online) Schematic diagram of the method to find the global minimum [72, 73]. 16 blocks with 16 random numbers in the level 1 represent 16 sites of an $1d$ lattice. Successive blocks form a pair. The block with minimum number of the pair is forwarded from level 1 to level 2 and a pointer is assigned between the two blocks. A similar procedure is repeated for higher levels up to the level where only one block with the global minimum number exists. Moving from the top level in the reverse direction of the pointer to the lowest level one can find the position of the site with global minimum.

have studied this model on four different graphs, namely, (i) $1d$ regular lattice with periodic boundary condition (ii) two dimensional ($2d$) square lattice with periodic boundary condition (iii) the Barabási - Albert (BA) scale-free graph [5] and on an (iv) N-clique graph. On every graph the nodes represent agents and the nearest neighbors of each node connected by direct links constitute the neighbor list of every agent. We report elaborately the results of our model on an $1d$ lattice and mention the key results of the same model studied on other graphs in tables.

The dynamics starts with N agents, each having an amount of wealth $w_i, \{i = 1, N\}$ drawn from a uniform distribution with the average $\langle w(N)\rangle$ =1 irrespective of the system size N. The discrete time t is the number of bipartite transactions. At an arbitrary time t first the site $i = i_{min}$ is searched out which has the minimal wealth w_{min}. The other agent j is selected randomly with uniform probability from the neighbors of i_{min}. Both agents invest their entire amount of wealth. Therefore the total invested amount by both the traders is: $\delta_{ij}(t) = w_i(t)+w_j(t)$. This amount is then randomly divided into two parts and received by them also randomly:

$$w_i(t+1) = \epsilon(t)\delta_{ij}(t) \quad w_j(t+1) = \bar{\epsilon}(t)\delta_{ij}(t). \tag{3.2}$$

where $\epsilon(t)$ is a freshly generated random fraction and $\bar{\epsilon} = 1 - \epsilon$. These transactions are repeated ad infinitum. After some relaxation time the system reaches a stationary state when the wealth distribution assumes a time independent form.

In $1d$ a linear chain of N sites with periodic boundary condition has been considered where the neighbor list of an arbitrary site i consists of the two nearest neighboring sites at $i \pm 1$. Therefore the second agent j is selected randomly with equal probability from this list. If w_{min} is very small then the probability that it

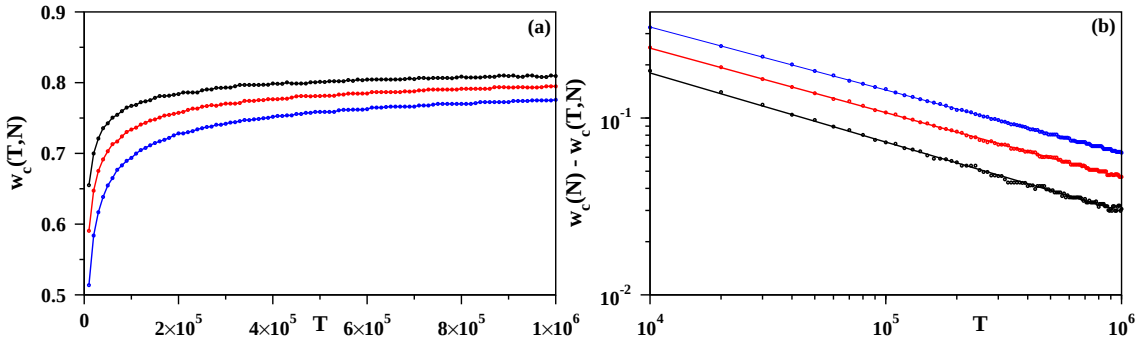


Figure 3.3: (Color online) The system is relaxing from initial state to the stationary state for $N = 2^{10}$ (black), 2^{11} (red) and 2^{12} (blue). (a)The position of the poverty line $w_c(T, N)$ has been plotted with the relaxation time T. (b) The same data has been plotted with deviation $w_c(N) - w_c(T, N)$ vs. T on a $\log - \log$ scale. Extrapolation of slopes of these curves gives an estimate for the dynamical exponent $z \approx 2.84$.

will be replaced by an even smaller wealth after trade is also small. However this probability gradually increases as w_{min} increases. As the sites with the minimal values of wealth are systematically replaced, very soon all nodes with small w values are replaced by larger values of w resulting a vacancy in the small w region. This is explained pictorially in Fig. 3.1. On an $1d$ lattice with $N = 512$ we plot the lattice positions i along the abscissa and the corresponding wealth w_i along the ordinate. As time increases a vacant region gradually forms for small values of w for all sites. If on the average the wealth of none of the agent is below a certain threshold value w, it is called the 'poverty line'. In Fig. 3.1 the poverty line gradually moves up with time and finally settles at a critical value $w_c^{half}(512) = 0.8070$ at the stationary state. This behavior is the same for all system sizes N but with different values of $w_c(N)$. Unlike the model in [69] here the critical poverty line is spontaneously determined by the dynamical evolution of the system where no fine tuning is necessary which is the distinctive signature of self-organization and we will see in the following that the model exhibits critical behavior as well.

To find the agent with minimal wealth a brute-force search takes cpu $\sim N$. A much faster algorithm to search for the globally extremal (minimal or maximal) site was proposed by Grassberger [72, 73] which stores the data in a hierarchical structure. This takes cpu $\sim \ln N$. For sites of a $1d$ lattice of size $N = 2^n$, one starts with N number of blocks in level 1. Each block contains the value of wealth w_i of each site. Successive blocks form $N/2$ pair of blocks in level 1. The minimal values of wealth from each pair of level 1 are stored in the next level (level 2) in $N/2$ blocks and a pointer is assigned between the corresponding blocks of the two levels. This procedure is repeated up to level the $(n + 1)$ where only one block containing the global minimal value of the wealth w_{min} exists. The location of the site with minimal wealth w_{min} can be found by going in reverse to the pointer direction from the top level to the bottom level. For each time if there is any change in the values of the wealth of the blocks or the sites in level 1 one has to update the the wealth values with pointer direction from ground to top level accordingly. A schematic diagram of the procedure is shown in Fig. 3.2. We have used this method for $1d$,

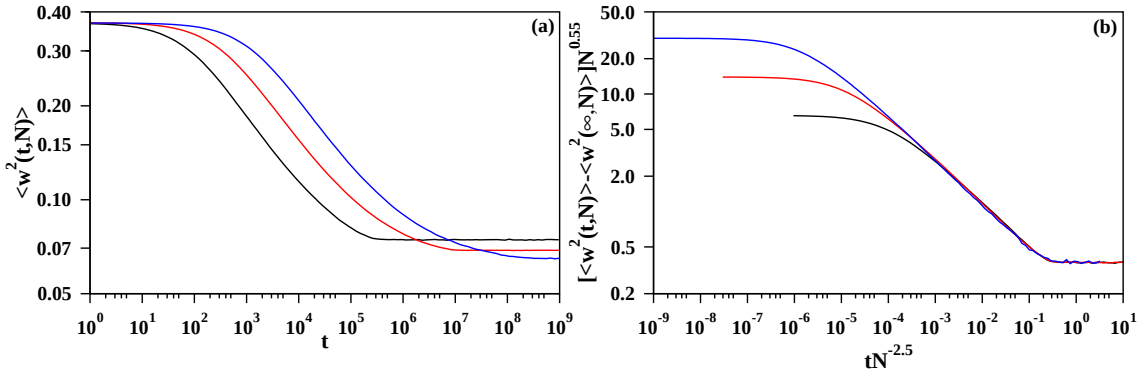


Figure 3.4: (Color online) Estimating the relaxation times. (a) The average wealth square $\langle w^2(t,N)\rangle$ per agent has been plotted with time t for $N = 2^8$ (black), 2^{10} (red) and 2^{12} (blue) and after a long time they approach the stationary state value $\langle w^2(\infty,N)\rangle$. (b) The scaled deviation $[\langle w^2(t,N)\rangle - \langle w^2(\infty,N)\rangle]N^{0.55}$ has been plotted with the scaled variable $tN^{-2.5}$. The relaxation exponent $z = 2.70$.

$2d$ and N-clique graphs. For BA graphs we used the brute-force algorithm.

3.4.1 Relaxation to Stationary State

We first estimate the relaxation time required for the system to reach the stationary state. During this relaxation period the wealth distribution gradually changes starting from its initial uniform distribution to its time independent form in the stationary state. The relaxation time has been estimated as a function of deviation of the poverty line from its critical value in two ways.

For a given system size N we have simulated our model up to 10^6 time steps and calculated the wealth distribution $P(T,w,N)$ at 100 time instants T at the interval of 10^4 steps. These distributions are calculated for a sample size of $\sim 10^6$ independent runs. The value of the poverty line $w_c(T,N)$ at time T is determined by the value of w where $P(T,w,N)$ is the maximum. This estimation is done for all values of T. In Fig. 3.3(a) we plot on a $lin-lin$ scale $w_c(T,N)$ vs. T vs. for $N = 2^{10}$, 2^{11} and 2^{12}. We see that in each case the $w_c(T,N)$ saturates to its stationary state value $w_c(N)$ as time T gradually increases. These data have been replotted in Fig. 3.3(b) using a $\log-\log$ scale as $w_c(N)-w_c(T,N)$ vs. T where we have used $w_c(N) = 0.8242$, 0.8375 and 0.8383 to obtain the best straight line plots. The slopes of these straight lines are 0.3910, 0.3647 and 0.3562 respectively which are then extrapolated with $N^{-1.626}$ to obtain the exponent z as $[w_c(N)-w_c(T,N)] \propto T^{-1/z}$ with $z \approx 2.84$.

In a second method we calculated $\langle w^2(t,N)\rangle$ with time t starting from its initial value when the distribution is uniform. After a long time this quantity saturates to its stationary value $\langle w^2(\infty,N)\rangle$. In Fig. 3.4(a) we show the plots of $\langle w^2(t,N)\rangle$ vs. t on a $\log-\log$ scale. In Fig. 3.4(b) $[\langle w^2(t,N)\rangle - \langle w^2(\infty,N)\rangle]N^{0.55}$ has been plotted with the scaled value of time $tN^{-2.5}$ using $\langle w^2(\infty,N)\rangle = 0.057$, 0.0607 and 0.061 for $N = 2^8, 2^{10}$ and 2^{12} respectively. A nice data collapse has been obtained with the following scaling form

$$[\langle w^2(t,N)\rangle - \langle w^2(\infty,N)\rangle]N^{0.55} \sim \mathcal{F}_z(tN^{-2.5}) \tag{3.3}$$

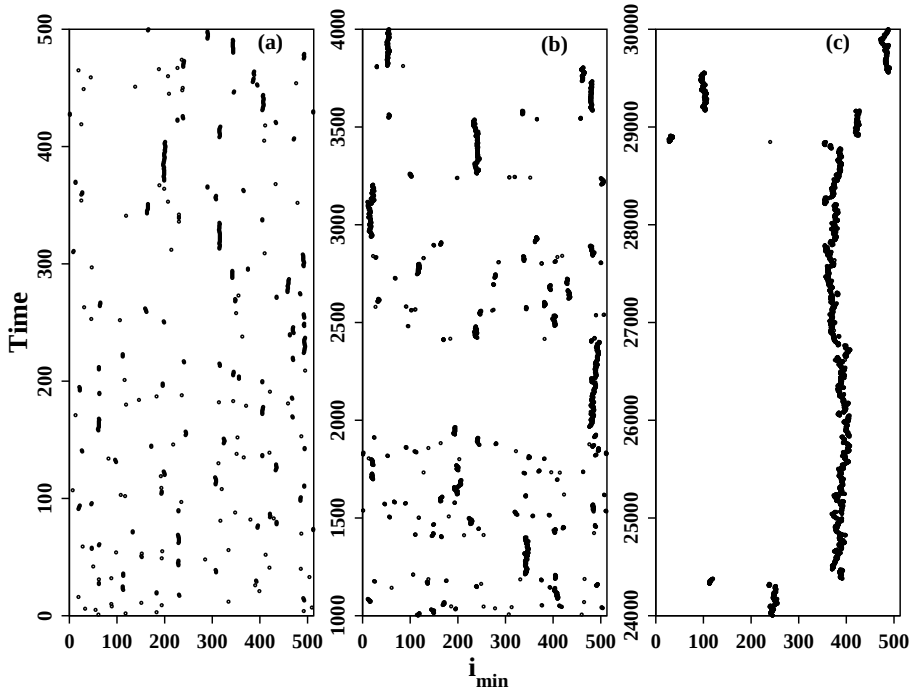


Figure 3.5: An exhibition of the correlation that sets in the system. In a single run the lattice sites i_{min} with globally minimal wealth in successive time steps are marked for a system of size $N = 512$ in $1d$. The gradually increasing duration of correlation has been exhibited by time windows of increasing lengths: (a) 500, (b) 3000 and (c) 6000.

where the scaling function $\mathcal{F}_z(x)$ varies as $x^{-1/z}$ for small x. A direct measurement of the slope of the scaled plot gives an estimate of $1/z = 0.37$ which corresponds to $z = 2.70$. We conclude an average value of $z = 2.77(7)$.

3.4.2 Correlation in the Stationary State

Starting from an uncorrelated wealth distribution the system becomes more and more correlated as time passes. This is reflected in the fact that the probability of occurrence of the minimal sites close to each other in successive time steps gradually increases. At the early uncorrelated stage the position of the minimal site at the next time step is likely to be anywhere in the lattice. However as time increases, the poverty line moves up, consequently w_{min} increases and the probability that the minimal site at the next time step occurring at the same site or at the neighboring updated site also increases. This is shown in Fig. 3.5 using a $1d$ lattice of $N = 512$. For a single run it is observed that the locations of w_{min} are quite random (Fig. 3.5(a)). However as time evolves these positions gradually become more correlated (Figs. 3.5(b) and 3.5(c)). In general one can consider the successive jumps of i_{min} positions constituting a Lévy flight random walk [74]. We see below that indeed their flight lengths follow a power law distribution.

The correlation in the stationary state is quantitatively measured by the probability distribution $P(\ell)$ of the distance of separation ℓ between successive minimal

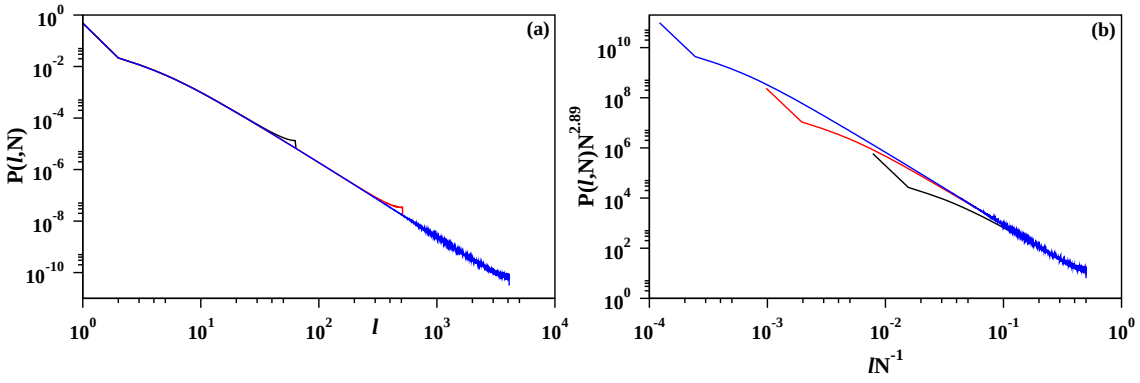


Figure 3.6: (Color online) (a) The probability $P(\ell, N)$ that sites with minimal wealth at successive time steps are separated by a distance ℓ has been plotted with ℓ for the system sizes $N = 2^7$ (black), 2^{10} (red) and 2^{13} (blue). The slopes are -2.64, -2.81 and -2.89.(b) The finite-size scaling of $P(\ell, N)N^{\eta_\pi}$ vs. $lN^{-\zeta_\pi}$ with $\eta_\pi = 2.89$ and $\zeta_\pi = 1$ and therefore $\pi = \eta_\pi/\zeta_\pi = 2.89(5)$.

sites using periodic boundary condition. This distribution measured in the stationary state has been plotted in Fig. 3.6(a) for different system sizes $N = 2^7, 2^{10}$ and 2^{13}. The value of $P(\ell)$ for $\ell = 0$ and 1 are approximately 0.4575(1) and 0.4820(1) and then it decreases as a power law $P(\ell) \sim \ell^{-\pi}$ with increasing ℓ. A direct measurement of slope gives $\pi \approx 2.89$. Fig. 3.6(b) exhibits the finite-size scaling of the same data when the ℓ and $P(\ell, N)$ axes are scaled as:

$$P(\ell, N)N^{\eta_\pi} \propto \mathcal{F}_\pi(\ell N^{-\zeta_\pi}). \tag{3.4}$$

where $\mathcal{F}_\pi(x)$ is a universal scaling function with the scaling exponents $\eta_\pi = 2.89$ and $\zeta_\pi = 1$. From this scaling analysis we get $\pi = \eta_\pi/\zeta_\pi = 2.89(5)$.

There exists a spatial correlation too in this model. A two point correlation function has been measured in the stationary state. The average correlation between two sites situated at a distance of separation x has been defined as:

$$\mathcal{C}(x) = \langle w(0)w(x)\rangle - \langle w\rangle^2 \tag{3.5}$$

where $\langle w\rangle$ is always set equal to unity. We assume a power law decay for the correlation, i.e., $\mathcal{C}(x) \sim x^{-\chi}$ for $x \to \infty$. For $1d$ a plot of (not shown) $\mathcal{C}(x)$ vs. x on a $\log - \log$ scale indicates a power law for large x values. However considerable variation of slopes exists for system sizes $N = 2^8, 2^{10}$ and 2^{12}. The slopes are: -1.17, -1.29. -1.34 respectively which extrapolates to a value of $\chi = 1.5(1)$ in the limit of $N \to \infty$. Our estimate for χ in $2d$ is 2.2(2).

3.4.3 Wealth Distribution in the Stationary State

Next we estimated the probability density distribution $P(w, N)$ of wealth in the stationary state of the system of size N. The distribution grows very rapidly near $w_c(N)$ for all values of N and then decays very fast. In Fig. 3.7(a) we show the plot of $P(w, N)$ vs. w for $N = 2^7, 2^{10}$ and 2^{13}. All of them have similar variations

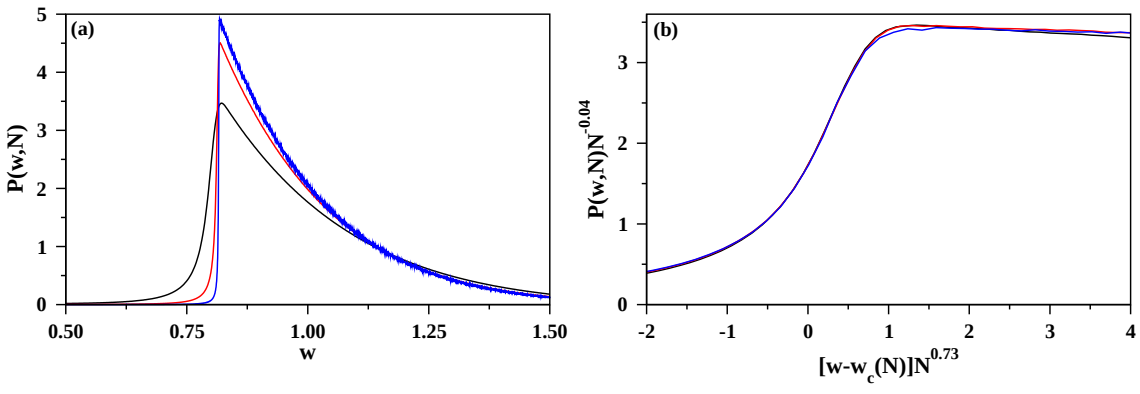


Figure 3.7: (Color online) For Minimal Wealth model in $1d$. (a) The wealth distribution $P(w, N)$ in the stationary state for $N = 2^7$ (black), 2^{10} (red) and 2^{13} (blue). (b) A finite-size scaling using $P(w, N)N^{-0.04}$ vs. $[w - w_c(N)]N^{0.73}$ for system sizes $N = 2^{11}$ (black), 2^{12} (red) and 2^{13} (blue).

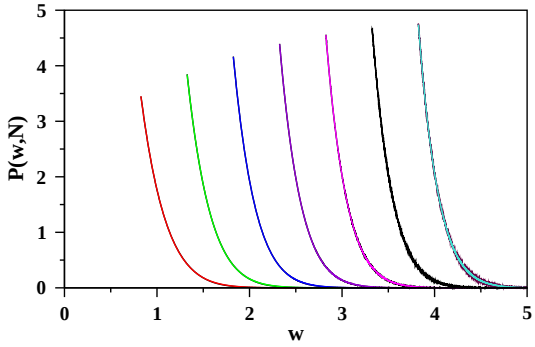


Figure 3.8: (Color online) Gaussian fit of the wealth distribution $P(w, N)$ in the stationary state for seven different system sizes $N = 2^7$ to 2^{13} in $1d$. The sequence starts with the distribution for $N = 2^7$ at the extreme left and is shifted to the right by 0.2 when system size is multiplied by a factor of 2.

but with increasing system size the curves gradually become sharper. Therefore we tried a finite-size scaling analysis in Fig. 3.7(b) for the growing region and for $N = 2^{11}, 2^{12}$ and 2^{13}. A nice data collapse is observed when axes are scaled and $P(w, N)N^{-0.04}$ has been plotted with $[w - w_c(N)]N^{0.73}$.

