Saulo-Davi Soares e Reis

Local and global effects on navigation in small-world networks and explosive percolation

Fortaleza - CE

Nov / 2012

Saulo-Davi Soares e Reis

Local and global effects on navigation in small-world networks and explosive percolation

Tese apresentada ao Curso de Pós-Graduação em Física da Universidade Federal do Ceará como parte dos requisitos para obtenção do título de doutor em Física.

Orientador:
Prof. José Soares de Andrade Júnior.

DEPARTAMENTO DE FÍSICA
CENTRO DE CIÊNCIAS
UNIVERSIDADE FEDERAL DO CEARÁ

Fortaleza - CE

Nov / 2012

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R312l Reis, Saulo-Davi Soares e.
Local and global effects on navigation in small-world networks and explosive percolation / Saulo-Davi Soares e Reis. – 2012.
86 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Física , Fortaleza, 2012.
Orientação: Prof. Dr. José Soares de Andrade Júnior.

1. Sistemas complexos. 2. Navegação em redes. 3. Percolação explosiva. 4. Complex systems. 5. Navigation in networks. I. Título.

CDD 530

Tese de doutorado sob o título *Local and global effects on navigation in small-world networks and explosive percolation*, apresentada por *Saulo-Davi Soares e Reis* e aprovada em 23 de Novembro de 2012, pela a banca de examinadores:

Prof. Dr. José Soares de Andrade Júnior
Departamento de Física - UFC (orientador)

Prof. Dr. André Auto Moreira
Departamento de Física - UFC

Prof. Dr. Humberto de Andrade Carmona
Departamento de Física - UFC

Prof. Dr. André de Pinho Vieira
Instituto de Física - Departamento de Física Geral - USP

Dr. Nuno Miguel Azevedo Machado de Araújo
Institut f. Baustoffe - ETH Zürich

To my family and friends,
without them, there is no life.

Acknowledgement

I would like to thank everyone who contributed in some way to this thesis, in particular:

– To my family, first;

– To Prof. José Soares de Andrade Jr., for the precious supervision, encouragement and friendship;

– To Prof. André Auto Moreira, for the enlightened discussions and friendship;

– To Prof. Hans Jürgen Herrmann, for the invitation to visit his group at the Federal Institute of Technology Zürich (ETHz), and the warm reception I found at the CompPhys;

– To Dr. Nuno Araújo, for the helpful discussions and friendship;

– To the members of the group led by Prof. Soares and all their friendship;

– To my personal friends and all their meaning in my life;

– To those Professors with whom I learned a lot, inside and outside the classroom;

– To the Department of Physics (UFC) and its staff;

– To CNPq for the financial support.

When the flush of a new-born sun fell first on Eden's green and gold,
Our father Adam sat under the Tree and scratched with a stick in the mould;
And the first rude sketch that the world had seen was joy to his mighty heart,
Till the Devil whispered behind the leaves,"It's pretty, but is it Art ?"

Rudyard Kipling

Resumo

Um número significativo de redes reais possui características locais ou não-locais bem definidas. Nós estudamos como estas características podem influenciar processos de navegação e processos percolativos que venham a ocorrer nas mesmas. Primeiramente, estudamos o problema de navegação em redes regulares com ligações de longo alcance e sujeitas a um vínculo de custo. Neste caso, a rede é construída a partir de uma rede regular de dimensão d a ser melhorada por meio da adição de ligações de longo alcance (atalhos) com uma probabilidade $P_{ij} \sim r_{ij}^{-\alpha}$, onde r_{ij} é a distância de Manhattan entre os sítios i e j. Mostramos que a condição de navegação ótima, $\alpha = d+1$, permanece ótima, independente da estratégia de navegação utilizada, seja ela baseada em um conhecimento local ou global da estrutura da rede. Em seguida, apresentamos um processo de crescimento de agregados que fornece uma clara conexão entre a Mecânica Estatística no equilíbrio e o processo percolativo não-local conhecido como Percolação Explosiva. Mostramos que dois ingredientes são suficientes para obter uma transição abrupta na fração do sistema ocupada pelo maior agregado: (i) os tamanhos de todos os agregados devem ser mantidos aproximadamente iguais durante o processo percolativo e (ii) a inclusão de ligações de *fusão* (i.e., ligações que conectam agregados diferentes) deve dominar o processo em detrimento de ligações redundantes (i.e., ligações que conectam sítios em um mesmo agregado). Por último, introduzimos um modelo que generaliza a regra do produto para Percolação Explosiva que revela os efeitos da não-localidade no comportamento crítico do processo de percolação. Mais precisamente, pares de ligações não ocupadas são escolhidos de acordo com uma probabilidade que decai em lei de potência com sua distância de Manhattan, e apenas a ligação que conecta agregados para os quais o produto de seus tamanho é o menor, é ocupada. Nossos resultados para redes regulares finitas em diversas dimensões sugerem que, na criticalidade, o expoente da lei de potência tem uma influência significativa nos expoentes de escala, onde observa-se uma transição nos expoentes da percolação tradicional para os expoentes da percolação explosiva (não-local) em determinados casos.

