BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

TRANSPORT AND PERCOLATION IN COMPLEX NETWORKS

by

GUANLIANG LI

B.S., University of Science and Technology of China, 2003

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2013

Approved by

First Reader

H. Eugene Stanley, Ph.D.
Professor of Physics

Second Reader

William Skocpol, Ph.D.
Professor of Physics

To my family.

Acknowledgments

In the path for pursuing my Ph.D, there are many people supporting me, guiding me and helping me.

My advisor Prof. H. Eugene Stanley, I am thankful to him for opening the door to my Ph.D, all the years offering the financial support to me along the path to achieving it and teaching me to prepare excellent talks. I feel lucky in his research group that I have the opportunities to exchange and cooperate with many other research groups from all over the world.

Prof. Shlomo Havlin is the main collaborator in my research career for these years. He is always guiding me in academic field and he has been an exemplary mentor to me. I express my deep thankfulness to him.

I also would like to thank other collaborators: Prof. Sergey V. Buldyrev, Prof. Lidia A. Braunstein and Prof. José S. Andrade, Jr. They are very helpful friends sharing knowledge and experience with me generously.

I am also very grateful to the members of my dissertation committee, Prof. William Skocpol, Prof. Karl Ludwig, Prof. William Klein and Prof Martin Schmaltz. Lastly, I thank the staff at Gene's Polymer Studies Group and BU Physics Department, especially Bob Tomposki and Mirtha Cabello, for helping me all these years.

TRANSPORT AND PERCOLATION IN COMPLEX NETWORKS

(Order No.)

GUANLIANG LI

Boston University Graduate School of Arts and Sciences, 2013

Major Professor: H. Eugene Stanley, Professor of Physics

ABSTRACT

To design complex networks with optimal transport properties such as flow efficiency, we consider three approaches to understanding transport and percolation in complex networks. We analyze the effects of randomizing the strengths of connections, randomly adding long-range connections to regular lattices, and percolation of spatially constrained networks.

Various real-world networks often have links that are differentiated in terms of their strength, intensity, or capacity. We study the distribution $P(\sigma)$ of the equivalent conductance for Erdős-Rényi (ER) and scale-free (SF) weighted resistor networks with N nodes, for which links are assigned with conductance $\sigma_i \equiv e^{-ax_i}$, where x_i is a random variable with $0 < x_i < 1$. We find, both analytically and numerically, that $P(\sigma)$ for ER networks exhibits two regimes: (i) For $\sigma < e^{-ap_c}$, $P(\sigma)$ is independent of N and scales as a power law $P(\sigma) \sim \sigma^{\langle k \rangle / a - 1}$. Here $p_c = 1/\langle k \rangle$ is the critical percolation threshold of the network and $\langle k \rangle$ is the average degree of the network. (ii) For $\sigma > e^{-ap_c}$, $P(\sigma)$ has strong N dependence and scales as $P(\sigma) \sim f(\sigma, ap_c/N^{1/3})$.

Transport properties are greatly affected by the topology of networks. We investigate the transport problem in lattices with long-range connections and subject to a cost constraint, seeking design principles for optimal transport networks. Our network is built from a regular d-dimensional lattice to be improved by adding long-range connections with probability $P_{ij} \sim r_{ij}^{-\alpha}$, where r_{ij} is the lattice distance between site i and j. We introduce a cost constraint on the total length of the additional links and find optimal transport in the system for $\alpha = d + 1$, established here for $d = 1, 2$ and 3 for regular lattices and d_f for fractals. Remarkably, this cost constraint approach remains optimal, regardless of the

strategy used for transport, whether based on local or global knowledge of the network structure.

To further understand the role that long-range connections play in optimizing the transport of complex systems, we study the percolation of spatially constrained networks. We now consider originally empty lattices embedded in d dimensions by adding long-range connections with the same power law probability $p(r) \sim r^{-\alpha}$. We find that, for $\alpha \leq d$, the percolation transition belongs to the universality class of percolation in ER networks, while for $\alpha > 2d$ it belongs to the universality class of percolation in regular lattices (for one-dimensional linear chain, there is no percolation transition). However for $d < \alpha < 2d$, the percolation properties show new intermediate behavior different from ER networks, with critical exponents that depend on α.

Contents

List of Tables

List of Figures

List of Abbreviations

ER	Erdős-Rényi
SF	scale-free
SW	small-world
BA	Barabási-Albert
SD	strong disorder
pdf	probability density function
IIPC	incipient infinite percolation cluster

Chapter 1

Introduction

Complex networks with substantial non-trivial topological features have been studied recently in many science fields such as biological, social and communication.[1–20] Transport and percolation are important properties of a complex network. By studying transport property, we understand the origin of flow efficiency in a network and obtain methods to improve it. The percolation property of a network provides us the underlying topological structure of a network. Percolation is extremely useful in oil fields finding, forest fire controlling, diffusion in disordered media, etc.[21]

Networks are simple because they contain only nodes and links which connect nodes. The number of links a node has is called the degree of this node. The degree distribution is the probability distribution of these degrees over the whole network. In the meantime networks are complex because of the countless degree distributions leading to varied forms. Our work is based on *random network* models, which are obtained by generating nodes and links at random. Different random network models produce different degree distributions.

The most classic random network model is called Erdős-Rényi (**ER**) model.[22,23] In ER model, each node has the same probability p to be connected to the other nodes. This results in that the average degree of each node is $\langle k \rangle = pN$, where N is the system size or number of nodes in a network. The degree distribution of an ER network is a Poisson distribution

$$P(k) = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}, \tag{1.1}$$

when $N \to \infty$.

In the last decade two well-studied classes of complex networks are scale-free (**SF**) networks[6,10,11,24,25] and small-world (**SW**) networks.[1–3,26,27] Both networks are characterized by specific structural features. The degree distribution of SF networks follows a power law distribution

$$P(k) \sim k^{-\lambda}, \tag{1.2}$$

where λ is a parameter typically in the range $2 < \lambda < 4$. The power law implies that the degree distribution of SF networks has no characteristic scale. The most known generative model of SF networks is Barabási-Albert (**BA**) model[25] with $\lambda = 3$. Since 1990s, the Internet and World Wide Web have been developed in a high speed and the interest of studying them as complex networks boomed after they were discovered that their degree distributions follow a power law over a few orders of magnitude.[28–31]

The study of small-world phenomena can be traced back to several experiments conducted by Stanley Milgram in 1960s.[32,33] The experiments suggested that human society is a small-world network and the average path length between any two selected people is six. This phenomena is called "six degrees of separation". Under the help of computer resource, small-world model was created and has been extensively studied in the recent ten years. Watts and Strogatz proposed a tractable model of small-world network with high transitivity in 1999[1–3] although the degree distribution of this model does not match most real-world networks very well.[9] The degree distribution of small-world models was presented analytically[34] and it belongs to the ubiquitous normal distribution.

In this work, we focus on the networks described above: **ER**, **SF** and **SW** networks. Inevitably our work is related with some other models, but they are derivate from these three basic models.

We understand the topological structures of complex networks through the models. However there exist many other important properties we need to know in the networks. A typical property of a network is transport which determines the flow efficiency in the network. Transport in many random networks is much different from regular networks.

Transport is anomalous in random networks.[35] To quantify transport of a network, average path length $\langle \ell \rangle$ defined as the average number of steps along the shortest path for all pairs of nodes in the network is the most important measure for the network. For the random models (ER, SF and SW networks), $\langle \ell \rangle$ has been found [6] and proven analytically[36] that it follows a logarithmic function

$$\langle \ell \rangle \sim \ln N. \tag{1.3}$$

The average path length is not the unique parameter to evaluate the transport of a network. For example, a power grid is mainly made up of metal wires which can be treated as resistors in complex network. In this power grid, the conductance is a more important parameter to evaluate the transport than path length because electric current can travel in parallel paths across the electrodes. For this kind of networks, we introduce weight to links and call these types of networks as weighted/disordered networks.[37,38] In a resistor network, each link is a resistor with different resistance or conductance. When the resistance represents the weight of each link, the larger the weight, the harder it is for electric current to transverse the link.

The first part of our work (in Chapter 2) focuses on conductance distribution of resistor networks. We use conductance not resistance in our work for convenience's sake because conductance represents the transport efficiency more intuitively. Let us consider a random resistor network with N nodes and average degree $\langle k \rangle > 1$. Our models are limited to ER and SF networks. Each link is assigned with conductance

$$\sigma_i \equiv e^{-ax_i}, \tag{1.4}$$

where x_i is a random variable with $0 < x_i < 1$ and the parameter a controls the strength of the disorder for each link. When $a = 0$, $\sigma_i = 1$, this returns to unweighted case which has been studied in Ref.[35]

We randomly select two nodes A and B from the resistor network, then place two electrodes on them (as source and sink) and load a voltage difference $V_{AB} \equiv V_A - V_B = 1$

between them. This results in electric current I_{AB} flowing from A to B in the system. The current I_{AB} is measurable and we get the conductance between A and B as

$$\sigma_{AB} = I_{AB}/V_{AB} = I_{AB}. \tag{1.5}$$

If we measure the conductance over all the pairs of nodes in the network, we can get conductance distribution $P(\sigma)$ of the network. $P(\sigma)$ provides the basic transport property of resistor network. However in our simulation of resistor networks, we cannot measure the current I_{AB} directly. Actually we solve Kirchhoff equations to get the current on each resistor link based on Kirchhoff's current law. And σ_{AB} can be calculated as sum of currents on all links connected to A or B. The method of solving Kirchhoff equations is a traditional and effective way to get the conductance of a resistor network with limited system size, i.e. N is finite. But it is very slow and cannot be applied to big system size because it solves a nearly $N \times N$ matrix while we have limited computer resources.

We provide an iterative fast algorithm to obtain $P(\sigma)$ and compare it with the traditional algorithm of solving Kirchhoff equations. The iterative algorithm can give us $N \to \infty$ approach and performs much faster than the method of solving Kirchhoff equations while it gives the corresponding results of $P(\sigma)$ with the traditional method of solving Kirchhoff equations.

Through the results of two methods, we find, both analytically and numerically, $P(\sigma)$ exhibits two regimes for ER networks:[39]

(i) A low conductance regime for $\sigma < e^{-ap_c}$, where $p_c = 1/\langle k \rangle$ is the critical threshold of the network. In this regime $P(\sigma)$ is independent of N and follows the power law $P(\sigma) \sim \sigma^{\langle k \rangle / a - 1}$.

(ii) A high conductance regime for $\sigma > e^{-ap_c}$ in which we find that $P(\sigma)$ has strong N dependence and scales as $P(\sigma) \sim f(\sigma, ap_c/N^{1/3})$.

For SF networks with degree distribution as Eq.(1.2), where $k_{min} \leq k \leq k_{max}$, we find numerically also two regimes, similar to those found for ER networks.

In Chapter 3, we investigate the design principles for optimal transport networks. This time, we focus on general complex networks, no longer resistors networks. Transport properties are greatly affected by the topology of networks. So basically, optimizing the transport of a network is to optimize the topological structure of the network.

Let us start from a simple 2-dimensional regular $N = L \times L$ square lattice. Every node is connected to its four nearest neighbors through short-range connections.[40] The transport property can be simply represented by the average path length $\langle \ell \rangle$ over all the pairs of nodes in the network. For this regular square lattice, $\langle \ell \rangle \sim L$ and transport is normal compared to random networks.[35] But the transport is inefficient in regular lattices.

To improve transport, a simple and effective way is to add some long-range connections (shortcuts) to the regular lattice to make it become a SW network.[1–3] With an appropriate number of additional long-range connections, the average path length can be dramatically reduced to $\langle \ell \rangle \sim \ln L$.[6,36] Our model is based on Kleinberg's navigation model.[41] A node i has a probability to connect to a remote node j through a long-range connection. The probability follows

$$P_{ij} \sim r_{ij}^{-\alpha}, \tag{1.6}$$

where r_{ij} is the lattice distance between node i and j and α is a variable exponent. The model is reasonable, for example, in a social network people are rich in structured short-range connections (knowing a lot of friends nearby) and have a few random long-range connections (knowing a limited number of friends faraway). The networks underlying the model follow the 'small-world' paradigm.[1]

It is not hard for us to conclude that the optimal $\langle \ell \rangle$ is achieved at $\alpha = 0$ in this model from a simple analysis. We understand that only when $\alpha = 0$, the long-range connection does not depend on the distance between the selected nodes, and thus the process of adding long-range connections is simply the same as the procedure used to create an ER network.[22] In this situation, the average path length $\langle \ell \rangle \sim \ln N$.

However in real-world networks, there is always a *cost* on a long-range connection. And the cost is usually proportional to the length of long-range connection. The conclusion

drawn above that optimal transport occurs at $\alpha = 0$ is valid only under the condition that there is no cost limit. Because when $\alpha = 0$, the average length of long-range connections $\langle r \rangle$ reaches the biggest value, this means it costs most in this condition. Now it is not the most effective way to improve transport obviously.

We introduce a cost constraint Λ on the total length of the additional long-range connections, i.e. $\Lambda \equiv \sum r_{ij} \sim N$. The constraint that the total cost of long-range connections is proportional to the system size N is acceptable. With cost limit, the optimal transport no long occurs at $\alpha = 0$ because in this situation $\langle r \rangle$ is large so that only a few long-range connections are available. In the meantime, if α is large, we have a lot of long-range connections but the average length of them is small, the optimizing becomes ineffective in this condition although we have a large number of long-range connections. Thus there is a trade-off between the length of long-range connections and the number of them. So there must exist an intermediate regime of α. In this intermediate regime, the balance between the length and the number of long-range connections is achieved and the optimal transport is gained.

Our numerical results show that the most efficient transport is attained at $\alpha = 3$ for two-dimensional lattice when total length of long-range connections is constrained. In a more general d-dimensional lattice, we find the optimal transport occurs at $\alpha = d + 1$ when system size $N \to \infty$. We also extend the conclusion to a fractal lattice with dimension d_f and find the optimal transport occurs at $d_f + 1$. This is in sharp contrast with the results obtained for unconstrained navigation, where the optimal condition is $\alpha = 0$.

To further understand the role that long-range connections play in optimizing the transport of complex systems, we study the percolation of spatially constrained networks. Usually the spatial constraints are neglected when studying the properties of networks, such as World Wide Web or citation network where the distance is not considered. However, in many other networks, the distance does play an important role, such as the Internet,[12,42] airline networks,[38,43] human travel networks,[15,44] wireless communication networks,[45] social networks,[46,47] etc. All these type of networks are embedded in two-dimensional space.

It has recently been shown that these spatial constraints are important and in certain cases can significantly alter the topological properties of the networks.[27,48–56]

Let us first consider the percolation on two simple network models: an ER network and a two-dimensional square lattice. The percolation process can be simply described as following: nodes are removed from a fully connected ER network or a regular square lattice one by one, the remaining nodes are still connected as a giant cluster until the ratio of removed nodes q exceeds a critical value q_c. When $q > q_c$, the giant cluster breaks into small clusters. The proportion of remaining nodes $p_c = 1 - q_c$ at this critical value is called the threshold of percolation. For an ER network, if $pN = 1$, then it will almost surely have a giant cluster whose size is in the order of $n^{2/3}$.[22] For a two-dimension lattice, when $p \geq 0.59$, one giant cluster extends from top to bottom and from left to right.[21,57]

In chapter 4, we study how spatial constraints are reflected in the percolation properties of networks embedded in d-dimensional lattices. We now consider originally empty lattices embedded in d dimensions ($d = 1$ and 2) by adding long-range connections with the same power law probability $p(r) \sim r^{-\alpha}$ as in chapter 3. Similar distribution has been found in spatially embedded real networks such as social and airline networks. We find for $\alpha \leq d$, the percolation transition belongs to the universality class of percolation in ER networks, while for $\alpha > 2d$, it belongs to the universality class of percolation in regular lattices. However for $d < \alpha < 2d$, the percolation properties show new intermediate behavior different from the two types of universality classes aforementioned, with critical exponents which depend on α.

Chapter 2

Transport and Percolation Theory in Weighted Networks

2.1 Overview

The networks are represented by nodes associated with individuals, organizations, or computers and by links representing their interactions. In many real networks, each link has an associated weight, the larger the weight, the harder it is to transverse the link. These networks are called "*weighted*" networks.[37,38]

Transport is one of the main functions of networks. In a resistor network, transport property is represented by the conductance distribution. The problem of conductance distribution in *unweighted* resistor networks, i.e. each link has unit resistance, was studied in Ref.[35] and it was found that $P(\sigma)$ of ER resistor networks has a narrow shape with exponential tail, while SF networks has a wide range shape with power law tail. It was concluded that SF networks exhibit larger values of conductance than ER networks, thus building SF networks is better for transport.

However the effect of disorder on transport in networks is still an open question. Based on the problem of conductance distribution in unweighted resistor networks, we introduce weights to the resistors. And we continue to study the distribution $P(\sigma)$ of the equivalent electrical conductance σ between two randomly selected nodes A and B on ER and SF

weighted networks. After we introduce weights to the resistors, most resistors have resistance much greater than 1, and the transport property of the network has been changed much. It is important and meaningful to solve the problem of transport in weighted resistor networks.

We first provide an iterative fast algorithm to obtain $P(\sigma)$ for disordered resistor networks, and then we develop a theory to explain the behavior of $P(\sigma)$. The theory is based on the percolation theory[57] for a weighted random network. We model a weighted network by assigning the conductance of a link connecting node i and node j as in Ref.[58]

$$\sigma_{ij} \equiv \exp[-a x_{ij}], \tag{2.1}$$

where the parameter a controls the broadness ("strength") of the disorder, and x_{ij} is a random number taken from a uniform distribution in the range [0,1]. We use this kind of disorder since a recent study of magnetoresistance in real granular materials systems[58] shows that the conductance is given by Eq. (2.1). Moreover, a recent study[59] shows that many types of disorder distributions lead to the same universal behavior. The range of $a \gg 1$ is called the strong disorder (**SD**) limit.[60,61] The special case of unweighted networks, i.e., $a = 0$ or $\sigma_{ij} = 1$ for all links have been studied earlier.[35]

To construct ER networks of size N, we randomly connect nodes with $\langle k \rangle N/2$ links, where $\langle k \rangle$ is the average degree of the network. To construct SF networks, in which the degree distribution follows a power law, we employ the Molloy-Reed algorithm.[62]

2.2 Iterative fast algorithm

The traditional algorithm to calculate the probability density function (pdf) $P(\sigma)$ is to compute σ between two nodes A and B by solving the Kirchhoff equations[63] with fixed potential $V_A = 1$ and $V_B = 0$ and compute $P(\sigma)d\sigma$, which gives the probability that two nodes in the network have conductance between σ and $\sigma + d\sigma$. However, this method is time consuming and limited to relatively small networks because it solves an $(N-2) \times (N-2)$

matrix (Refer to Appendix A) while we have very limited computer resources. Here we also use an iteration algorithm proposed by Grimmett and Kesten[64] to calculate $P(\sigma)$ and show that it gives the same results as the traditional Kirchhoff method.