The functional form of the decay of the probability distribution $P(w, N)$ has been studied right after the maximal jump at $w_c(N)$. This part fits very well with the Gaussian form:

$$P(w, N) = \frac{A(N)}{\sqrt{2\pi}\sigma(N)} \exp[-\frac{(w - \mu(N))^2}{2\sigma^2(N)}] \tag{3.6}$$

In Fig. 3.8 we showed $P(w, N)$ vs. w on a $lin - lin$ scale for seven different system sizes: $N = 2^7, 2^8, ..., 2^{13}$. In each case the fitting curve is indistinguishable from the data. We observe a systematic variation of $A(N)$, $\mu(N)$ and $\sigma(N)$ with system size N. For example $A(N) \approx 511.5 - 1828N^{-0.294}$, $\mu(N) \approx -1.10 + 62.5N^{-1.215}$ and $\sigma(N) \approx 0.64 + 0.429N^{-0.436}$.

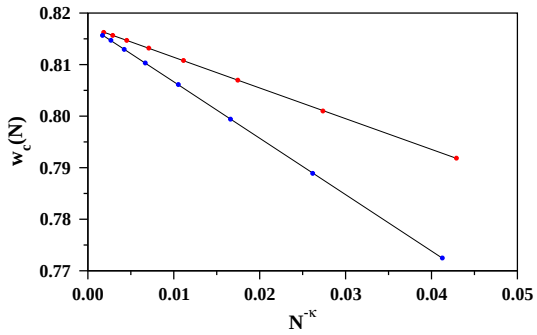


Figure 3.9: (Color online) Estimation of the critical poverty line $w_c(\infty)$. The $w_c^{half}(N)$ are the upper plot (in red) and $w_c^{slope}(N)$ are the lower plot (in blue). When they are extrapolated with $N^{-\kappa}$ with $\kappa(half) = 0.649$ and $\kappa(slope) = 0.657$, $w_c(\infty)$ is obtained as 0.8174 and 0.8176 respectively. We conclude $w_c(\infty) = 0.8175(2)$.

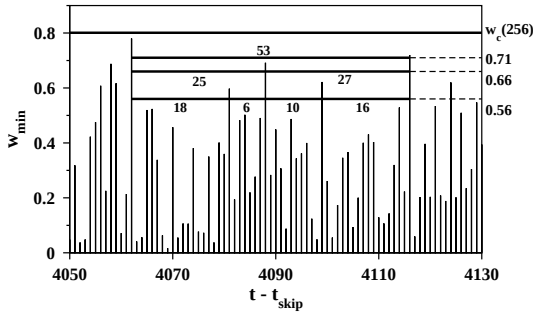


Figure 3.10: A portion of the time series of minimal wealth values in successive time steps is shown for a $1d$ system with $N = 256$ and $w_c(256) = w_c^{half}(256) = 0.8010$ in the stationary state. It shows that depending on the value of w_o an avalanche can be broken into a hierarchy of avalanches. For this run the system has been relaxed for the initial $t_{skip} = 5 \times 10^8$ time steps.

The precise value of $w_c(\infty)$ is calculated by extrapolating $w_c(N)$ values which are calculated using the following two methods. We have seen in Fig. 3.7(a) that the probability distribution of $P(w, N)$ becomes increasingly steeper with increasing N. For a certain size N we defined $w_c^{half}(N)$ as the value of w for which $P(w, N)$ is half of its maximum value. In a second method the $w_c(N)$ has been calculated in the following way. A pair of successive points on the $P(w, N)$ vs. w curve which has the largest slope is found out. A straight line joining these two points is then extrapolated to meet the w axis at $w_c^{slope}(N)$. The pair of values of $w_c^{half}(N)$ and $w_c^{slope}(N)$ for eight N values 2^7 to 2^{14} are then extrapolated with $N^{-\kappa}$. A least square fit of straight line has been done for trial values of κ starting from 0.20 to 1.20 at an interval of 0.001 and the errors have been calculated. The errors are minimal for $\kappa(half) = 0.649$ and $\kappa(slope) = 0.657$. Using these two values of the exponent κ we extrapolate $w_c(N)$ values with $N^{-\kappa}$ in Fig. 3.9 to meet the $w_c(N)$ axis at 0.8174 and 0.8176 respectively. We conclude a value for $w_c(\infty)$ for $1d$ model as 0.8175(2).

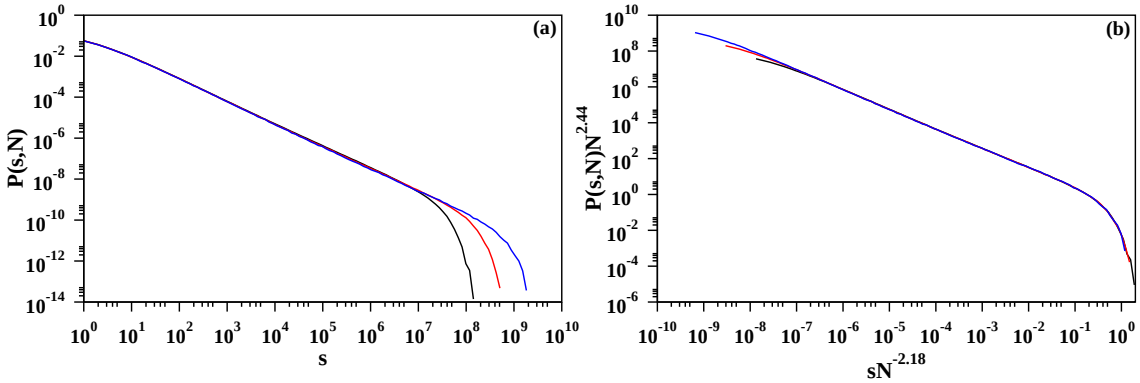


Figure 3.11: (Color online) For $1d$ Minimal Wealth model. (a) The avalanche size distribution for $N = 2^{12}$ (black), 2^{13} (red) and 2^{14} (blue). (b) A finite-size scaling of the data in (a) with scaling exponents $\eta_\tau = 2.44$ and $\zeta_\tau = 2.18$ giving the avalanche size exponent $\tau = \eta_\tau/\zeta_\tau \approx 1.12(1)$.

3.4.4 Avalanche Size Distribution

In the stationary state successive values of minimal wealth w_{min} fluctuates with time. If a certain reference wealth is fixed by hand at $w = w_o$ then the successive w_{min} appear below and above w_o line. One defines a w_o-avalanche as the sequence of successive bipartite trades whose w_{min} values are smaller than w_o. The size s of the avalanche measures the duration of the avalanche i.e., at times t and $t + s + 1$ the $w_{min} > w_o$, whereas at every time step from $t+1$ to $t+s$ the $w_{min} < w_o$. When w_o is set equal to $w_c(N)$ it is called a critical avalanche. This is explained in Fig. 3.10 where part of the w_{min} time series for $N = 256$ and $w_c(256) = w_c^{half}(256) = 0.8010$ is displayed discarding the initial $t_{skip} = 5 \times 10^8$ time steps. For $w_o = 0.71$ an avalanche of size 53 breaks into two avalanches of sizes 25 and and 27 when w_o is reduced to 0.66. On further reduction of w_o to 0.56 these two avalanche break into even smaller avalanches of sizes 18, 6 and 10, 16 respectively. Thus any avalanche can be splitted into a hierarchy of smaller avalanches if w_o value is lowered [75]. On the other hand if w_o is raised the average avalanche size increases and becomes infinite at certain value of w_o.

At the critical point the distribution of the avalanche life-times has a power law tail in the limit of $N \to \infty$: $P(s, \infty) \sim s^{-\tau}$. In the stationary state we used $w_o = w_c^{half}(N)$ and measured life-times of a large number of avalanches for different system sizes and plot the probability distributions $P(s, N)$ vs. s using a $\log - \log$ scale in Fig. 3.11(a). Each curve has a straight portion in the intermediate regime of the avalanche sizes and this regime becomes gradually larger on increasing N. The direct measurement of slopes in the scaling region gives $\tau(N) = 1.086$, 1.091 and 1.096 for $N = 2^{12}, 2^{13}$ and 2^{14} respectively. A finite-size scaling is very much suitable with the following scaling form:

$$P(s, N)N^{\eta_\tau} \propto \mathcal{F}_\tau(sN^{-\zeta_\tau}) \tag{3.7}$$

where the scaling function $\mathcal{F}_\tau(x) \sim x^{-\tau}$ in the limit of $x \to 0$ and $\mathcal{F}_\tau(x)$ approaches zero very fast for $x >> 1$. The exponents η_τ and ζ_τ fully characterize the scaling of

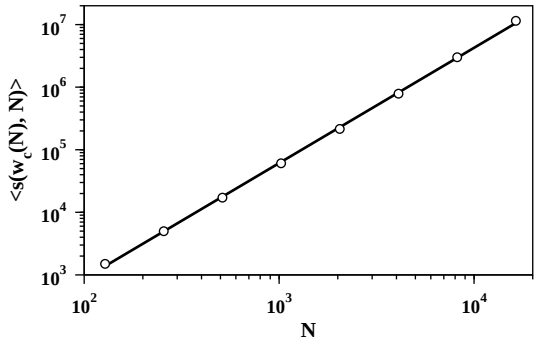


Figure 3.12: The average size of the avalanches $\langle s(w_c(N), N)\rangle$ right at the poverty line $w_c(N)$ has been plotted with system size N on a $\log - \log$ scale. The slope of this curve gives the exponent $\beta = 1.92(2)$.

$P(s, N)$ in this case. An immediate way to check the validity of this equation is to attempt a data collapse by plotting $P(s, N)N^{\eta_\tau}$ vs. s/N^{ζ_τ} with trial values of the scaling exponents. For $1d$ the values for obtaining the best data collapse are found to be $\eta_\tau = 2.44$ and $\zeta_\tau = 2.18$ (Fig. 3.11(b)). The life-time distribution exponent for $1d$ is therefore $\tau = \eta_\tau/\zeta_\tau \approx 1.12(1)$.

Next we calculated the average value of avalanche life-times $\langle s(w_c, N)\rangle$ right at the critical poverty line. In Fig. 3.12 we plot this quantity with system size N on a $\log - \log$ scale. The plot fits excellent to a straight line and its slope gives the value of the exponent $\beta \approx 1.92(2)$ in: $\langle s(w_c, N)\rangle \sim N^\beta$. Assuming the distribution $P(s, N)$ of avalanche sizes holds good up to a cut-off $s_{max} \sim N^{\zeta_\tau}$ one gets a scaling relation $\beta = \zeta_\tau(2 - \tau)$ and our estimates of $\beta = 1.92$, $\zeta_\tau = 2.18$ and $\tau = 1.12$ satisfy this relation very closely.

The size of the w_o-avalanches are smaller when $w_o < w_c(N)$ and we have studied how the average avalanche size grows as the deviation $(w_c(N) - w_o)$ decreases. Similar to the BS model we assume $\langle s(w_o)\rangle \sim [w_c - w_o]^{-\gamma}$ for $N \to \infty$. We measured the average size $\langle s(w_o, N)\rangle$ of the w_o-avalanches for different system sizes N and plotted them with $w_c(N) - w_o$ in Fig. 3.13(a) with $w_c(N) = w_c^{half}(N)$. For all plots on $\log - \log$ scale the curves are horizontal as deviation $w_c(N) - w$ is very small. However as the deviation increases the curves become straight with negative slopes -1.98, -2.15, -2.28 and -2.38 for for $N = 2^8, 2^{10}, 2^{12}$ and 2^{14} respectively. These values when extrapolated with $N^{-0.208}$ give $\gamma = 2.67$ for $N \to \infty$. Again a finite-size scaling has been possible as shown in Fig. 3.13(b):

$$\langle s(w_o, N)\rangle N^{-\eta_\gamma} \propto \mathcal{F}_\gamma([w_c(N) - w_o]N^{\zeta_\gamma}) \tag{3.8}$$

where $\mathcal{F}_\gamma(x)$ is another scaling function. From this data collapse the scaling exponents $\eta_\gamma = \beta = 1.92$ and $\zeta_\gamma = 0.73$ with $\gamma = \eta_\gamma/\zeta_\gamma \approx 2.63$ is obtained.

For every system size N there is a value of $w_o = w_c^{inf}(N)$ so that when w_o is raised to this value the avalanche size becomes infinite. This implies that if we plot the data in Fig. 3.13(a) with respect to $w_c^{inf}(N) - w_o$ then we should be able to see the divergence of average avalanche size instead of saturation of the avalanche sizes. We plot this in Fig. 3.13(c) using $w_c^{inf}(N) = 0.8167$, 0.8169, 0.8170, 0.8172 for $N = 2^8$, 2^{10}, 2^{12} and 2^{14} respectively again on a $\log - \log$ scale. Each curve is

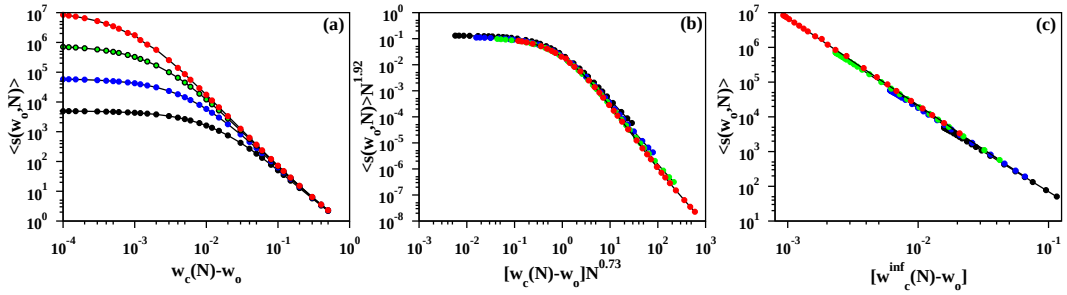


Figure 3.13: (Color online) (a) The average value of the avalanche size $\langle s(w_o, N)\rangle$ has been plotted with deviation $w_c(N)-w_o$ from the poverty line $w_c(N)$. The system sizes used are $N = 2^8$ (black), 2^{10} (blue), 2^{12} (green) and 2^{14} (red). The value of γ obtained by extrapolation of slopes is 2.67. (b) Finite size scaling analysis of the data in (a) is shown. The scaling exponents $\eta_\gamma = \beta = 1.92$ and $\zeta_\gamma = 0.73$ gives $\gamma = \eta_\gamma/\zeta_\gamma \approx 2.63$. (c) Plot of the data in (a) but with $w_c(N) = w_c^{inf}(N)$. The slopes for four different system sizes in (a) on extrapolation gives $\gamma = 2.66$. We conclude $\gamma = 2.65(5)$.

a straight line but with different slopes: -2.31, -2.43, -2.51 and -2.56 respectively. When these slopes are extrapolated with $N^{-0.31}$ the extrapolated value for $N \to \infty$ is -2.66. Our conclusion for the value of the exponent $\gamma = 2.65(5)$.

3.4.5 Persistence of Wealth in the Stationary State

The time interval between two successive updates of wealth of an agent is known as the persistence time t_p. Different agents have to wait for different amounts of times in general. More specifically agents having small amount of wealth have to wait for very little times. On the other hand potentially rich agents have to wait long enough times. In the stationary state we measure the persistence times for all sites of the lattice and use this data to plot their probability distribution. More precisely we set a clock to each site. Whenever there is a change of wealth at this site the time is noted and the clock time is reset to zero. At the stationary state we collect a large number of persistence time data and use these data to calculate the persistence time distribution.

We assume a power law variation of the persistence time distribution as $P(t_p) \sim t_p^{-\theta}$ in the limit of $N \to \infty$. For finite size systems the distributions $P(t_p, N)$ vs. t_p are plotted on a $\log - \log$ scale (not shown) and the direct measurement of slopes give the $\theta(N)$ values for finite size systems. These values are extrapolated as $N^{-0.494}$ to obtain $\theta = 1.539$ for $N \to \infty$ in $1d$. In a similar analysis for a $2d$ square lattice of size L using an extrapolation with respect to $L^{-0.565}$ we get $\theta = 1.25$ for $L \to \infty$.

Persistence exponents are also obtained by the finite-size scaling analysis. In Fig. 3.14(a) we show the scaling plot of $P(t_p, N)N^{\eta_\theta}$ vs. $t_p N^{-\zeta_\theta}$ with $\eta_\theta = 3.41$ and $\zeta_\theta = 2.21$. This gives $\theta = \eta_\theta/\zeta_\theta = 1.543$ in $1d$. Similar scaling analysis in terms of the system size L in $2d$ square lattice has been performed with $\eta_\theta = 3.49$ and $\zeta_\theta = 2.74$ which gives $\theta = 1.274$ (Fig. 3.14(b)). Averaging θ values obtained from direct measurement and scaling analysis we conclude $\theta = 1.541(10)$ in $1d$ and

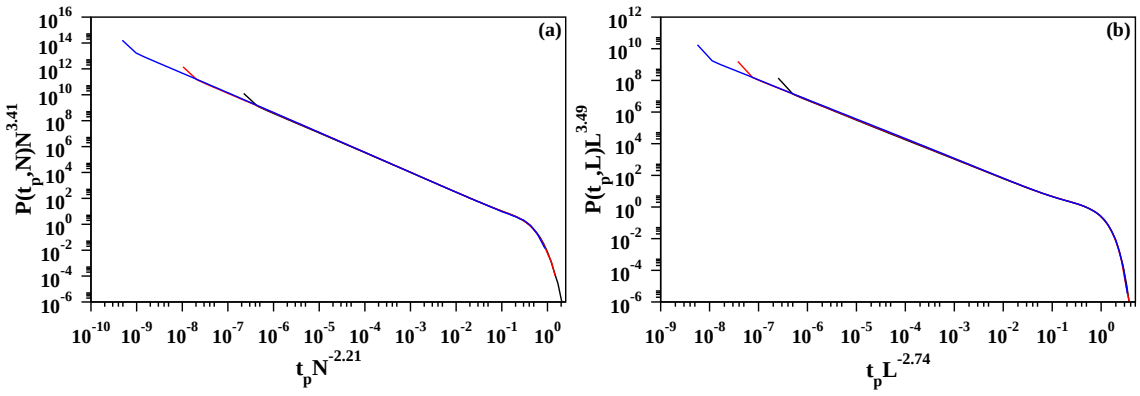


Figure 3.14: (Color online) The probability distribution $P(t_p, N)$ of persistence times t_p at the stationary state. (a) The finite-size scaling of the distribution in $1d$ for $N = 2^{10}$ (black), 2^{12} (red) and 2^{14} (blue). From the scaling exponents $\eta_\theta = 3.41$ and $\zeta_\theta = 2.21$ the persistence exponent $\theta = \eta_\theta/\zeta_\theta \approx 1.543$ is obtained. (b) The finite-size scaling of the distribution in $2d$ for $L = 2^8$ (black), 2^9 (red) and 2^{10} (blue). From the scaling exponents $\eta_\theta = 3.49$ and $\zeta_\theta = 2.74$ the persistence exponent $\theta = \eta_\theta/\zeta_\theta \approx 1.274$ is obtained.

	Minimal Wealth model				BS model		Manna model	
	$1d$	$2d$	BA graph	N-clique	$1d$	$2d$	$1d$	$2d$
w_c	0.8175(2)	0.6887(2)	0.6444(2)	0.6076(2)	0.66702(8) [72]	0.328855(4) [75]	0.89199(5) [76]	0.68333(3) [76]
τ	1.12(1)	1.29(1)	1.50(1)	1.50(1)	1.073(3) [72]	1.245(10) [75]	1.112(6) [77]	1.273(2) [77]
γ	2.65(5)	1.58(5)	1.02(5)	1.00(5)	2.70(1) [75]	1.70(1) [75]		
π	2.89(5)	3.94(5)	-	-	3.23(2) [75]			
z	2.77(7)							
β	1.92(2)	0.95(2)	0.52(2)	0.50(2)			2	2
θ	1.541(10)	1.262(10)	1.00(1)	1.00(1)				

Table 3.1: Values of different exponents of Minimal Wealth model are compared with those of existing models in the literature.

$\theta = 1.262(10)$ in $2d$. Estimates of all exponents measured in this chapter have been quoted in Table 3.1 and are compared with their corresponding values in BS and Manna model.

3.5 The Maximal Wealth model

Next we studied the Maximal Wealth model where one agent is necessarily the agent with maximal wealth. The other agent being selected randomly with uniform probability from the neighbors of the first agent. Random re-shuffling of wealth takes place in the same way as in the Minimal Wealth model.

In Fig.15(a) we plot again for the Maximal Wealth model the values of wealth w_i of different agents at a certain instant of time in the stationary state with their positions i along an $1d$ lattice of size $N = 512$. In contrast to the similar plot of the Minimal Wealth model in Fig. 3.1 here an upper cut-off for the wealth has been visible at $w_c^{half}(512) = 1.3924$.