Abstract

A significant number of real networks have well-defined local and nonlocal features. We investigate the influence of these features in the navigation through small-world networks and in explosive percolation. First, we investigate the navigation problem in lattices with long-range connections and subject to a cost constraint. Our network is built from a regular d-dimensional lattice to be improved by adding long-range connections (shortcuts) with probability $P_{ij} \sim r_{ij}^{-\alpha}$, where r_{ij} is the Manhattan distance between nodes i and j, and α is a variable exponent. We find optimal transport in the system for $\alpha = d+1$. Remarkably, this condition remains optimal, regardless of the strategy used for navigation being based on local or global knowledge of the network structure. Second, we present a cluster growth process that provides a clear connection between equilibrium statistical mechanics and the nonlocal explosive percolation process. We show that the following two ingredients are sufficient for obtaining an abrupt transition in the fraction of the system occupied by the largest cluster: (i) the size of all growing clusters should be kept approximately the same, and (ii) the inclusion of merging bonds (i.e., bonds connecting nodes in different clusters) should dominate with respect to the redundant bonds (i.e., bonds connecting nodes in the same cluster). Finally, we introduce a generalization of the product rule for explosive percolation that reveals the effect of nonlocality on the critical behavior of the percolation process. Precisely, pairs of unoccupied bonds are chosen according to a probability that decays as a power law of their Manhattan distance, and only that bond connecting clusters whose product of their sizes is the smallest becomes occupied. Our results for d-dimensional lattices at criticality shows that the power law exponent of the product rule has a significant influence on the finite-size scaling exponents for the spanning cluster, the conducting backbone, and the cutting bonds of the system. For all these types of clusters, we observe a clear transition from ordinary to (nonlocal) explosive percolation.

Contents

List of Figures

List of Tables

1 Introduction

Networks are systems composed by nodes, sometimes called vertices or sites, which have a connection or link between them. Therefore, it is natural to represent *complex systems* usually composed by a large number of constituents, which typically interact between them, by a network. In this case, the nodes play the role of the constituents, and the links represent their interaction. These systems are complex because certain features of their global behavior emerge nontrivialy from their microscopic dynamics, in other words, the whole system has properties that are not verified when one analyzes its constituents individually.

Real networks are found in several branches of science, such as in Computer Science, with the World Wide Web (WWW) [1, 2], in Information Technology with the Internet [3], the cellular network of metabolic process [4] in Biology, and in Social Sciences in networks of acquaintances of individuals in social groups [5, 6]. In these systems typically one finds a network of nontrivial interactions, however the study of real network data reveals that there are common properties between presumably distinct networks [7, 8]. Therefore, the study of the topological features of these networks is important to the understanding of complex systems [9].

In this thesis, we first study the effects of locality and nonlocality on the problem of navigation in spatially embedded small-world networks, following the construction of Kleinberg [10, 11], where shortcuts between two nodes are added to a regular lattice according to the probability, $P(r_{i,j}) \sim r_{i,j}^{-\alpha}$, where i and j are distinct nodes of the regular lattice. After that, we show that the nonlocal features of the explosive percolation is the main key to postpone the percolation threshold, and consequently, obtain an abrupt transition in a percolation process. Finally, we make use of a power law probability to generalize the product rule (PR) of explosive percolation, introduced by Achlioptas, D'Souza, and Spencer [12], in order to study the effect of nonlocal features in the scaling exponents of such percolation process.

This thesis is organized as follows: In Chapter 2, we present briefly some concepts of network theory, such as the degree and the degree distribution, the shortest path and clustering. After that, we present the Erdős-Rényi network model [13], discussing its structural evolution

and properties. At the end of this chapter, we present the small-world phenomenon explained in terms of the Milgram's experiment [14], and the paradigm of Watts and Strogatz for small-world networks [15, 16].