In the limit $N \to \infty$ we ignore the loops between 2 randomly chosen nodes because the probability to have loops is very small. Hence the resistivity R_i of a randomly selected branch i connecting a node with infinitely distant nodes satisfies $R_i = r_i + 1/(\sum_{j=1}^{k-1} R_j^{-1})$, where $r_i = e^{ax_i}$ is the random resistance of the link outgoing from this node and k is a random number taken from the distribution $\tilde{p}_k = p_k \cdot k / \langle k \rangle$, which is the probability that a randomly selected link ends in a node of degree k, where p_k is the original degree distribution. In Fig. 2.1, we show the schematic iteration method. The randomly selected nodes A and B are connected to the infinitely distant nodes C. When we calculate R_{AC}, the resistance between A and C, we perform the iterative steps as follows:

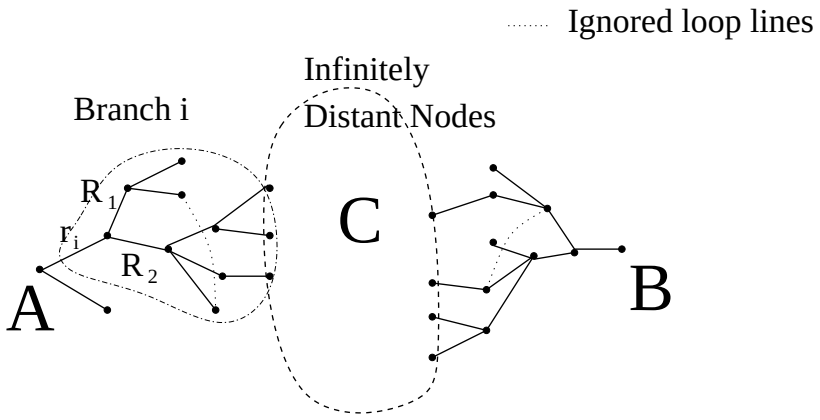

Figure 2.1: Schematic Iteration model. In this example R_1 is infinite, so it is not taken into account in the sum in R_i of Eq. (2.2).

First we calculate the distribution of resistivities of the branches connecting node A with C. We start with $\mathcal{N}$ branches having resistivities $R_i^{(0)} = 0 \quad (i = 1, 2, ..., \mathcal{N})$, where $\mathcal{N}$ is an arbitrary large number. Thus, initially the histogram of these resistivities $P_0(R) = \delta(R)$. At the iterative step $n + 1$, we compute a new histogram $P_{n+1}(R)$ knowing the histogram $P_n(R)$. In order to do this we generate a new set of resistivities $R_i^{(n+1)}$ by connecting in parallel $k-1$ outgoing branches coming from a randomly selected node of degree k obtained

from the distribution $\tilde{p}_k = p_k \cdot k / \langle k \rangle$. Then we compute the resistivity of a branch going through this node via an incoming link with a random resistivity $r_i^{(n)}$ taken from the link resistivity distribution,

$$R_i^{(n+1)} = r_i^{(n)} + \frac{1}{\sum_{j=1}^{k-1} 1/R_j^{(n)}}. \tag{2.2}$$

In Eq. (2.2), if at least one of the terms $R_i^{(n)} = 0$, we take $R_i^{(n+1)} = r_i^{(n)}$. Thus after the first iterative step $P_1(R)$ coincides with the distribution of link resistivities.

According to the theorem proved in,[64] as $n \to \infty$, $P_n(R)$ converges to a distribution of the resistivities of a branch connecting a node to the infinitely distant nodes. The resistivity between a randomly selected node of degree k and the infinitely distant nodes is defined by

$$\tilde{R}_{(i)} = \frac{1}{\sum_{j=1}^{k} 1/R_j}, \tag{2.3}$$

where k is selected from the original degree distribution p_k and R_j is selected from $P_{n\to\infty}(R)$.

Finally, to compute the resistivity R_{ij} between two randomly selected nodes i and j (for example A and B in Fig. 2.1), we compute $R_{ij} = \tilde{R}_{(i)} + \tilde{R}_{(j)}$, where $\tilde{R}_{(i)}$ and $\tilde{R}_{(j)}$ are randomly selected resistivities between a node and the infinitely distant nodes. If $\mathcal{N}$ is a sufficiently large number, we find the conductance distribution $P(\sigma)$ between any two randomly selected nodes.

2.3 Results

In Figs. 2.2(a) and 2.2(b) we show the results of $P(\sigma)$ using the traditional method of solving Kirchhoff's equations for different values of N and the iterative method with $N \to \infty$ for both ER and SF networks. We see that the main part of the distribution (low conductances) does not depend on N, and only the high conductance tail depends on N.

The behavior of the two regimes, low conductance and high conductance, can be understood qualitatively as follows: For strong disorder $a \gg 1$ all the current between two nodes follows the optimal path between them. The problem of the optimal path in a random graph

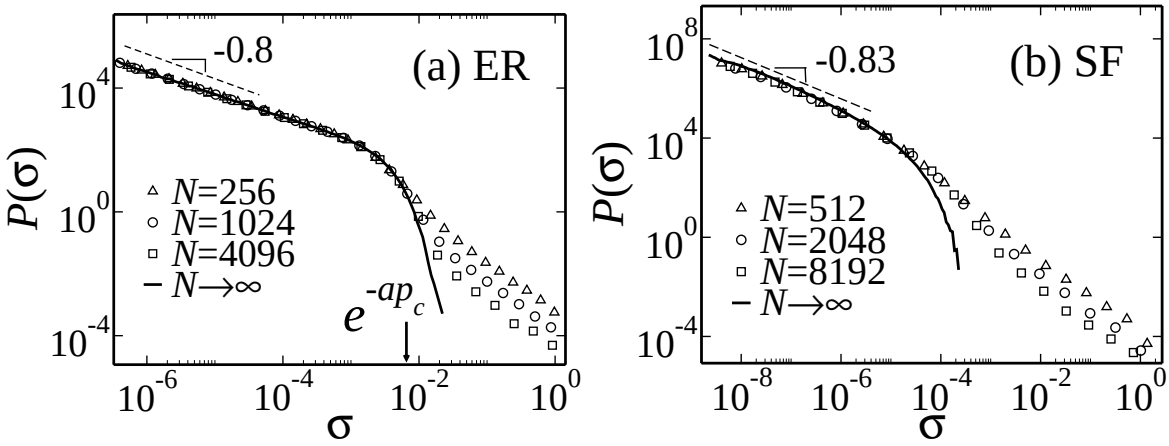

Figure 2.2: Plots of $P(\sigma)$ for several values of N. The symbols are for the Kirchhoff method and the solid line is for the iterative method with $N \to \infty$. (a) ER networks with fixed $\langle k \rangle = 3$ and $a = 15$. (b) SF networks with fixed $\lambda = 3.5$, $k_{\min} = 2$, $\langle k \rangle \approx 3.33$ and $a = 20$. The dashed line slopes are from the prediction of Eq.(2.11) or (2.13).

in the strong disorder limit can be mapped onto a percolation problem on a Cayley tree with a degree distribution identical to the random graph and with a fraction p of its edges conducting.[65] However, the conductance on this path is determined by the bond with the lowest conductance $e^{-ax_{max}}$, where x_{max} is the maximum random number along the path. In the majority of cases $x_{max} > p_c$, where p_c is the critical percolation threshold of the network, and only when the two nodes both belong to the incipient infinite percolation cluster (IIPC),[57] $x_{max} < p_c$. Since the size of the IIPC scales as $N^{2/3}$, the probability of randomly selecting a node inside the IIPC is proportional to $N^{2/3}/N = N^{-1/3}$.[22,57] Then the probability of randomly selecting a pair inside the IIPC is proportional to $(N^{-1/3})^2 = N^{-2/3}$. These nodes contribute to the high conductance range $\sigma > e^{-ap_c}$ of $P(\sigma)$. The low conductance regime is determined by the distribution of x_{max}, that follows the behavior of the order parameter $P_\infty(p)$ (for $p > p_c$) in the percolation problem which is independent of N.[65] (This will be explained later in the theoretical approach for the low conductance regime.)

We call the low conductance regime a *non-percolation regime* and the high conductance regime a *percolation regime*. In contrast, the property of existing two regimes does not show up in the optimal path length[66,67] and only the scaling regime with N appears. This is since the path length for almost all pairs is dominated by the IIPC.[67]

In Figs. 2.3(a) and 2.3(b) we plot for a given N only the non-percolation part of $P(\sigma)$ as a function of σ for fixed values of $\langle k\rangle/a$ and different $\langle k\rangle$ and a values for ER networks. We see that it obeys a power law with the slope $\langle k\rangle/a - 1$ for $\sigma < e^{-ap_c}$. Note that for ER networks $p_c = 1/\langle k\rangle$.[22] In Fig. 2.3(c), we plot the conductance distribution for SF networks for fixed values of $\langle k\rangle/a$. We can see the non-percolation part seems to obey a power law as do the ER networks.

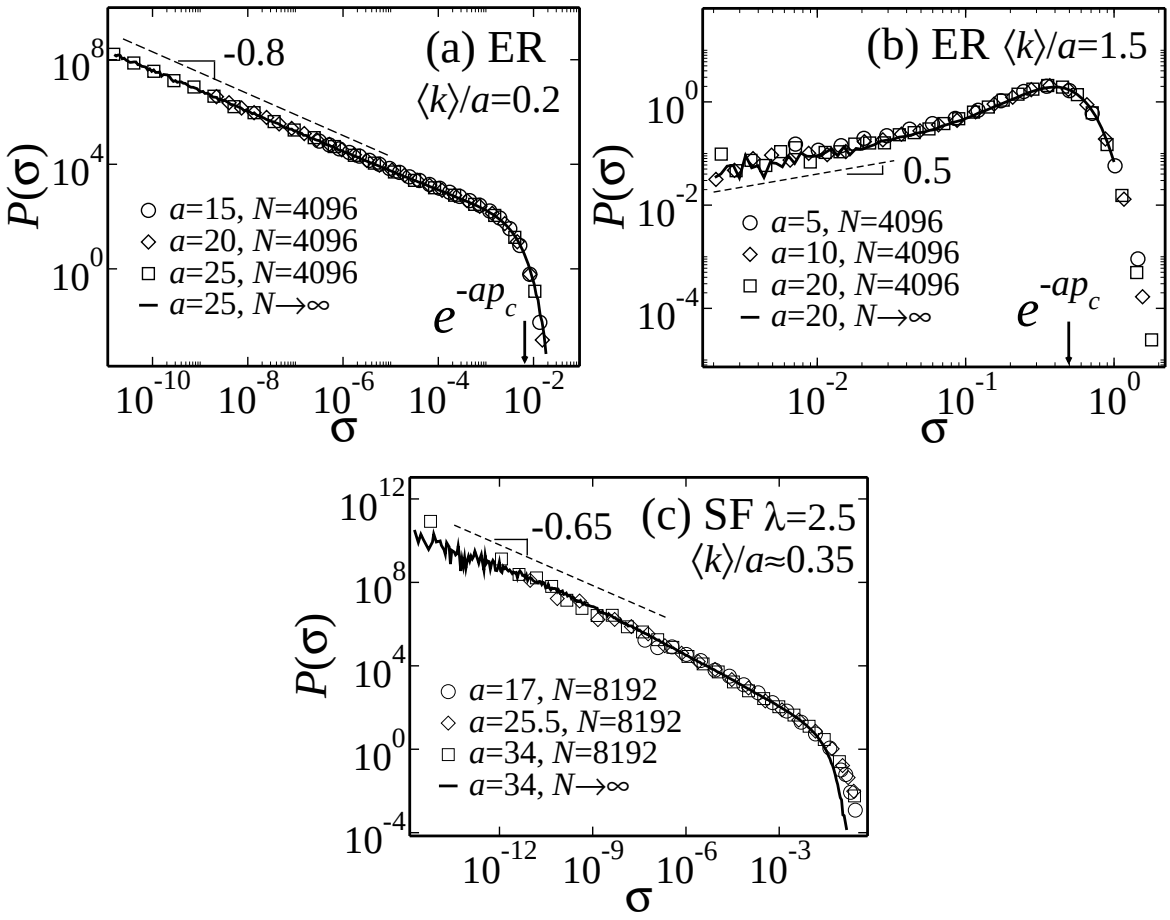

Figure 2.3: Plots of $P(\sigma)$ for fixed $\langle k\rangle/a$. The symbols are for the Kirchhoff method and the solid line is for the iterative method. For the same $\langle k\rangle/a$, the iterative method for different a shows the same $P(\sigma)$ except that the lower cutoff is different. (a) ER network with $\langle k\rangle/a = 0.2$. (b) ER network with $\langle k\rangle/a = 1.5$. (c) SF network with $\langle k\rangle/a \approx 0.35$, $\lambda = 2.5$. The dashed line slopes are from the prediction of Eq.(2.11) or (2.13).

2.4 Analytical approach for $P(\sigma)$

Next we present an analytical approach for the form of $P(\sigma)$ for low conductance regime. The distribution of the maximal random number $x_{\max}$ along the optimal path can be expressed in terms of the order parameter $P_\infty(p)$ in the percolation problem on the Cayley tree, where $P_\infty(p)$ is the probability that a randomly chosen node on the Cayley tree belongs to the IIPC.[65] For a random graph with degree distribution p_k, the probability to arrive at a node with k outgoing branches by following a randomly chosen link is $(k+1)p_k/\langle k\rangle$.[68] The probability that starting at a randomly chosen **link** on a Cayley tree one can reach the ℓth generation is

$$f_\ell(p) \equiv f_\ell = 1 - \sum_{k=1}^{\infty} \frac{p_k k \left(1 - p f_{\ell-1}\right)^{k-1}}{\langle k\rangle}, \tag{2.4}$$

where $f_0 = 1$. Slightly different from f_ℓ is the probability that starting at a randomly chosen **node** one can reach the nth generation,

$$\tilde{f}_n = 1 - \sum_{k=0}^{\infty} p_k (1 - p f_{n-1})^k. \tag{2.5}$$

In the asymptotic limit f_ℓ converges to P_∞ for a given value of p,

$$f_\ell \to P_\infty(p) = 1 - \sum_{k=1}^{\infty} \frac{p_k k \left(1 - p P_\infty\right)^{k-1}}{\langle k\rangle}. \tag{2.6}$$

In this limit we have a pair of nodes on a random graph separated by a very long path of length n. The probability that two nodes will be connected (conducting) at given p, can be approximated by the probability that both of them belong to the IIPC:[64]

$$\Pi(p) = \left[\frac{\tilde{P}_\infty(p)}{\tilde{P}_\infty(1)}\right]^2, \tag{2.7}$$

where $\tilde{P}_\infty(p) \equiv \lim_{n\to\infty} \tilde{f}_n = 1 - \sum_{k=0}^{\infty} p_k (1 - p P_\infty)^k$. Note that the negative derivative of $\Pi(p)$ with respect to p is the distribution of x_{max} and thus gives $P(\sigma)$ in the SD limit. In our case $\sigma = e^{-ap}$, so replacing p by $p = -\ln\sigma/a$ in Eq. (2.7) and differentiating with

respect to σ, we obtain the distribution of conductance in the SD limit when the source and sink are far apart ($n \to \infty$),

$$P(\sigma) = -\frac{d}{d\sigma}\Pi(\sigma) = \frac{2\tilde{P}_\infty(p)}{\sigma a[\tilde{P}_\infty(1)]^2} \cdot \frac{\partial \tilde{P}_\infty(p)}{\partial p} \; |_{p=-\ln\sigma/a} \; . \tag{2.8}$$

For ER networks the degree distribution is a Poisson distribution with $p_k = \langle k\rangle^k e^{-\langle k\rangle}/k!$[22] and thus $P_\infty(p)$ satisfies

$$P_\infty(p) = 1 - e^{-\langle k\rangle p P_\infty(p)}, \tag{2.9}$$

which has a positive root P_∞ for $p > p_c = 1/\langle k\rangle$. And $\tilde{P}_\infty(p) = P_\infty(p)$, thus

$$P(\sigma) = \frac{2P_\infty(p)}{\sigma a[P_\infty(1)]^2} \cdot \frac{\partial P_\infty(p)}{\partial p} \; |_{p=-\ln\sigma/a}, \tag{2.10}$$

where $P_\infty(p)$ and $P_\infty(1)$ are the solutions of Eq. (2.9).

We test the analytical result Eq. (2.10) by comparing the numerical solution of Eqs. (2.9) and (2.10) with the simulations on actual random graphs by solving Kirchhoff equations (Figs. 2.2 and 2.3). The agreement between the simulations and the theoretical prediction is perfect in the SD limit, i.e. when $\langle k\rangle/a$ is small.

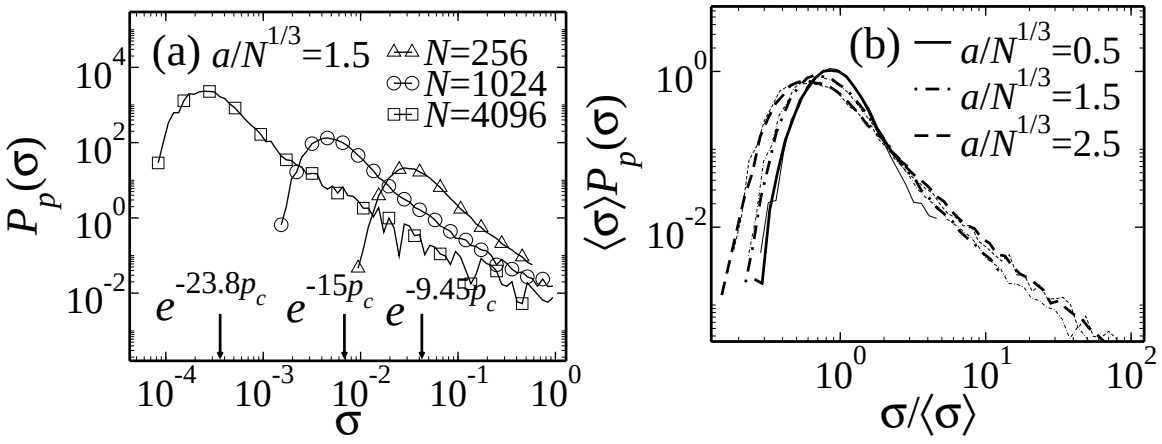

Figure 2.4: Kirchhoff method results of the percolation part of ER networks with the same value of $p_c = 1/\langle k\rangle = 0.33$. (a) Normalized $P_p(\sigma)$ for fixed $a/N^{1/3} = 1.5$. (b) Scaled plot of $\langle\sigma\rangle P(\sigma)$ as function of $\sigma/\langle\sigma\rangle$ for three values of $a/N^{1/3}$. For each value of $a/N^{1/3}$, the thick line is for $N = 256$ and the thin line is for $N = 1024$.

Next we simplify $P(\sigma)$ from Eq. (2.10). Assuming that $P_\infty(1) \approx 1$ which is true for large $\langle k \rangle$ and approximating a slow varying function $P_\infty(p)$ by $P_\infty(1)$ we obtain

$$P(\sigma) \approx 2\frac{\langle k \rangle}{a}\sigma^{\langle k \rangle/a-1}, \tag{2.11}$$

for the range $e^{-a} \leq \sigma \ll e^{-ap_c}$ with $p_c = 1/\langle k \rangle$. In Figs. 2.2 and 2.3 we also show the results predicted by Eq. (2.11). For an infinite network, for $p \leq p_c = 1/\langle k \rangle$, $P_\infty(p) = 0$, and hence, the distribution of conductances must have a cutoff at $\sigma = e^{-ap_c}$. Indeed, in Fig. 2.2(a) and Figs. 2.3(a) and 2.3(b) we see that the upper cutoff of the iterative curves for ER networks is close to e^{-ap_c}.