In this case the stationary state wealth distribution $P(w, N)$ takes an opposite

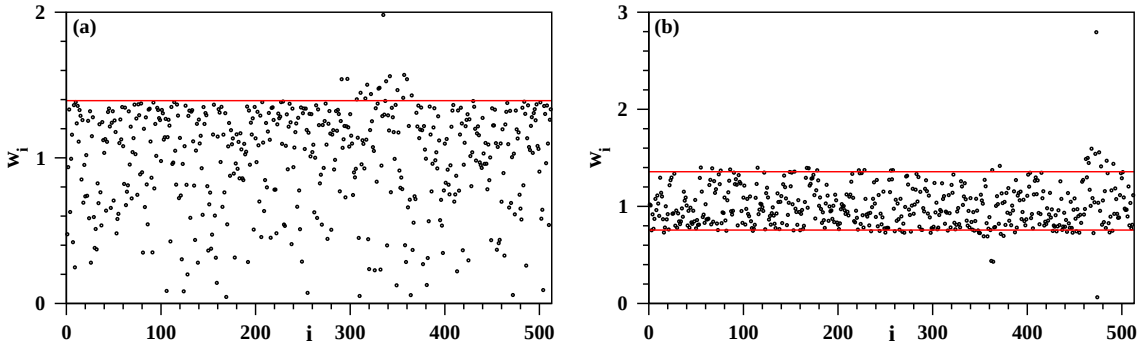


Figure 3.15: (Color online) Values of wealth w_i of different agents in the stationary state are plotted with their positions i along a $1d$ lattice for the (a) Maximal Wealth model, the red line is at $w_c^{half}(512) = 1.3924$. (b) mixture of the Minimal Wealth and Maximal Wealth models with probability $p = 1/2$, the red lines correspond to peaks at $w_c(512) = 0.7559$ and at 1.3574.

shape (Fig. 3.16(a)). A critical wealth $w_c(N)$ exists here as well. $P(w, N)$ takes a Gaussian form elevated by a constant term $c(N)$ for $w < w_c(N)$, whereas for $w > w_c(N)$ it sharply decreases to zero. The parameters of the Gaussian function (Eqn. (3.6)) are different for different N and they vary very systematically with N as: $A(N) \approx 30.48 - 133.35N^{-0.361}$, $\mu(N) \approx 3.74 - 6.66N^{-0.85}$ and $\sigma(N) \approx 1.047 + 0.572N^{-0.321}$ and the constant $c(N) \approx 0.031 + 4.265N^{-0.73}$.

The critical value of wealth $w_c(\infty)$ in the asymptotic limit of system sizes has been estimated by the same method as used for the Minimal Wealth model. The $w_c^{half}(N)$ and $w_c^{slope}(N)$ values have been calculated for $N = 2^8$, 2^{10}, 2^{12} and 2^{14}, extrapolated with $N^{-0.586}$ and $N^{-0.620}$ and the asymptotic values are 1.3610 and 1.3608 respectively. We conclude $w_c(\infty) = 1.3609(2)$ for $1d$. A similar analysis gives $w_c(\infty) = 1.7076(2)$ for $2d$ square lattice, 1.8895(2) for the BA graph and 1.9998(2) for the N-clique. It may appear that the Minimal Wealth and Maximal Wealth models should be symmetric about the average wealth per agent which we have set at $\langle w \rangle = 1$. We have seen above that this indeed not the case since w_c values are 0.8175 and 1.3609 for the Minimal Wealth and Maximal Wealth models respectively. The symmetry between these two models are broken by the presence of a rigid wall at $w = 0$ which means that negative value of wealth of an agent is not allowed.

The avalanche size distributions have been studied as well. A finite-size scaling of these distributions has been done and are plotted in Fig. 3.16(b) using $\log - \log$ scale as before for $N = 2^8, 2^{10}$ and 2^{12}. The scaling exponents are $\eta_\tau = 2.39$ and $\zeta_\tau = 2.10$ respectively giving the value of the avalanche size exponent $\tau = \eta_\tau/\zeta_\tau \approx 1.14(1)$.

Finally we have studied a mixture of the Minimal Wealth and Maximal Wealth models. At every bipartite trade the first agent with minimal wealth is selected with probability p or with maximal wealth with probability $1 - p$. The second agent is selected with uniform probability from the neighbors of the first agent. A snapshot of the individual wealth for $p = 1/2$ at the stationary state for different agents has been shown in Fig. 3.15(b). Here the wealth values are restricted within a 'wealth-band' with sharp cut-offs at a high and a low end. Consequently the shape of the wealth

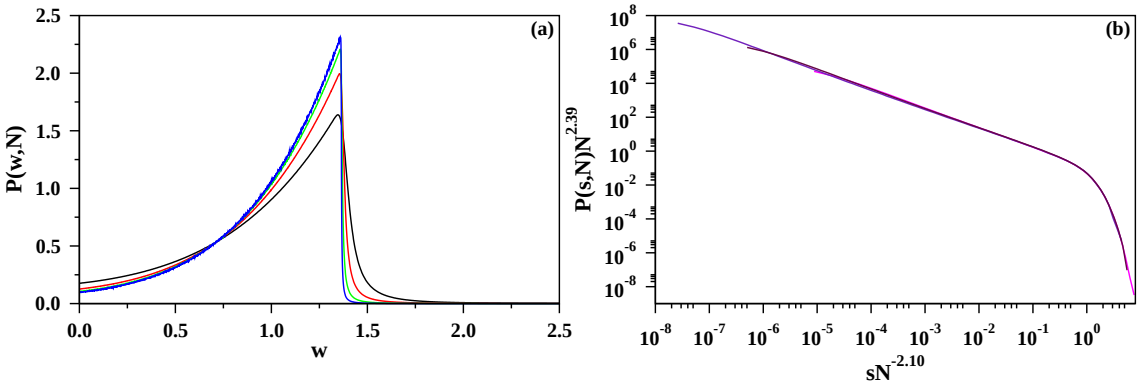


Figure 3.16: (Color online) (a) The wealth distribution $P(w, N)$ vs. w for the Maximal Wealth model for $N = 2^8$ (black), 2^{10} (red), 2^{12} (green) and for 2^{14} (blue). (b) Finite-size scaling of the avalanche size distribution $P(s, N)$ at the critical threshold $w_c(N)$ for $N = 2^8$ (black), 2^{10} (red) and 2^{12} (blue). The scaling exponents are $\eta_\tau = 2.39$ and $\zeta_\tau = 2.10$ which gives the exponent $\tau = \eta_\tau/\zeta_\tau \approx 1.14(1)$.

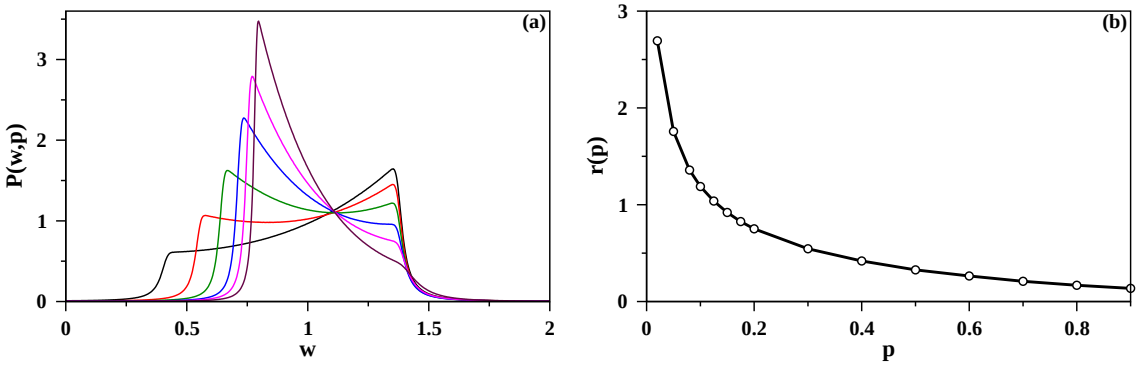


Figure 3.17: (Color online) (a) Mixture of Minimal Wealth and Maximal Wealth models with probabilities p and $1-p$ respectively for $N = 2^9$ and for $p = 0.02$ (black), 0.08 (red), 0.2 (green), 0.4 (blue), 0.6 (magenta) and 0.9 (maroon). (b) Variation of the ratio $r(p)$ of heights of the right peak and the left peak with probability p.

distribution $P(w,p)$ has double peaks for all N and we plot the distribution in Fig. 3.17(a) for $N = 2^9$. Portion of the distribution between the peaks fits excellent to elevated Gaussian distributions with different parameter values for different values of p in the range between 1/2 and 1. On the two sides of this region the distribution decays to zero very fast. The figure shows the plot of $P(w,p)$ vs. w for $p = 0.02$, 0.08, 0.2, 0.4, 0.6 and 0.9. In Fig. 3.17(b) we plot the ratio $r(p)$ of the heights of right peak and the left peak with the probability p.

3.6 Comparison with Other Models

Our model can be compared with the similar models in the literature. (i) In the PIAV model an agent can have negative value of wealth, i.e., this model allows an agent to run into debt. In comparison our model has positive values of wealth for all agents. There is a wall at the zero value of wealth and no agent can have debt. Consequently we will see in the following that the wealth distributions in the Minimal and Maximal Wealth versions of the PIAV model are symmetric with respect to the average money $\langle w \rangle = 1$ where as in our case the wealth distributions are asymmetric. (ii) We have shown for the first time that the persistence time distribution in Minimal Wealth model has power law distributions with very non-trivial exponents in both $1d$ and $2d$. (iii) As an extension of this model we have studied the Maximal Wealth model where the first agent selected is necessarily the richest agent. Further we have studied a probabilistic mixture of the Minimal and Maximal Wealth models. (iv) PIAV model has been claimed to be of the same universality class as that of Bak-Sneppen model where as precise calculation of critical exponents of our model shows that our model is very unlikely to be in the BS universality class.

(i) In the model of Ghosh et al. the cut-off w_o in the wealth is continuously tuned to arrive at the critical poverty line denoted by w_c. In comparison the dynamical process in our model self-organizes the system so that the critical poverty wealth w_c is spontaneously reached. The estimated values of the critical points are close but distinctly different in two models. (ii) Consequently the Ghosh *et al.* model is an example of continuous phase transition where as our model gives a Self-organized critical state. As a result there is no super-critical phase in our model in comparison to the model of Ghosh *et al.* (iii) On the other hand a number of exponents measured in our model are not measured in Ghosh *et al.* model, e.g., avalanche size distribution exponent etc. However the relaxation exponent measured in both models are wide different, i.e., in 1d, $z = 2.77$ in our model compared to 1.9 in Ghosh *et al.*

3.7 Conclusion

Social inequality in terms of economic strengths is ubiquitous for the people of all countries. Perhaps this inequality acts as the major driving force behind the advancement of society. Consequently the mechanism that establishes this inequality in a society is an important issue and attracted the attention of researchers over the last century. Here we have studied a modified version of the conservative self-organized extremal model introduced by Pianegonda *et. al.* which is motivated

by the wealth distribution in a society. In this model the entire wealth of the society is strictly conserved. It evolves by a trade dynamics that takes the society from equality (or any other initial wealth distribution) to a stationary state where strong social inequality is present. The dynamics is an infinite sequence of stochastic bipartite tradings where one of the agents has the globally minimal value of wealth, the other one being selected randomly from the neighbors of the first agent. Our numerical study reveals that this model is one of the simplest models of Self-organized Criticality where the stationary state is non-ergodic. This model is very similar to the self-organized critical Bak-Sneppen model for the ecological evolution of interacting species. Using numerical simulation we have estimated a number of critical exponents for this model on an $1d$ regular lattice, $2d$ square lattice, the Barabási - Albert scale-free graph and on the N-clique graph. We present evidences which suggest that this model does not belong to the universality class of either the Bak-Sneppen model or the Manna model of Self-organized Criticality. This model belongs to a new universality class perhaps because of strict conservation of wealth is maintained in its dynamical rules.

Chapter 4

Disease Spreading Model with Partial Isolation

4.1 Introduction

An infectious disease spreads from an infected (sick) person to a susceptible (healthy) person through some kind of contact. For different diseases the nature of contacts are in general different. For example the infection may spread in physical contact, by air borne viruses or through some intermediate carrier like mosquitoes or other insects. In practice different individuals have different number of people in their contact neighborhood. Thus a closed community with a set of population and a set of their neighbors constitute a contact graph. With time the infected individuals become cured but other individuals may become infected. Thus the disease propagates from one set of individuals to another set of individuals and this process continues. For an individual the infection persists for certain duration of time. For some diseases an individual recovers after infection but he may remain prone to further infection. In this case, there may be an endemic state where a non-zero fraction of population remains always infected. For some other diseases one may become completely immune and cannot have the same disease for a second time [78-80].

In this chapter we present numerical evidence that when infected individuals are partially isolated i.e., have a limited capacity of spreading infection, the threshold rate for propagation of the disease to attain an endemic state is enhanced, always non-zero on any contact graph, including the scale-free networks. In the known models of disease spreading all infected individuals are assumed to be infinitely strong in spreading diseases because they can infect all their neighbors. In comparison our work is likely to be relevant in modeling more realistic situations when each infected individual has a maximum but finite capacity of spreading.

A well-known model of disease spreading in a community is the susceptible-infected-susceptible (SIS) model [78]. Community members are assumed to be positioned at the nodes of an arbitrary graph. An individual can be in only two possible states: susceptible or infected. The number k of edges meeting at an infected node is known as the degree of the node and the infection is transmitted along these edges. In general if a susceptible node has one or more infected neighbors, it gets infected at a continuously tunable rate β. At the same time an infected node recovers at a rate γ. Usually SIS model is studied with an effective spreading rate $\lambda = \beta/\gamma$

choosing $\gamma = 1$. This implies that the node is infected only for a unit time and then recovers to become susceptible in the next time step. In a mean-field analysis the time evolution is described by the following equations:

$$\frac{d\rho_s}{dt} = -\beta\rho_s\rho_i + \gamma\rho_i, \quad \frac{d\rho_i}{dt} = \beta\rho_s\rho_i - \gamma\rho_i, \tag{4.1}$$

where ρ_s and ρ_i are the susceptible and infected population densities respectively so that: $\rho_s + \rho_i = 1$. In a discrete time dynamics all infected nodes at time t become susceptible at time $t + 1$ where as some of the susceptible nodes in time t become infected in the next time step. Starting from an arbitrary initial density of infected nodes the community evolves to a time independent stationary state. We define 'prevalence' as the average density of infected individuals $\rho(\lambda)$ in the stationary state. The SIS model exhibits a non-equilibrium transition from a completely healthy phase with $\rho(\lambda) = 0$ to a partially infected phase with $\rho(\lambda) > 0$ at a critical value λ_c of the infection rate [78-80].

In the susceptible-infected-recovered (SIR) model an individual can be in three possible states: susceptible, infected and recovered [78]. Once infected, an individual is recovered in the next time step and becomes immune to further infection and continues in the recovered state for ever after. One usually starts with all susceptible nodes except one infected node. It is observed how the disease from a single infected node spreads to a large number of nodes of the system. Here also beyond a certain critical rate λ_c the outbreak reaches a finite fraction of the population. Since the infection spreads along the edges of the contact network, SIR model is equivalent to the bond percolation where λ is the bond occupation probability. In a mean-field analysis the time variations of the population densities are governed by

$$\frac{d\rho_s}{dt} = -\beta\rho_i\rho_s, \quad \frac{d\rho_i}{dt} = \beta\rho_i\rho_s - \gamma\rho_i, \quad \frac{d\rho_r}{dt} = \gamma\rho_i, \tag{4.2}$$

where ρ_i, ρ_s and ρ_r are the fractions of the population in the infected, susceptible and recovered states and $\rho_i + \rho_s + \rho_r = 1$ [78-80].

The disease spreading process depends on the spreading rate λ as well as the nodal degrees k. The disease spreads in a branching process. In this case the branching ratio is the average number of susceptible people who got infected by a single infected individual. It is known that when the branching ratio is larger than unity the epidemic persists, otherwise the disease vanishes from the society. When the average number of contact neighbors is large, spreading is ensured even when the spreading rate λ is small but $\langle k\rangle\lambda > 1$. In other words the condition for disease spreading is the spreading rate $\lambda > \lambda_c = 1/\langle k\rangle$ [81-83].

4.2 Model

We argue that a partial isolation measure adopted by the infected individuals affects the disease spreading process [84]. An infected person takes precaution and therefore goes to some kind of isolation to cut-off his contacts from others. However in reality often such an isolated process is not perfect but only partial. As a result effectively the number of susceptible people who are still exposed to his infection turns out to be less, but yet non-zero, than his actual number of susceptible neighbors.

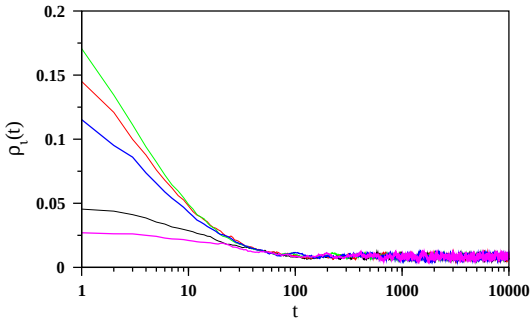


Figure 4.1: (Color online) In the stationary state of the SISI model on BA networks the prevalence does not depend on the densities of the infected nodes at the initial stage. For $N = 2^{14}$, $n = 2$, $m = 2$ and $\lambda = 0.55$ the time evolutions of the densities of the infected nodes $\rho_i(t)$ against the time t have been plotted for different initial densities of infected nodes: $\rho_i(0) = 0.05$ (black), 0.25 (red), 0.50 (green), 0.75 (blue) and 0.95 (magenta).

The effect of partial isolation is incorporated in a modified SIS model by introducing an integer parameter 'n' which is a measure of the spreading capacity. We propose that under partial isolation, an individual's capacity of spreading disease is limited up to a maximum number of n neighbors, irrespective of his actual number of contacts. The capacity n is assumed to be the same for all individuals. Therefore on an arbitrary graph, disease spreads probabilistically from a node at a rate λ to only n susceptible neighbors, randomly selected from his k neighbors. Like in annealed disorder, at every time step t, each infected node randomly selects the set of n susceptible neighbors. We refer this model as the susceptible-infected-susceptible model with partial isolation (SISI).

More specifically the simulation procedure used in studying the SISI is the following. On an arbitrary graph each node represents an individual, who can be in two possible states: infected or susceptible. At an arbitrary intermediate time the microstate of the system is described by specifying the state of each node. The system evolves from one microstate to another by a synchronous updating procedure. In general each node has few infected and few susceptible neighbors. For each infected node n neighboring nodes are randomly selected from its subset of susceptible neighbors and these nodes are marked. In case the total number of susceptible neighbors is less than n, then all of them are marked. In general a marked susceptible node may be the neighbor of multiple infected nodes, yet it is marked only once. Therefore all marked susceptible nodes are on equal footing irrespective of the number of their infected neighbors. When the marking process is complete for all nodes, there are only three types of nodes: susceptible and marked nodes, susceptible and un-marked nodes and infected nodes. All nodes are then updated synchronously using the following procedure: (i) Each susceptible and marked node is infected with a probability λ or is left susceptible with probability $(1 - \lambda)$, (ii) each susceptible but un-marked node is left susceptible and (iii) each infected node becomes susceptible with probability one. The following numerical evidence suggests that assigning a fixed capacity n of disease spreading results non-zero critical

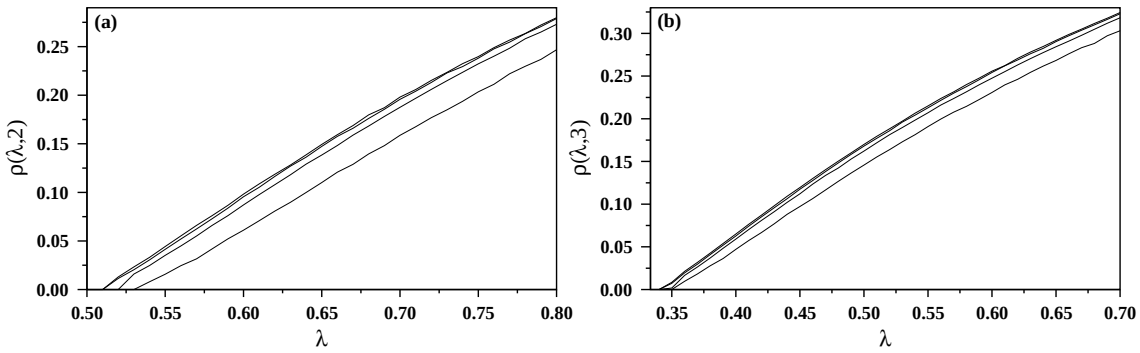


Figure 4.2: Disease spreading model SISI with maximal capacity of spreading n on BA network of size $N = 2^{16}$ with m links per incoming node. The prevalence $\rho(\lambda, n)$ has been plotted with the disease spreading rate λ for different n: (a) $n = 2$ for $m = 2, 3, 4, 5$ (b) $n = 3$ for $m = 3, 4, 5, 6$; parameter m increases from right to left.

spreading rates depending on n for both the SIS and SIR models.

Usually real-world networks have strong heterogeneous structures. One way to characterize these networks is studying the nodal degree distributions. For example the degree distributions of a large number of networks have power law tails. These networks are called the scale-free networks (SFN). It is known that the Internet, World-Wide-Web (WWW), biological networks like protein interaction networks, social networks like collaboration networks or contact networks of sexually transmitted diseases are the important examples of SFNs [85, 86]. Recently the SIS and SIR have been studied on different random and correlated graphs. It has been shown that on SFNs the threshold value λ_c for the endemic state is very small and the prevalence vanishes exponentially as the network size increases to infinity [87].