In Chapter 3, we present the Kleinberg's model for small-world networks and his decentralized algorithm for local navigation [10, 11]. After that we introduce a global navigation process in Kleinberg's network model [17], and show the sharp difference between global and local knowledge in the navigation process. In Chapter 4, we introduce a network navigation model with a cost constraint [18], conceptually inspired on Kleinberg's model. We show that this limitation changes substantially the navigation process in the network for the decentralized local algorithm and for global navigation.

Chapter 5 deals with percolation. In this chapter we briefly introduce the percolation model and the main features of the traditional percolation process. After that, we present the Product Rule (PR) for explosive percolation [12] and its consequences in the percolation processes taking place on the Erdős-Rényi network and on a square lattice. In Chapter 6, we introduce a Hamiltonian approach for explosive percolation [19]. We study the asymptotic behavior of the proposed Hamiltonian model and analyze the simulation results. Later, we present an equilibrium solution for the Hamiltonian model. In chapter 7, we introduce a generalized PR for explosive percolation, and present results of this generalized PR for finite d-dimensional lattice. Finally, we leave the final remarks of the thesis for Chapter 8.

2 *Basic concepts of network science*

Complex systems have been the focus of several relevant studies by the scientific community. In particular, the investigation on complex networks has been shown to be fruitful among fields traditionally distinct as Physics, Chemistry, Biology, and Ecology, with other subjects as Social Science and Economy [7, 8]. Complex systems typically have a huge number of constituents, which interact in such a way that certain phenomena emerge from the ensemble, but they are absent when we analyze each constituent individually. One of the basic approaches used in network science, inspired by the seminal work of Watts and Strogatz [20], is the comparative study of networks from different fields of science. Therefore, many network models are proposed based on the observations of properties of real systems. Basically, these networks are identified from natural and artificial systems, from the Internet to biological and social systems. Statistically, they share similar mathematical characteristics [8, 7, 9]. Academically, Social Science has a long history studying *real networks*, specifically, since the work of Moreno during the 1920's and 1930's, where he and his colleagues presented patterns of friendship relations on small groups of people. [21].

In this chapter, we present some basic concepts about complex networks. For this, we introduce firstly, concepts of graph theory, as the degree, shortest path, and clustering. Then, we present the so called Milgram's experiment and small-world phenomenon, followed by the Watts-Strogatz [20] paradigm and the Kleinberg model [10, 11].

2.1 Network theory

The first real problem to be solved by network theory, called by mathematicians as graph theory, is now regarded as the celebrated solution of the Swiss mathematician Leonhard Euler (1707-1783), at the year of 1735, to the problem of the *seven bridges of Königsberg* (nowadays, the city is known as Kaliningrad, an exclave of Russian Federation, located between Poland and Lithuania on the Baltic Sea). Figure 2.1(a) shows the structure of all possible paths in Königsberg at Euler's time. The nodes represent separated land masses, and the links the