As discussed above, the range of high conductivities corresponds to the case where both the source and the sink are on the IIPC. Previously we found this percolation part to scale as $N^{-2/3}$. Using Fig. 2.2(a), we compute the integral for each $P(\sigma)$ from e^{-ap_c} to ∞, and find that indeed $\int_{e^{-ap_c}}^{\infty} P(\sigma)\mathrm{d}\sigma \sim N^{-2/3}$, in good agreement with the theoretical approach. To show how the percolation part of $P(\sigma)$ is related to the parameters N, a and p_c, we analyze the conductance between pairs on the IIPC, i.e., each link on the optimal path from source to sink has x less than p_c. We compute $P_p(\sigma)$ of these pairs on the IIPC. When we simulate this process, we have only $N^{-2/3}$ probability to find this part from the original normalized distribution $P(\sigma)$. Thus, we normalize $P_p(\sigma)$ by dividing by $N^{-2/3}$. Figures 2.4(a) and 2.4(b) show the normalized $P_p(\sigma)$ of pairs on the IIPC. In this range, we see that $P_p(\sigma)$ is dominated by high conductivities and we find $\langle \sigma \rangle \approx e^{-ap_c}$ and

$$\langle \sigma \rangle P_p(\sigma) = f\left(\frac{\sigma}{\langle \sigma \rangle}, \frac{ap_c}{N^{1/3}}\right), \tag{2.12}$$

that is, for fixed $ap_c/N^{1/3}$, $\langle \sigma \rangle P_p(\sigma)$ scales with $\sigma/\langle \sigma \rangle$ as seen in Fig. 2.4(b). The scaled distributions have the same shape for the same $ap_c/N^{1/3}$ which specifies the strength of disorder similarly to the behavior of the optimal path lengths.[59,66,67,69] The explanation of this fact for the distribution of conductances is analogous to the arguments presented in Refs.[65,66] for the distribution of the optimal path. Thus the position of the maximum

of the scaled curves in Fig. 2.4(b), and the whole shape of the distributions, depend on $ap_c/N^{1/3}$.

We also find that the extreme high conductivities correspond to the case where source and sinks are separated by only one link. In this case, $P(\sigma) = \frac{\langle k \rangle}{aN\sigma} \sim \sigma^{-1}$, $(\sigma < 1)$.

2.5 Conclusions

In summary, we find that $P(\sigma)$ exhibits two regimes. For $\sigma < e^{-ap_c}$, we show both analytically and numerically that for ER networks $P(\sigma)$ follows a power law,

$$P(\sigma) \sim \sigma^{\langle k \rangle / a - 1}. \tag{2.13}$$

We also find that for SF networks, Eq. (2.13) seems to be a good approximation, consistent with numerical simulations. The distributions of optimal path length and the path length of the electrical currents in complex weighted networks[66,67] have been found to depend on N for all length scales and all types of networks studied. In contrast, here we find that the *low* conductance tail of $P(\sigma)$ does not depend on N for both ER and SF networks. However, the *high* conductance regime ($\sigma > e^{-ap_c}$) of $P(\sigma)$ does depend on N, in a similar way to the optimal path length and current path length distributions.[66,67]

Chapter 3

Towards Design Principles for Optimal Transport Networks

3.1 Overview

Complex web-like structures have been the subject of intensive research in several fields, including those focusing on social relationships, biological resources, and transportation systems.[6,9,10,12,19,20,70] The topological features of these systems, which go beyond the standard regular lattice geometry, have been described in terms of complex network structures. In this way, the theory of random graphs as well as concepts like "small-worldliness and "scale-freeness" have been consistently used to characterize and classify the diverse complex networks found in nature,[1,2,22,25,32] providing interesting insights about their underlying structure and functionality. Generally speaking, the geometrical features of complex networks are not necessarily associated or restricted to a given topological dimension in space. However, a large number of real transport networks can be geographically represented or spatially embedded,[71–75] such as the US airport network,[43] networks of streets and highways,[76] physical systems,[77] mobile agents[78] and also the network of activity in the brain.[79,80]

In science, nature and technology, the transport of information, energy, or even people, can be optimized by adding long-range connections (shortcuts) to an underlying geograph-

ical network. In recent studies,[27,81] it has been shown that the optimal design of transport networks can be associated with the presence of special "critical" correlations between the local structure and the long-range connections, added in such a way as to generate gradients that permit the information to flow efficiently from source to target in the network. In several real systems, however, transport is usually constrained by some involved cost. In a subsequent study,[82] it has been shown that, without loss of generality, transport with local and global knowledge on a square lattice with cost limitation on the additional links can be optimized through the inclusion of long-range connections between pairs of nodes following a probability distribution that decays as a power law of their Manhattan distance, namely the the distance counted as the number of connections separating nodes in the regular lattice. Efficient transport is then obtained when the exponent α of the power law distribution is tuned to 3 for a two-dimensional lattice, in sharp contrast with the previous results for unconstrained local[27,81] and global[83] navigation, where the optimal value is $\alpha = 2$ and 0, respectively. Considering the results for a one-dimensional and a two-dimensional lattices, it is then conjectured that optimal transport is achieved when $\alpha = \alpha_e = d + 1$ for a d-dimensional lattice. Subsequently, power law distributions of long-range connections on geographical networks have also been used to study navigation and other types of processes.[48,76,84–88]

In this work, we propose a new *real-time* algorithm for efficient study of the global navigation. Using this algorithm, simulation results could be obtained for large system sizes, up to the order of 10^9 nodes. This is carried out by memorizing only the neighbors of a particular node at each time step of the algorithm. Initially, optimal exponents have been obtained using this algorithm, reproducing previous numerical estimates of both unconstrained and constrained global navigation for one and two dimensions.[27,81–84] In addition, we present here results of simulations for three-dimension global navigation and constrained global navigation processes on a fractal lattice of dimension d_f, reinforcing the conjectured optimal navigation exponent, $\alpha_e = d + 1$.

3.2 Navigation without cost

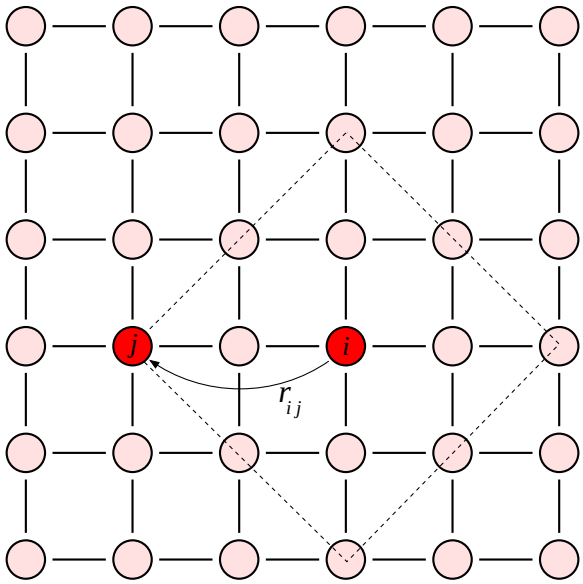

Figure 3.1: Two-dimensional square lattice with long-range connections. Each node has 4 short-range connections to its nearest neighbors. A long-range connection is added to a randomly chosen node i with probability proportional to $r^{-\alpha}$. Here $r = 2$, there are 8 nodes (on dashed square box) with the same lattice distance r to node i, and we randomly choose the node j from these 8 nodes to be connected to node i.

Using local information and a *decentralized* algorithm, the problem of efficient navigation in small-world networks was recently studied by Kleinberg.[27] Figure 3.1 shows a regular two-dimensional square lattice with $N = L \times L$ nodes, where L is the linear size of the lattice. Accordingly, each node i has a random long-range connection to a node j with probability $P(r_{ij}) \sim r_{ij}^{-\alpha}$, where r_{ij} is the lattice (Manhattan) distance between node i and j. This model follows the small-world paradigm, i.e., it is rich in short-range connections and but has only few long-range connections. The optimal navigation by a *decentralized* algorithm based on the local information occurs when the exponent $\alpha = 2$.[27]

The probability $P(r_{ij})$ that nodes i and j will have a long-range connection can be mapped on a density distribution $p(r)$, where $r = r_{ij}$. The number of nodes separated by a lattice distance r from a given node in a d-dimensional lattice is proportional to r^{d-1} (see

Fig. 3.1). Thus we have

$$p(r) \sim r^{-\alpha} r^{d-1}. \tag{3.1}$$

The normalization factor of Eq. (3.1) scales as

$$\int_1^L r^{-\alpha} r^{d-1} dr \sim \begin{cases} L^{d-\alpha}, & \alpha < d \\ \ln L, & \alpha = d \\ (\alpha - d)^{-1}, & \alpha > d \end{cases} \tag{3.2}$$

where $L = N^{1/d}$.

In order to improve the transport property of a lattice network, it is not necessary to assign every node a long-range connection – which would be a high-cost strategy. Instead, we assign a long-range connection to a small fraction of randomly selected nodes. This leads to a dramatic improvement in the transport properties of the network, but at a much lower cost. This model can be generated using the following steps:

(i) A regular d-dimensional lattice with N nodes is created with each node connected to its $2d$ nearest neighbors.

(ii) A node i is randomly selected from the total N nodes to receive a long-range connection. The length of the long-range connection r is randomly generated using Eqs. (3.1) and (3.2).

(iii) Another node j is also randomly selected from those nodes with the same lattice distance r to node i. We then connect node i and node j. For example, Fig. 3.1 shows eight nodes (on the dashed square box) that have the same lattice distance to node i ($r = 2$). We randomly take node j from these eight nodes and connect it to node i.

(iv) Repeat steps (ii) and (iii) until the total number of long-range connections N_l reaches a preset value, e.g., 10% of the total number of nodes N.

3.3 The Real-Time Algorithm for global navigation

In many real-world optimal navigation problems, one has access to global information when designing the optimal transport network. With global information, between any two randomly selected nodes a and b, we can compute the shortest path length ℓ_{ab}. In the next section we will show that ℓ follows approximately a Gaussian distribution. The average shortest path length $\langle \ell \rangle$ is thus the most important parameter when evaluating the transport efficiency of the entire network. The usual method of calculating $\langle \ell \rangle$ is first to build the model with long-range connections, and then measure the shortest path length ℓ for every pair of nodes. This method is fast and effective for small systems, however, given the present-day computer resources, it is not practical nor efficient to pre-build such a big system and then add long-range connections.

Hence, we introduce a *real-time* algorithm to calculate $\langle \ell \rangle$ for large systems, which gives us the same results as the common method previously described, but much faster. To evaluate $\langle \ell \rangle$ for a network, the usual procedure is to calculate the shortest path length for every pair of nodes. If we have many realizations of a network, however, we can randomly pick one pair of nodes in one realization and calculate ℓ of this pair. Then, in another realization, we randomly pick another pair of nodes and calculate ℓ again. After a large number of realizations, we are then able to determine $\langle \ell \rangle$. Since for every realization the network is created by using the same parameters, $\langle \ell \rangle$ from many different realizations reflects the result from different pairs of nodes in a single network.

For a single realization of the *real-time* algorithm, it is not necessary to create the entire network. Starting from a randomly-selected node a, we generate its neighbors in real time (time step). For example, in a square lattice, node a always has 4 nearest neighbors (or $2d$ nearest neighbors on a d-dimensional lattice), if node a is on the boundary, we generate its neighbors using a periodic boundary condition. After that, we consider the additional long-range connections of node a, the one that it receives according to step (ii), and the others that eventually start from another node, according to step (iii). As long as we have

all the available neighbors of node a (including connections to both the nearest- and long-range neighbors), we record and classify them as shell one nodes, meaning that they are one link apart from node a. After that, we generate all the neighbors of the shell one nodes not generated yet and classify them as shell two nodes. We repeat this process until we reach the destination node b, which is also randomly selected. We count the number of steps from node a to b during this process. In this way, we find ℓ_{ab} between node a and b. We repeat many realizations to find the average $\langle \ell \rangle$, until we generate a smooth curve $\langle \ell \rangle$ vs α for different values of α.

A crucial step of this algorithm is when we consider the additional long-range connections of node a. When we add long-range connections to the network we always randomly select two nodes with the lattice distance r, thus the number of additional long-range connections k_l for each node obeys a Poisson distribution,

$$f(k_l) = \frac{\lambda^{k_l} e^{-\lambda}}{k_l!}, \tag{3.3}$$

where λ is the average number of long-range connections for each node, and is calculated as

$$\lambda = \langle k_l \rangle = \frac{2N_l}{N}, \tag{3.4}$$

where N_l is the total number of long-range connections, which can be a preset value.

Using the Poisson distribution (3.3), we generate the number of long-range connections k_l for node a. It must be noted that k_l can be greater than one, this means we do not limit the number of long-range connections for each node. This is a little different from the original model of Kleinberg. From Eqs.(3.1) and (3.2), we assign r to each long-range connection of node a. Finally, as described in step (iii), we choose all the long-range neighbors for node a.

In this *real-time* algorithm, we do not build the entire network, but only generate the neighbors needed for each step. This algorithm saves computer resources and produces exactly the same results as those from the more usual method of building the entire network.

This *real-time* algorithm makes the simulation of very large systems possible (e.g., 10^9 nodes).

3.4 Navigation with cost constraint

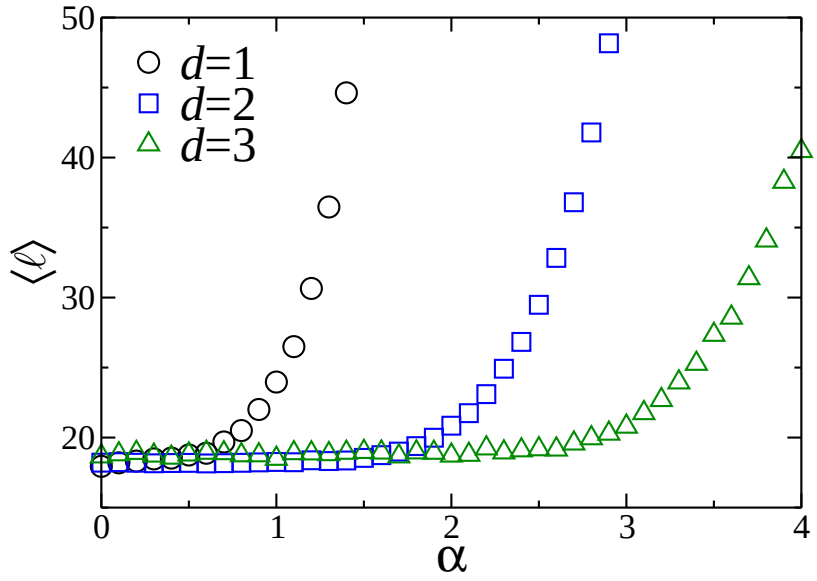

Figure 3.2: A fraction of 10% of nodes in a regular d-dimensional ($d = 1$, 2 and 3) lattice with linear size $L = 1000$ are randomly selected to receive long-range connections with different α. As seen, the optimal $\langle \ell \rangle$ is achieved for $\alpha = 0$. The results are averaged over 4,000 realizations for each network. Note, when α increases above the value of d, $\langle \ell \rangle$ increases dramatically. For the dependence of $\langle \ell \rangle$ on L for different α, see Refs.[51,87]

We note that when $\langle \ell \rangle$ is based on the global information of the network it does not demonstrate the uniqueness of the navigation based on local information described by Kleinberg. In this case, the optimal $\langle \ell \rangle$ is achieved at $\alpha_e = 0$[51] which can be understood from a simple analysis. When $\alpha = 0$, the length of the long-range connections does not depend on the distance, and thus the process of adding long-range connections is simply the same as the procedure used to create a Watts- Strogatz (WS) network[1] or even an Erdős-Rényi (ER) network.[22] In this situation, the average shortest path length scales as a logarithm function of the network size, $\langle \ell \rangle \sim \ln N$.

Figure 3.2 shows $\langle \ell \rangle$ for three different lattices ($d = 1$, 2, and 3) with the same linear size

$L = 1000$, in which a fixed fraction (10%) of nodes receive long-range connections. We find that the optimal $\langle \ell \rangle$ indeed emerges at $\alpha_e = 0$. Moreover, when $\alpha < d$ (for $d = 1$, 2, and 3), $\langle \ell \rangle$ appears almost the same as the results for $\alpha = 0$, which means in this regime $\langle \ell \rangle \sim \ln N$, and that the transport property behaves as small-world networks, see also Ref.[51,87]

In real-world situations, however, the financial cost of adding links always plays an important role when improving the transport in an existing network. Consider the case of an existing transport network which needs improvements.[89] The financial cost to build up a large number of new direct connections between distant stops (i.e., non-neighboring sites) can make it prohibitive, since only limited resources are normally available for this task. These types of problems can be modeled in the following way, which is similar to the formulation presented in Ref.[27] In a two-dimensional regular square lattice, with all $N = L^2$ sites present, each site i is connected with its four nearest neighbors. The sites represent the stops and the bonds represent the routes of the transport system (see Fig. 3.1). The distinctness from Kleinberg navigation is that the addition of long-range connections to the system stops when their total length (cost), $\sum r_{ij}$, reaches a given value Λ. We further assume that the total cost $\mathit{\Lambda}$ will be proportional to the size N of the network, i.e., $\mathit{\Lambda} = AN$, where A is a constant.[82] This assumption is justified since bigger systems should obtain proportionally larger budgets for improvement. Moreover, the total length of the links in the original lattice is proportional to N (number of nodes).

Since α controls the average length of the long-range connections, we obtain that, for a fixed value of Λ, and small values of α, longer connections, but fewer in number can be added due to the imposed total length limit. We therefore expect that an optimal navigation condition must be revealed as a trade-off between the length and the number of long-range connections N_l added to the system. Here we show that a rather different behavior can be observed when realistic constraints on total length are imposed on the process of adding long-range connections, regardless if navigation is based on local or global knowledge of the network structure.

When the total cost is fixed at $\mathit{\Lambda} = AN = AL^d$,[82] the available number of long-range

connections will be

$$N_l = \Lambda / \langle r \rangle \,, \tag{3.5}$$

where $\langle r \rangle$ is the average length of long-range connections, which can be calculated from (3.1) for a given α as,

$$\langle r \rangle \sim \int_1^L r^{d-\alpha} dr \sim \begin{cases} L, & \alpha < d \\ L/\ln L, & \alpha = d \\ L^{d+1-\alpha}, & d < \alpha < d+1 \\ \ln L, & \alpha = d+1 \\ 1, & \alpha > d+1. \end{cases} \tag{3.6}$$

Note that in Eq. (3.6), $\langle r \rangle$ decreases and N_l increases, when α increases. When α is small ($\alpha \leqslant d$), $\langle r \rangle \sim L$ and $N_l = \Lambda / \langle r \rangle \sim L^{d-1}$, only a small fraction (in fact zero fraction in the limit $L \to \infty$) of nodes are needed to have long-range connections in order to improve the transport of the network. When $\alpha > d+1$, however, a large number of long-range connections are available ($N_l \sim L^d$) to improve the transport of the network, but each long-range connection is short, thus, these long-range connections do not efficiently improve transport. Thus, the intermediate regime of α can be expected to be useful and optimize the transport on the network, i.e., $d < \alpha \leqslant d+1$.

Figure 3.3 shows the distribution $P(\ell)$ of shortest path length ℓ for different α on a two-dimensional lattice with additional long-range connections of total length $N = L^2$. Note that $P(\ell)$ follows an approximate Gaussian distribution for different α. Consequently, since $\langle \ell \rangle$ has its minimum value when $\alpha = d+1 = 3$, the optimal navigation is achieved at that α value. Figure 3.4 shows the normalized distribution $P(\ell)$ for different α. The different curves in Fig. 3.3 approximately collapse to a single curve when scaled appropriately.