4.3 SISI Model on Barabási-Albert Network

First we studied the SISI model on the Barabási-Albert (BA) network [85]. For asymptotically large BA networks the degree distribution is a power law: $P(k) \sim k^{-\gamma}$ with $\gamma = 3$ [85]. However for the networks of finite size N the upper cut-off of the nodal degrees is known to grow like $k_{max} \sim N^x$ with $x = 1/2$. The BA networks are constructed using the 'rich get richer' principle. Such a network grows by adding new nodes one by one. A new node gets connected to m distinct nodes of the existing network with probabilities proportional to their degrees. For simulation of BA network a simpler version is often useful: A link of the existing network is randomly selected with uniform probability and one of its two end nodes is connected with probability 1/2 to the new node [88]. Once the BA network has been constructed, each node is assigned the status of a susceptible or an infected individual with probability 1/2. For each infected node n susceptible neighbors are randomly selected and are marked. Therefore one susceptible node may be marked by more than one infected node. However we ignore the multiple markings of a node, i.e., a node can be either marked or unmarked. When marking is complete for all nodes of the graph, a synchronous process runs through the three steps mentioned

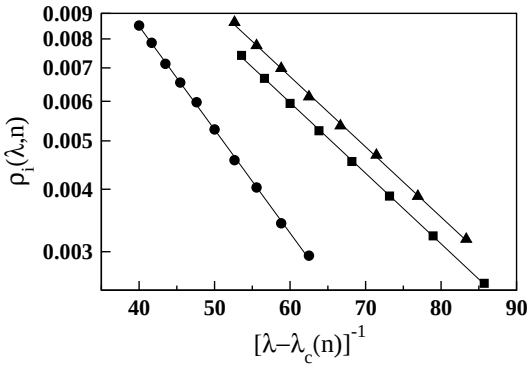


Figure 4.3: Plot of the prevalence $\rho(\lambda, n)$ with the inverse deviation $[\lambda - \lambda_c(n)]^{-1}$ from the critical point λ_c for different maximal capacities n. Calculations are done on a BA network of size $N = 2^{18}$ where every new node is connected to m distinct nodes of the existing network. Symbols used for different (n, m) are: (2,2) (circle), (3,3) (square) and (4,4) (triangle) using the $\lambda_c(n)$ values 0.5, 0.3383 and 0.25 respectively.

above to update the status of every node to obtain the microstate in the next time step.

We first exhibit that in the stationary state the density of infected nodes $\rho_i(t)$ fluctuates with time but its average assumes a fixed time independent value which does not depend on the initial density of infected nodes. We simulated the SISI model on the BA network starting with different densities of initially infected nodes, namely $\rho_i(0) = 0.05$, 0.25, 0.50, 0.75 and 0.95. The density of infected nodes against time plots have been shown in Fig. 4.1 for these five initial densities for $N = 2^{14}$, $n = 2$, $m = 2$ and $\lambda = 0.55$. It is seen that beyond a certain relaxation time of the order of 1000 the average values of prevalence are nearly the same for all $\rho_i(0)$. We conclude that the stationary state is independent of the densities of the infected nodes at the initial stage.

In Fig. 4.2 we plot the prevalence $\rho(\lambda, n)$ for BA networks of size $N = 2^{16}$. Fig. 4.2(a) shows the plots for $n = 2$ and $m = 2$, 3, 4 and 5. As m increases the link density in the BA network increases and the curves gradually converge to a m independent curve. The prevalence monotonically decreases with decreasing λ and vanishes as $\lambda \to 0.5$. Similarly we plot the prevalence against the infection rate for the capacity values $n = 3$ with $m = 3$, 4, 5, 6 in Fig. 4.2(b). These plots are quite similar, only difference being they approach to different values of $\lambda_c(n)$ for different n.

Approach to the threshold value is found to be exponential like $\rho(\lambda, n) \propto \exp(-C/[\lambda - \lambda_c(n)])$. This is seen in Fig. 4.3 where we have plotted $\rho(\lambda, n)$ against $[\lambda - \lambda_c(n)]^{-1}$ on a semi-log scale for $N = 2^{18}$ which fitted nicely to a straight line and we estimated in Fig. 4.3 that $\lambda_c(2) = 0.5$, $\lambda_c(3) = 0.3383$ and $\lambda_c(4) = 0.25$. Therefore our numerical results indicates that for SISI on BA networks the critical threshold for the infection rate $\lambda_c(n) \approx 1/n$.

The SISI model can be generalized in the following way. A susceptible node having multiple infected neighbors may have higher probabilities to become infected. We have further studied this modified model on BA network when a susceptible node

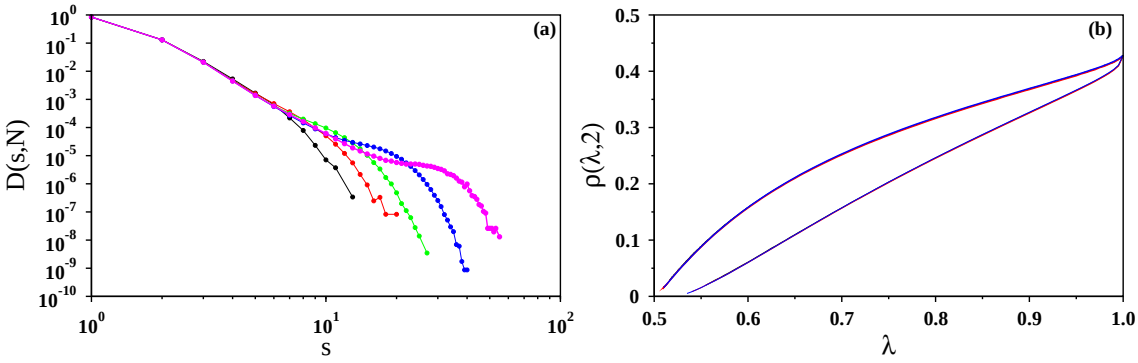


Figure 4.4: (Color online) (a) Probability distribution $D(s)$ for the number of times s a susceptible site can be marked in gSISI model on BA network. Using $n = 2$ and $m = 2$ the system size N has been varied from 2^{10} to 2^{18}, increased by a factor of 4 (from left to right). For large values of N the variation approaches a power law with the exponent ≈ 4.8 followed by a hump at the tail. (b) Plot of the prevalence $\rho(\lambda, 2)$ against the infection rate λ for the SISI (lower curves) and for the gSISI (upper curves). The colors black, red and blue in both sets correspond to $N = 2^{12}, 2^{14}$ and 2^{16} respectively.

is marked s times by its multiple infected neighbors. It is assumed that such a node has the probability proportional to s to become infected. This model will be referred as the generalized SISI or briefly gSISI. To get an idea how frequently such nodes occur in the stationary state we first calculate the probability distribution $D(s)$ that a node is marked s times. In Fig. 4.4(a) we plot the $D(s)$ against s on a double logarithmic scale for five different system sizes. For the large system sizes the variation approaches a power law with an exponent value ≈ 4.8 followed by a hump at the tail. It is also noted that $D(1) \approx 0.84$ and the rest of the multiply marked nodes with $s > 1$ values have the total probability ≈ 0.16. This 16 percent multiply marked susceptible nodes makes the difference between the SISI and the gSISI. In Fig. 4.4(b) we present the prevalence $\rho(\lambda, 2)$ against λ for BA networks of size $N = 2^{12}, 2^{14}$ and 2^{16} with $n = 2$, $m = 2$ and observe that the prevalence for gSISI is always larger than that of SISI for the whole range of $\lambda > 1/2$. We understand this in the following way. Since $\lambda > 1/2$ the 16% nodes marked $s > 1$ times are always infected. This enhances the prevalence in gSISI compared to that in SISI. Further we observe numerically that variation of prevalence for gSISI has little dependence on the network size N. We plotted $\rho(\lambda, 2)$ against λ curve for $N = 2^{12}, 2^{14}$ and 2^{16}. They nearly overlapped and approach the zero prevalence at $\lambda_c(N) = 0.516, 0.509$ and 0.510 respectively. This is to be compared with SISI whose corresponding values of $\lambda_c(N)$ are 0.555, 0.540 and 0.535 respectively. So we conclude that $\lambda_c(2)$ for the gSISI approaches the asymptotic value of $1/2$ faster and perhaps has less finite size dependence.

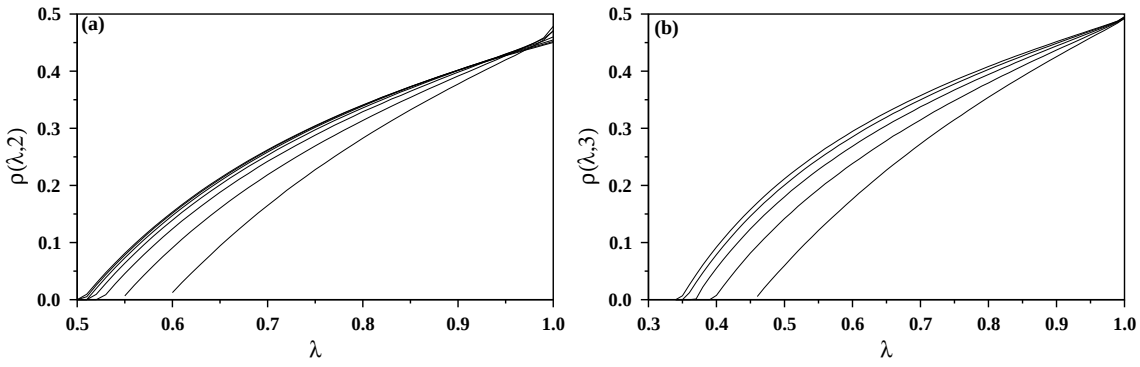


Figure 4.5: SISI model studied on the largest component of the Random graphs of different average degrees $\langle k \rangle$. The prevalence $\rho(\lambda, n)$ at the steady state has been plotted for two different maximal spreading capacities n: (a) For $N = 2^{16}$, $n = 2$, $\langle k \rangle = 2$ to 8 and (b) For $N = 2^{16}$, $n = 3$, $\langle k \rangle = 2$ to 6; the parameter $\langle k \rangle$ increases from right to left.

4.4 SISI Model on Random Graphs

The SISI model has also been studied on Random Graphs (RG) [3]. To generate a RG one starts with N nodes with each node being a component of size unity and no links. At an arbitrary intermediate stage a pair of nodes is selected randomly with uniform probability and are linked if they are not connected already. As more and more links are dropped sizes of different components increase. As the link density p = (number of links)/N is increased, different components merge among themselves and thus larger components are formed. The Order Parameter (OP) is measured by the fractional size of the largest component. It is known that at the critical link density of $p_c = 1/2$ the OP increases very rapidly indicating a continuous transition. At p_c correlations appear in the system and the system makes a transition from a phase of local connectivity to a phase of global connectivity. We studied SISI on RGs using average degree $\langle k \rangle = 2p$ as a parameter. After generating a Random Graph with a pre-assigned value of $\langle k \rangle$ we label different components by the burning method. We study SISI model only on the largest connected component.

As before, initially the nodes are randomly assigned the susceptible and infected status with probability 1/2. The dynamics is then switched on with a pre-assigned value of the infection rate λ. A specific value of the maximal capacity n of spreading is assigned. After some initial relaxation period the time variation of prevalence becomes stationary. At this stage the prevalence has been averaged over a large number of time steps. This is repeated for different initial Random Graph configurations, different values of average degrees $\langle k \rangle$ and n. The average prevalence $\rho(\lambda)$ is then plotted in Fig. 4.5(a) and Fig. 4.5(b) for $n = 2$ and 3 respectively. Here also the critical infection rate $\lambda_c(n, \langle k \rangle)$ tends to $1/n$ as $\langle k \rangle$ increases.

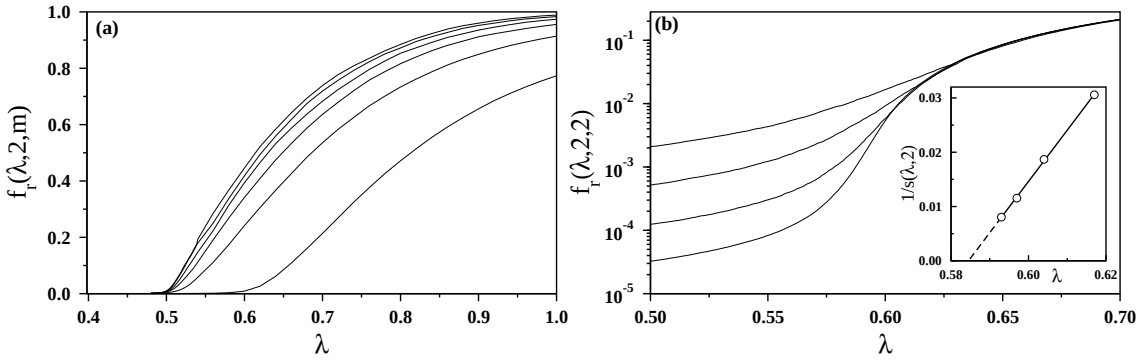


Figure 4.6: (a) SIRI model on the BA network of size $N = 2^{14}$. Plot of the fraction $f_r(\lambda, 2, m)$ of nodes that recovered in the SIRI model of disease spreading on BA network. The maximal capacity parameter used is $n = 2$ and m values are increased from 2 to 7 from right to left. (b) Plot of the same for $(n, m) = (2, 2)$ but for $N = 2^{12}, 2^{14}, 2^{16}$ and 2^{18}, N increases from top to bottom. Inset shows the plot of inverse of the largest slope $s(\lambda, 2)$ of these four curves with λ which extrapolates to $\lambda_c(2, 2) = 0.585$.

4.5 SIRI Model on Barabási-Albert Network

Next we study the Susceptible-Infected-Recovered model with partial Isolation (SIRI) with limited capacity of spreading. For example, on a BA network of size N initially all nodes are made susceptible and then a seed node is randomly selected and infected at time $t = 0$, the disease spreads from this node. For each BA network the spreading process is simulated from 100 to 800 such initial seeds. The data is averaged over 100 independent BA networks. At an intermediate time t, an infected node i with degree k_i has some infected and some susceptible neighbors in general. As before a subset of n nodes are then randomly selected from the set of susceptible neighbors and marked. If n is larger than the number of susceptible neighbors, all of them are marked. When the marking is complete for all infected sites the system is updated as: (i) All susceptible and marked nodes are infected with probability λ, left susceptible with probability $1 - \lambda$, (ii) all susceptible and unmarked nodes are left susceptible and (iii) all infected nodes are recovered with probability one. The recovered nodes remain recovered ever after and never become susceptible again. Nodes which got infected at time t become the source of infection at time $t + 1$ and this process continues. We first fix the spreading capacity n of the infected individuals and then increase the average number of links m that come out of a newly added node of the BA network. The whole simulation is repeated for $n = 2$, 3 and 4. The data for $n = 2$ only is shown in Fig. 4.6(a). Here we plot $f_r(\lambda, 2, m)$ versus λ. Here f_r is the fraction of nodes which get infected and subsequently recovered till the system is completely disease free. The size of the BA network used is $N = 2^{14}$ and the parameter m is varied from 2 to 7. It is observed that for all curves with $m > 2$ the critical threshold $\lambda_c(2, m) \approx 0.5$ but not for the case $(n, m) = (2, 2)$. To study this case in more detail we used four different network sizes $N = 2^{12}, 2^{14}, 2^{16}$ and 2^{18} and plotted $f_r(\lambda, 2, 2)$ against λ on a log-lin scale in Fig. 4.6(b). The slope of the plot gradually becomes larger as the system size increased from top to

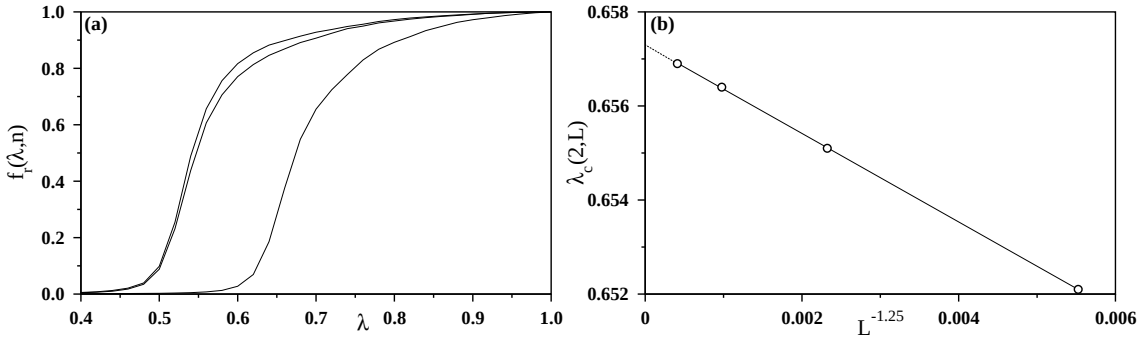


Figure 4.7: SIRI model on the square lattice. (a) The fraction $f_r(\lambda, n)$ of the recovered sites has been plotted with the disease spreading rate λ for the capacity of spreading $n = 2$, 3 and 4 (from right to left). (b) Plot of the variation of $\lambda_c(2, L)$ against $L^{-1.25}$ which extrapolates to $\lambda_c(2, \infty) = 0.657$ as $L \to \infty$.

bottom. We calculated the maximum of the slope $s(\lambda, 2) = d \log f_r(\lambda, 2, 2)/d\lambda|_{max}$ which is expected to diverge as $N \to \infty$. Therefore we plot inverse of $s(\lambda, 2)$ with λ in the inset, and extrapolate to $s(\lambda, 2) \to \infty$ which intersects the λ axis at 0.585. We conclude that the particular case of SIRI with maximal capacity $n = 2$ on BA networks with $m = 2$ has a different epidemic threshold $\lambda_c = 0.585$. Further we repeated these calculations for $n = 3$ and 4 on BA networks of $m = n$ and higher. The threshold values obtained are approximately $1/n$ for all m.

4.6 SIRI Model on Square Lattice

SIRI model on square lattice is also found to be interesting. In Fig. 4.7(a) the fraction of recovered sites $f_r(\lambda, n)$ is plotted against λ for the capacities $n = 2$, 3 and 4 for the square lattices of size $L = 128$. It is observed that the variation for $n = 2$ is much different from those of $n = 3$ and 4. The $n = 4$ case is exactly the ordinary bond percolation case and its critical threshold is found to be very close to 0.5 [27]; a similar value is obtained for $n = 3$ as well. However the threshold for $n = 2$ is found to be approximately 0.66 which is much different from 0.5. To find out the precise value of $\lambda_c(2, \infty)$ we studied larger lattice sizes 64, 128, 256 and 512. To measure the critical disease spreading threshold in the thermodynamic limit the $\lambda_c(2, L)$ values are extrapolated with $L^{-1/\nu}$ in Fig. 4.7(b) to its infinite size threshold value $\lambda_c(2, \infty) = 0.657$ as:

$$\lambda_c(2, L) = \lambda_c(2, \infty) + AL^{-1/\nu} \tag{4.3}$$

where ν is found to be 0.80(2).

We plot in Fig. 4.8(a) only $f_r(\lambda, 2)$ for these system sizes and observe that as the system size increases the curves become sharper. Numerically the coordinates $\lambda_c(2, L), f_r(\lambda_c(2, L), 2)$ of the points with largest slopes are estimated and a finite size scaling is done around these points using the following scaling form as shown in Fig. 4.8(b).

$$f_r(\lambda, 2)L^{0.18} \sim \mathcal{G}\{[\lambda - \lambda_c(2, L)]L^{0.66}\}. \tag{4.4}$$

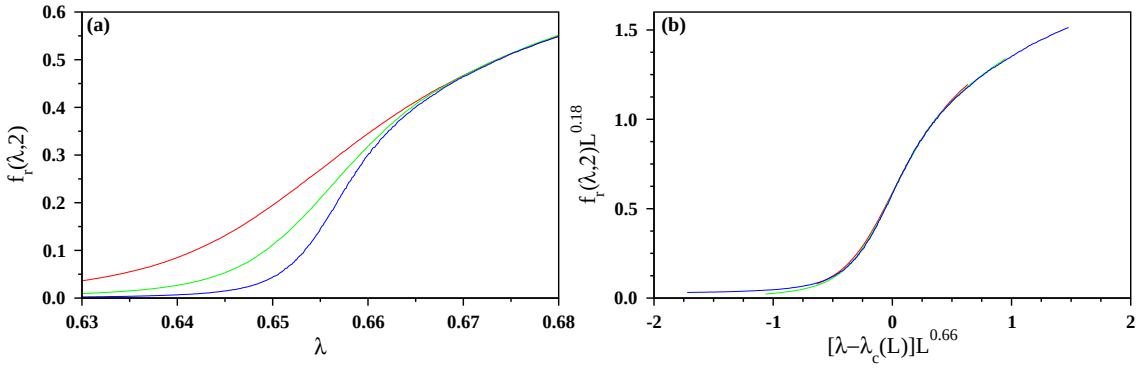


Figure 4.8: (Color online) SIRI model on the square lattice. (a) Plot of $f_r(\lambda, 2)$ for $n = 2$ only but for three different system sizes $L = 128$ (red), 256 (green) and 512 (blue), the slope increases with the system size. (b) For each of these curves the points where the slope is largest have been identified and a finite size scaling described by Eqn. (4.4) has been done around these points in (b).