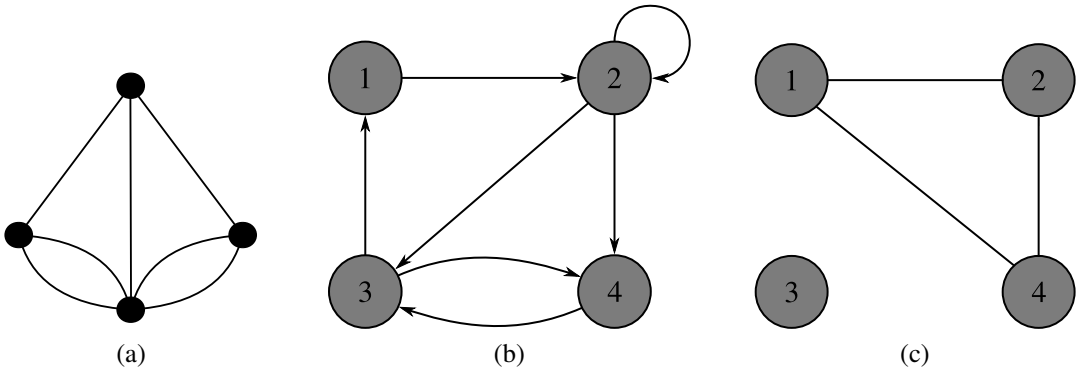


Figure 2.1: (a) Euler's network for the Königsberg bridge problem, (b) directed and (c) undirected networks. (a) The links represent the bridges of the old Königsberg's city connecting four different land masses, represented by the nodes. The nodes are the Kneiphof island and the banks on the Pregel River. (b) A directed network $G = (V,E)$, as defined in Subsec. 2.1.1, where $V = \{1,2,3,4\}$ and $E = \{(1,2),(2,2),(2,3),(2,4),(3,1),(3,4),(4,3)\}$. The link $(2,2)$ is a loop, in this case an 1-loop. (c) An undirected network $G = (V,E)$, where $V = \{1,2,3,4\}$ and $E = \{(1,2),(1,4),(2,4)\}$. The node 3 is an isolated node. The Euler's network is an example of an undirected network, since there is no preferential way to cross any bridge.

bridges connecting these pieces of land in old Köningsberg. Euler presented the answer to the following question: Is it possible to walk through Köningsberg crossing each bridge only once? Euler proved that such a walk was impossible to the pedestrians[1].

In this Section, we present some properties that characterize the structure of a network, and their definitions on the framework of network theory.

2.1.1 Definitions and notations

Mathematically, we can represent a network, complex or not, by a graph. A graph is a pair of sets $G = (V,E)$ [22], where V is finite set of N nodes, and E is a binary relation of V, which determines the links between two nodes of V. Each link is represented by a pair of nodes (i, j), and it can be classified in two types of links, a *directed* link, or an *undirected* link.

In directed networks, at least some fraction of the links are directed, in other words, the network has at least a link that has a specific direction, from one node i to other node j. The network of citations of scientific papers is an example of a directed network [23], where the direction of the links is well defined, namely they are directed from newer papers towards older ones. Figure 2.1(b) is a pictorial representation of a directed network G, with a set of nodes $V = \{1,2,3,4\}$. The circles represent the nodes, and the arrows the links. Directed networks

[1]Many years later, in 1873, Carl Hierholzer proved that such kind of walk exists if, and only if, each node has an even number of links.

may have loops, meaning links that begin and end at the same node.

An undirected network is represented by the set G, as a directed network, however the set of links E is an unordered set, meaning that $(i, j) = (j, i)$, as depicted in Fig. 2.1(c), which represents an example of undirected network. In undirected networks $i \neq j$, which prohibits the presence of self-loops, implying that the link (i, j) consists of a pair of distinct nodes.

2.1.2 Degree and degree distribution

The degree of a node i on an undirected network is the number of adjacent nodes to the node i [8], in other words, the degree k_i of i is equal to its number of neighbors, or its number of links. If we are dealing with a directed network, the degree of a node is composed of two distinct portions: k_i^{out} which is the number out-going links, and the number k_i^{in} of in-going links to the node i. Then, the net degree of node i is

$$k_i = k_i^{in} + k_i^{out}. \tag{2.1}$$

One can define the fraction p_k of nodes with degree k, hence, p_k is the probability to choose, at random, a node with degree k. The graphical representation of p_k can be made by a histogram which is known as degree distribution $P(k)$. The degree distribution is the simplest statistical signature of a network, and is only the first step to its analysis, but generally, the degree distribution is significant for the understanding of a network and the processes taking place on it, additionally, its low- and high-degree parts are important for different network properties and functions.

Usually, in classical random graph theory, the degree distribution has a fast decay form for large k, e.g., $P(k) \sim 1/k!$. As a consequence, all the moments of k, $\sum_k k^n P(k)$ are finite, even when reaching the thermodynamic limit, i.e., $N \to \infty$. In this situation, the average degree, $\langle k \rangle = \sum_N kP(k)$ is a characteristic scale of the network. In contrast, for many real networks, the degree distribution typically decays slowly, with the presence of hubs (highly connected nodes). The power law, $P(k) \sim k^{-\gamma}$, is an example of slow decay distribution, usually found in real networks. Since the form $k^{-\gamma}$ allows a rescaling of k by a constant, which only has the effect as a multiplicative factor, $(ck)^{-\gamma} = c^{-\gamma}k^{-\gamma}$, the power law distributions are called *scale-free* distributions, and networks that present a power law form on its degree distribution are then called *"scale-free networks"*. For a scale-free distribution, the value of the moment $\langle k^n \rangle = \sum_k k^n P(k)$ is bounded by the upper limit of the sum. Therefore, for an infinite network, if $\gamma \leq n+1$, the nth and higher moments of its distribution diverge.