We extract more quantitative information about this navigation problem by performing extensive simulations for different values of α and different system sizes N. We assume that the total length (cost) is proportional to the total length of the links in the underlying

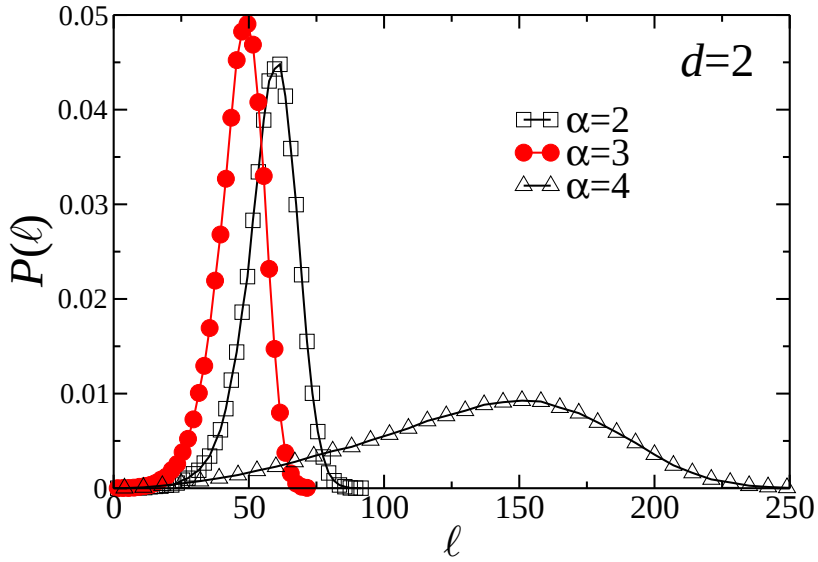

Figure 3.3: Distribution of the shortest path length ℓ for two-dimensional lattice ($L = 1000$) with additional long-range connections where the total length Λ of the added long-range connections is limited to $N = L^2$. Note, the non-monotonic behavior with respect to α. For $\alpha = d + 1 = 3$ the location of the peak of the distribution is the smallest. We sampled 100,000 network realizations for each α.

network, i.e., $\Lambda = AL^d$, where A is a constant. That is, the budget to improve the system is a fraction of the cost of the current network (without long-range connections).[90] We first focus on identifying the optimal transport conditions on regular lattices, i.e., $d = 1$, $d = 2$ and $d = 3$. In each case, we simply add long-range connections to the regular lattice. The procedure is almost the same as in the Kleinberg navigation model, except that in step (iv) we stop adding the long-range connections when the total length of long-range connections $\sum r_{ij}$ reaches a preset value $\Lambda = AN$, instead of a fixed number of long-range connections. After that, we calculate the average shortest path $\langle \ell \rangle$ over all realizations of pairs of nodes.

From the results in Figs. 3.5 (a), (b) and (c), we see the presence of a minimum $\langle \ell \rangle$ for different system sizes at the same value of the exponent $\alpha = d + 1$, when $N \to \infty$. Thus, based on the global knowledge of the network structure, the most efficient navigation occurs at $\alpha_e = d + 1$.

To further test the optimal navigation condition, $\alpha_e = d+1$, we plot $\langle \ell \rangle$ vs L for different α. Figures 3.6 (a), (b) and (c) clearly show that for $\alpha \neq d + 1$ the shortest path length $\langle \ell \rangle$

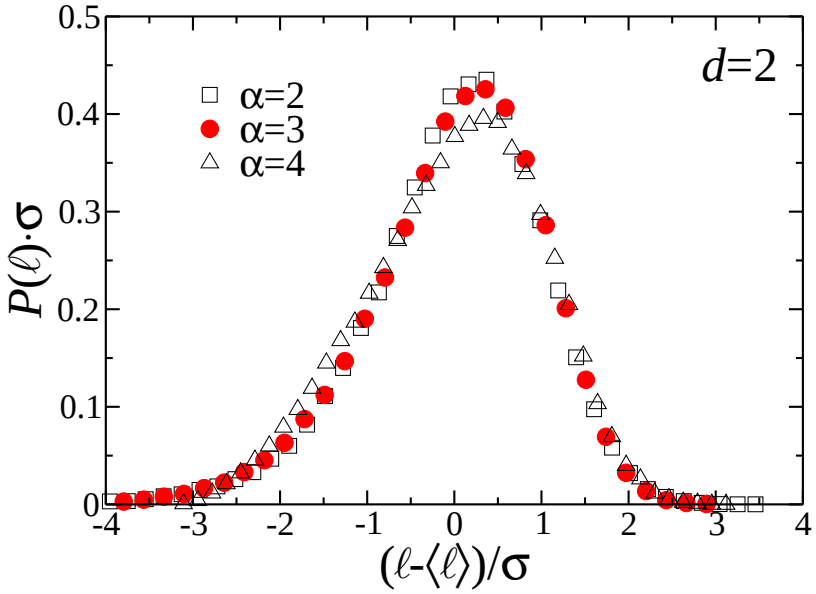

Figure 3.4: Normalized distribution of the shortest path length ℓ for two-dimensional lattice ($L = 1000$) with additional long-range connections where the total length Λ of the added long-range connections is limited to $N = L^2$. In the figure, $\langle \ell \rangle$ is the mean ℓ and σ is the standard deviation for each curve. We sampled 100,000 network realizations for each α.

follows a power law with L. For $\alpha = d + 1$, $\langle \ell \rangle$ follows power law with a smaller exponent when $d = 1$ (Fig. 3.6 (a)), and it appears to be less than a power law for $d > 1$ (Figs. 3.6 (b) and (c)).

Figures 3.5 (d) and 3.6 (d) show the analogous optimal navigation results when the substrate is a fractal. Specifically, the fractal is generated from an original two-dimensional regular lattice. We randomly remove the nodes with a probability $1 - p = q$. We increase q from 0 until a critical percolation occurs.[21,57] In this critical condition, $q = q_c \cong 0.4$ a giant cluster extends from top to bottom and from left to right across the lattice, which is a fractal. The dimension of this giant cluster is $d_f \cong 1.9$. In Ref.[81] it is shown that, for local navigation and without constraints on total length, the optimal transport is for $\alpha = d_f$. In order to improve the transport with global knowledge on this fractal, we append additional long-range connections to the cluster using the same procedure as in a two-dimensional regular lattice. The difference here is that the total length of long-range connections is fixed to be AN_f, where N_f is the number of nodes in the giant component of the fractal,

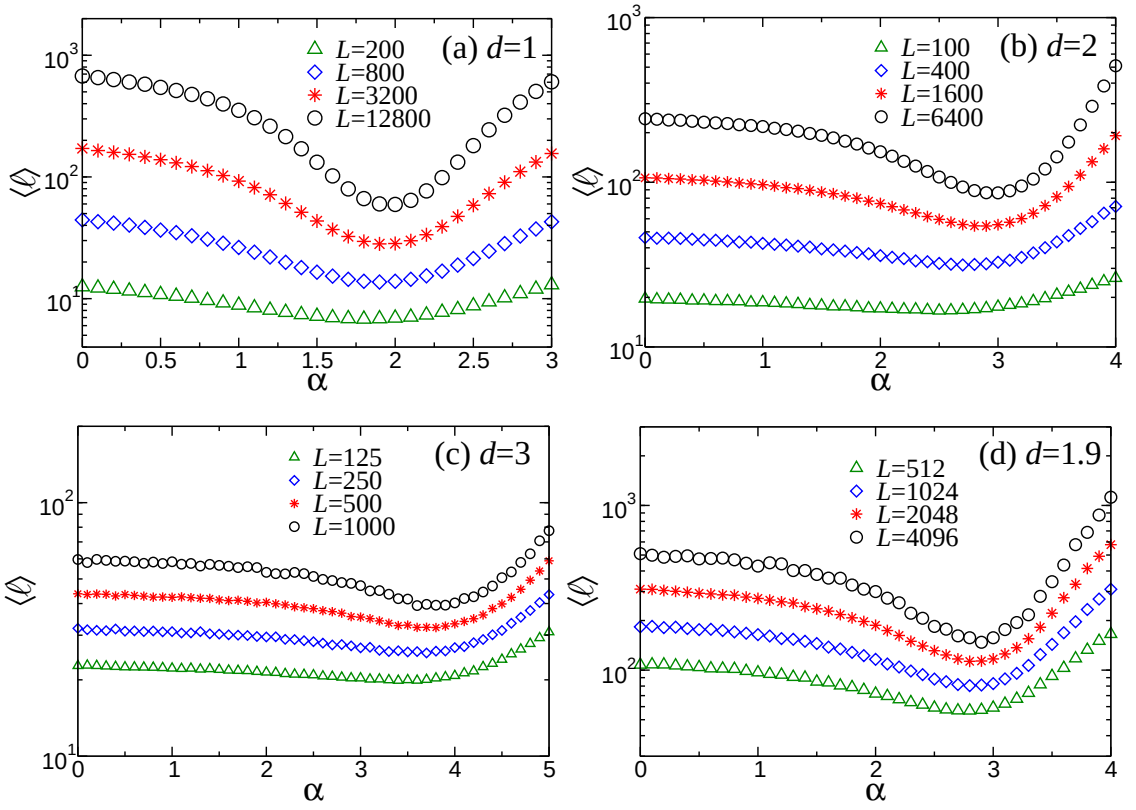

Figure 3.5: Average shortest path length $\langle \ell \rangle$ as a function of α for (a) one-, (b) two-, (c) three-dimensional lattices and (d) fractal ($d = d_f \cong 1.9$) with additional long-range connections taken from the power law distribution, Eq.(3.1) as a function of α. The total length Λ of the added long-range connections is limited to $10N$ for one-dimensional lattice, N for two- and three-dimensional lattices and N_f for the fractal. The plots suggest that the optimal shortest path length is achieved at $\alpha = d+1$ for regular lattices and $\alpha = d_f+1$ for the fractal. Note that (b) is similar to Fig. 3 in Ref.[82] but with larger system sizes. The results are averaged over 4000 realizations for the three smaller L and 400 realizations for the largest L.

which can be calculated as $N_f = L^{d_f}$. From Figs. 3.5 (d) and 3.6 (d), we see that optimal navigation occurs at $d_f + 1 \cong 2.9$ when $N \to \infty$. We also tested another fractal lattice, i.e., the Sierpinski carpet[91] with $d_f = 1.89$ and find that the optimal navigation is achieved at $d_f + 1 \cong 2.89$ when $N \to \infty$. Note that the *real-time* algorithm can not be used on these critical percolation lattices because the shortest path length is calculated on the giant cluster which must be pre-built up.

Figure 3.6 (a) suggests that in a one-dimensional lattice $\langle \ell \rangle$ always follows a power law

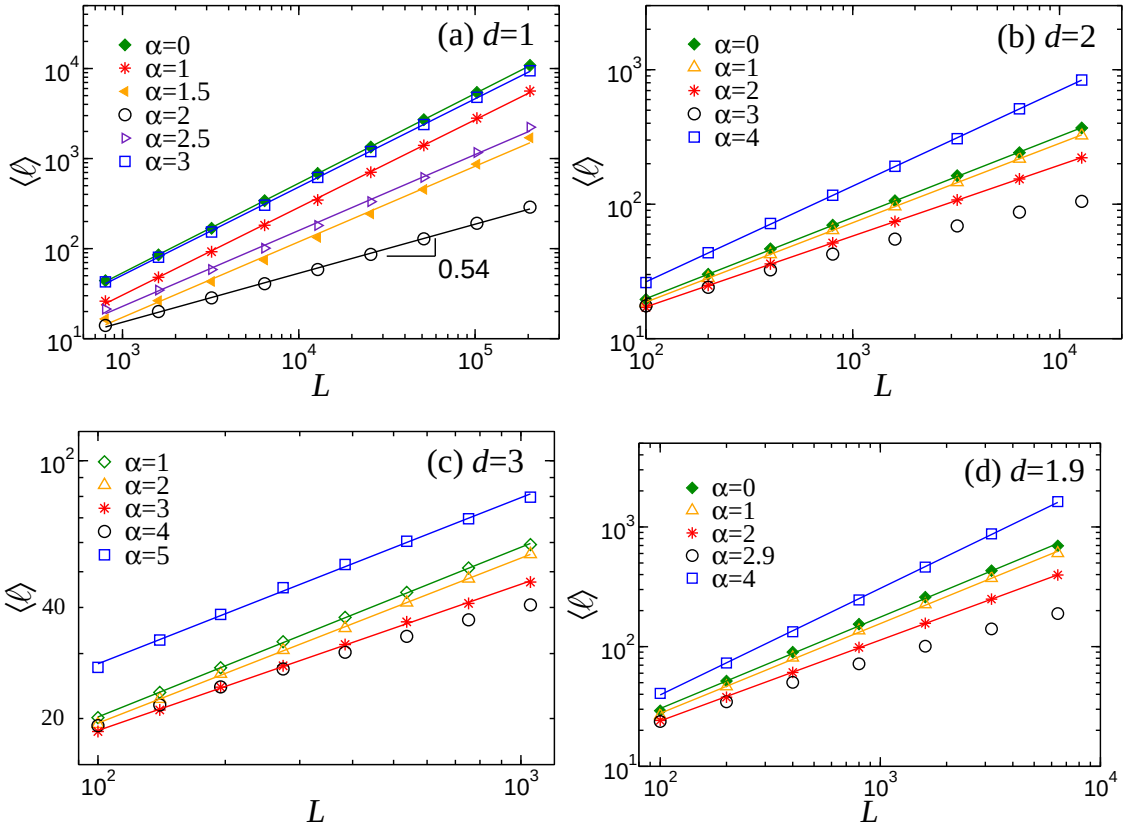

Figure 3.6: Average shortest path length $\langle \ell \rangle$ as a function of system linear size L with different α for (a) one-, (b) two-, (c) three-dimensional and (d) fractal ($d_f \cong 1.9$) lattices with additional long-range connections taken from the power law distribution, Eq.(3.1). The total length Λ of the added long-range connections is limited to $10N$ for one-dimensional lattice, N for two- and three-dimensional lattices and N_f for the fractal. The plots suggest that the optimal shortest path length is achieved at $\alpha = d + 1$ for regular lattices and $\alpha = d_f + 1$ for the fractal. For $d = 1$, the slope of the fitting line $\delta_s \cong 0.54$ for $\alpha = d+1 = 2$, $\delta_s \approx 0.84$ for $\alpha = 1.5$ and 2.5, and $\delta_s \approx 1$ for $\alpha = 0$, 1 and 3. For $d = 2$, $\delta_s \approx 0.60$ for $\alpha = 0$ and 1, $\delta \cong 0.71$ for $\alpha = 4$, however for $\alpha = d + 1 = 3$, $\langle \ell \rangle$ seems to follow a weaker dependence from a power law, more likely a logarithmic law (see Fig. 3.7 (b)). For $d = 3$, $\delta_s \approx 0.46$ for $\alpha = 0, 1$ and 2, $\delta_s \cong 0.40$ for $\alpha = 3$, $\delta_s \cong 0.45$ for $\alpha = 5$, however for $\alpha = d + 1 = 4$, $\langle \ell \rangle$ seems to follow a logarithmic law. For $d = 1.9$, $\delta_s \approx 0.75$ for $\alpha = 0$ and 1, $\delta \cong 0.89$ for $\alpha = 4$, however for $\alpha = d_f + 1 = 2.9$, $\langle \ell \rangle$ seems to follow a logarithmic law. The results are averaged over 4000 realizations for each α for $d = 1, 2$ and 3 and 1000 realizations for $d = 1.9$.

dependence as a function of system sizes. For $d > 1$, as seen in Figs. 3.6 (b), (c) and (d), however, $\langle \ell \rangle$ scales as a power law with the system linear size L for all values of α except for $\alpha = d + 1$ for which the scaling seems to be less than a power law. For $d > 1$, we test two

possible forms for $\langle \ell \rangle$ vs L, (i) a power law and (ii) a logarithmic law. Figure 3.7 (a) shows the successive slopes δ_s obtained from $\ln \langle \ell \rangle$ vs $\ln L$ for $d = 2$, testing whether $\langle \ell \rangle$ follows a power law or not. Here we assume $\langle \ell \rangle \sim L^{\delta_s}$ and see that δ_s remains approximately constant when $\alpha \neq d+1$, but decreases when $\alpha = d+1 = 3$. This suggests that $\langle \ell \rangle$ follows a power law only when $\alpha \neq d+1$. Similar results have been obtained for $d = 3$ and $d = d_f$ in Fig. 3.8 (a) and (b) respectively. We next assume that $\langle \ell \rangle$ vs L follows a logarithmic law with exponent γ_s, i.e., $\langle \ell \rangle \sim \ln^{\gamma_s} L$. In Fig. 3.7 (b) we plot the data assuming this function in a double logarithmic plot. As can be seen, apart from the case $\alpha = 3$, which fits quite well as a straight line, for the other values of α, $\langle \ell \rangle$ increases faster. Indeed, we plot in Fig. 3.7 (c) the successive slopes γ_s obtained from the plot of $\ln \langle \ell \rangle$ vs $\ln \ln L$ for $d = 2$ (Fig. 3.7 (b)). We see that γ_s keeps almost a constant value when $\alpha = d+1 = 3$ but it increases when $\alpha \neq d+1$. This suggests that $\langle \ell \rangle$ follows a power of a logarithmic dependence when $\alpha = d+1$. Similar results have been obtained for $d = 3$ and $d_f = 1.9$ in Fig. 3.8 (c) and (d) respectively.

3.5 Analytic Arguments

Besides the support from simulation results, we present analytic arguments suggesting that for $N \to \infty$ the optimal navigation is achieved for $\alpha = d+1$. Figure 3.6 (a) shows the one-dimensional case in which the scaling of $\langle \ell \rangle$ with L is a power law for different α, and that the power law is smallest when $\alpha = 2$. Li *et al.*[88] provides an exact solution for the optimal navigation with a total cost restriction for the one-dimensional case. They conclude that, for $d = 1$, the optimal navigation occurs at $\alpha = 2$ when $N \to \infty$.