The universal scaling function $\mathcal{G}(x)$ should have the L independent form x^τ in the thermodynamic limit of $L \to \infty$. For such a relation to hold good one should have

$$L^{-0.18+0.66\tau} \sim 1 \tag{4.5}$$

which implies $\tau = 0.18/0.66 \approx 0.27$. This value of τ is quite consistent with our directly measured numerical value of $\tau = 0.30$ from a replot of the data in Fig. 4.8(b) on a $\log - \log$ scale.

4.7 Summary

It has been observed already in the literature that the critical rate of disease spreading tends to vanish on the scale-free graphs as the system size is increased [87]. In this article we have looked this problem in further detail. We introduced a parameter n which measures the capacity of an infected individual to spread the disease i.e., the maximum number of susceptible people an infected individual possibly can infect. When an infected individual becomes susceptible again in the next time step, we call this model as SISI. SISI model with different values of n have been studied on the BA network for different values of m, the number of links with which a new node is connected. For a given n the critical infection rates $\lambda_c(m, n)$ gradually approaches $\lambda_c(n)$ as m increases. Using extensive numerical calculations we find $\lambda_c(n)$ is non-zero for all n and approaches to zero as $1/n$.

The SISI model has also been studied on the Random Graphs. Since RGs have many components, we have sorted out the giant component at the critical link density and studied the SISI on the giant component. Here also different nodes have different degrees k with the average $\langle k \rangle = 1$. Like the SISI on BA network we have imposed a maximal cut-off n of the number of susceptible neighbors an infected node can possibly infect. Our numerical results show that here also the critical infection rate $\lambda_c(n, \langle k \rangle)$ tends to $1/n$ as $\langle k \rangle$ increases.

The SIRI version has been studied on the square lattice. Here we studied the three possible values of the maximal capacity $n = 2$, 3 and 4. For $n = 4$ it is the ordinary SIR model on the square lattice. For $n = 3$ we see that the fraction of recovered sites $f_r(\lambda, 3)$ against λ curve almost coincides with that for the $f_r(\lambda, 4)$ against λ. However the same plot for $f_r(\lambda, 2)$ is quite different (Fig. 4.7) and we have obtained $\lambda_c(2, \infty) = 0.657$ after finite size extrapolation. The corresponding correlation length exponent ν is found to be 0.80(2) which is much different from the exact value of 4/3 in the case of bond percolation.

4.8 Conclusion

To conclude we have studied SIS and SIR type disease spreading model where each infected individual has a finite maximal capacity n of infection spreading. Numerical study of this model on a number of different substrate graphs indicate that the critical infection rate $\lambda_c(n)$ does depend on n and is always non-zero. A similar modification of the SIR model on square lattice indicates that for $n = 2$, may belong to a new universality class compared to the ordinary bond percolation [27].

Chapter 5

Space-filling Percolation

5.1 Introduction

We have already described in the introductory chapter that the phenomenon of Percolation deals with the appearance of long range correlations in a disordered system. Random occupation of space by some elementary objects in a system, extended over a finite region of space, gives rise to a correlation length. This length diverges as the density of the objects, in other words the probability of occupation, increases to certain critical density. In this chapter we have studied a percolation problem, whose building units are two dimensional growing circular disks. These disks come in contact with one another and in this way form an assembly of touching disks which we look upon as a contact network. Apart from studying the well known features of this percolation problem we have studied the network characteristics of the contact network of the assembly of touching disks.

After its introduction by Hammersley [23] a number of different variations of the percolation problem have been introduced and studied in the literature. Almost all of these models have exhibited continuous transitions across their percolation transition points. Only once exception that shows a discontinuous transition is the Bootstrap Percolation [89] modeling the competition between exchange and crystal field interaction in some magnetic materials. In this percolation problem, depending on the lattice structure, it may happen that culling of even a single spin evacuates a globally connected system in a recursive process. While this was the background picture, only recently a new concept of "Explosive Percolation" (EP) has been introduced which suggested that the nature of transition may indeed be discontinuous in some percolation models [90] as per some general prescription. This immediately implies that the associated Order Parameter, estimated by the size of the largest cluster, should undergo a discontinuous change at the point of transition. In the context of percolation theory such a discontinuous transition can happen only when the largest cluster merges with the maximal of the second largest cluster, which also has a macroscopic size [91, 92].

The original model of Explosive Percolation had been studied on complete graphs [90], later different versions of EP have been studied on the square lattices [93, 94], on scale-free networks [95, 96] and also for real-world networks [97]. In most of these cases, on the basis of numerical results, it has been claimed that the sudden jumps in their Order Parameters are indeed the signatures of the discontinuous transitions.

However, recently it has been shown that, although this class of EP models exhibit very sharp changes in their Order Parameters for finite size systems and appear to exhibit discontinuous transitions, they actually have continuous transitions in the asymptotic limit of large system sizes [98].

Here, we propose and study a variant of the Continuum Percolation (CP) model of growing circular two-dimensional disks that finally led us to arrive at a similar conclusion, i.e., the proposed model exhibited a similar sharp but continuous transition [99]. Briefly, here we study the percolation problem in an assembly of growing circular disks. These disks are released one at a time with their centers located at random positions in the uncovered region, the 'uncovered region' being the space which is not covered by any of the disks. These disks grow at a uniform rate so that at any arbitrary intermediate stage different disks have different radii. In general, a slight overlap among them is allowed when a disk grows to overlap with another one for the first time. This mutual overlap ensures that the global connectivity is achieved at a certain density of disks. In the long time the pattern of disks cover the entire space. We are interested in the study of percolation properties of this space-filling pattern which, to our knowledge, has not been studied yet.

Various models of space-filling patterns have been studied in the literature characterized by their fractal dimensions d_f. In the Apollonian gasket, disks are placed iteratively in the curvilinear triangular spaces between the sets of three mutually touching disks. Consequently the area of the uncovered space gradually vanishes and has the fractal dimension $d_f = 1.305686729(10)$ [100]. In Space-filling bearing patterns a region of two dimensional space is covered by an infinite set of mutually touching disks which can rotate without slipping with a fixed peripheral speed. Different patterns have different fractal dimensions which assume values between 1.3057 and 1.4321 [101]. Due to their deterministic algorithms the global connectivity of these patterns is guaranteed. On the other hand, in one model of random space-filling pattern of touching disks, such a global connectivity is not ensured. Here the nucleation centers of the disks are selected at random locations in the uncovered region one after another. After introduction when a disk grows all other disks remain frozen. Such a disk grows till it gets in contact with another disk for the first time when it stops. Such a pattern has the fractal dimension $d_f \approx 1.64$ [102]. Recently space-filling patterns in three dimensional random bearings have been studied in [103]. All these patterns, in the limit of infinite number of generations, are space-filling. It is known that, while this limit is being taken, the dust of remaining uncovered pore spaces form a Fractal set. The fractal dimension d_f can be estimated following [104].

5.2 The Model

The pattern of circular disks is generated within a unit square box placed on the $x - y$ plane. The radii of all disks grow uniformly at a continuously tunable rate δ. To generate a pattern, a specific value of δ is assigned for all disks. The time t is a discrete integer variable that counts the number of disks released. Therefore at time t the pattern has exactly t disks of many different radii. Initially the square box is completely empty. Then at each time step a new disk with zero radius is

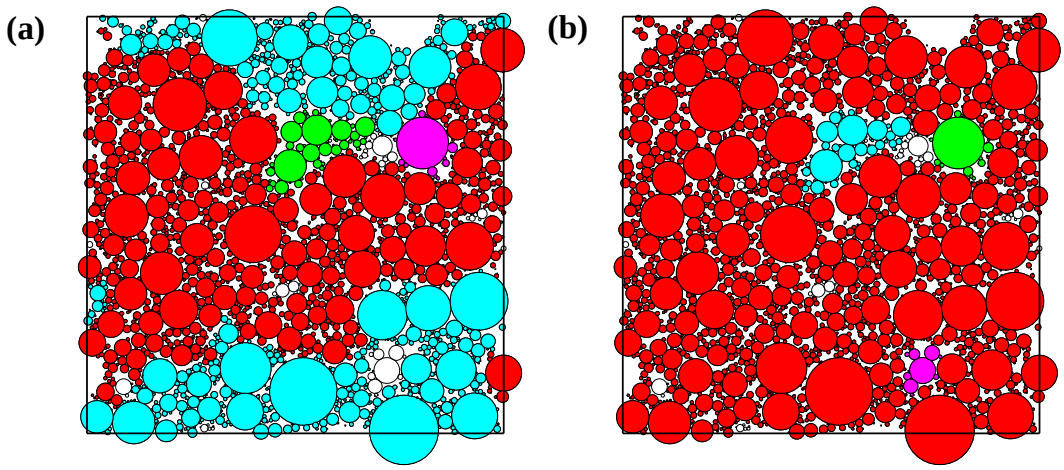


Figure 5.1: (Color online) Top four largest clusters of the disk patterns (with $\delta = 0.001$) right before (a) and after (b) the maximal jump in the largest cluster. Different clusters have been shown by circles filled using different colors and rest of the circles are kept unfilled. (a) At time $t = 1973$ the number of disks in the top four large clusters are 1167 (Red); 667 (Cyan); 57 (Green); 16 (Magenta), (b) and at time $t = 1974$ they are 1836 (Red); 57 (Cyan); 16 (Green); 13 (Magenta).

introduced. Within the square box the assembly of disks cover a region of space which is called the 'covered region', the remaining space being the uncovered region. While growing, once a disk overlaps with another disk, it stops immediately and does not grow any further. Such a disk is called a 'frozen' disk.

At any arbitrary intermediate time step the following activities take place: (i) A point with coordinates $\{x_i, y_i\}$ is randomly selected with uniform probability anywhere within the uncovered region. More specifically a pair of coordinates $(x_i, y_i), 0 \leq x_i, y_i < 1$, with uniform probabilities is selected. If it represents a point within the covered region, it is rejected and another trial is made. This process continues until a random point within the uncovered region is reached. A circular disk of radius δ is placed with its center fixed at this point. (ii) Simultaneously, the radii of all other non-frozen disks are also increased synchronously by the same amount δ. Every growing disk is checked to see if it has overlapped with any other disk; if it has, its growth is stopped and it is declared to be a frozen disk.

Gradually, the space within the square box is increasingly covered by the disks and therefore the amount of uncovered area decreases with time. We define the control variable p as the sum of the areas of all disks. It may be noted that p is slightly larger than the actual "area coverage" since the overlapped areas are doubly counted in the total sum of disk areas. However it has been observed that the total overlap area tends to vanish in the limit of $\delta \to 0$ and we refer to p as the area coverage in the following discussion.

In a particular run, the simulation is stopped only when the area coverage p reaches a pre-assigned value or some pre-defined condition becomes valid. For example, to reach the percolation point, the run is terminated only when a global connectivity appears through the overlapping disks from the top to the bottom for the first time. If one continued further, a stage would come when the different pieces

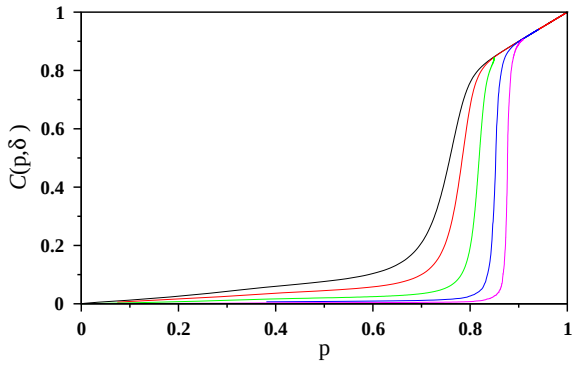


Figure 5.2: (Color online) The Order Parameter $\mathcal{C}(p, \delta)$ has been plotted against the area coverage p for growth rates $\delta = 0.002$ (black), 0.0008 (red), 0.0002 (green), 0.00004 (blue) and 0.00001 (magenta); the value of δ decreases from the left to the right.

of uncovered regions would be so tiny that any newly introduced disk would freeze immediately at the first time step. This would be the natural exit point of the simulation. However in most cases of our calculation simulations were run up to the percolation point.

We are interested in studying the percolation process of this growth model. Multiple overlapping disks form different clusters. Discs of a specific cluster are connected among themselves through overlaps. The size s of a cluster is determined by the sum of the areas of all disks of the cluster. It has been observed, that in an arbitrary pattern near the percolation point, typically there are two large clusters. For example in Fig. 5.1(a) we exhibit a typical disk pattern just before the percolation point at time $t = 1973$ grown at a rate $\delta = 0.001$. The four top largest clusters are shown by disks, filled using different colors. Right at the next time step $t = 1974$ two small disks connect the top two largest clusters so that the size of the largest cluster jumps from 0.439 to 0.747 (Fig. 5.1(b)). This behavior is typical of the percolation process studied here. Prior to the percolation point the largest and second largest clusters have a tendency to compete and grow simultaneously while maintaining their comparable sizes. We define the percolation point to be the time when the Order Parameter jumps by a maximum amount in a single time step. This happens only when the largest cluster merges with the maximal of the second largest cluster. In the following we present simulation results exhibiting this behavior.

Looking at the structures of different clusters in Fig. 5.1, it may be possible to understand qualitatively the reason behind the observed discontinuous jump in the size of the largest cluster. The main difference between the ordinary percolation problem and our model is, the disks here are not static objects, they actually grow uniformly till they overlap with other disks. When two large clusters are close to each other, there are small vacant regions at the surface of these clusters which separate them, but there are many more small vacant regions within the bulk of the clusters. Since new disks are dropped in the vacant regions with uniform probabilities it is much more likely that a new disk would start growing within the bulk. When such a disk grows, it would be a member of this particular cluster with probability one

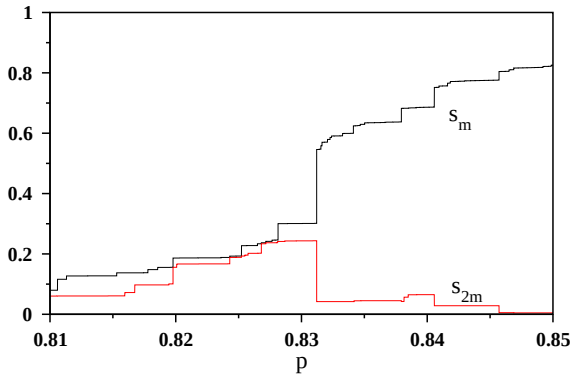


Figure 5.3: (Color online) Variation of the sizes of the largest (black) cluster s_m^α and the second largest (red) cluster s_{2m}^α with the area coverage p for a single run and for $\delta = 0.0001$. While s_m^α increases monotonically, s_{2m}^α increases to a maximum and then drops to a small value when the second largest cluster merges with the largest cluster. This corresponds to the maximal jump in the largest cluster and is identified as the percolation point for the α-th run.

since these vacant spaces are surrounded by disks of the same cluster. In this way, the vacant regions inside a cluster gradually tend to fill up and the cluster tends to increase its size (the area covered by it) by becoming more and more compact in the bulk, and in the limit of the growth rate $\delta \to 0$ and $p_c(\delta) \to 1$ the entire pattern becomes compact. Since the bulk growth rate is faster than the surface growth, two neighboring clusters take more time to merge, but when they merge they already become nearly compact and macroscopic in their sizes. We argue that qualitatively this is the reason that the largest cluster has a macroscopic jump in its size when it merges with the second largest cluster, and thus the possibility of a discontinuous transition arises.

Given a specific growth rate δ one generates the disk assembly until the percolation point. At this stage the distribution $n(r, \delta)$ of radii r of the disks in the pattern is defined. Similarly, the total number of such disks in the subset whose radii are at least r, is denoted by $N(r, \delta)$; the total perimeter of all disks in the subset is denoted by $P(r, \delta)$ and the total remaining uncovered area external to all disks in the subset is denoted by $A(r, \delta)$. It is assumed that at the percolation point, in the limit of $r \to 0$ all these quantities vary as some powers of r as follows [104]:

$$\begin{aligned} n(r,\delta) &= \Sigma_{r_i=r} 1 \sim r^{-(d_f+1)} \\ N(r,\delta) &= \Sigma_{r_i \geq r} 1 \sim r^{-d_f} \\ P(r,\delta) &= 2\pi \Sigma_{r_i \geq r} r_i \sim r^{1-d_f} \\ A(r,\delta) &= 1 - \pi \Sigma_{r_i \geq r} r_i^2 \sim r^{2-d_f} \end{aligned} \tag{5.1}$$

In each case, the summation is taken over all disks whose radii are larger than r.

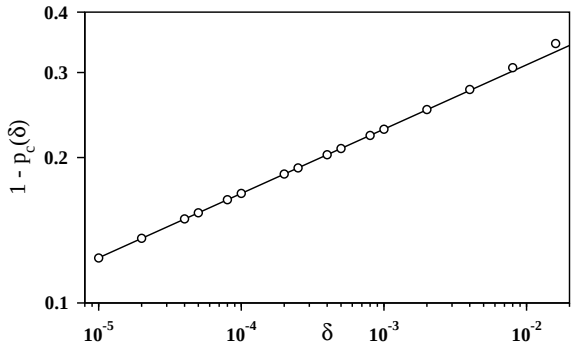


Figure 5.4: The deviation $1 - p_c(\delta)$ of the percolation threshold from unity has been plotted with the growth rate δ on a $\log - \log$ scale. It is observed that as $\delta \to 0$ the deviation vanishes as a power of δ as: $1 - p_c(\delta) \sim \delta^{0.133}$.

5.3 The results

Let $s_m(p, \delta)$ denote the size (i.e., the maximal covered area) of the largest cluster. Then the Order Parameter $\mathcal{C}(p, \delta)$ of the growth process is determined by the average size of the largest cluster of the pattern for an area coverage p,

$$\mathcal{C}(p, \delta) = \langle s_m(p, \delta) \rangle \tag{5.2}$$

the average being taken over a large number of uncorrelated growth processes. Since no checking of the global connectivity is required in this part of the simulation, we have used the periodic boundary condition along both the x and y axes. In Fig. 5.2 we plotted $\mathcal{C}(p, \delta)$ versus p for five different values of the growth rate δ. It has been observed that for every plot around a specific value of $p = p_c(\delta)$ the growth of the Order Parameter is very sharp. This happens because for a typical run α the maximal jump in $s_m^\alpha(p, \delta)$ takes place at p_c^α.

A closer look into the growth process reveals that this maximal jump in the largest cluster occurs only when the maximum of the second largest cluster merges with the largest cluster. This has been exhibited explicitly in Fig. 5.3 where we plot the sizes of the largest cluster s_m^α and that of the second largest cluster s_{2m}^α with the area coverage p for a single run α. While s_m^α grows monotonically, the growth of s_{2m}^α is not so because it reaches a maximum and then falls to a much lower value. We assume that always the smaller cluster merges with the larger cluster. Therefore, when s_{2m}^α merges with s_m^α it is the third largest cluster s_{3m}^α that becomes the second largest cluster s_{2m}^α. This may happen a few times and we mark that particular value of $p = p_c^\alpha$ where the maximal jump in s_m^α takes place.

We define $p_c^\alpha(\delta)$ as the percolation threshold of the α-th run [92]. This value is then averaged over a large number of un-correlated runs and the percolation threshold $p_c(\delta)$ is defined as:

$$p_c(\delta) = \langle p_c^\alpha(\delta) \rangle. \tag{5.3}$$

It may be observed in Fig. 5.2 that the values of percolation thresholds $p_c(\delta)$ are gradually shifting towards unity as $\delta \to 0$. To see this approach quantitatively, we plotted $1 - p_c(\delta)$ against δ on a $\log - \log$ scale in Fig. 5.4. For small values of δ

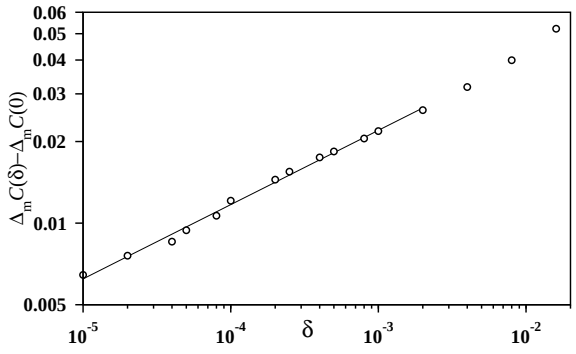


Figure 5.5: How the average value of the maximal jump in the Order Parameter $\Delta_m\mathcal{C}(\delta)$ approaches to $\Delta_m\mathcal{C}(0)$ has been shown. After some initial curvature the plot fits to a straight line as $\delta \to 0$. This implies that Eqn. (5.7) indeed holds good with $\Delta_m\mathcal{C}(0) = 0.16$ and $\mu = 0.274(5)$.

the data points indeed fit nicely to a straight line with slope 0.133(5). Therefore we may write that as $\delta \to 0$,

$$1 - p_c(\delta) \sim \delta^{0.133}. \tag{5.4}$$

This implies that in the limit of $\delta \to 0$ the area coverage at the percolation point becomes unity. In other words when the growth rate is infinitely slow, global connectivity appears for the first time when the entire space is covered by the disks and this limiting pattern is therefore space-filling. We therefore call this problem "space-filling percolation". The percolation threshold has also been determined using the usual definition, i.e., when the global connectivity appears for the first time in the system. In this case, the periodic boundary condition along the x-axis and open boundary condition along the y-axis have been used. For a specific run, as more and more disks are released, we keep track as to whether connectivity between the top and the bottom boundaries of the unit square box through the system of overlapping disks has appeared. When such a connectivity appears for the first time, we refer to the corresponding pattern of disks as the percolation configuration and use the total area coverage as the second definition of the percolation threshold for this particular run. As before, an average of these threshold values for a large number of independent runs gives us the value of $p_c(\delta)$. It has been observed that the percolation thresholds measured using the two methods differ by small amounts, e.g., 0.039 and 0.019 for $\delta = 0.001$ and 0.0001 respectively and this difference gradually diminishes as $\delta \to 0$.