2.1.3 Shortest path and clustering

A *path* from a node i to another node i' in a network defined by a graph $G = (V,E)$ is a sequence of links $\{(i_0,i_1),(i_1,i_2),\ldots,(i_{l-1},i_l)\}$, where $i = i_0$, $i' = i_l$, and $i_m \in V$ for $m = 0,1,2,\ldots,l$ [22]. Therefore, the path length $l_{i,i'} = l$ is the number of links that connects the nodes i and i' throughout this path. In this way, one can define the shortest path length $\ell_{i,i'}$ between two nodes of a network as the sequence of links that minimizes the path length $l_{i,i'}$. As a consequence, the *average shortest path length*, $\langle \ell \rangle$, of the entire network is defined as [24]

$$\langle \ell \rangle = \frac{1}{N(N-1)} \sum_{i,j \in G; i \neq j} \ell_{ij}. \tag{2.2}$$

The definition 2.2 is not convenient for networks that are composed of more than one component, because of the existence of nodes belonging to different clusters, for which a path length is undefined. Usually, we consider that the shortest path between this kind of pair of nodes diverges, which leads to a divergent average shortest path length. To avoid this problem, it is common to neglect pairs of nodes placed in disconnected clusters, in other words, only pairs of nodes that have paths between them are considered to compute $\langle \ell \rangle$. Figure 2.2(a) explains the notion.

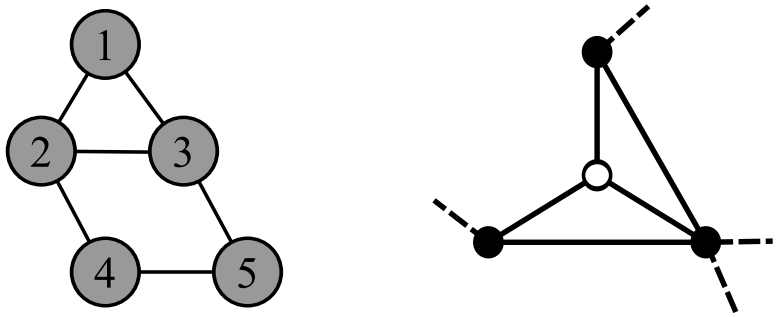
Another important topological measure to characterize networks is the *clustering*[2], which is defined as the probability of three nodes being pairwise connected. On social networks, we can say that the clustering is the probability for two distinct friends of a person to be also friends of each other. In network theory, the clustering probability measures the probability to find a triangle in the network, in other words, the existence of sets of three nodes, where each node is connected to the other two. One can quantify the clustering by the clustering coefficient C, following Watts and Strogatz [20], which can be computed using the *local clustering coefficient* (see Fig. 2.2(b)) defined as,

$$c_i = \frac{2e_i}{k_i(k_i-1)}. \tag{2.3}$$

The quantity e_i is the number of links between the neighbors of i, and $k_i(k_i-1)/2$ represents all possible links between the neighbors of i. The clustering coefficient is the mean value of c_i over all nodes of the network,

$$C = \frac{1}{N} \sum_{i \in G} c_i. \tag{2.4}$$

[2]The concept of clustering was adapted from the social science, where it is called *transitivity*.



(a) (b)

Figure 2.2: (a) The sequences of links $\{(1,2),(2,3),(3,5)\}$, $\{(1,2),(2,4),(4,5)\}$, and $\{(1,3),(3,5)\}$ are all paths between nodes 1 and 5, however, only the latter is the shortest path of length 2. Here, the average shortest path of the entire network is $\langle \ell \rangle = 1.4$. (b) The local clustering coefficient of the central node is $2/3$.