Next, we suggest a simple analysis showing that, for $d > 1$, $\alpha = d+1$ is indeed the only case for which a logarithmic scaling of $\langle \ell \rangle$ with L can occur while for $\alpha \neq d+1$ a power law with L must hold. For a fixed total cost $\Lambda = AL^d$, the density of long-range connections is defined as $\rho = N_l/N$, where N_l is the available number of long-range connections in the lattice. From Eq. (3.5) and $\Lambda = AN$, we find that $\rho = A \langle r \rangle^{-1}$. From Eq. (3.6), it follows

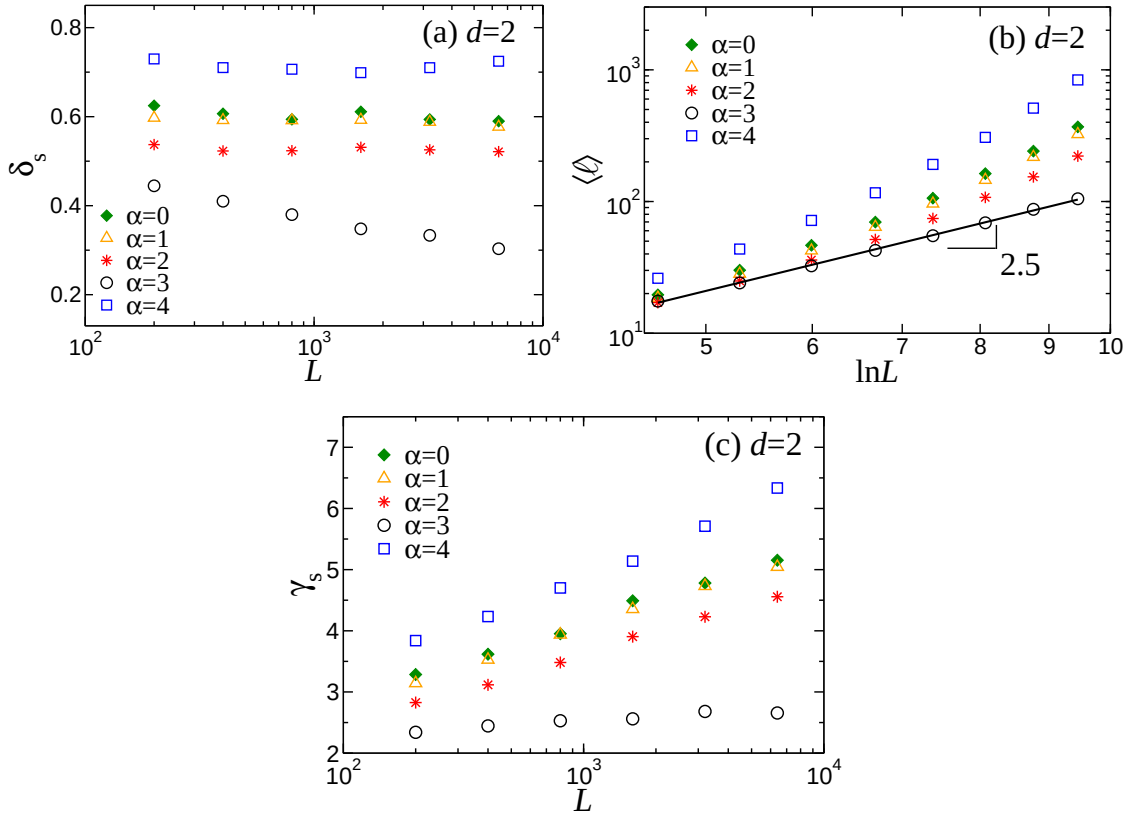

Figure 3.7: Successive slopes for $d = 2$ of (a) δ_s obtained from $\ln\langle\ell\rangle$ vs $\ln L$ (of Fig. 3.6 (b)), (b) $\langle\ell\rangle$ as a function of $\ln L$ in a double logarithmic plot and successive slopes of (c) γ_s obtained from $\ln\langle\ell\rangle$ vs $\ln\ln L$ taken from (b). The total length Λ of the added long-range connections is limited to $N = L^2$. Note that in (a) for $\alpha = d + 1 = 3$, δ_s decreases with L while for other values of α, δ_s is roughly constant. In (c) for $\alpha = d+1 = 3$, γ_s keeps roughly constant and for other values of α, γ_s increases with L. This suggests that for $\alpha = 3$ the relation between $\langle\ell\rangle$ and L is a function that increases less than a power law and more likely that $\langle\ell\rangle$ increases logarithmically with L.

that for $d \le \alpha < d+1$, $\rho \sim L^{\alpha-d-1}$ and, for $\alpha < d$, $\langle r\rangle \sim L$, leading to $\rho \sim L^{-1}$. So when $\alpha < d+1$, the density ρ of long-range connections decreases as a power law with L. As a consequence of this power law decrease in density, $\langle\ell\rangle$ must increase as a power of L. To verify this, we bound $\langle\ell\rangle$ by the relation $\langle\ell\rangle > \rho^{-1/d}$. Here $\rho^{-1/d}$ is from the small world model in which $\alpha = 0$, with a fixed concentration of links $\langle\ell\rangle \sim \rho^{-1/d}\ln L$.[92] Since, for $0 < \alpha < d+1$, $\langle\ell\rangle$ decreases with increasing α, the bound $\langle\ell\rangle > L^{(d+1-\alpha)/d}$ is rigorous and $\langle\ell\rangle$ in this range must scale as a power of L. For $\alpha > d+1$ and $N \to \infty$, from Eq. (3.6) $\langle r\rangle \sim 1$ and the density becomes independent of the system size, i.e., $\rho \sim 1$. When this

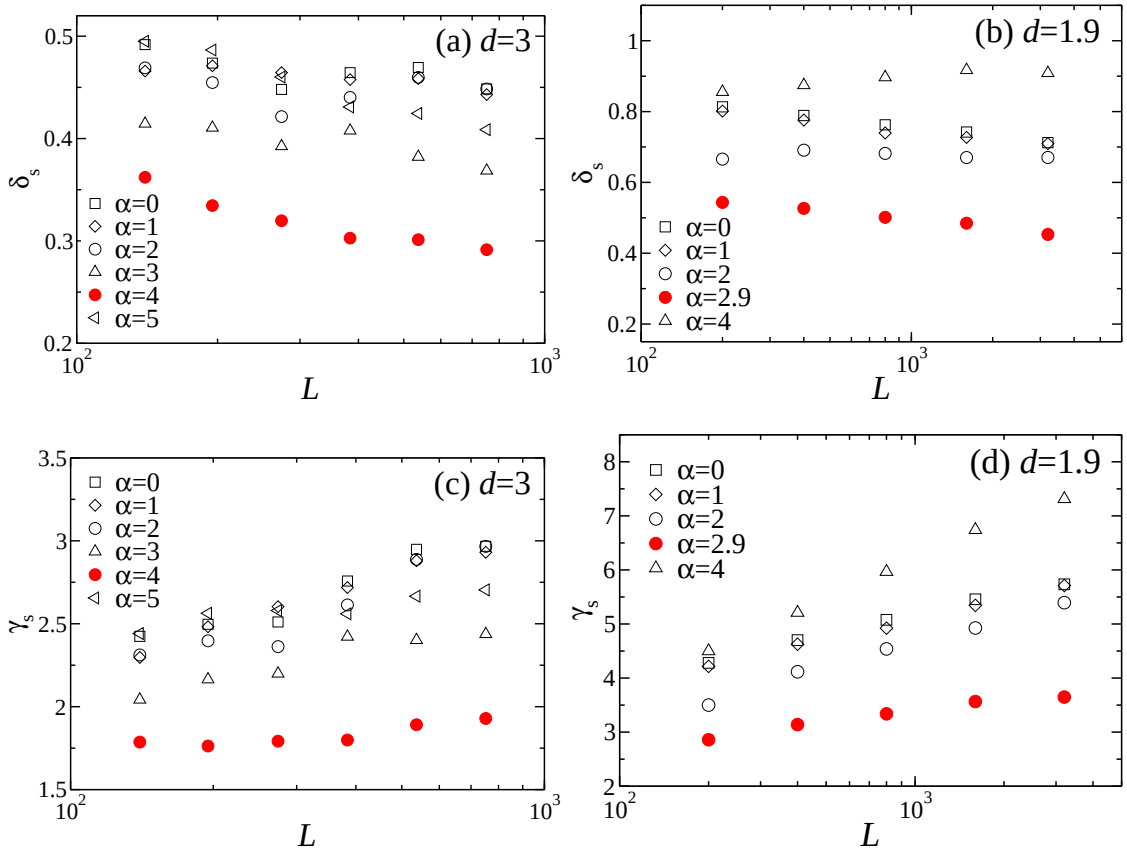

Figure 3.8: Successive slopes of δ_s obtained from $\ln\langle\ell\rangle$ vs $\ln L$ for (a) three-dimensional lattice $d = 3$ and (b) fractal $d = 1.9$, and successive slopes of γ_s obtained from $\ln\langle\ell\rangle$ vs $\ln\ln L$ for (c) three-dimensional lattice $d = 3$ and (d) fractal $d = 1.9$. The total length Λ of the added long-range connections is limited to N for (a) and (c), and N_f for (b) and (d).

is the case, the effect of the constraint Λ on navigation is negligible. Thus the navigation on the networks is similar to the original lattice without additional long-range connections, therefore $\langle\ell\rangle \sim L$. Thus we conclude that, as Figs. 3.6 (b), (c) and (d) show, when $\alpha = d+1$, the increase with L of $\langle\ell\rangle$ is less rapid than a power law and may scale logarithmically with L.

It is important to note that our global navigation scheme with $\langle\ell\rangle$ can be considered as a lower bound to any other transport navigation process. For example, a strategy based on purely local knowledge of the network structure will necessarily perform worse than any other with global information. In Ref.,[27] for example, the greedy (decentralized) algorithm is introduced as a paradigm based on local information, where the traveler, when leaving a

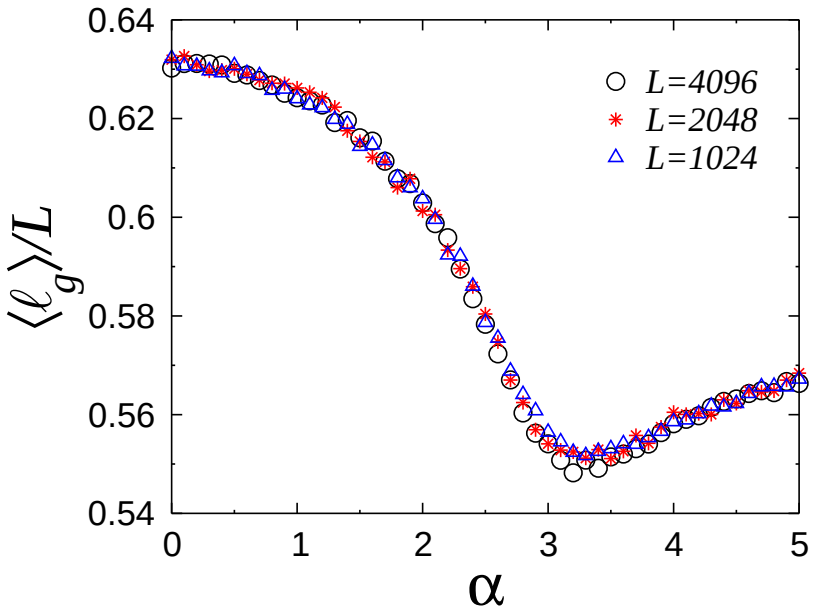

Figure 3.9: The characteristic average delivery time $\langle \ell_g \rangle / L$ as a function of α for navigation with the greedy (decentralized) algorithm in two-dimensional lattice. The cost Λ involved to add long-range connections changes the behavior of the density of long-range connections. As a result of that, a minimum is observed at $\alpha \approx 3$. Each data point is a result of 4000 simulations and the cost Λ is fixed at L^2.

node, chooses to move to the one among its neighbors which has the smallest Manhattan distance to the target. Kleinberg found that $\alpha = 2$ is the optimal value in the navigation with the greedy algorithm[27] for two-dimensional lattice. We next ask, what would be the optimal α for the greedy algorithm when cost restriction $\Lambda = AL^2$ is imposed? We find also for the greedy algorithm that the optimal value is $\alpha = 3$. This is shown in Fig. 3.9, where we plot the average delivery length $\langle \ell_g \rangle$ that a message travels with only local information of the system geometry. The message is sent from the source node s to the target node t through a network generated with the constraint $\Lambda = AL^2$. Remarkably, the presence of a minimum also at $\alpha \approx 3$ shows that the type of information (local or global) used by the message holder to pass it through the system during the navigation process becomes unimportant if the network is constructed under length (cost) limitations. However, the two mechanisms display very different and distinct behaviors regarding the scaling with system size. While we observe logarithmic growth for the optimal condition $\alpha = 3$, in the case of

global information, the time to reach the source, with the greedy algorithm and with cost constraint, appears to increase linearly with size for all values of α. The linearity of $\langle \ell_g \rangle$ with L is observed in the scaling collapse (Fig. 3.9) of the curves of $\langle \ell_g \rangle / L$ vs α.

3.6 Conclusions

In summary, our results suggest that, regardless of the strategy used by the traveler, based on local or global knowledge of the network structure, the best transportation condition is obtained with an exponent $\alpha = d+1$, where d is the topological dimension of the underlying lattice. Our results hold for $d = 1, 2$ and 3 for regular d-dimensional lattices and d_f for fractal. The results recently reported by Bianconi *et al.* [43] on the U.S. airport network yield an exponent $\alpha = 3$, which is similar to our optimal exponent for $d = 2$. The fact that the probability of a flight connection within U.S. decays as a power law with the distance between airports, $r^{-\alpha}$, where $\alpha = 3.0 \pm 0.2$, seems to reveal the optimized aspect of the network under the conditions of geographical availability (for customer satisfaction) and cost limitations (for airline company profit). Furthermore, recent studies by Gallos *et al.* measured empirically α in the brain and found $\alpha = d_f + 1 \approx 3.1$ which may suggest that the brain is optimizing connections with a cost constraint.[79,80]

The result $\alpha = d + 1$ is in sharp contrast with the results obtained for unconstrained systems with global and local information, where the optimal conditions are $\alpha = 0$[83] and $\alpha = d$,[27,81] respectively. The contrast between the optimal results is even more dramatic. While in the unconstrained case the mean length of a link diverges, we find that when cost is considered the mean length is finite. In the case where the traveler has global knowledge of the network, and is able to identify the shortest path for navigation, we obtain a slow (logarithmic) growth with size for the transit time at the optimal condition. A different picture is obtained if the traveler has only local knowledge of the network. For example, in the case where the transportation path is decided based on the Manhattan distance to the target, we obtain a linear growth of the transit time with system size, for all values of the

exponent α. Finally, our results suggest that the idea of introducing a cost constraint in the navigation problem offers a different theoretical framework to understand the evolving topologies of other important complex network structures in nature, such as subways, trains, or the Internet. Of course, at this point we must emphasize that our approach represents only one specific model within a larger family of models where design principles can be tested to improve the performance of the transport system. In the case of airport networks, for example, other variables than the particular cost function that we adopted can be used for realistic optimization purposes.

Chapter 4

Percolation of Spatially Constrained Networks

4.1 Overview

In Kleinberg's model, we assign additional long-range connections to a regular d-dimensional lattice. Our objective is clear, i.e. to improve the transport of the network. However we have not considered the underlying structure of the new formed small-world like network. We know that the original d-dimensional lattice has a definite dimension. Obviously after we introduce additional long-range connections, the dimension has been changed to an non-integer dimension. And with adjustable number and length of long-range connections, we are able to create any non-integer dimension networks.

Moreover, we want to understand the robustness of this new type of spatially constrained network if it is under random attacks, which are like the percolation process. In complex networks, percolation is another extremely important distinguishing feature besides transport. In order to obtain percolation process based on Kleinberg's model, we now consider originally *empty* lattices embedded in d dimensions by adding long-range connections with the same power law probability $p(r) \sim r^{-\alpha}$, where $\alpha \geq 0$ is a variable which controls the strength of spatial constant, and r is the Euclidean distance between the nodes, which is a little different from the lattice distance r_{ij} in Chapter 3 because in a spatial embedded network Euclidean distance is well defined.[51] We choose a power law for the distance distribution because it is supported recently from the findings in real networks, such as In-

ternet,[12,42] airline networks,[38,43] human travel networks[15,44] and other social networks.[46–48] This model is simplified so that all the links are long-range connections.

In the beginning we add enough number of long-range connections until the average degree of the network reaches $\langle k \rangle \sim 4$. In this condition, almost all the nodes are connected. Then we randomly remove $qN = (1-p)N$ nodes, leaving pN nodes remained. The remaining nodes are still connected as a giant (spanning) cluster until q exceeds a critical point q_c. The proportion of the remaining nodes $p_c = 1 - q_c$ at this critical point is called the *threshold* of percolation. At the threshold, we have

$$M \sim N^{\beta}, \tag{4.1}$$

where M is the number of nodes in the giant cluster and β is a constant independent of the system size N. As mentioned in Chapter 1, for the special case, the percolation of ER networks at the threshold, $\beta = 2/3$. And below the threshold, the giant cluster breaks into small clusters, and it does not exist in the network. This percolation process of removing nodes is called *site percolation*.

In this chapter, we study the percolation properties of spatially constrained networks embedded in one- or two-dimensional space and find how spatial constraints affect the percolation properties of the networks, i.e. how the parameter α influences the giant cluster size M, the threshold p_c or q_c, cluster size distribution $n(S)$ and even the *chemical dimension* d_ℓ.[93,94] We find that, for $\alpha \leq d$, the percolation transition belongs to the universality class of percolation in ER networks as mean field, while for $\alpha > 2d$ it belongs to the universality class of percolation in regular lattices (only for two-dimensional lattice, for one-dimensional lattice, there is no percolation transition found because network is a regular linear chain). However for $d < \alpha < 2d$, the percolation properties show new intermediate behavior different from ER networks, with critical exponents that depend on α.

4.2 Percolation Properties

4.2.1 Build the network

The algorithm to build the networks is almost the same as the *pre-build* algorithm introduced in Chapter 3, however we can not use the *real-time* algorithm because we investigate the percolation over the whole network and we need to pre-build the whole network. The difference from Chapter 3 is that now we do not create the short-range connections which connect a node to its nearest neighbors.

First we arrange the nodes in a d-dimensional regular lattice ($d = 1$ or 2), so we have well-defined Euclidean distances. Then we randomly add $k_t N/2$ long-range connections to the network through the steps in Chapter 3. And at last we remove the duplicate links. If there are no duplicate links, we are supposed to obtain the average degree of the network $\langle k \rangle = k_t$. But for different α, especially when $\alpha > 2d$, most of long-range connections are very short so that it is easier for each node to have duplicate links to its nearest neighbors. Due to the generation process, the final network average degree $\langle k \rangle$ is actually less than k_t. To make all the nodes connected, we set $k_t = 4$ and find that in the two-dimensional case, $3.6 < \langle k \rangle \leq 4$ for all α, while in one-dimensional case, $3.1 < \langle k \rangle < 4$ for all α.

4.2.2 Critical threshold of percolation

Critical threshold is obtained by measuring the sizes of the giant (spanning) cluster and 2nd biggest cluster in the percolation process, i.e. M and M_2 as function of p. Below the critical threshold p_c there is no spanning cluster, but when $p > p_c$, M/N becomes finite. At the threshold, the size of 2nd biggest cluster M_2 reaches a maximum value according to percolation theory.[93,94]

Figure 4.1 shows the critical thresholds for several typical values of α by determining the size of 2nd biggest cluster. In table 4.1, p_c is shown for different α in linear chain (one-dimensional lattice) and two-dimensional square lattice percolation respectively. We can see that when $\alpha \leq d$ ($d = 1$ and 2), p_c is very close to the known result $p_c = 1/k_t = 0.25$

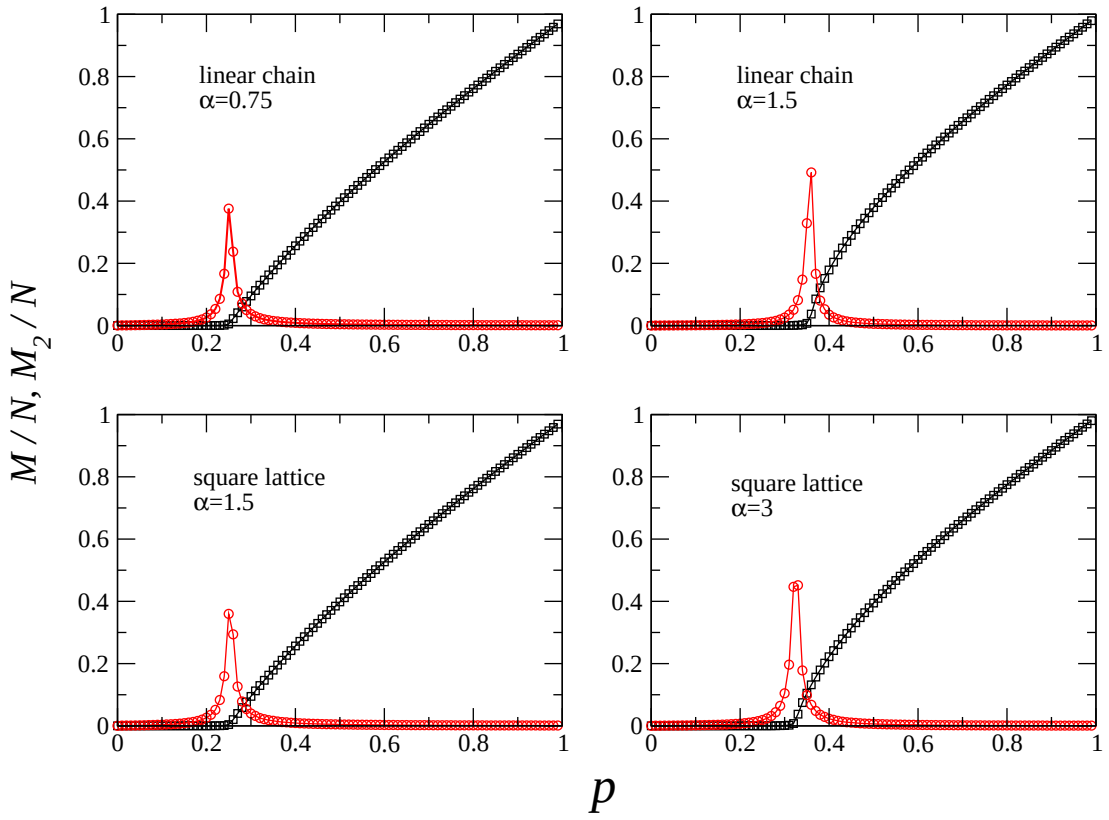

Figure 4.1: Sizes of the giant cluster M (black squares) and the second biggest cluster M_2 (red circles) as function of the fraction of remaining nodes p for several typical values of α. M_2 is re-scaled by appropriate factor for better visibility. In the linear chain plots ($N = 10^6$), $p_c = 0.25$ and 0.35 for $\alpha = 0.75$ and 1.5 respectively, and in the square lattice plots ($L = 10^3$), $p_c = 0.25$ and 0.33 for $\alpha = 1.5$ and 3 respectively.

for percolation in ER networks. For $\alpha > 2d$ in linear chain, percolation transition is not observed because the network is a regular one-dimensional lattice, there is no percolation in this case. While for $\alpha > 2d$ in square lattice, e.g. $\alpha = 5$, $p_c = 0.57$ is very close to the known value of site percolation in the regular square lattice, $p_c = 0.59$. However for $d < \alpha < 2d$, the percolation properties show new intermediate behavior different from the percolation in ER networks or regular lattices, which depends on α.