Right at the percolation threshold $p_c^\alpha(\delta)$ of the α-th run the disk pattern has at least one or, may have even more than one disk which have the largest radius $r_m^\alpha(\delta)$. The averaged radius $\langle r_m(\delta)\rangle$ of the largest disk, averaged over many independent runs, decreases as the growth rate δ decreases. Numerically, a power law form associated with a logarithmic correction has been observed to be the best representation of this variation:

$$\langle r_m(\delta)\rangle \sim \{\delta \log(1/\delta)\}^{0.331}. \tag{5.5}$$

The logarithmic correction in Eqn. (5.5) is supported by a detailed numerical analysis. A plot of $\langle r_m(\delta)\rangle$ vs. δ on a $\log - \log$ scale (not shown here) exhibits slow but

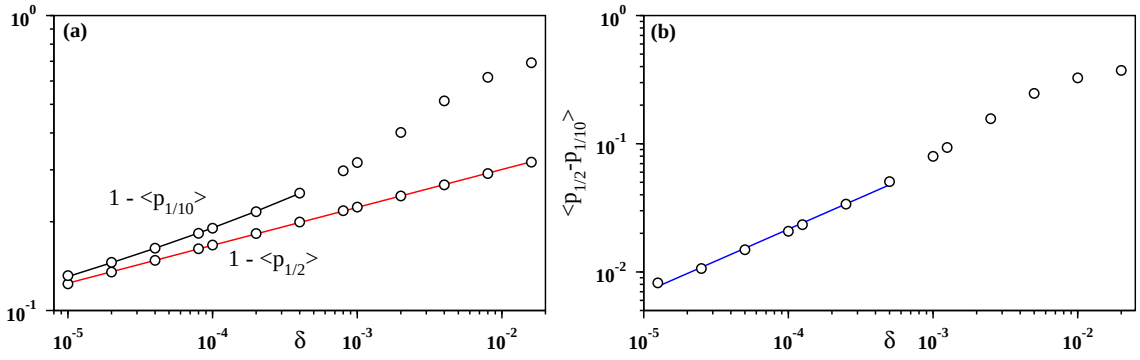


Figure 5.6: (Color online) (a) The quantity $1 - \langle p_{1/2} \rangle$ grows as $0.556\delta^{0.131}$ and $1 - \langle p_{1/10} \rangle$ fits very well to the form $0.556\delta^{0.131} + 3.00\delta^{0.523}$. (b) Across the window $\Delta p = p_{1/2} - p_{1/10}$ the Order Parameter $\mathcal{C}(p, \delta)$ jumps from 1/10 to 1/2. The average size of this window has been plotted with δ on log − log scale and is observed to vanish as $2.45\delta^{0.50}$ as $\delta \to 0$.

systematic increase of the local slopes as $\delta \to 0$. On the other hand, plotting the same data of $\langle r_m(\delta) \rangle$ against $\delta \ln(1/\delta)$ gives all values of local slopes very close to 1/3 and no systematic variation is observed. We therefore conclude that the presence of logarithmic correction as in Eqn. (5.5) is a better possibility.

In a similar way, the average radius $\langle r(\delta) \rangle$ of all disks has been calculated in the following way. For a particular percolation configuration we first calculate the arithmetic mean of the radii of all disks and then average this quantity over a large number of configurations. A power law dependence of $\langle r(\delta) \rangle$ on δ has been observed.

$$\langle r(\delta) \rangle \sim \delta^{0.666}. \tag{5.6}$$

It may be noted that, though $\langle r(\delta) \rangle < \langle r_m(\delta) \rangle$, the exponent of the former is larger, since as $\delta \to 0$ the value of $\langle r(\delta) \rangle$ decreases much faster than $\langle r_m(\delta) \rangle$. The errors in the exponents are of the order of 0.005.

Let us denote that the maximal jump in the Order Parameter by $\Delta_m \mathcal{C}(\delta)$. This is the average of the maximal jumps in the size of the largest clusters over a large number of independent runs, i.e., $\Delta_m \mathcal{C}(\delta) = \langle \Delta_m s_m(p, \delta) \rangle$. In Fig. 5.5 we plot $\Delta_m \mathcal{C}(\delta) - \Delta_m \mathcal{C}(0)$ with δ on a log − log scale. Since $\Delta_m \mathcal{C}(0)$ cannot be estimated directly we tried with different trial values of $\Delta_m \mathcal{C}(0)$ to make the plot which fits best to a straight line. Though there is an initial curvature for large values of δ, the latter points obtained as $\delta \to 0$ fit nicely to a straight line. This implies that the variation can be termed as a power law like:

$$\Delta_m \mathcal{C}(\delta) = \Delta_m \mathcal{C}(0) + A\delta^{\mu} \tag{5.7}$$

with $\Delta_m \mathcal{C}(0) = 0.16$, $A = 0.15$ and $\mu = 0.274(5)$. This relation can be interpreted that even in the limit of $\delta \to 0$ the average maximal jump in the Order Parameter i.e., the area coverage of the largest cluster, is a finite fraction of the entire area of the disk pattern. Therefore this is also another signature of the discontinuous percolation transition in our model.

The rapidity with which the Order Parameter increases in Fig. 5.2 at the percolation threshold can also be quantified by measuring the width of the window around

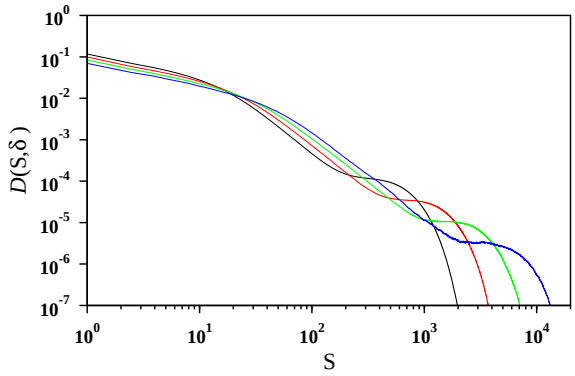


Figure 5.7: (Color online) Display of the binned data for the cluster size distribution $D(S)$ measured by the number of disks S in a cluster. From the slopes in the intermediate regime we obtained $\tau = 2.16$, 2.13, 2.14 and 2.13 respectively for growth rates $\delta = 0.002$ (black), 0.001 (red), 0.0005 (green) and 0.00025 (blue), δ decreases from left to right. The intermediate power law regime gets elongated as δ decreases and we conclude a value of the associated exponent $\tau = 2.14(2)$.

the percolation threshold following the method used in [90]. For a single run, we define $p_{1/10}$ as the minimum value of the area coverage p for which $\mathcal{C} > 1/10$. Similarly $p_{1/2}$ is the minimum value of p for which $\mathcal{C} > 1/2$. The difference in these two area coverages $p_{1/2} - p_{1/10}$ is the size of the window through which a 40 percent jump in the Order Parameter takes place. Averaging over a large number of uncorrelated runs we estimated $\langle p_{1/10} \rangle$, $\langle p_{1/2} \rangle$ and $\langle p_{1/2} - p_{1/10} \rangle$. In Fig. 5.6(a) we plotted $1 - \langle p_{1/10} \rangle$ and $1 - \langle p_{1/2} \rangle$ using a log $-$ log scale. It is observed that $1 - \langle p_{1/2} \rangle$ fits very well to a nice straight line over the entire range of δ implying a power law variation $0.556\delta^{0.131}$; where as $1 - \langle p_{1/10} \rangle$ fits quite well to the form $0.556\delta^{0.131} + 3.00\delta^{0.523}$. In Fig. 5.6(b) we plotted $\langle p_{1/2} - p_{1/10} \rangle$ against δ. Here also, apart from some initial curvature for large δ the curve fits to a power law

$$\langle p_{1/2} - p_{1/10} \rangle = 2.45\delta^{0.50}. \tag{5.8}$$

Therefore as $\delta \to 0$ a 40 percent increase in the Order Parameter requires a vanishingly small change in the area coverage. This is again another evidence that the percolation transition in this model is likely to be a discontinuous transition. The error in the exponent is of the order of 0.02.

The percolation transition in our model is analyzed by yet another method, this time studying the cluster size distribution at the percolation point. For this study we defined the cluster size S in a different way, this time it is the number of disks belonging to a specific cluster. In Fig. 5.7 we have presented the binned data for the probability distributions of the cluster sizes for four different values of δ. The cluster size distribution data have been collected only when the global connectivity appears for the first time. All four $D(S, \delta)$ vs. S plots on the log - log scale have similar nature. After some initial slow variation, the $\log D(S, \delta)$ decreases linearly with $\log S$ in the intermediate power law regime. Finally at the tail of the distribution there is a hump, meaning an enhanced probability for the large clusters which connects the two ends of the system. It is assumed and which seems to be very likely that as

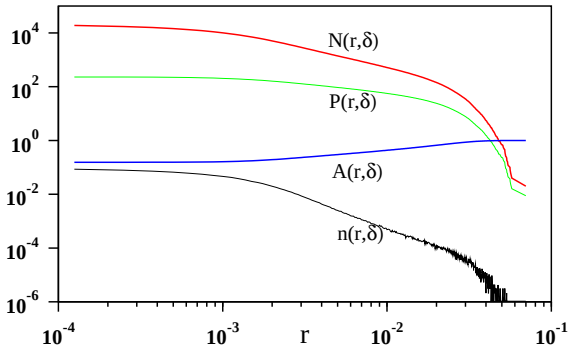


Figure 5.8: (Color online) The distribution $n(r, \delta)$ of the number of disks of radii r, the total number of disks $N(r, \delta)$, the total perimeter $P(r, \delta)$ and the remaining uncovered area $A(r, \delta)$ for disk patterns grown at a rate $\delta = 0.000125$. The slopes of each curves gives an estimate of the fractal dimension d_f and an average value of $d_f = 1.42(10)$ has been obtained.

δ decreases the position of the hump shifts systematically to large values of S and therefore in the limit of $\delta \to 0$ the entire intermediate regime would fit to a power law of the form: $D(S) \sim S^{-\tau}$. We conclude an average value of $\tau = 2.14(10)$ which is to be compared with $\tau = 187/91$ for ordinary percolation [27]. In the context of percolation theory, it is well known that a power law distribution of the cluster sizes is a signature of a continuous percolation transition. Therefore, the power law distribution of the cluster sizes obtained for our model also can be interpreted as an indication of a possible continuous transition, but this interpretation is in contradiction to the discontinuous transition indicated by the Eqns. (5.7) and (5.8).

Here we recall that the original model of Explosive Percolation, which goes by the name of 'Achlioptas Process' [90] has a similar story. For this model, most of the numerical results indicated that the associated percolation transition is discontinuous. However, recently Riordan and Warnke have rigorously proved that a class of Explosive Percolation models defined on complete graphs which use Achlioptas type processes are in fact continuous in the asymptotic limit of large size graphs [98]. While saying this, we also like to note, for the Explosive Percolation models defined in Euclidean space no such result is available and it is, therefore, still possible that Explosive Percolation in the plane has a discontinuous transition. To sum up, we find it proper to conclude, on the basis of our numerical study that the percolation transition in our model behaves similar to the Achlioptas Process, i.e., though apparently it exhibits the behavior alike to a discontinuous transition, it is indeed a continuous transition.

Finally we measured the fractal dimension of the dust of pore spaces right at the percolation threshold using the Eqn. (5.1). We considered a large sample of uncorrelated disk patterns that have been grown at a rate δ. A periodic boundary condition has been imposed along the x direction and the pattern is grown till a global connection appears along the y axis when further growth is terminated. For each pattern we estimated the following quantities: the distribution $n(r, \delta)$ of the number of disks of radii r, its cumulative distribution $N(r, \delta)$, total perimeter $P(r, \delta)$ and the remaining uncovered area $A(r, \delta)$. We plot all four quantities in Fig. 5.8

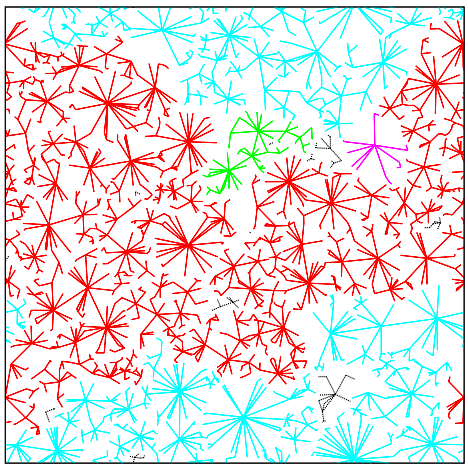


Figure 5.9: (Color online) The picture of the contact network corresponding to the disk pattern in Fig. 5.1(a). A link between a pair of disks that overlapped is drawn by a straight line joining their centers. The pattern has many clusters. The sub-graphs corresponding to the top four largest clusters are displayed by the same colors as used in Fig. 5.1(a).

using a $\log - \log$ scale for $\delta = 0.000125$. For each curve the scaling appeared in the intermediate regime of disk radii. Estimation of slopes of these curves in their scaling regions and using Eqn. (5.1) we have obtained the values of the fractal dimensions as 1.42, 1.41, 1.40 and 1.46 respectively. Clearly a large scatter of the estimated value of d_f is present, yet this data indicates that the d_f is like to be around 1.42 a with rather large error of around 0.10. A more accurate estimation needs patterns to be generated using even smaller value of the growth rate δ.

5.4 Contact Network

A contact network for the assembly of overlapping disks may be defined identifying the centers of the disks as nodes. In addition a link between a pair of nodes is introduced if and only if their corresponding disks overlap [105]. As time passes the contact network grows in the number of nodes as well as links. In Fig. 5.1(a) we have exhibited the disk pattern at the percolation threshold where different disks are of different radii. In general the large disks have overlaps with many other disks and therefore in the contact network these nodes form the hubs of the network. In the same way smaller disks have fewer links but their numbers are more. The contact network corresponding to the Fig. 5.1(a) has been exhibited in Fig. 5.9. Since there are many clusters, the network is not a singly connected graph. The four top large clusters are represented by four sub-graphs of the network.

The degree k_j of a node is the number of links meeting at node j. Here that would be equal to the number of other k_j disks that have overlap with the j-th disk. Typically such networks are scale-free networks which have power law degree distributions. Similarly we expect that for the contact network of the disk pattern, the largest component of the graph, right at the percolation threshold has the degree

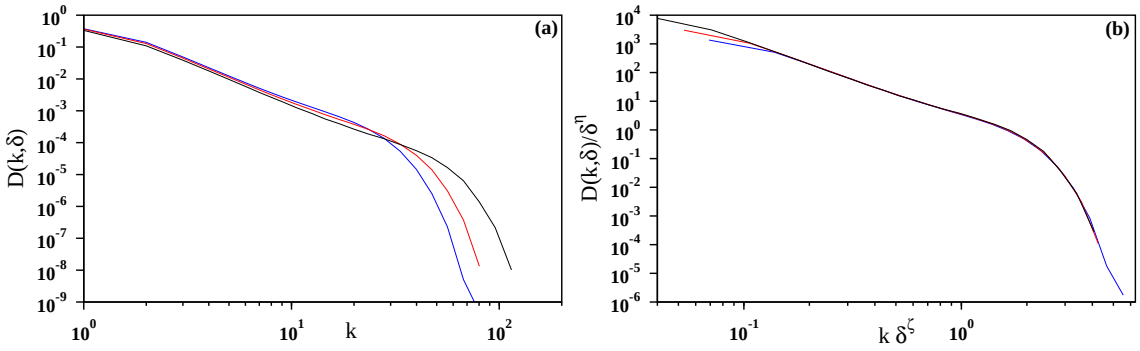


Figure 5.10: (Color online) (a) The nodal degree distribution $D(k,\delta)$ of the contact network of the pattern at the percolation threshold for three different values of the growth rate $\delta = 0.0001$ (blue), 0.00004 (red) and 0.00001 (black); the value of δ decreases from the left to the right. Binned data have been used to reduce fluctuation. (b) The finite size scaling analysis of the data in (a) where $D(k,\delta)/\delta^{0.89}$ scales as $k\delta^{0.29}$ which gives $\gamma = 3.07$.

distribution $D(k) \sim k^{-\gamma}$, in the limit of $\delta \to 0$ where γ is the degree distribution exponent to be estimated. For finite δ we have calculated the probability distribution $D(k,\delta)$ which is the probability that a randomly selected node has degree k. In Fig. 5.10(a) we plot $D(k,\delta)$ against k on a $\log-\log$ scale for the three different values of δ. Apart from the very small and large values of k the curves are quite straight in the intermediate regimes indicating that the degree distributions indeed have power law variations in the intermediate range of degree values. A direct measurement of slopes gives the values of $\gamma(\delta)$ actually depend on δ and extrapolate like $\gamma(\delta) = \gamma(0) - 5.8\delta^{0.39}$ with the extrapolated value of $\gamma(0) = 2.80$. This analysis is supported by a finite size scaling analysis of the same data. In Fig. 5.10(b) we plot the same data used in Fig. 5.10(a) but scale both the axes with suitable powers of δ. The best data collapse corresponds to

$$D(k,\delta)/\delta^{\eta} \sim \mathcal{G}(k\delta^{\zeta}) \tag{5.9}$$

where the values of $\eta \approx 0.89$ and $\zeta \approx 0.29$ are obtained; $\mathcal{G}(x)$ being the universal scaling function. This implies that from the scaling analysis the estimate for the exponent γ in the limit of $\delta \to 0$ is $\gamma = \eta/\zeta \approx 3.07$. Averaging the two estimates, we quote a value of $\gamma = 2.94(14)$.

This estimated value of the degree distribution exponent close to 3, prompted us to compare this contact network with the Barabási-Albert network [5]. When a new disk is introduced, its center is selected randomly with uniform probability anywhere within the uncovered region. For an already existing disk i of radius r_i if the center of the new disk is selected within the annular ring of radius $(r_i + \delta)$ then the new disk becomes its neighbor immediately after introduction. Therefore it is more probable that a newly added disk becomes the neighbor of a larger disk than a smaller one and its probability is proportional to r_i. But this observation holds only while the disk i is growing, i.e., till its degree $k_i = 0$. The moment it gets its first neighbor i.e., the first link, it's growth stops and it becomes frozen. Thereafter the degree k_i of i increases but its radius does not, and gradually the annular space

fills up. Therefore unlike the the 'rich gets richer' principle in the Barabási-Albert network [5] here the attachment probability is not proportional to the degree k. We studied how the average radius of all disks whose degrees are equal to k depends on k. We have plotted (not shown here) for a small value of δ using the log - log scale $\langle r(k)\rangle / \log(k)$ vs. k on a log - log scale which fit very well to a straight line indicating that the following form with logarithmic correction may be valid

$$\langle r(k)\rangle \sim k^{0.72} \log k \tag{5.10}$$

for the growth of the average radius of a disk with the their degree k.

5.5 Conclusion

Signature of discontinuous jumps in the Order Parameter has been observed in a continuous percolation transition of an assembly of growing circular disks which overlap slightly before becoming frozen. Extrapolation of numerical results indicate that in the limit of the extremely slow growth rate of $\delta \to 0$ the percolation transition occurs when the area coverage is unity, i.e., the disk pattern is space-filling even at the percolation point. Surprisingly, within our numerical accuracy, it has been observed that the percolation transition of such a system has a discontinuous macroscopic jump in the Order Parameter and also associated with a vanishing width of the transition window. On the contrary, the cluster size distribution has been found to have a power law decaying functional form. We conclude that though the transition in our model is actually continuous it exhibits certain features of a discontinuous transition. We conclude that our model is yet another example like the Achlioptas Process [90], the original model of Explosive Percolation, where a similar sharp but continuous transition has been observed.

Chapter 6

Weighted Network Analysis of Earthquake Seismic Data

6.1 Introduction

Earthquake is probably one of the most important natural phenomenon affecting human life and property, hence its study and quest of a preventive mechanism has a long history. Different laws have been formulated from empirical observations e.g., the Gutenberg-Richter law [106] relates the frequency of tremors of a given magnitude and the Omori law [107] describes the temporal rate of decay of aftershocks corresponding to a main shock. More recently Physicists have tried to explain the earthquake dynamics as a scale invariant process. Correlation among different shocks as well as the recurrence time distributions [108-111] have been studied. The spatial positions of earthquake epicenters have been claimed to form fractal sets [112, 113]. A number of models have been proposed and studied. For example, the well known Burridge-Knopoff model describes the slow creeping of the continental plates along the fault lines as a stick-slip process [114]. Very importantly Bak et al. suggested that the phenomenon of earthquakes may be looked upon as a Self-Organized Critical process which spontaneously generates long range spatio-temporal correlations or scale-invariance [34, 115]. However the actual mechanism of the underlying dynamics or efficient forecasting of this complex phenomenon have not been possible yet.