2.1.4 Lattices and fractals

In finite-dimensional lattices, which are supposed to have no long-range connections, the average shortest path length, $\langle \ell \rangle$, depends on the lattice size, N, as a power law

$$\langle \ell \rangle \sim N^{1/d}, \tag{2.5}$$

where d is the dimensionality of the lattice, an integer number. In contrast, *fractals* may present non-integer dimensionalities. To a fractal, the form of the dependence with N is unchanged, meaning that, $\langle \ell \rangle \sim N^{1/d_f}$, where d_f is its fractal dimension. One can notice that there is no big difference between a lattice and a fractal, in terms of the scaling behavior of $\langle \ell \rangle$. For example, for a square lattice of 10^4 nodes, $\langle \ell \rangle \sim 10^2$. This is not true when d or d_f tends to infinity. In a finite size system, is possible to compute its dimension (or its fractal dimension), by counting down the number o nodes $n(\ell)$ within a distance ℓ from an arbitrary node. In both cases this number is $n \sim \ell^{d_f}$.

2.2 The Erdős-Rényi model

The interest of mathematicians on graphs started with the work of Paul Erdős and Alfred Rényi [13] on random networks. In it, they proposed a network model of N nodes connected by m links randomly chosen with probability p. The Erdős-Rényi (ER) model defines an ensemble of networks, $G(N,p)$, where m links exist with probability $\binom{M}{m} p^m (1-p)^{M-m}$. The quantity $M = N(N-1)/2$ is the number of possible nonredundants[3] links of a network of size N. Many

[3] A redundant link is defined as a link connecting two nodes already connected by another link.

analytical results can be obtained from the ER model in the limit of large N. In this limit, the ER network has a characteristic average degree, given by

$$\langle k \rangle = p(N-1) \simeq pN. \tag{2.6}$$

In this section, we present some characteristics of the ER model.

2.2.1 The structural evolution of the network

Many properties of the ER model depend on p. One can assume $p(N) \sim N^{-z}$, where z is a tunable variable in the range $[0, \infty)$. Several values of z are threshold to different arrangements of the ER model. For values of z larger than $3/2$, the presence of isolated clusters is dominant, however, when $z \to 1$, the network presents increasingly larger clusters. A simple analysis reveals a phase transition on the connectivity of the network, when $z = 1$, corresponding to a critical probability of $p_c = 1/N$, and a critical average degree of $\langle k \rangle_c = 1$, so that

- For $p < p_c$, the network does not present clusters larger than $O(\ln N)$, and any cluster has more than one loop.
- When $p = p_c$, a giant cluster of size $O\left(N^{2/3}\right)$ emerges; at the neighborhood of the transition, the cluster size distribution follows a power law $P(s) \sim s^{-\mu}$, with exponent $\mu = 5/2$ (or $3/2$ when we are dealing with the cluster size distribution of randomly chosen nodes)[4].
- For $p > p_c$, the network has a cluster of size $O(N)$.

This corresponds to a second-order phase transition, which belongs to the universality class of percolation on infinite dimension.

2.2.2 Degree distribution of the ER model

The degree distribution $P(k_i)$ of the ER model with connecting probability p follows a binomial distribution

$$P(k_i = k) = C_{N-1}^{k} p^k (1-p)^{N-1-k}. \tag{2.7}$$

Since the node i is allowed to have up to $N-1$ links, the probability of node i to have k_i links is p^{k_i}, and the probability of the other links to be absent is $(1-p)^{N-1-k_i}$. The factor $C_{N-1}^{k_i}$ is the

[4] It is easy to see that the probability $P'(s)$ of a node to be in a connected cluster of s nodes is proportinal to the size of the cluster s times the cluster size distribtuion $P(s)$. Therefore, $P'(s) \sim sP(s) = s^{-3/2}$.

number of ways that the k_i links can be arranged in all $N-1$ possibilities. In the thermodynamic limit, $N \to \infty$, the Poisson distribution asymptotic behavior is valid, so that

$$P(k) \simeq e^{-pN} \frac{(pN)^k}{k!} = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \tag{2.8}$$

Therefore, although there are many ways to arrange the links, the ER model is homogeneous, meaning that the majority of nodes has approximately the same degree $\langle k \rangle$.

2.2.3 Average shortest path length of the ER model

One way to characterize and study diffusion, scattering, the navigation of a property or the propagation of information on a random network, as the ER model, is computing the average shortest path length $\langle \ell \rangle$ between two arbitrary pair of nodes in the network.