Next we analyze how the mass (size) of giant cluster M scales with the system size N at critical threshold. Generally the giant cluster has been found to be fractal so that its mass M dependence with size is governed by a non-integer exponent,[93,94] i.e. $M \sim N^{\beta}$ (eq. (4.1)). For the percolation of ER networks, $\beta = 2/3 = 0.67$.[5,22] Table 4.1 and shows that in

α	p_c	β	d_ℓ	ν_ℓ	d_f	d_e	$\frac{d_f}{d_e}$	$1+\frac{1}{\beta}$	τ
0.75	0.25	0.67	1.79	0	∞	∞	/	2.50	2.50
1	0.25	0.67	1.80	0	∞	∞	/	2.50	2.50
1.25	0.27	0.70	1.81	0.60	3.05	4.45	0.69	2.43	2.44
1.5	0.35	0.77	1.82	1.14	1.63	2.12	0.77	2.30	2.30
1.75	0.49	0.86	1.79	1.46	1.26	1.45	0.87	2.16	2.15

α	p_c	β	d_ℓ	ν_ℓ	d_f	d_e	$\frac{d_f}{d_e}$	$1+\frac{1}{\beta}$	τ
1.5	0.25	0.67	1.78	0	∞	∞	/	2.50	2.50
2	0.25	0.67	1.79	0	∞	∞	/	2.50	2.49
2.5	0.27	0.70	1.81	0.46	3.92	5.65	0.69	2.43	2.44
3	0.33	0.76	1.81	0.88	2.12	2.76	0.76	2.31	2.32
3.5	0.41	0.87	1.79	0.92	1.92	2.18	0.88	2.15	2.14
4	0.49	0.93	1.78	0.95	1.89	2.00	0.94	2.07	2.06
5	0.57	0.94	1.79	0.96	1.87	1.99	0.94	2.06	2.05

Table 4.1: Critical exponents and thresholds at percolation transition for the networks embedded in linear chains (upper table) and square lattice (lower table). The results are measured with system size $N = 10^6$ for linear chain and $L = 10^3$ for square lattice, except that β is obtained from different system sizes.

both one- and two-dimensional lattices percolation, when $\alpha \leq d$, $\beta = 0.67$, the percolation shows mean-field-like behavior like the percolation of ER networks. And for $\delta > 2d$ (only in two-dimensional lattice percolation), where long-range connections are very rare, nodes are connected locally, $\beta \approx 0.94$, the percolation is a regular square lattice percolation.[93,94]

Figure 4.2 shows M vs N in a double-logarithmic plot for the percolation of one- and two-dimensional lattices with several different α and p around p_c. It is observed that only when $p = p_c$, M vs N shows a power law relation (straight line in a double-logarithmic plot). While for $p < p_c$ and $p > p_c$, the curve of M vs N bends down and up respectively. These feature is a characteristic of the percolation transition.[93,94]

4.2.3 Fractal dimension of percolation

Let us consider the giant cluster at the critical threshold. Fractal dimension d_f describes how the mass S within a sphere of radius r scales with r:

$$S(r) \sim r^{d_f}, \tag{4.2}$$

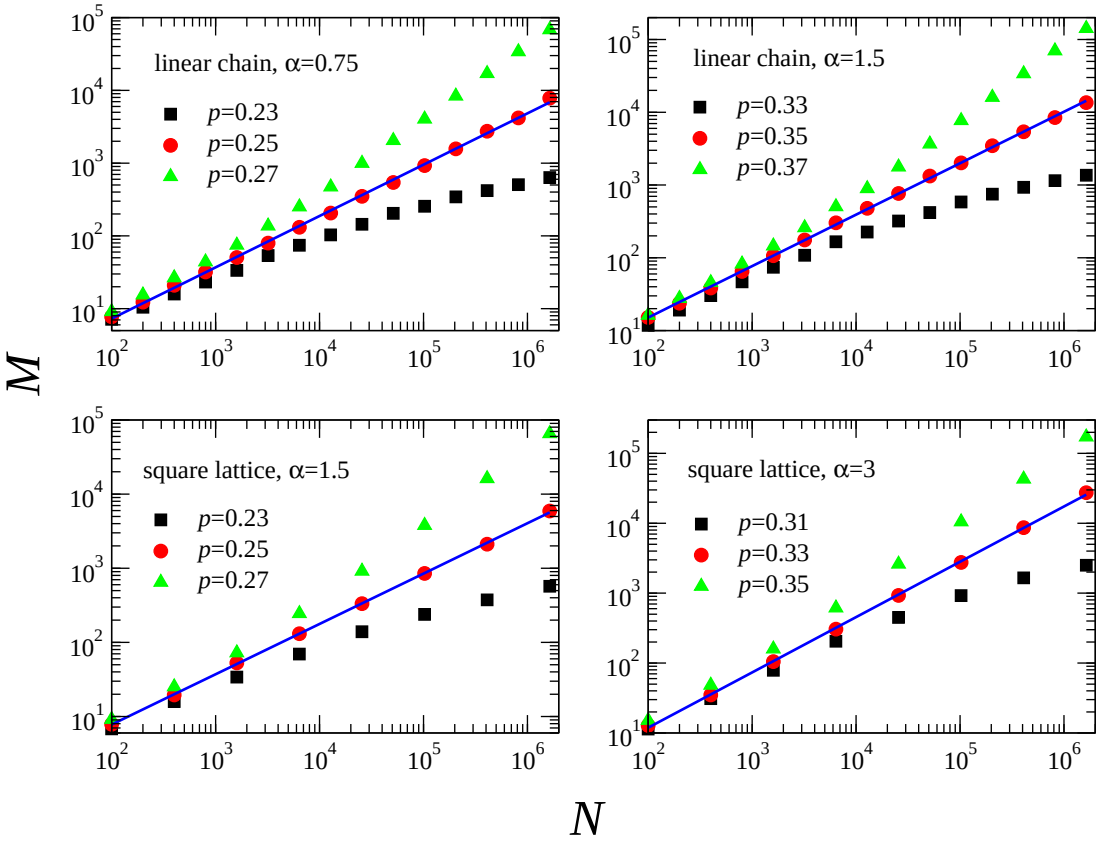

Figure 4.2: Size of the giant cluster M as a function of the system size N for several typical values of α and different fractions p of the remaining nodes in one-dimensional linear chain and two-dimensional square lattice. In the plots, curves bend down for $p < p_c$ (black squares), bend up for $p > p_c$ (green triangles) and follow a power law relation (straight line in log-log plot) for $p = p_c$ (red circles). The slopes β of M vs N are: 0.67 and 0.77 for linear chain when $\alpha = 0.75$ and 1.5 respectively, and 0.67 and 0.76 for square lattice when $\alpha = 1.5$ and 3 respectively.

where d_f is the fractal dimension of the giant cluster. $S(r)$ is obtained by averaging over many cluster realizations. And above criticality $S(r) \sim r^{d_e}$, where d_e is the dimension of the embedded network. Therefore, we have

$$M \sim N^{\frac{d_f}{d_e}}. \tag{4.3}$$

Combine with eq.(4.1), we have

$$\beta = \frac{d_f}{d_e}. \tag{4.4}$$

d_f and d_e can be obtained from the network at criticality ($p = p_c$) and above criticality ($p = 1$). Then we compare $\frac{d_f}{d_e}$ with the results of β from figure 4.2 so that we can evaluate the accuracy of exponents at percolation.

Actually in the process of obtaining d_f and d_e, we first obtain the *chemical dimension* d_ℓ,[93,94] which describes how the mass of the cluster within a chemical length, i.e. length of the shortest path, ℓ scales with ℓ:

$$S(\ell) \sim \ell^{d_\ell}. \tag{4.5}$$

Comparing with Eqs. (4.2), we can infer the relation between regular Euclidean distance r and chemical distance ℓ:

$$r \sim \ell^{d_\ell/d_f} \equiv \ell^{\nu_\ell}. \tag{4.6}$$

The relation is often written as $\ell \sim r^{d_{\min}}$, where $d_{\min} \equiv 1/\nu_\ell$ can be regarded as the fractal dimension of the minimal path.

So d_f is obtained by the following steps:

(i) A node i is randomly selected from the giant cluster.

(ii) Find all the nodes with chemical distance ℓ ($\ell = 1, 2, 3...$) from node i. Record the number of these nodes as M_ℓ.

(iii) Compute the radius R_ℓ of these M_ℓ nodes from step (ii).

(iv) Repeat steps (i) to (iii) for many realizations and get the average of M_ℓ and R_ℓ at the same chemical distance ℓ over many networks.

We plot M_ℓ vs ℓ from the results of step (ii) and R_ℓ vs ℓ from step (iii) to obtain d_ℓ and ν_ℓ respectively. We plot M_ℓ vs R_ℓ from step (iv) to get d_f. For the dimension of embedded network d_e, we follow the same steps above to obtain it but steps are performed on the network without removing any nodes ($p = 1$).

Figure 4.3 shows M vs ℓ and R vs ℓ in percolation at the critical threshold. From the plots, we can conclude that the chemical dimension, also known as the *graph dimension*

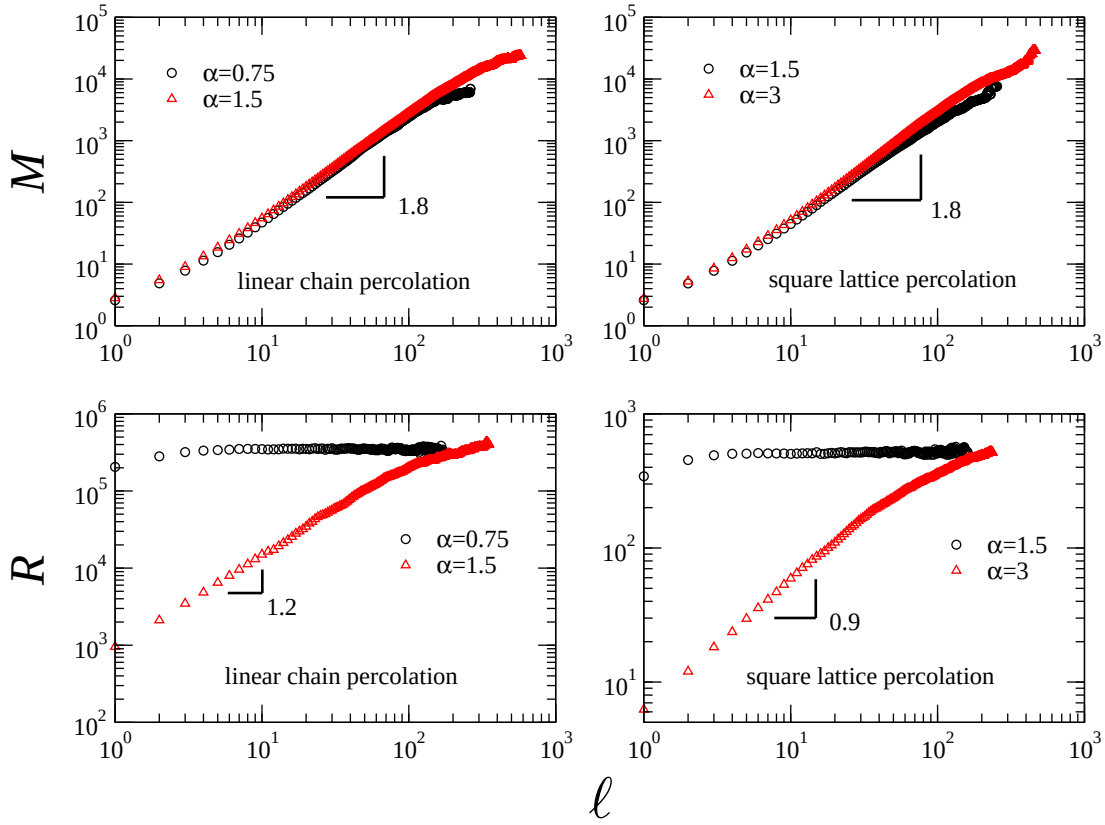

Figure 4.3: Size of the giant cluster M and radius R as a function of the chemical distance ℓ for several typical values of α in percolation at the critical threshold in one-dimensional linear chain (left 2 plots) and two-dimensional square lattice (right 2 plots). Chemical dimension d_ℓ is obtained from M vs ℓ (upper 2 plots): $d_\ell \approx 1.8$ for all α in percolation, and ν_ℓ is determined from R vs ℓ (lower 2 plots).

or *topological dimension*, $d_\ell \approx 1.8$ for all α. However, in the plots of R vs ℓ, ν_ℓ can be 0 because when $\alpha \to 0$, the radius R of giant cluster is independent of the chemical distance.

From figure 4.4 M vs R, we obtain d_f and d_e, which can also be determined from d_ℓ and ν_ℓ, i.e. $d_f \approx d_\ell/\nu_\ell$. The difference between d_f and d_e is that d_f is fractal dimension which is obtained in the percolation process at the critical threshold, while d_e is the dimension of the embedded network with $\langle k \rangle \approx 4$ and it is obtained before removing the nodes. Table 4.1 shows all the values of d_f, d_e and $\frac{d_f}{d_e}$ for different α. Eq. (4.4) claims that $\beta = \frac{d_f}{d_e}$, and from table 4.1, we can compare both values and find that they agree with each other completely.

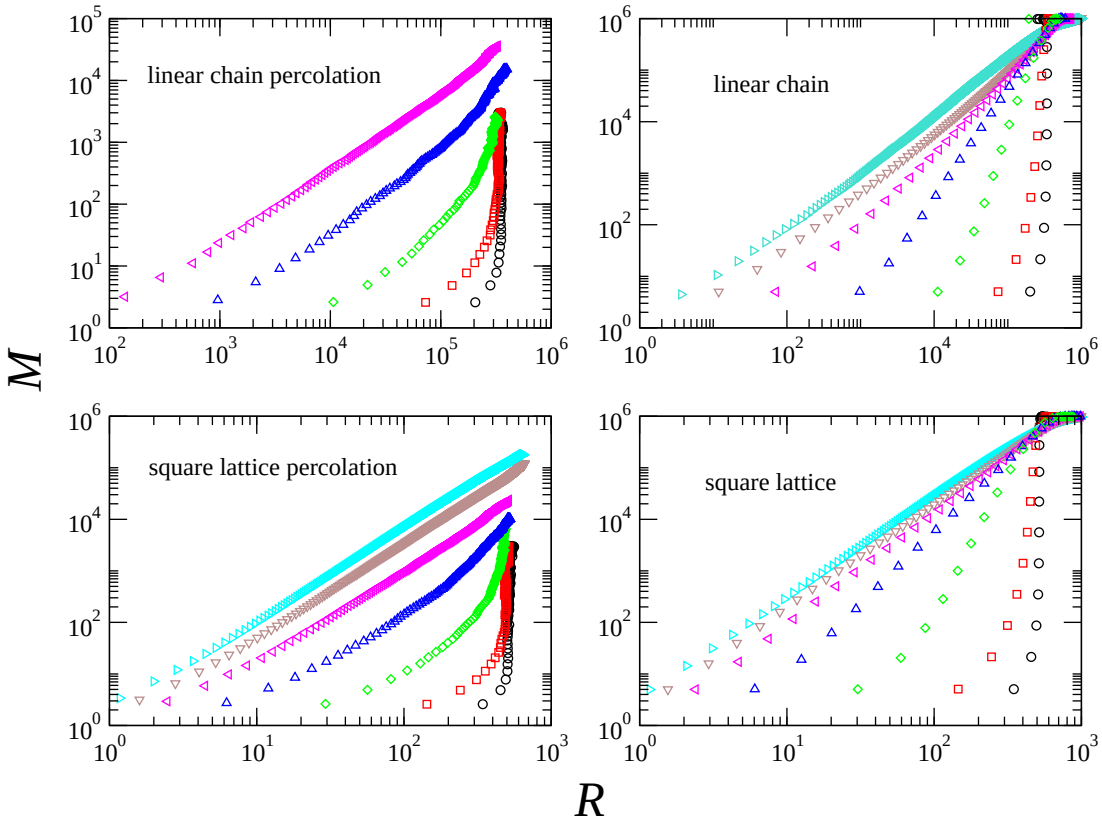

Figure 4.4: Size of the giant cluster M as a function of radius R for different values of α in percolation at the critical threshold (left 2 plots) and above criticality on the embedded networks (right 2 plots) in one-dimensional linear chain (upper 2 plots) and two-dimensional square lattice (lower 2 plots). d_f and d_e are determined from M vs R in percolation and above the criticality in the embedded networks respectively. The values of d_f and d_e are in table 4.1. For linear chain, the symbols represent different α as followings: ($\circ$) 0.75, ($\square$) 1, ($\diamond$) 1.25, ($\triangle$) 1.5, ($\triangleleft$) 1.75, (∇) 2 and ($\triangleright$) 2.25. For square lattice, the symbols represent different α as followings: ($\circ$) 1.5, ($\square$) 2, ($\diamond$) 2.5, ($\triangle$) 3, ($\triangleleft$) 3.5, (∇) 4 and ($\triangleright$) 5.