Recently the time sequence of occurrence of different tremors and the positions of their epicenters in different earthquake catalogs have been studied using the tools of complex network theory. Baiesi and Paczuski considered different tremor events as the nodes of a network where a pair of nodes are linked if the correlation between them exceeds a certain threshold value [116, 117]. In another method, Abe and Suzuki considered a link between every pair of successive tremor events [118-125]. Both groups suggested that the earthquake network has scale-free structures with small-world properties. In addition, there are other studies which deal with correlations, recurrence times and modeling of earthquakes [126-129].

To construct the earthquake network, Abe and Suzuki digitized the entire earthquake region into a rectangular grid, using the lattice constant as a tunable parameter [118]. Here, a cell is considered as a node if at least one tremor has its epicenter within this cell. A pair of nodes is defined to be connected by a link if

Region	SC	JU	CAN
Period	1973 - 2011	1985 - 1998	1950 - 1992
θ_{min}	32	25.73	0
θ_{max}	37	47.964	87.75
ϕ_{min}	-122	126.43	-99.86
ϕ_{max}	-114	148.0	177.1
n	572601	200910	25970
$(L_{lat}L_{lon})^{1/2}$	638.33 Km	2178.01 Km	14715.41 Km

Table 6.1: All angles are measured in degrees. **SC**: Southern California Earthquake Data Center, http://www.data.scec.org/ **JU**: Japan University Network Earthquake Catalog, http://wwweic.eri.u-tokyo.ac.jp/db/junec/ **CAN**: Canada's National Earthquake Database, http://earthquake.usgs.gov/earthquakes/eqarchives/

and only if at least one pair of successive seismic events occur whose epicenters are located within these two cells. It has been observed that the associated network is hierarchical, it has the assortative property and it's clustering co-efficient follows a scaling relation with cell size as well as the size of the earthquake network. It is to be noted that, in this network analysis of earthquakes, all links are treated on identical footing irrespective of (i) the strengths of the tremors and (ii) the number of successive pair of events that correspond to a particular link which implies that the network introduced by Abe and Suzuki is an unweighted network.

On the other hand, a more detailed analysis reveals that roles played by different links have different importance - which demands the introduction of a new variable called 'weight' associated with a link. In many networks link weights represent the strength of ties between nodes which are not similar and in fact quite often they are highly heterogeneous. There are several well known examples. For example, the passenger traffic between a pair of airports in airport networks [13,130], strength of the pair-interaction between two species in ecological network [131], the volume of trade between two countries in the international trade network [15,16] etc. A number of new properties of these networks have come to light when they are analyzed considering the link weights.

The method of Baiesi and Paczuski is crucially based on the work of Bak et al. in [108]. However, careful examinations of [108] have revealed that their work ambiguously contains several artificial assumptions regarding e.g., precision of time separation for detecting events etc. Later, it has been found that the unified scaling law in [108] for waiting times fails to hold for data sets other than the Californian one [132]. This is the reason why one cannot see universal properties in networks of the Baiesi-Paczuski type. On the other hand, the method of Abe and Suzuki is certainly associated with a universal law [133].

In this chapter, we study the earthquake network as a weighted network to gain better insights about the structural properties and correlations present in the network [134]. In section 6.2, we define the weighted network and describe the quantities observed. In section 6.3, we present the results and their analysis. In section 6.4, we describe the rich-club analysis. The chapter is summarized in section 6.5.

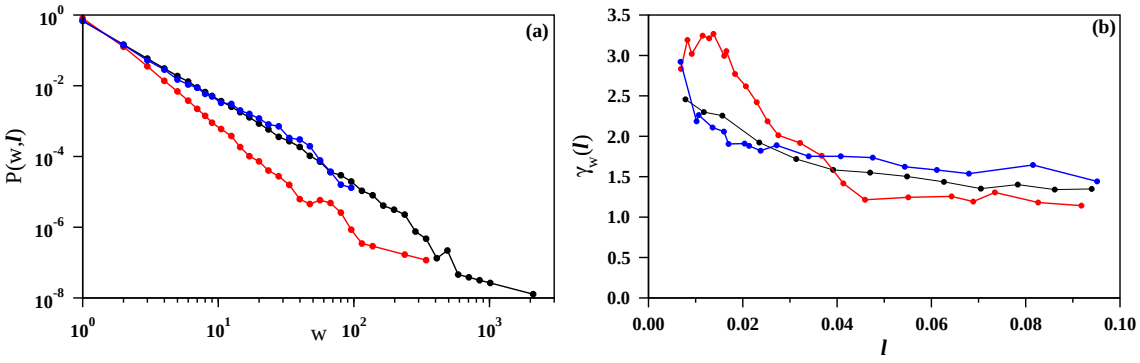


Figure 6.1: (Color online) (a) Plot of the link weight distributions $P(w, \ell)$ against weight w using $\ell = 0.01$. (b) Plot of the variation of the link weight distribution exponent $\gamma_w(\ell)$ against ℓ. Colors used: SC (black), JU (red) and CAN (blue).

6.2 The Weighted Earthquake Network

Given the time sequence of occurrence of seismic events in a particular earthquake catalog, the network is constructed following Abe and Suzuki in [118]. We define the weight w associated with a specific link as the total number of successive events between its two end nodes. It is observed that these weights are not similar at all, in fact they vary over a wide range.

Three different earthquake catalogs have been analyzed, namely, the Southern California Earthquake Data Center catalog (SC), Japan University Network Earthquake catalog (JU) and Canada's National Earthquake Database catalog (CAN). Different specifications of these catalogs are given in Table 1. Each catalog contains the data for every seismic event within a specified period: geographical positions of the epicenters described by their latitudes (θ) and longitudes (ϕ) and the precise times of occurrence. During the period from 1932-1972 the number of events in the SC data is very small compared to rest of the period. Therefore we have analyzed the SC data for the period from 1973-2011 containing a total of 572601 events. The Japan University catalog has the total of 200910 seismic events between 1985 and 1998; where as the Canadian catalog has the record of 25970 seismic events within the interval between 1950 and 1992.

These epicenters are extended over a region which we refer as the 'earthquake region'. The earthquake region corresponding to a particular earthquake catalog is characterized by the minimum and the maximum values of these coordinates, i.e., $(\theta_{min}, \theta_{max})$ and (ϕ_{min}, ϕ_{max}). Following the method in [118] the earthquake regions have been digitized into different grids to analyze the data. The main difference is we have used two dimensional cells on the surface of the earth and ignored the depth coordinate of the epicenter where as Abe and Suzuki mainly used three dimensional cells. To make the size L of the cells dimensionless we define the parameter $\ell = L/(L_{lat}L_{lon})^{1/2}$ [124]. Here, L_{lat} and L_{lon} are the total extensions along the north-south and the east-west directions respectively of the entire earthquake region. The total number of data in the catalog is denoted by n. The North-South distance between (θ_i, ϕ_i) and $(\theta_{min}, \phi_{min})$ is $d_{NS} = R(\theta_i - \theta_{min})$ and the East-West distance is $d_{EW} = R(\phi_i - \phi_{min})cos\theta_{av}$, where the radius of the earth is $R = 6370$ Km and θ_{av}

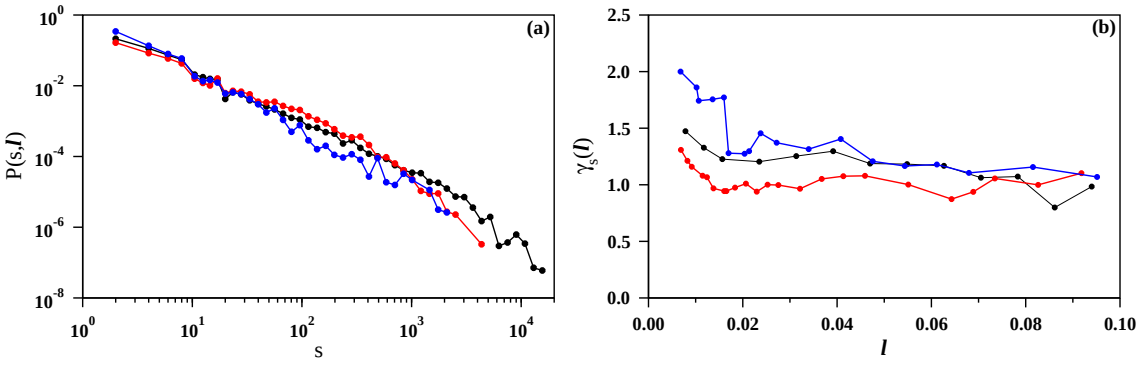


Figure 6.2: (Color online) (a) Plot of the nodal strength distributions $P(s,\ell)$ against strength s using $\ell = 0.01$. (b) Plot of the variation of the nodal strength distribution exponent $\gamma_s(\ell)$ against ℓ. Colors used: SC (black), JU (red) and CAN (blue).

is $(\theta_{min} + \theta_{max})/2$.

The weighted earthquake network is constructed in the following way. A cell is considered to be a node if at least one tremor of any magnitude has its epicenter located within the cell. Two successive tremors occurred at nodes i and j is referred by a 'bond' between the two nodes. Self-loops, i.e., successive events occurring at the same node, are not considered. Any number of successive events may occur between the same pair of nodes which implies that a number of bonds may exist between the nodes. If there is at least one bond between the nodes i and j, the nodes are said to be linked. Therefore, the number of distinct links of this earthquake network is much less than the total number of bonds. The weight w_{ij} of the $i-j$ link is the number of bonds between them. Strength of a node i is the total amount of weight supported by the node: $s_i = \Sigma_j w_{ij}$, the sum runs over the k_i neighbors of i. When different link weights are uncorrelated, the average strength of a node of degree k is $s(k) \simeq \langle w \rangle k$. On the other hand, in the presence of correlation $s(k) \propto k^{\beta}$ with $\beta > 1$.

The clustering coefficient is a measure of the three point correlation among the neighboring nodes. For many real-world networks, the clustering coefficient $\mathcal{C}(k)$, for k-degree nodes, decreases as $\mathcal{C}(k) \sim k^{-\beta_k}$ with $\beta_k \approx 1$. The weighted clustering coefficient [13] is defined as $\mathcal{C}^w(i) = (1/(s_i(k_i-1)))\sum_{j,h}((w_{ij}+w_{ih})/2)a_{ij}a_{ih}a_{jh}$ where a_{ij}s are the elements of the adjacency matrix. The nodal degree-degree correlation is measured by the average degree of the neighbors of a node i is defined as $k_{nn,i} = \Sigma_j k_j/k_i$ where j runs over the k_i neighbors of i [13]. For the nodes of degree k, $\langle k_{nn}(k) \rangle = \Sigma_{k_i=k} k_{nn,i}/N_k = \Sigma_{k_1} k_1 P(k_1|k)$ where N_k is the number of nodes having degree k. In absence of correlation, $\langle k_{nn}(k) \rangle$ is a constant. If $\langle k_{nn}(k) \rangle$ increases or decreases with k, the network is said to be assortative or disassortative respectively. For weighted networks one defines, $k^w_{nn,i} = \Sigma_j w_{ij} k_j/s_i$ and similarly the $\langle k^w_{nn}(k) \rangle$. The heterogeneity in link weights is measured by the 'disparity' $Y(i) = \sum_{k_i} [w_{ij}/s_i]^2$ [19, 20]. When weights are of the same order $Y(k) \sim 1/k$ (for $k \gg 1$) and if few weights dominate then $Y(k) \sim 1$ [21].

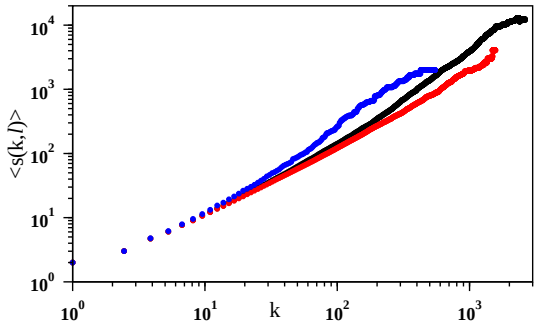


Figure 6.3: (Color online) The average nodal strength $\langle s(k,\ell)\rangle$ plotted against the degree k for SC (black), JU (red) and CAN (blue) data for for $\ell = 0.01$.

6.3 Results

In Fig. 6.1(a) we plotted the probability distribution $P(w,\ell)$ of link weights on a double-logarithmic scale graph for the three countries for $\ell = 0.01$. The data for South California and Japan appear to be described by power laws of the form $P(w,\ell) \propto w^{-\gamma_w}$. Though the catalog of Canada is quite short, yet its distribution may also be described by a power law. Slopes of the curve give the estimation of the exponent $\gamma_w = 2.38, 3.28$ and 1.87 for SC, JU and CAN respectively. These exponents $\gamma_w(\ell)$ have been measured by varying the cell size ℓ. In Fig. 6.1(b) we have shown the variation of $\gamma_w(\ell)$ with ℓ. Evidently $\gamma_w(\ell)$ approaches a fixed value as ℓ increases.

Similarly the probability distribution of nodal strengths $P(s,\ell)$ are plotted in Fig. 6.2(a) for $\ell = 0.01$. Again a power law decay function $P(s,\ell) \sim s^{-\gamma_s}$ have been observed to fit well for all three earthquake regions. Slopes of the curve estimate the value of the exponent $\gamma_s = 1.40, 1.24$ and 1.86 for SC, JU and CAN respectively. In a similar way as before the variation of $\gamma_s(\ell)$ against ℓ have been shown in Fig. 6.2(b) which also saturate as ℓ increases.

In Fig. 6.3 the average strength of all k degree nodes $\langle s(k,\ell)\rangle$ are plotted against k for $\ell = 0.01$. It has been observed that the average strength scales almost linearly for all the three earthquake zones up to $k \sim 100$ exhibiting almost no strength-degree correlation. But beyond this regime the variation of $\langle s(k,\ell)\rangle$ is no more linear and in this regime the strength-degree exponent $\beta > 1$. β is found 1.53, 1.26 and 1.23 for SC, JU and CAN respectively. It has also been observed that the value of β shows increasing trend with the increasing box sizes. This signifies that weights associated with links meeting at small degree nodes are uncorrelated where as those for large degree nodes are correlated. Therefore the large degree nodes not only have large number of links but the weight of their links are also large.

In Fig. 6.4 we plot unweighted and weighted average clustering coefficient of all k degree nodes and plotted them with k. In Fig. 6.4(a) the nature of variation of $\mathcal{C}(k,\ell)$ with k shows a decaying tail for larger degree values for all three earthquake regions implying a signature of hierarchical structure of the network as found by Abe and Suzuki [120]. However in Fig. 6.4(b) the plot of weighted clustering coefficient $\mathcal{C}^w(k,\ell)$ this trend is almost neutralized showing no visible signature of hierarchy.

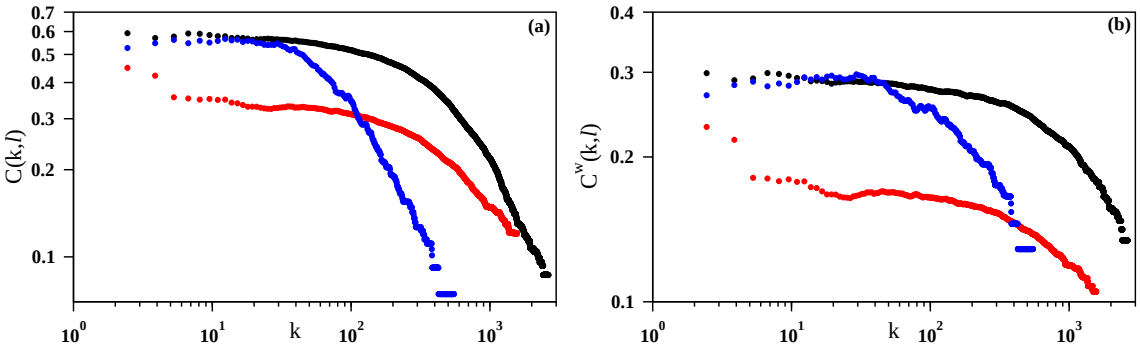


Figure 6.4: (Color online) (a) Plot of the clustering coefficient $C(k)$ for nodes with degree k. (b) The weighted clustering coefficients $C^w(k)$ against nodal degree k. Colors used: SC (black), JU (red) and CAN (blue) for $\ell = 0.01$.

This may be the reflection of the fact that though the neighbors of the large degree nodes are not very densely connected but links connecting them are strong links bearing large weights.

The average degree $\langle k_{nn}(k,\ell)\rangle$ of an arbitrary neighbor of a node of degree k has been plotted with k in Fig. 6.5(a) and its weighted counterpart $\langle k^w_{nn}(k,\ell)\rangle$ has been shown in Fig. 6.5(b). This quantity also represents the mixing pattern of the nodes. Decay of $\langle k_{nn}(k,\ell)\rangle$ with k at large degrees indicates that the mixing is disassortative, that is large degree nodes have no general preference to connect with large degree nodes. On the other hand the plot of $\langle k^w_{nn}(k,\ell)\rangle$ in Fig. 6.5(b) displays almost uniform variation instead of decay at large degrees signifying the absence of degree-degree correlation. The same feature is prominent for all the three earthquake regions. It may be noted that in [120] an assortative dependence of $\langle k_{nn}(k,\ell)\rangle$ for large values of degree k has been observed, which may be due to the fact that the elementary grid cells used in this work are three dimensional in comparison to two dimensional cells used by us.

In Fig. 6.6(a) the average disparity $\langle kY(k,\ell)\rangle$ of link weights connected to a node of degree k is plotted with k. It is evident from the plot that for small values of k, $\langle kY(k,\ell)\rangle$ is in the order of unity showing almost no disparity. But in the regime of $k > 100$ disparity does exist among the link weights and the $\langle kY(k,\ell)\rangle$ grows gradually with k. Since the earthquake network has been constructed using the information of the time sequence of tremor events, this network may be considered as a directed network as well. Here a bond is directed from i to j, if the corresponding tremor in the cell j occurred after that in the cell i. This implies that a typical node may have any number of inward as well as outward bonds. A weighted version of this directed network may be defined by introducing the directed weight $w_{i\to j}$ which is the number of bonds that are directed from the node i to node j. Therefore, the weighted in-degree, which measures the frequency of being a successor, may be defined as $k^w_i(in) = \Sigma_j w_{i\leftarrow j}$ and similarly for $k^w_i(out) = \Sigma_j w_{i\to j}$. A detailed look however reveals, since the successive tremors have constructed the network, every inward bond to a node must have an outward bond from that node, with the exception of the first and last nodes of the tremor sequence. This implies that for every node the weighted in-degree and out-degree are equal. Now, since the

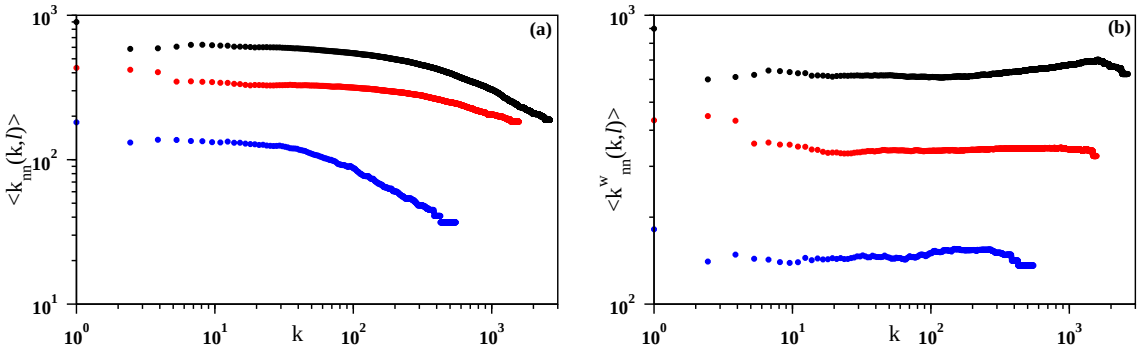


Figure 6.5: (Color online) (a) Plot of average neighbor degree $\langle k_{nn}(k,\ell)\rangle$ and (b) the weighted average neighbor degree $\langle k^w_{nn}(k,\ell)\rangle$ with k. Colors used: SC (black), JU (red) and CAN (blue) for $\ell = 0.01$.

strength s_i of a node i is the sum of weights of all k_i links meeting at i we have $k_i^w(in) = k_i^w(out) = s_i/2$.

Another important measure of a directed network is known as the 'reciprocity'. In general, a number of links in a directed network is bi-directional and the rest are uni-directional. The reciprocity r is the ratio of the number of bi-directional links and the total number of links. In the context of the earthquake network, the reciprocity r measures the probability if one node at the end of a directed link is an epicenter of a tremor then what is the probability that the other end node would also be the epicenter of another tremor. In the Fig. 6.6(b) we plot the variation of $r(\ell)$ with the cell size ℓ. It is observed that $r(\ell)$ gradually grows with the cell size. Increasing cell size is a kind of coarse-graining. Therefore, as the cell size increases, a cell accommodates a larger number of epicenters. This implies that if a pair of successive tremors occur in two different cells, another pair of successive tremors would take place in the opposite direction within the same cells with an enhanced probability.