For the ER model, from Eq. (2.8), one can say that each node i has, on average, $\langle k \rangle$ first neighbors, and through all first neighbor an average of $\langle k \rangle^2$ second neighbors can be reached. So, an arbitrary chosen node i will have on average $\langle k \rangle^\ell$ nodes apart by a length ℓ. If the number N of nodes of the network is sufficiently large, therewith $\ln(N) \gg \langle k \rangle$, the number of nodes apart from node i by a shortest path length of ℓ_i will grow exponentially, so the number of nodes on the ℓth shell will be $O(N)$. Equating $\langle k \rangle^\ell$ with N by the logarithm function, leads us to the following equation:

$$\langle \ell \rangle \sim \frac{\ln(N)}{\ln(\langle k \rangle)}, \tag{2.9}$$

which gives the average shortest path length of the ER model, therefore $\langle \ell \rangle$ is asymptotically dependent on the logarithm of N. In the next section, we will see that this type of behavior of the average shortest path length characterizes the so-called "small-world" phenomenon. Using Eq. (2.6), we can write $\langle \ell \rangle$ as:

$$\langle \ell \rangle \sim \frac{1}{1-z}. \tag{2.10}$$

The average shortest path length increases from $\langle \ell \rangle = 1$ with $z = 0$ (the graph is a fully connected graph with $p = 1$) to higher values when $0 < z < 1$, and eventually diverges when $z \to 1$ ($p \to p_c$), meaning that exists a tree-like spanning cluster at the critical threshold $p = p_c$. One can found negative values of $\langle \ell \rangle$ when $1 < z < \infty$, meaning that the graph is basically disconnected.

2.2.4 Clustering coefficient

Considering a node i of a random network generated using the ER model, and its first neighbors, the probability of two neighbors of node i to be connected with each other is equal to the probability of two randomly chosen nodes to be connected. Therefore, the clustering coefficient of the ER model is

$$C_{ER} = p = \frac{\langle k \rangle}{N}. \tag{2.11}$$

Thus, from Eq. (2.11), the fraction $C_{ER}/\langle k \rangle$ decreases with N^{-1}. This property reveals that the ER model is not a good choice to model real networks, which are usually large, but still present a very strong clustering [25].

2.3 The small-world phenomenon

2.3.1 The Milgram's experiment

The term "small-world" is very common in our daily vocabulary. In 1967, the psychologist Stanley Milgram, introduced the term in the scientific community [14]. Milgram proposed an experiment which consists in mailing a box between two different individuals in the United States of America (USA). Using the jargon of network science, these two people correspond to different nodes of a real social network, where the links represent the connection between then, i.e., friendship. For the experiment, one person was labeled as a "target", who lived in Boston, Massachusetts; and, at random, "source" people were chosen in Omaha, Nebraska[5] (Fig. 2.3).

The source individuals received a box with a brief summary of the experiment, together with some information about the target, namely the complete name, a photography of, and the address of the target. The basic procedure is described as follows:

Basic procedure of the Milgram's experiment

1. The participants should include their complete name and the name of the next person that would receive the box to a list.

2. They should send a postcard, which were contained in large quantity in the box, back to Milgran. This part of the procedure allowed Milgram to follow the "path" of the box.

[5] In fact, it was the second attempt of Milgram. The first one, with starting points in Wichita, Kansas, and a target person in Sharon, Massachusetts, resulted only in three finished chains.

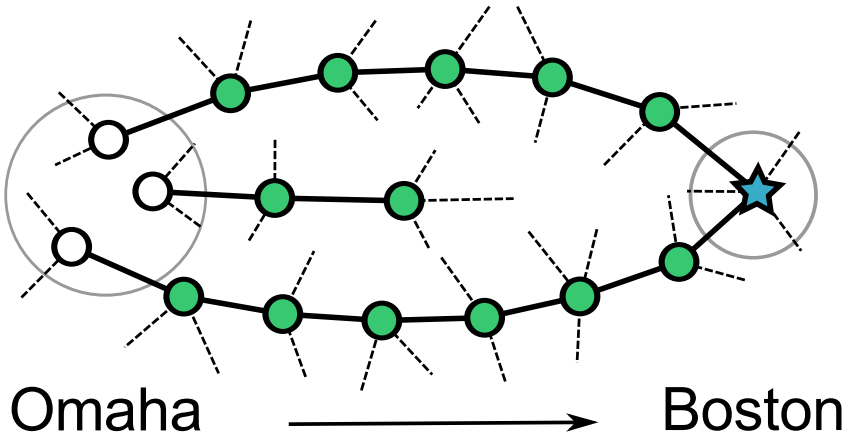


Figure 2.3: Stanley Milgram built a social network of acquaintances in USA. Source people (white circles) were asked to send a letter to a target person at Boston (blue star) through chains of acquaintances (green circles). Note that some chains were broken off.