To support these results more, we apply an independent approach to analyze the cluster size density distribution $n(S)$ at the critical threshold shown in figure 4.5. At criticality, $n(S)$ is expected to scale as $n(S) \sim S^{-\tau}$,[93,94] where

$$\tau = 1 + \frac{1}{\beta}. \tag{4.7}$$

We are able to measure τ directly from the plots in figure 4.5 (slopes are $-\tau + 1$), which are complementary cumulative distributions of the cluster size S. We also have the values of β

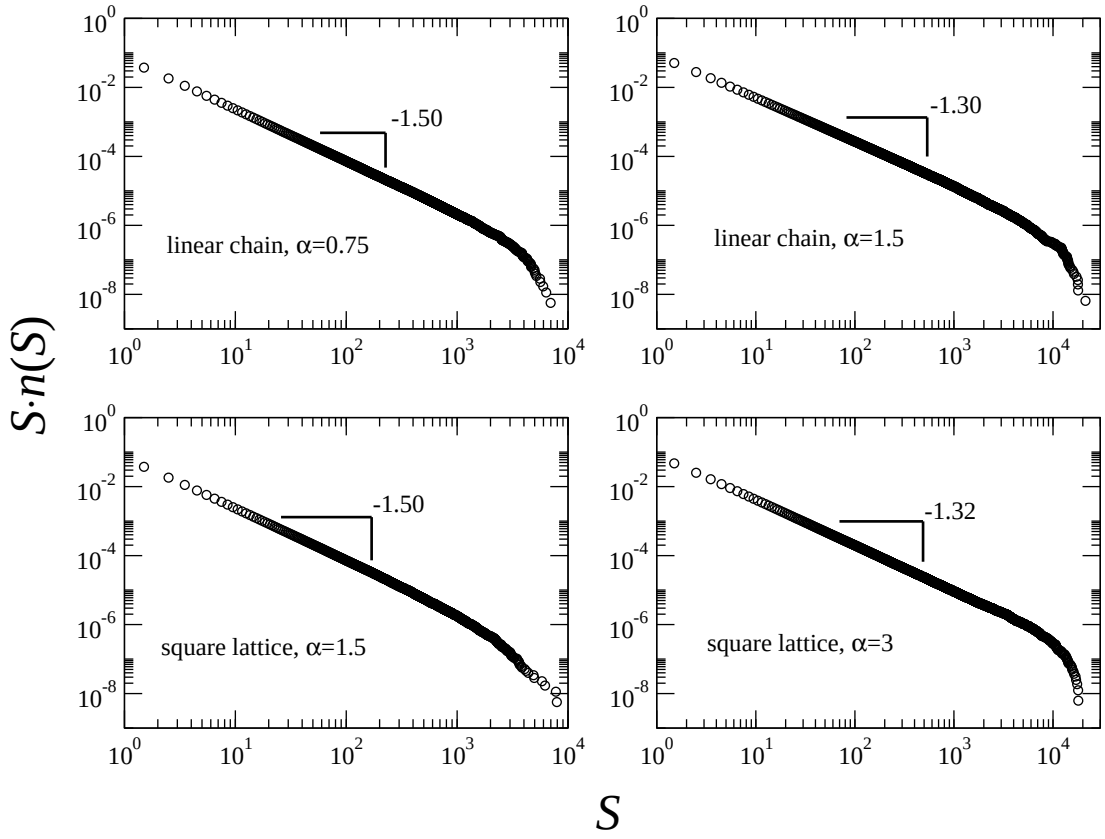

Figure 4.5: Density distribution $n(S)$ of cluster sizes multiplied by S for several typical values of α in percolation at the critical threshold in one-dimensional linear chain and two-dimensional square lattice.

for different α, therefore we can compare both values in left and right side of Eq. (4.7). Table 4.1 shows that they are in very good agreement with the values of both sides in Eq. (4.7). For linear chain and square lattice when $\alpha \leq d$, we obtain the classical mean-field value $\tau = 2.5$ known for ER networks.[93,94] And for square lattice when $\alpha > 2d$, we obtain the same exponent as for percolation in regular two-dimensional lattices, $\tau = 2.05$. For linear chain when $\alpha > 2d$, no percolation is observed. In the intermediate range ($d < \alpha < 2d$), τ varies with α, this leads to a new universality classes due to the competition between the spatial constraints and the long-range connections.

4.3 Conclusions

We have studied the percolation of spatially constrained networks embedded in one-dimensional linear chain and two-dimensional square lattice with the length of links following a power law probability $p(r) \sim r^{-\alpha}$. We have identified that there are three distinct regimes where different types of percolation transitions with different critical exponents exist. For linear chain and square lattice in the first regime when $\alpha \leq d$, the giant cluster is characterized by nodes connected with long links comparable to the system size and the transition belongs to the universality class of percolation in ER networks. And in the second intermediate regime when $d < \alpha < 2d$, the giant cluster is comprised of localized cliques which are connected by few long links and the critical exponents seem to change with α in a continuous way. Finally, in the last regime when $\alpha > 2d$, the giant cluster has only short-range connections and the transition belongs to the universality class of percolation in two-dimensional square lattice with no percolation observed in one-dimensional linear chain. We have compared the percolation exponents at the critical threshold: β vs $\frac{d_f}{d_e}$ and τ vs $1 + \frac{1}{\beta}$. The results show that they are in very good agreement with the values. And recently analogous effects have been found for long-range links on fractal networks.[95] Finally, we like to note that although our analysis has been performed on the linear chain and square lattice, for reasons of universality we expect that the results will not change for different two-dimensional lattice or continuum structures. Moreover, we expect that similar three regimes will also appear for percolation in ER networks embedded in three dimensions.

Chapter 5

Summary

We have mainly finished **THREE** projects related to transport and percolation in complex networks:

(i) The conductance distribution on weighted resistor networks.

(ii) Optimal transport with cost constraint.

(iii) Percolation of spatially constrained networks.

In the first project, we study the conductance distribution $P(\sigma)$ of the equivalent conductance for ER and SF weighted resistor networks based on the study on unweighted resistor networks by Lopéz. In our model of weighted networks, we introduce the disorder on the links with conductance $\sigma = e^{-ax}$, where a controls the strength of the disorder. For $a = 0$, the networks simply returns to the unweighted case. We find that the conductance distribution and the transport property have been significantly changed after we introduce the disorder. With disorder, most links have large resistance and the transport efficiency has been significantly decreased. $P(\sigma)$ has been characterized by percolation transition properties for different regimes. We find, both analytically and numerically, that $P(\sigma)$ for ER networks exhibits two regimes: (i) For $\sigma < e^{-ap_c}$, $P(\sigma)$ is independent of N and scales as $P(\sigma) \sim \sigma^{\langle k \rangle / a - 1}$. Here $p_c = 1/\langle k \rangle$ is the critical percolation threshold of the network and $\langle k \rangle$ is the average degree of the network. (ii) For $\sigma > e^{-ap_c}$, $P(\sigma)$ has strong N dependence and scales as $P(\sigma) \sim f(\sigma, ap_c/N^{1/3})$.

In the second project, we investigate the average shortest path length as the transport property for a regular lattice after adding additional long-range connections to improve the transport property. The model is base on Kleinberg's navigation model in which that the long-range connections follows a power law distribution: $p(r) \sim r^{-\alpha}$. The purpose is to find the optimal α so that in this condition the average shortest path length $\langle \ell \rangle$ of the whole network reaches minimum. In Kleinberg's model, each link has the same chance to receive a long-range connection. It is not unexpected that in this model, $\langle \ell \rangle$ gets minimum when $\alpha = 0$ because in this condition, the network is characterized by nodes connected with long links comparable to the system size so that the network become a small-world network with $\langle \ell \rangle \sim \ln N$, which is dramatically reduced compared to $\langle \ell \rangle \sim L$ in the original lattice without additional long-range connections. Kleinberg's navigation is in the condition without considering cost constaint. We introduce a cost to the total length of long-range connections because in real-world networks, there is always a cost to build additional connections and it is reasonable that the cost is proportional to the length of the connection. After introduce the cost, there is a trade-off between the number of long-range connections and the length of them. This leads the optimal transport no longer occurs at $\alpha = 0$. Our results show that it is shifted to $\alpha = d + 1$, established here for $d = 1, 2$ and 3 for regular lattices and d_f for fractals.

In the last project, we examine the percolation transition of spatially constrained networks embedded in one-dimensional linear chain and two-dimensional square lattice with the length of links following the same power law distribution $p(r) \sim r^{-\alpha}$ as in the second project. The difference is that in this project we build the networks on the originally empty lattices, i.e. the nodes have no local connections to their nearest neighbors. Because all the long-range connections randomly connect the nodes with appropriate distance ruled by the power law distribution, actually, the networks form an ER-like networks. We study the percolation process by removing nodes from the pre-build ER-like networks embedded in one-dimensional linear chain and two-dimensional square lattice. We find that there are three distinct regimes where different types of percolation transitions with different critical

exponents exist. For $\alpha \leq d$, the percolation belongs to the universality class of percolation in ER networks. For $\alpha > 2d$ (only for two-dimensional square lattice), the percolation belongs to the universality class of percolation in regular lattices. But for the intermediate regime $d < \alpha < 2d$, the percolation shows new behavior different from ER networks, with critical exponents that depend on α. According to the limited computer resources and the pre-build networks needed (so that we can not perform the real-time algorithm in the second project), we have not tested our results in three-dimensional lattice. However we expect that similar three regimes will also appear for percolation in the networks embedded in three dimensions.

Appendices

Appendix A: Solving Kirchhoff Equations

Consider a resistor network with N nodes and all the nodes are connected together so that there are no isolated nodes. We impose an electric potential difference between any two selected nodes A and B, i.e. $V_A = 1$ and $V_B = 0$, there must exist an electric current I flowing from node A to B. We measure I_{AB} and get the conductance between A and B,

$$\sigma_{AB} = I_{AB}/(V_A - V_B) = I_{AB}. \tag{5.1}$$

If we know the topological structure of the network and the resistance of each link r_{ij} which connecting node i and j, actually we are able to analytically solve a set of Kirchhoff linear Equations based on Kirchhoff's circuit law,[63] i.e. the sum of all the currents flowing out from a node i is equal to 0 (except node A or B):

$$\sum_{j=1}^{k} I_{ij} = 0, \tag{5.2}$$

where j sums over all the links connecting to node i and k is the the number of links connecting to node i, i.e. the degree of the node. From Ohm's law, the current form node i to node j along link ij with resistance r_{ij} is given by

$$I_{ij} = \frac{V_i - V_j}{r_{ij}}. \tag{5.3}$$

Therefore we have

$$\sum_{j=1}^{k} \frac{V_i - V_j}{r_{ij}} = I_{AB}(\delta_{iA} - \delta_{iB}), \tag{5.4}$$

where δ_{ij} is the Kronecker delta

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \tag{5.5}$$

this means the sum of all the currents flowing out from a node is equal to 0 except node A or B. Since V_A and V_B are known, there are $N-2$ unknown potential V_i, leading to $N-2$ linear equations which need to be solved. Therefore, we can use an $(N-2)\times(N-2)$ matrix equation to represent the linear equations

$$\mathbf{CV} = \mathbf{D}, \tag{5.6}$$

where $\mathbf{C}, \mathbf{V}$ and $\mathbf{D}$ are given by

$$\mathbf{C} = \begin{pmatrix} \sum_{j=1,j\neq 1}^{N} \frac{1}{r_{1j}} & -\frac{1}{r_{12}} & \cdots & -\frac{1}{r_{1N}} \\ -\frac{1}{r_{21}} & \sum_{j=2,j\neq 2}^{N} \frac{1}{r_{2j}} & \cdots & -\frac{2}{r_{1N}} \\ \vdots & \vdots & \vdots & \vdots \\ -\frac{1}{r_{N1}} & -\frac{1}{r_{N2}} & \cdots & \sum_{j=1,j\neq N}^{N} \frac{1}{r_{Nj}} \end{pmatrix}, \tag{5.7}$$

$$\mathbf{V} = \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{pmatrix}, \tag{5.8}$$

and

$$\mathbf{D} = \begin{pmatrix} \frac{V_A}{r_{1A}} + \frac{V_B}{r_{1B}} \\ \frac{V_A}{r_{2A}} + \frac{V_B}{r_{2B}} \\ \vdots \\ \frac{V_A}{r_{NA}} + \frac{V_B}{r_{NB}} \end{pmatrix}. \tag{5.9}$$

Note that none of the matrices have A or B row and $r_{ij} = r_{ji}$, thus $\mathbf{C}$ is a symmetric matrix $\mathbf{C} = \mathbf{C^T}$. And $r_{ij} = \infty$ if node i and j are not connected. This results in that matrix $\mathbf{C}$ is typically a sparse matrix and the matrix inversion is relatively fast.

Finally, we get potential of each node by

$$\mathbf{V} = \mathbf{C^{-1}D}, \tag{5.10}$$

and the conductance between node A and B

$$\sigma_{AB} = I_{AB} = \sum_{j=1}^{k} \frac{V_A - V_j}{r_{Aj}}. \tag{5.11}$$

Appendix B: Remarks "Transport and Percolation Theory of Weighted Network"

I. EQUATIONS FOR DISTRIBUTION FUNCTIONS

In Chapter 2, the following equation is derived for the random variables r, r_i, or, equivalently, for $\sigma = 1/r$, $\sigma_i = 1/r_i$. (Notations are changed a bit for convenience):

$$r = r_0 + \left[\sum_{i=1}^{q-1} \sigma_i\right]^{-1}. \tag{5.12}$$

The meaning of the notations: $r_0 = 1/\sigma_0$ is a resistivity of a link, connecting two nodes of the network. It is a predefined random variable, for which a concrete distribution function is chosen. $r = 1/\sigma$ and $r_i = 1/\sigma_i$ are resistivities of the branches of some random Bethe lattice. All of them have the same distribution function $p_R(r)$, which connects with the distribution of conductivities p_Σ as $p_R(r)dr = p_\Sigma(\sigma)d\sigma$. Conductivities σ_i, $i = 1, \cdots q-1$, are *independent* random variables. And $q = 1, 2, \cdots$ is a discrete random variable, the degree of a node in our Bethe lattice, whose distribution $\Pi_1(q)$ is connected with distribution of node degree in the infinite network with independent nodes $\Pi(q)$ as $\Pi_1(q) = (q/\bar{q})\Pi(q)$, where $\bar{q}$ is the average degree in the network, $\bar{q} = \sum_q q\Pi(q)$. Replacement of the network with Bethe lattice is determined by the thermodynamic limit: in an infinite random graph with independent nodes (almost) any finite subgraph is a tree. The problem is set as follows: given the distribution of node degrees $\Pi(q)$, and the one of link resistivities, $p_R^{(0)}(r)$, to find, using relation (5.12), the distribution $p_R(r)$. If this problem is solved, one can find the distribution of conductivities $\sigma_i^{(\infty)}$, $p_\Sigma^{(\infty)}(\sigma)$, between a randomly chosen node i and a "ground", consisting of infinitely far nodes, using the relation:

$$\sigma_i^{(\infty)} = \sum_{j=1}^{q_i} \sigma_{ij}, \tag{5.13}$$

where discrete random variable q_i has the has the distribution function $\Pi(q)$, and q_i in-

dependent statistically equal variables r_{ij}, distributed with probability density $p_\Sigma(s)$, are the conductivities of q_i branches, connecting the node i with the "ground". Finally, the resistivity between two randomly chosen nodes of a graph is:

$$r_{ij} = r_i^{(\infty)} + r_j^{(\infty)} = 1/\sigma_i^{(\infty)} + 1/\sigma_j^{(\infty)}, \tag{5.14}$$

where $r_i^{(\infty)}$ and $r_j^{(\infty)}$ are two random variables, independently distributed with probability density $p_R^{(\infty)}(r)$. Eqs. (5.12-5.14) constitute a sequence, which allows to obtain the distribution function of inter-node conductivity in the random graph $P(R)$, provided, that the distributions of links' resistivities $p_R^{(0)}(r)$ and node degrees are given.

The main disadvantage of the above scheme is that the formulas (5.12-5.14) are not equations in the commonly accepted in statistical physics sense. They contain random variables, while in the statistical mechanics people usually dealing with relations between averaged quantities, such as e.g. probability densities, for example $p_R(r) = \langle \delta(r - r_i) \rangle$, where the random variable r_i is a resistance of some branch between its root site and a ground, and the average $\langle \cdots \rangle$ is the one over a chosen graphs ensemble (in our case it is an ensemble of configuration model).

It appears to be possible to obtain the closed set of equations, resulting in the distribution function of inter-node resistivities, if to use the Laplace representation for distributions of resistivities and conductivities. Assume, for example, that $R(x)$ and $S(x)$ are resistivity r and conductivity σ distribution functions of a branch, resp., i.e.:

$$R(x) = \left\langle e^{-xr_i} \right\rangle = \int_0^\infty dr e^{-xr} p_R(r), \tag{5.15}$$

$$S(x) = \left\langle e^{-x\sigma_i} \right\rangle = \int_0^\infty d\sigma e^{-x\sigma} p_\Sigma(\sigma). \tag{5.16}$$

Relation between these two functions can be established, if to use integral identity:

$$e^{-x/\alpha} = 1 - \sqrt{x} \int_0^\infty \frac{dy}{\sqrt{y}} J_1(2\sqrt{xy}) e^{-\alpha y}, \tag{5.17}$$

where J_1 is the Bessel's function. For example, if to replace α with the random variable $r_i = 1/\sigma_i$, and to average both parts, we arrive at the relation:

$$S(x) = 1 - \sqrt{x}\int_0^\infty \frac{dy}{\sqrt{y}} J_1(2\sqrt{xy}) R(y). \tag{5.18}$$

The reverse relation is exactly the same, of course. From Eq. (5.12), using the identity (5.17), it is easy to obtain:

$$\begin{aligned} R(x) = \left\langle e^{-xr_i} \right\rangle &= R_0(x) \left\langle \exp\left[-x \left(\sum_{j=1}^{\bar{q}-1} \sigma_j \right)^{-1} \right] \right\rangle \\ &= R_0(x) \left\{ 1 - \sqrt{x}\int_0^\infty \frac{dy}{\sqrt{y}} J_1(2\sqrt{xy}) \left\langle [S(y)]^{\tilde{q}-1} \right\rangle_1 \right\}, \end{aligned}$$

where the average $\langle \cdots \rangle_1$ in the latter integral is the one over the random variable $\tilde{q}$, which have the distribution function $\Pi_1(q) = \langle \delta(q - \tilde{q}) \rangle_1$. Here $R_0(x) = \langle e^{-xr_0} \rangle$ is the distribution function of link's resistivity in the Laplace representation. As a result, we have:

$$R(x) = R_0(x) \left\{ 1 - \sqrt{x}\int_0^\infty \frac{dy}{\sqrt{y}} J_1(2\sqrt{xy}) \varphi_1[S(y)] \right\}, \tag{5.19}$$

where we have introduced:

$$\begin{aligned} \varphi_1(x) = \left\langle x^{\tilde{q}-1} \right\rangle_1 &= \sum_q \Pi_1(q) x^{q-1} = \frac{1}{\langle A_{kl} \rangle} \left\langle \sum_{i,j=1}^N A_{ij} x^{q_j - 1} \right\rangle = \frac{1}{2L} \sum_{j=1}^N \left\langle q_j x^{q_j - 1} \right\rangle \\ &= \frac{1}{\bar{q}} \varphi'(x), \;\; \varphi(x) = \langle x^{q_i} \rangle = \sum_q \Pi(q) x^q. \end{aligned} \tag{5.20}$$

The meaning of φ and φ_1 are Z-representations of the distribution functions of a node degree q_i, and of a branching coefficient $q_j - 1$ of a randomly chosen link, respectively. Eqs. (5.28) and (5.19) form a closed set of ones for the function $R(x)$.

In the same manner, one can express the distribution $S_\infty(x) = \left\langle \exp(-x\sigma_i^{(\infty)}) \right\rangle$ of the conductivities $\sigma_i^{(\infty)}$ between a node i and the "ground" through $S(x)$, using the relation (5.13) as:

$$S_\infty(x) = \varphi[S(x)],$$

and the corresponding distribution of resistivities $R_\infty(x) = \left\langle \exp\left(-xr_i^{(\infty)}\right)\right\rangle$ as:

$$R_\infty(x) = 1 - \sqrt{x}\int_0^\infty \frac{dy}{\sqrt{y}} J_1(2\sqrt{xy})\varphi[S(y)]. \tag{5.21}$$

Finally, the distribution function $R_2(x) = \langle \exp(-xr_{ij})\rangle$ of the resistivity r_{ij} between randomly chosen pair of nodes i and j may be written, using Eq. (5.14) simply as:

$$R_2(x) = R_\infty^2(x). \tag{5.22}$$

In principle, it is possible to replace Eqs. (5.18), (5.19) and (5.21) with equivalent ones, using instead of Eq. (5.17) another (but closely related) integral identity:

$$e^{-x/\alpha} = \frac{\sqrt{x}}{2}\int_{-i\infty+\delta}^{+i\infty+\delta} \frac{dy}{\sqrt{x}} H_1^{(2)}(2\sqrt{xy})e^{-\alpha y}, \tag{5.23}$$

where $H_1^{(2)}$ is the second Hankel function. These alternative form of equations are more suitable, when one wants to use the saddle point approximation.