6.4 The Rich-Club effect

Analysis of different real-world networks exhibits non-trivial correlations among the nodes. An implication of such correlation is, for an unweighted network the large degree nodes are connected among themselves forming a club. Such a club is defined by a subset of n_k nodes whose degree values are at least k. A rich-club coefficient (RCC) is measured as $\psi(k) = 2E_k/[n_k(n_k-1)]$ where E_k is the number of links that actually exists in the club and $[n_k(n_k-1)]/2$ is the number of maximum possible links in the club [135]. A high value of $\psi(k)$ implies that members are indeed tightly connected. However gradually it has been realized that only this definition is not sufficient, since with this measure even uncorrelated random graphs exhibit some rich-club effect. It has been suggested that one needs to define a 'null model' or the maximally random network (MRN) maintaining the nodal degree values $\{k_i\}$ preserved, measure the corresponding RCC $\psi_{ran}(k)$ and observe the variation of ratio $\rho_{ran}(k) = \psi(k)/\psi_{ran}(k)$. One method to generate MRN is the pairwise link

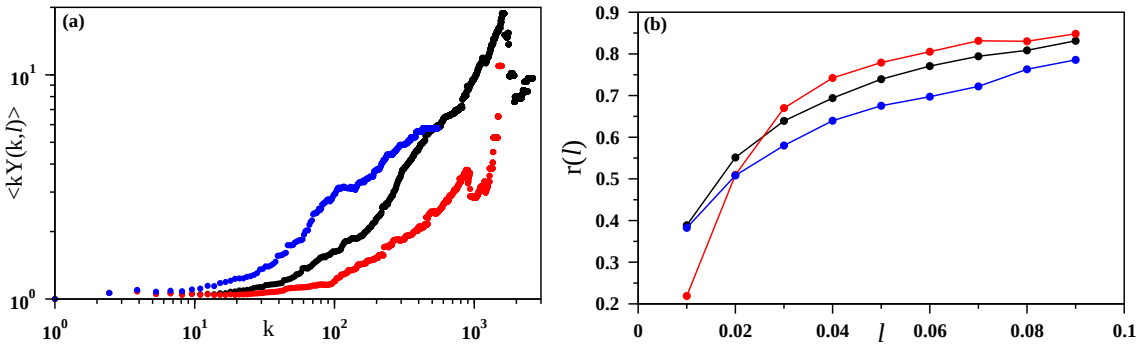


Figure 6.6: (Color online) (a) Average disparity $\langle kY(k,\ell)\rangle$ of links connected to a node of degree k has been plotted with degree k. (b) Reciprocity $r(\ell)$ has been plotted against the cell size ℓ. For both figures SC (black), JU (red) and CAN (blue) data for $\ell = 0.01$.

end exchange method [135].

The concept of rich-club can be applied to weighted networks as well by slightly modifying the definition. The rich-club in weighted network is defined as the subset of nodes whose strengths are at least s. Therefore in our case the nodes of the rich-club control a major portion of the world's earthquake dynamics. The RCC of the weighted network is modified as: $R_w(s) = 2\Sigma_{i,j}w_{ij}/[n_s(n_s-1)]$. Consequently the maximally random weighted network (MRWN) can be generated by preserving both the nodal degrees $\{k_i\}$ as well as the nodal strengths. To generate the MRWN a self-consistent iteration procedure is used to obtain the link weight distribution consistent with the nodal strength list $\{s_i\}$. Initially arbitrary random values of the weights w_{ij} are assigned to all links maintaining that the link weights are symmetric i.e., $w_{ij} = w_{ji}$. For an arbitrary node i the difference $\delta_i = s_i - \Sigma_j w_{ij}$ is calculated. Weights of all k_i links meeting at the node i are then updated as $w_{ij} \rightarrow w_{ij} + \delta_i\,(w_{ij}/\Sigma_j w_{ij})$, to balance s_i and $\Sigma_j w_{ij}$. By repeated iterations the link weights quickly converge and attain consistency with nodal strengths $\{s_i\}$. It is checked that $\langle s_i s_j\rangle \sim w_{ij}$ relation is well satisfied for this MRWN.

In Fig. 6.7(a) we show the variation of the ratios $\rho_{ran}(k,\ell)$ for earthquake network and MRN with k. Since $\rho_{ran}(k) \sim 1$ for all k, the unweighted network does not show the presence of rich-club phenomenon with respect to MRN. In Fig 7(b) the ratios $\rho_{ran}(s,\ell) = R_w(s)/{R_w}^{ran}(s)$ of the weighted rich-club coefficient of the earthquake network and MRWN plotted with s/s_{max} for $l = 0.01$. where s and s_{max} are the strength and maximum strength of the nodes. Since $\rho_{ran}(s) > 1$ for large s, we identify a strong rich-club ordering in the weighted earthquake network.

6.5 Summary

We have analyzed the seismic data sets of three different earthquake regions using the framework of the weighted networks. It has been observed that the link weights assume highly heterogeneous values and the probability distributions of link weights and nodal strengths follow power law decay forms. Un-weighted version of the

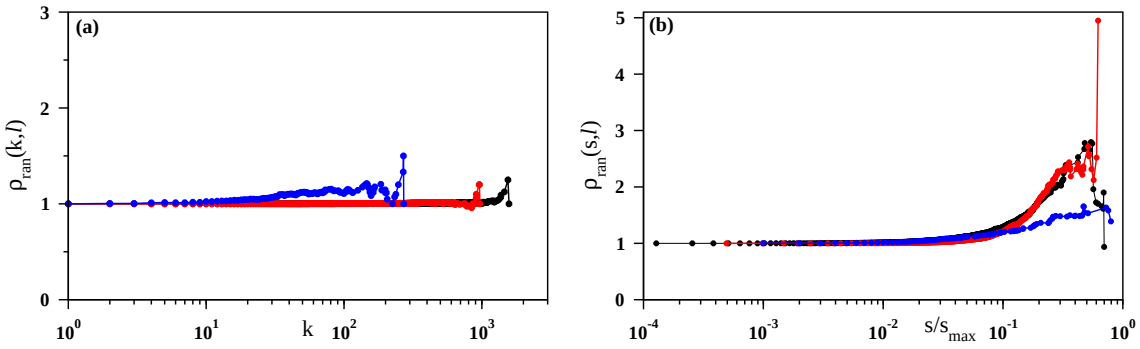


Figure 6.7: (Color online) For the earthquake network, plot of the (a) ratio $\rho_{ran}(k,\ell)$ of unweighted rich-club coefficient $\psi(k,\ell)$ and the same $\psi_{ran}(k,\ell)$ of MRN has been plotted against the nodal degree k; (b) ratio $\rho_{ran}(s,\ell)$ of weighted rich-club coefficient $R_w(s,\ell)$ and the same $R_w^{ran}(s,\ell)$ of MRWN has been plotted against the fractional nodal strength s/s_{max}. Colors used: SC (black), JU (red) and CAN (blue) for $\ell = 0.01$.

network shows the hierarchical nature, disassortative mixing and absence of rich-club phenomenon. On the other hand, the weighted version of the network exhibits absence of disassortative mixing and strong rich-club ordering.

Bibliography

[1] L. Euler, Academimae Petropolitance **8**, 128 (1736).

[2] J. L. Moreno, *Who shall survive?* (Nervous and Mental Disease Publishing Co., Washington, D.C., 1934).

[3] P. Erdős and A. Rényi, Publ. Math. Debrecen **6**, 290 (1959).

[4] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).

[5] A. L. Barabási and R. Albert, Science **286**, 509 (1999).

[6] M. Faloutsos, P. Faloutsos and C. Faloutsos, ACM SIGCOMM 99. Comput. Commun. Rev. **29**, 251 (1999).

[7] R. Guimera, S. Mossa, A. Turtschi, and L. A. N. Amaral, Proc. Natl. Acad. Sci. USA **102**, 7794 (2005).

[8] R. Albert, Journal of Cell Science **118**, 4947 (2005).

[9] L. A. N. Amaral, A. Scala, M. Barthélémy and H. E. Stanley, Proc. Natl. Acad. Sci. USA **97**, 11149 (2000).

[10] R. Albert, H. Jeong and A. L. Barabási, Nature **401**, 130 (1999).

[11] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee and S. S. Manna , Phys. Rev. E **67**, 036106 (2003).

[12] A. L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert and T. Vicsek, Physica A **311**, 590 (2002).

[13] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, Proc. Natl. Acad. Sci. USA **101**, 3747 (2004).

[14] A. E. Krause, K. A. Frank, D.M. Mason, R. E. Ulanowicz, and W.W. Taylor, Nature (London) **426**, 282 (2003).

[15] K. Bhattacharya, G. Mukherjee, J. Saramaki, K. Kaski and S. S. Manna, J. Stat. Mech. P02002 (2008).

[16] D. Garlaschelli and M. Loffredo, Phys. Rev. Lett. **93**, 188701 (2004).

[17] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, 2004).

[18] J. Saramäki, M. Kivelä, J. Onnela, K. Kaski and J. Kertész, Phys. Rev. E **75**, 027105 (2007).

[19] B. Derrida, H. Flyvbjerg, J. Phys. A **20**, 5273 (1987).

[20] M. Barthélemy, B. Gondran, E. Guichard, Physica A **319**, 633 (2003).

[21] E. Almaas, B. Kovács, T. Vicsek, Z. N. Oltvai and A.-L. Barabási, Nature **427**, 839 (2004).

[22] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **93**, 268701 (2004).

[23] S. R. Broadbent and J. M. Hammersley, Proc. Cambridge Philos. Soc. **53**, 629 (1957).

[24] H. Kesten, Commun. Math. Phys. **74**, 41 (1980).

[25] G. Grimmett, *Percolation* (Springer, Berlin, 1999).

[26] H. Mohammadi, E. N. Oskoee, M. Afsharchi, N. Yazdani and M. Sahimi, Int. J. Mod. Phys. C **20**, 1871 (2009).

[27] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1994).

[28] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, New York, 2010).

[29] Béla Bollobás, *Random Graphs*, 2nd Edition (Cambridge University Press, UK, 2001).

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

[31] C. D. Lorenz and R. M. Ziff, Phys. Rev. E **57**, 230 (1998).

[32] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford and New York, 1971).

[33] S. J. Gould and N. Eldredge, Paleobiology **3**, 115 (1977).

[34] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).

[35] P. Bak, S. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993).

[36] C. Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life* (John Murray, London, 1859).

[37] V. Pareto, *Cours d'economie Politique* (F. Rouge, Lausanne, 1897).

[38] W. W. Badger, *An entropy-utility model for the size distribution of income* (B. J. West(Ed) Mathematical models as a tool for the social science. Gordon and Breach, New York, pp 87-120, 1980).

[39] A. A. Drăgulescu and V. M. Yakovenko, Physica A **299**, 213 (2001).

[40] A. A. Drăgulescu and V. M. Yakovenko, *Modeling of Complex Systems: Seventh Granada Lectures*, edited by P. L. Garrido and J. Marro, AIP Conf. Proc. No. **661** (AIP, New York), 180 (2003).

[41] J. R. Iglesias, Science and Culture **76**, 437 (2010).

[42] M. N. Saha and B. N. Srivastava, *A Treatise on Heat* (Indian Press, Allahabad, 1931), p. 105.

[43] S. Sinha and B. K. Chakrabarti, Physics News **39**, 33 (2009).

[44] E. Majorana, Scientia **36**, 58 (1942).

[45] R. N. Mantegna and H. E. Stanley, *An Introducation to Econophysics: Correlation and Complexity in Finance* (Cambridge University Press, cambridge, 2000).

[46] L. P. Kadanoff, Simulation **16**, 261 (1971).

[47] E. W. Montroll and W. W. Badger, *Introduction to Quantitative Aspects of Social Phenomena* (Gordon and Breach, New York, 1974).

[48] H. E. Stanley, V. Afanasyev, L. A. N. Amaral, S. V. Buldyrev, A. L. Goldberger, S. Havlin, H. Leschhorn, P. Maass, R. N. Mantegna, C. -K. Peng, P. A. Prince, M. A. Salinger, M. H. R. Stanley, G. M. Viswanathan, Physica A **224**, 302 (1996).

[49] V. M. Yakovenko and J. B. Rosser, Rev. Mod. Phys. **81**, 1703 (2009).

[50] A. Chatterjee and B. K. Chakrabarti, Eur. Phys. J. B **60**, 135 (2007).

[51] S. Solomon, G. Weisbuch, L. de Arcangelis, N. Jan and D. Stauffer, Physica A **277**, 239 (2000).

[52] S. Pianegonda, J. R. Iglesias, G. Abramson, J. L. Vega, Physica A **322**, 667 (2003).

[53] A. A. Drăgulescu, V. M. Yakovenko, Eur. Phys. J. B **17**, 723 (2000).

[54] A. Chakraborti, B. K. Chakrabarti, Eur. Phys. Jour. B **17**, 167 (2000).

[55] M. Patriarca, A. Chakraborti, K. Kaski, Phys. Rev. E **70**, 016104 (2004).

[56] M. Á. Serrano and M. Boguñá, Phys. Rev. E **68**, 015101 (2003).

[57] A. Chatterjee, B. K. Chakrabarti, S. S. Manna, Physica A **335**, 155 (2004).
A. Chatterjee, B. K. Chakrabarti, S. S. Manna, Physica Scripta T **106**, 36 (2003).

[58] K. Bhattacharya, G. Mukherjee and S. S. Manna, *Econophysics of Wealth Distributions*, (Springer Verlag, Milan, 2005).

[59] P. K. Mohanty, Phys. Rev. E **74**, 011117 (2006).

[60] U. Basu and P. K. Mohanty, Eur. Phys. J. B **65**, 585 (2008).

[61] A. Chakraborty and S. S. Manna, Phys. Rev. E **81**, 016111 (2010).

[62] M. Patriarca, A. Chakraborti, K. Kaski, G. Germano, *Econophysics of Wealth Distributions*, Eds. A. Chatterjee, S. Yarlagadda, B. K. Chakrabarti (Springer, Milan, 2005) pp 93-110.

[63] J. Hoshen and R. Kopelman, Phys. Rev. B **14**, 3428 (1976).

[64] W. Souma, Fractals **9**, 463 (2001).

[65] S. Sinha, Physica A **359**, 555 (2006).

[66] J. R. Iglesias, S. Concalves, S. Pianegonda, J. L. Vega and G. Abramson, Physica A **327**, 12 (2003).

[67] P. Bak, *How Nature Works: The Science of Self-Organized Criticality*, (Copernicus, New York, 1996).

[68] S.S. Manna, J. Phys. A **24**, L363 (1991).

[69] A. Ghosh, U. Basu, A. Chakraborti and B. K. Chakrabarti, Phys. Rev. E **83**, 061130 (2011).

[70] A. Chakraborty, G. Mukherjee and S. S. Manna, Fractals **20**, 163, (2012).

[71] S. Lübeck, Int. Jour. Mod. Phys. B **18**, 3977 (2004).

[72] P. Grassberger, Phys. Lett. A **200**, 277 (1995).

[73] S. S. Manna, Phys. Rev. E **80**, 021132 (2009).

[74] B. B. Mandelbrot, *The Fractal Geometry of Nature*, (W. H. Freeman & Co., San Francisco, 1982).

[75] M. Paczuski, S. Maslov and P. Bak, Phys. Rev. E **53**, 414 (1996).

[76] S. Lübeck and P. C. Heger, Phys. Rev. Lett. **90**, 230601 (2003).

[77] H. N. Huynh, G. Pruessner and L. Y. Chew, J. Stat. Mech. 09024 (2011).

[78] R. M. Anderson and R. M. May, *Infectious diseases in humans* (Oxford University Press, Oxford, 1992).

[79] J. Marro and R. Dickman, *Non-equilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 1999).

[80] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases* (Griffin, London, 1975).

[81] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[82] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).

[83] M. Boguñá, R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).

[84] A. Chakraborty and S. S. Manna, Fractals **21**, 1350015 (2013).

[85] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[86] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).

[87] Y. Moreno and A. Vázquez, Europhys Lett. **31**, 265 (2003).

[88] G. Mukerjee and S. S. Manna, Phys. Rev. E **74**, 036111 (2006).

[89] J. Chalupa, P. L. Leath and G. R. Reich, J. Phys. C: Solid State Phys. **12**, L31 (1979).

[90] D. Achlioptas, R. M. D'Souza, and J. Spencer, Science **323**, 1453 (2009).

[91] H. K. Lee, B. J. Kim and H. Park, Phys. Rev. E **84**, 020101(R) (2011).

[92] S. S. Manna, Physica A **391**, 2833 (2012).

[93] R. M. Ziff, Phys. Rev. Lett. **103**, 045701 (2009).

[94] R. M. Ziff, Phys. Rev. E **82**, 051105 (2010).

[95] Y.S. Cho, J. S. Kim, J. Park, B. Kahng and D. Kim, Phys. Rev. Lett. **103**, 135702 (2009).

[96] F. Radicchi and S. Fortunato, Phys. Rev. Lett. **103**, 168701 (2009).

[97] R. K. Pan, M. Kivelä, J. Saramaki̇, K. Kaski, J. Kertész, Phys. Rev. E **83**, 046112 (2011).

[98] O. Riordan and L. Warnke, Ann. Appl. Prob. **22**, 1450 (2012).

[99] A. Chakraborty and S. S. Manna, Phy. Rev. E **89**, 032103 (2014).

[100] P. B. Thomas and D. Dhar, J. Phys. A **27**, 2257 (1994).

[101] S. S. Manna and H. J. Herrmann, J. Phys. A **24**, L481 (1991).

[102] S. S. Manna, Physica A **187**, 373 (1992).

[103] R. M. Baram and H. J. Herrmann, Phys. Rev. Lett. **95**, 224303 (2005).

[104] H. J. Herrmann, G. Mantica and D. Bessis, Phys. Rev. Lett. **65**, 3223 (1990).

[105] J. S. Andrade, Jr., H. Herrmann, R. F. S. Andrade and L. R. da Silva, Phys. Rev. Lett. **94**, 018702 (2005).

[106] B. B. Gutenberg and C. F. Richter, Bull. Seismol. Soc. Am. **34**, 185 (1944).

[107] F. Omori, J. Coll. Sci. Imp. Univ. Tokyo **7**, 111 (1895).

[108] P. Bak, K. Christensen, L. Danon, T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002).

[109] K. Christensen, L. Danon, T. Scanlon and, P. Bak, Proc. Natl. Acad. Sci. USA **99**, 2509 (2002).

[110] Á. Corral, Phys. Rev. Lett. **92**, 108501 (2004).

[111] Á. Corral, Phys. Rev. E **68**, 035102(R) (2003).

[112] Y. Y. Kagan, Physica D **77**, 162 (1994).

[113] P. G. Okubo and K. J. Aki, J. Geophys. Res. **92**, 345 (1987).

[114] R. Burridge and L. Knopoff, Bull. Seismol. Soc. Am. **57**, 341 (1967).

[115] P. Bak and C. Tang, J. Geophys. Res. **94**, 15635 (1989).

[116] J. Davidsen and M. Paczuski, Phys. Rev. Lett. **94**, 048501 (2005).

[117] M. Baiesi and M. Paczuski, Phys. Rev. E **69**, 066106 (2004).

[118] S. Abe and N. Suzuki, Euro. Phys. Lett. **65**, 581 (2004).

[119] S. Abe and N. Suzuki, Physica A **337**, 357 (2004).

[120] S. Abe and N. Suzuki, Phys. Rev. E **74**, 026113 (2006).

[121] S. Abe and N. Suzuki, Euro. Phys. J. B **59**, 93 (2007).

[122] S. Abe and N. Suzuki, Physica A **388**, 1917 (2009).

[123] S. Abe and N. Suzuki, Physica A **388**, 2511 (2009).

[124] S. Abe and N. Suzuki, Euro. Phys. Lett. **87**, 48008 (2009).

[125] S. Abe, D. Pasten and N. Suzuki, Physica A **390**, 1343 (2011).

[126] E. Lippiello, C. Godano and L. de Arcangelis, Phys. Rev. Lett. **98**, 098501 (2007).

[127] E. Lippiello, L. de Arcangelis and C. Godano, Phys. Rev. Lett. **100**, 038501 (2008).

[128] M. Bottiglieri, L. de Arcangelis, C. Godano and E. Lippiello, Phys. Rev. Lett. **104**, 158501 (2010).

[129] E. Lippiello, L. de Arcangelis and C. Godano, Euro. Phys. Lett. **72**, 678 (2005).

[130] R. Guimera and L. A. N. Amaral, Eur. Phys. Jour. B **38**, 381 (2004).

[131] J. W. Lee and S. E. Maeng, J. Phys.: Conf. Ser. **410**, 012067 (2013).

[132] V. Carbone, L. Sorriso-Valvo, P. Harabaglia and I. Guerra, Euro. Phys. Lett. **71**, 1036 (2005).

[133] S. Abe and N . Suzuki, Euro. Phys. Lett. **97**, 49002 (2012).

[134] A. Chakraborty, G. Mukherjee and S. S. Manna, Submitted to Physica A.

[135] V. Colizza, A. Flammini, M.A. Serrano, and A. Vespignani, Nature Physics **2**, 110 (2006).