Of course, the system of equations (5.18,5.19), which one have to solve to reach the final result is a very complicated one, — it has a nonlinear integral structure. So, there is a little hope to find explicit form of $R(x)$ or of $S(x)$ for some nontrivial choice of $\Pi(q)$ and $R_0(x)$. However, we hope, that many interesting asymptotic results may be extracted. As a simple example, let us show, how the asymptotic form of $R_2(x)$ at large positive x can be found (which corresponds to the probability density of inter-node resistivity r_{ij}, $p_2(r) \equiv \langle\delta(r_{ij} - r)\rangle$, at small values of r) for power law distributions of node degrees, $\Pi(q) \to q^{-\gamma}$ at $q \to \infty$.

II. LOW-RESISTIVITY ASYMPTOTIC

It appeared to be useful to include yet another representation of degree distribution, — let us call it ψ-representation. It may be defined as an inverse Laplace transform of the

distribution in Z-representation, namely of $\varphi(1-x)$:

$$\psi(z) = \int_{-i\infty+\delta}^{+i\infty+\delta} \frac{dx}{2\pi i} e^{xz} \varphi(1-x). \tag{5.24}$$

For a wide enough class of degree distributions the function $\varphi(x)$ is analytic at $x < 0$, and is such, that $e^{\epsilon x}\varphi(x) \to 0$ as $x \to -\infty$ for any $\epsilon > 0$, which implies $\psi(z) = 0$ at $z < 0$. If the degree distribution has an exponential tail, $\Pi(q) \sim a^q$, $a < 0$, at $q \to \infty$, $\varphi(x)$ has a singularity at $x_s = 1/a$. For a fat-tailed degree distributions $\Pi(q)$, decaying slower than any exponent, $\varphi(x)$ has a singularity at $x = 1$. In particular, if $\Pi(q) \sim q^{-\gamma}$ at $q \to \infty$, its Z-transform $\varphi(x)$ has a singular part $\sim (1-x)^{\gamma-1}$ as $x \to 1$. The inverse relation to Eq. (5.24) is:

$$\varphi(x) = \int_0^\infty dz\ e^{-(1-x)z} \psi(z). \tag{5.25}$$

Taking into account the relation, which serves as an inverse of the large equality in Eq. (5.20), we have:

$$\Pi(q) = \oint_{|x|<1} \frac{dx}{2\pi i} x^{-1-q} \varphi(x) = \frac{1}{\Gamma(q+1)} \int_0^\infty dz\ z^q e^{-z} \psi(z), \tag{5.26}$$

which may serve also as an analytic continuation of $\Pi(q)$ to non-integer q. Then, the relation, inverse to Eq. (5.26) is:

$$z e^{-z} \psi(z) = \int_{-i\infty}^{+i\infty} \frac{dq}{2\pi i} z^{-q} \Gamma(q+1) \Pi(q). \tag{5.27}$$

If the function $\Pi(q)$ decays slow enough, e.g. slower, then any exponent, the integral at large z may be calculated by the saddle point method, which gives simply: $\psi(z) \approx \Pi(z)$.

Using the integral expression for $\varphi_1(x) = \varphi'(x)/\bar{q}$, $\bar{q} = \varphi'(1)$, one can rewrite Eq. (5.19) as:

$$R(x) = \frac{R_0(x)}{\bar{q}} \int_0^\infty dz\ z\psi(z) \left\{ 1 - \sqrt{x} \int_0^\infty \frac{dy}{\sqrt{y}} J_1(2\sqrt{xy}) \exp[-z(1-S(y))] \right\}. \tag{5.28}$$

Now we make two crucial assumptions: 1) The random variable σ_j (conductivity of a branch) has a finite average $\bar{\sigma}$, i.e. $S(y) = 1 - \bar{\sigma}y + o(y)$, and the main contribution to the integral on z in Eq. (5.28) gives the region $z \gg 1$. Then, replacing $1 - S(y)$ with $\bar{\sigma}y$, and using the identity (5.17), we get:

$$R(x) \approx \frac{R_0(x)}{\bar{q}} \int_0^\infty dz\ z\psi(z) \exp\left(-\frac{x}{\bar{\sigma}z}\right). \tag{5.29}$$

Indeed, one can see, that if $\psi(z)$ varies not very fast, e.g. slower than an exponent, the main contribution to the integral gives the region $z \gtrsim x/\bar{\sigma}$. Assume, that at large enough z we have: $\psi(z) \approx \Pi(z) \sim z^{-\gamma}$. Then from Eq. (5.29) we immediately obtain:

$$R(x) \sim R_0(x) \left(\frac{x}{\bar{\sigma}}\right)^{2-\gamma}. \tag{5.30}$$

Assume, for example, $R_0(x) \sim x^{-\alpha}$, $\alpha > 0$, at $x \to \infty$. Then $R(x) \sim x^{2-\gamma-\alpha}$. But we assumed (see Eq. (5.18)):

$$\bar{\sigma} = \int_0^\infty dx\ R(x) < \infty, \tag{5.31}$$

which implies $\gamma > 3-\alpha$. However, it is well known, that the graph can be treated as a *locally* tree-like, if the second moment of its degree distribution converges, which automatically implies $\bar{\sigma} < \infty$. In principle, all the calculations above were not necessary and served simply as an illustration. It is enough to rewrite the expression (5.21) for R_∞, using ψ-representation of thee degree distribution as:

$$R_\infty(x) = \int_0^\infty dz\ \psi(z) \left\{1 - \sqrt{x} \int_0^\infty \frac{dy}{\sqrt{y}} J_1(2\sqrt{xy}) \exp[-z(1 - S(y))]\right\}. \tag{5.32}$$

Again, making the replacements: $1 - S(y) \to \bar{\sigma}y$, $\psi(z) \sim z^\gamma$, we obtain $R_\infty(x) \sim x^{3-\gamma}$, and:

$$R_2(x) = R_\infty^2(x) \sim \left(\frac{x}{\bar{\sigma}}\right)^{6-2\gamma} \text{ as } x \to +\infty. \tag{5.33}$$

Because the distribution function of inter-node resistances $P(R)$ is connected with $R_2(x)$

through the Laplace transformation:

$$R_2(x) = \int_0^\infty dR\ e^{-xR} P(R), \tag{5.34}$$

our result (5.33) is equivalent to:

$$P(R) \sim (\bar{\sigma} R)^{2\gamma - 7} \text{ as } R \to 0. \tag{5.35}$$

This result is almost trivial, and may be obtained simply by the replacement in Eq. (5.13):

$$\sum_{j=1}^{q_i} \sigma_{ij} \to \bar{\sigma} q_i, \tag{5.36}$$

which is justified for large q_i. Then we arrive at the result (5.35) in a quite straightforward way.

List of Journal Abbreviations

Phys. Rev. Lett.	Physics Review Letter
Rev. Mod. Phys.	Reviews of Modern Physics
Proc. Natl. Acad. Sci.	Proceedings of the National Academy of Sciences
SIAM Rev.	Society for Industrial and Applied Mathematics Review
Phys. Rep.	Physics Reports
Publ. Math.	Publicationes Mathematicae
Amer. J. Sociol.	American Journal of Sociology
Phys. Rev. E	Physics Review E
Europhys. Lett.	Europhysics Letters
Eur. Phys. J. B	The European Physical Journal B

Bibliography

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

[2] D. J. Watts, *Small Worlds* (Princeton University Press, Princeton, NJ, 1999).

[3] D. J. Watts, Networks, dynamics, and the small world phenomenon, Amer. J. Sociol., **105** 493 (1999).

[4] R. Cohen and S. Havlin, Phys. Rev. Lett. **90**, 058701 (2000).

[5] B. Bollobás, *Random Graphs* (Cambridge University Press, Cambridge, 2001).

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

[7] M. E. J. Newman, D. J. Watts and S. H. Strogatz, Proc. Natl. Acad. Sci. U.S.A. **99**, 2566 (2002).

[8] R. Cohen, S. Havlin and D. ben-Avraham, in *Handbook of Graphs and Networks*, edited by Bornholdt S. and Schuster H. G. (Wiley-VCH) 2002, Chapt. 4.

[9] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[10] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, 2003).

[11] R. Cohen and S. Havlin, Phys. Rev. Lett. **90**, 058701 (2003).

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

[13] L. K. Gallos *et al.*, Phys. Rev. Lett. **94** 188701 (2005).

[14] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[15] D. Brockmann *et al.*, Nature **439**, 462 (2006).

[16] S. N. Dorogovtsev, A. V. Goltsev and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).

[17] A. Barrat, M. Barthélemy and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, 2008).

[18] Y. Hu, Y. Fan and Z. Di, arXiv: 0902.3329 (2009).

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

[20] R. Cohen and S. Havlin, *Complex networks: Structure, Robustness and Function* (Cambridge University Press, 2010).

[21] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, 2003)

[22] P. Erdős and A. Rényi, Publ. Math. **6**, 290 (1959); Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[23] B. Bollobas, *Random Graphs* (Academic, London, 1985).

[24] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: a statistical physics approach* (Cambridge University Press, 2006).

[25] A.-L. Barabasi and R. Albert, Science **286**, 509 (1999).

[26] J. Kleinberg, Cornell Computer Science Technical Report 99 (unpublished, 1999).

[27] J. Kleinberg, Nature **406**, 845 (2000).

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

[29] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).

[30] R. Albert, H. Jeong and A.-L.Barabási, Nature(London) **406**, 378 (2000); **409**, 542(E) (2001).

[31] R. Kumar, P. Raghavan, S. Rajalopagan and A. Tomkins, *Proceedings of the 9th ACM Symposium on Principles of Database Systems*, p.1 (1999).

[32] S. Milgram, The Small World Problem, Psychology Today **1** (1), 60 (1967).

[33] J. Travers and S. Milgram, An Experimental Study of the Small World Problem, Sociometry **32**, 425 (1969).

[34] A. Barrat and M. Weigt, Eur. Phys. J. B **13**, 547 (2000).

[35] E. López, S. V. Buldyrev, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **94**, 248701 (2005).

[36] A. Fronczak, P. Fronczak and J. A. Hołyst, Phys. Rev. E **70**, 056110 (2004).

[37] L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **91**, 168701 (2003).

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

[39] G. Li, L. A. Braunstein, S. V. Buldyrev, S. Havlin and H. E. Stanley, Phys. Rev. E **75**, 045103(R) (2007).

[40] A periodical boundary condition is applied on the square lattice, this ensures that the nodes on the boundary still have four nearest neighbors. The length of each link is 1, we call such type of link as short-range connection.

[41] J. M. Kleinberg, Nature (London) **406**, 845 (2000); *in Proceedings of the 32nd ACM Symposium on Theory of Computing* (ACM, New York, 2000), pp. 163-170.

[42] S.-H. Yook, H. Jeong and A.-L.Barabási, Proc. Natl. Acad. Sci. **99**, 13382 (2002).

[43] G. Bianconi, P. Pinb and M. Marsilia, Proc. Natl. Acad. Sci. **106**, 11433 (2009).

[44] D. Balcan ,V. Colizza , B. Goncalves, H. Hu, J. R. Ramasco and A. Vespignani, Proc. Natl. Acad. Sci. **106**, 21484 (2009).

[45] H. Hu, S. Myers, V. Colizza and A. Vespignani, Proc. Natl. Acad. Sci. **106**, 1318 (2009).

[46] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan and A. Tomkins, Proc. Natl. Acad. Sci. **102**, 11623 (2005).

[47] R. Lambiotte, *et al.* Physica A **387**, 5317 (2008).

[48] Y. Hu, Y. Wang, D. Li, S. Havlin and Zengru Di, Phys. Rev. Lett. **106**, 108701 (2011).

[49] D. J. B. Soares, C. Tsallis, A. M. Mariz and L. R. da Silva, Europhys. Lett. **70**, 70 (2005).

[50] X. Xu, X. Zhang and J. F. F. Mendes, Phys. Rev. E **76**, 056109 (2007).

[51] K. Kosmidis, S. Havlin and A. Bunde, Europhys. Lett. **82**, 48005 (2008).

[52] M. Medo, Physica A **360**, 617 (2006).

[53] M. Kaiser and C. Hilgetag, Phys. Rev. E **69**, 036103 (2004); F. Nisbach and M. Kaiser, Eur. Phys. J. B **58**, 185 (2007).

[54] M. Boguna *et al.*, Phys. Rev. E **70**, 056122 (2004).

[55] M. A. Serrano *et al.*, Phys. Rev. Lett. **100**, 078701 (2008).

[56] A. F. Rozenfeld, R. Cohen, D. ben-Avraham and S. Havlin, Phys. Rev. Lett. **89**, 218701 (2002).

[57] A. Bunde and S. Havlin, *Fractals and Disordered Systems* (Springer-Verlag, Heidelberg, 1995).

[58] Y. M. Strelniker *et al.*, Phys. Rev. E **69**, 065105(R) (2004).

[59] Y. Chen *et al.*, Phys. Rev. Lett. **96**, 068702 (2006).

[60] M. Cieplak *et al.*, Phys. Rev. Lett. **72**, 2320 (1994); **76**, 3754 (1996).

[61] M. Porto *et al.*, Phys. Rev. E **60**, R2448 (1999).

[62] M. Molloy and B. Reed, Random Structures and Algorithms **6**, 161 (1995); Combin. Probab. Comput. **7**, 295 (1998).

[63] G. Kirchhoff, Ann. Phys. Chem. **72**, 497 (1847); N. Balabanian, *Electric Circuits* (McGraw-Hill, New York, 1994).

[64] G. Grimmett and H. Kesten, *Random electrical networks on complete graphs II: Proofs*, 1983 (http://arxiv.org/abs/math.PR/0107068).

[65] L. A. Braunstein *et al.*, in *Lecture Notes in Physics: Proceedings of the 23rd CNLS Conference, "Complex Networks," Santa Fe 2003*, edited by E. Ben-Naim, H. Frauenfelder, and Z. Toroczkai (Springer, Berlin, 2004).

[66] T. Kalisky *et al.*, Phys. Rev. E **72**, 025102(R) (2005).

[67] Z. Wu *et al.*, Phys. Rev. E **71**, 045101(R) (2005).

[68] T. E. Harris, *The Theory of Branching Processes* (Dover Publication Inc., New York, 1989).

[69] S. Sreenivasan *et al.*, Phys. Rev. E **70**, 046133 (2004).

[70] G. Caldarelli and A. Vespignani, *Large Scale Structure and Dynamics of Complex Webs* (World Scientific, 2007).

[71] M. Barthélemy, Physics Reports **499**, 1-101 (2010).

[72] P. Expert, T. S. Evans, V. D. Blondel, and R. Lambiotte, Proc. Natl. Acad. Sci. USA **108**, 7663 (2010).

[73] D. Li *et al.*, Nature Physics **7**, 481-484 (2011).

[74] W. Li *et al.*, Phys. Rev. Lett. **108**, 228702 (2012).

[75] A. Bashan *et al.*, The extreme vulnerability of interdependent spatially embedded networks, arXiv:1206.2062v1.

[76] M. P. Viana and L. da F. Costa, Phys. Lett. A **375**, 1626-1629 (2011).

[77] P. A. Morais, J. S. Andrade, E. M. Nascimento, and M. L. Lyra, Phys. Rev. E **84**, 041110 (2011).

[78] H.-X. Yang, W.-X. Wang, Y.-B. Xie, Y.-C. Lai, and B.-H. Wang, Phys. Rev. E **83**, 016102 (2011).

[79] L. K. Gallos, H. A. Makse, and M. Sigman, Proc. Natl. Acad. Sci. USA **109**, 2825 (2011).

[80] L. K. Gallos, H. A. Makse, M. Sigman, Front. Physiol. **3**, 123 (2012).

[81] M.R. Roberson and D. ben-Avraham, Phys. Rev. E **74**, 017101 (2006).

[82] G. Li, S. D. S. Reis, A. A. Moreira, S. Havlin, H. E. Stanley, and J. S. Andrade, Phys. Rev. Lett. **104**, 018701 (2010).

[83] K. Komidis, S. Havlin and A. Bunde, Europhys. Lett. **82**, 48005 (2008).

[84] H. Yang, Y. Nie, A. Zeng, Y. Fan, Y. Hu, and Z. Di, Europhys. Lett. **89** 58002 (2010).

[85] D. Li, G. Li, K. Kosmidis, H. E. Stanley, A. Bunde, and S. Havlin, Europhys. Lett. **93**, 68004 (2011).

[86] S. D. S. Reis, A. A. Moreira, J. S. Andrade, Phys. Rev. E **85**, 041112 (2012).

[87] T. Emmerich, A. Bunde, S. Havlin, G. Li and D. Li, Complex networks embedded in space: Dimension and scaling relations between mass, topological distance and Euclidean distance, arXiv:1206.5710

[88] Y. Li, D. Zhou, Y. Hu, J. Zhang and Z. Di, Europhys. Lett. **92**, 58002 (2010).

[89] H. Youn, M. T. Gastner, and H. Jeong, Phys. Rev. Lett. **101**, 128701 (2008).

[90] Here, we show the case $A = 1$ but we obtained similar results for several values in the range $0 < A < 1$.

[91] B. B. Mandelbrot, *Fractals: Form, Chance and Dimension* (Freeman, San Francisco, 1977).

[92] M. Barthélemy and L. A. N. Amaral, Phys. Rev. Lett. **82**, 15 (1999).

[93] A. Bunde and S. Havlin (Editors), *Fractals and Disordered Systems* (Springer, 1991).

[94] D. ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Media* (Cambridge University Press, 2000)

[95] D. H. Rozenfeld, C. Song and A. H. Makse, Phys. Rev. Lett. **104**, 025701 (2010).

Curriculum Vitae

Guanliang Li

Education:

Ph. D., Physics, Boston University 2012.

B.S., Modern Physics, University of Science and Technology of China, 2003.

Publications:

1. G. Li, L. A. Braunstein, S. V. Buldyrev, S. Havlin and H. E. Stanley, Transport and percolation theory in weighted networks, Phys. Rev. E **75**, 045103(R) (2007).
2. G. Li, S. D. S. Reis, A. A. Moreira, S. Havlin, H. E. Stanley and J. S. Andrade, Jr., Towards Design Principles for Optimal Transport Networks, Phys. Rev. Lett. **104**, 018701 (2010).
3. D. Li, G. Li, K. Kosmidis, H. E. Stanley, A. Bunde and S. Havlin, Percolation of spatially constrained networks, Europhys. Lett. **93**, 68004 (2011).
4. T. Emmerich, A. Bunde, S. Havlin, G. Li and D. Li, Complex networks embedded in space: Dimension and scaling relations between mass, topological distance and Euclidean distance, arXiv:1206.5710 (2012).
5. G. Li, S. D. S. Reis, A. A. Moreira, S. Havlin, H. E. Stanley and J. S. Andrade, Jr., Optimal transport on spatially embedded networks, Submitted to Phys. Rev. E.

Referee service:

Europhysics Letters.

Conference:

LEET '11, 8th USENIX Symposium on Networked Systems Design and Implementation, Boston, MA, *"Andbot: Towards Advanced Mobile Botnets"*, Mar 29, 2011.