3. If the holder of box knew the target[6], the box should be addressed directly to the target.

4. If he/she did not know the target, the holder should address the box to a known person, whom he/she believed would speed up the delivering process.

Although many boxes did not reach the target person, the fraction of boxes that approached the target indicated that it was necessary, on average, only 5.5 social links; a surprising result to Milgram. This is what is known as the "six degrees of separation". The experiment received many criticisms, since Milgram had to work with (i) poorly defined and subjective material, and (ii) poor statistics. However, the experiment has been recently reproduced by D. Watts and his collaborators [15], using e-mail messages with 18 target people from 13 countries, and 61163 source people spread across 166 countries, which resulted in 21163 distinct chains of acquaintances. Although only 384 of these chains were completed, they turned out to have a average length of 4, even less the the six degrees of separation.

2.3.2 The Watts-Strogatz model

The Milgram experiment revealed a type of network, very common in real situations, namely the *small-world network*. The first small-world network model was proposed by D. Watts and S. Strogatz [20], which incorporate two important characteristics, evidenced by the analysis of real networks, and by the Milgram experiment. First, as said before, real networks present very strong clustering, represented by a high value on the clustering coefficient C when

[6]In the experiment, by cultural reasons, to know a person means the existence of a friendship relation, characterized by treat him by the first-name.

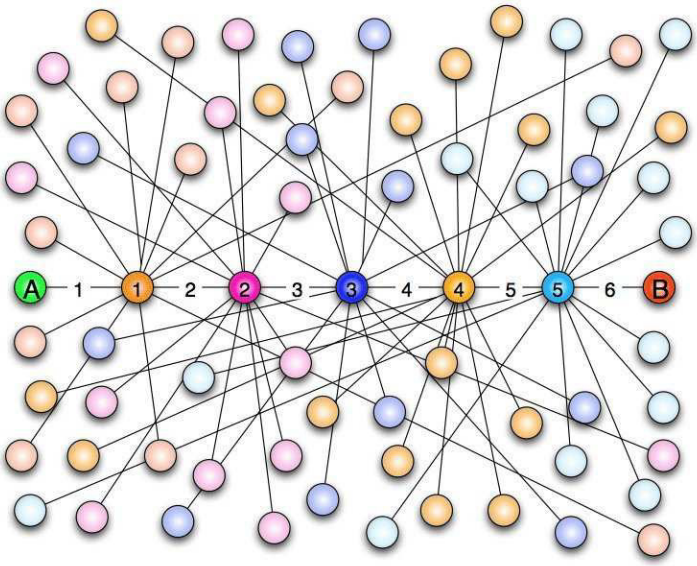


Figure 2.4: Six degrees of separation. Pictorial representation showing that for the information to be exchanged between nodes A and B, it is not necessary more than six links. Figure extracted from *wikipedia.org*.

compared with N. Second, the Milgram experiment showed that social networks have very small $\langle \ell \rangle$, even when the network is large. So, the small-world phenomenon should be characterized by a slow (logarithmically) increase of $\langle \ell \rangle$ with N.

Classical random networks display values of the clustering coefficient (Eq. 2.11) too small to model real small-world networks. Meanwhile, lattices (where the average degree is constant) present high shortest path lengths [7], what makes them unable to reproduce the six degrees of separation. However, due to its dealing with lattices, depending on their local structure, there is a high randomness, these networks present high $\langle \ell \rangle$; in counterpart, probability to find loops. In other words, depending of the lattice topology, C would present values compatible with the ones found for real networks. With this information at hand, Watts and Strogatz conjectured, and proved, that a model able to reproduce the small-world phenomenon, should have a topology in between the topology of a lattice, and the one of a completely random network.

The Watts-Strogatz (WS) model consists on adding random links to a lattice of any dimension or topology. Following the WS model, the addition of these *shortcuts* can be made by two different procedures. One can simply add them between pairs of nodes of the lattice with probability p, or by changing links of the lattice using shortcuts with probability p. The most widely studied network based on the WS paradigm is built starting from an one-dimensional network with periodic boundary conditions with N nodes (basically, a ring, or a loop of length $N-1$)

[7] An example of regular lattice is the square lattice, very used to model magnetic systems, where $\langle \ell \rangle \sim \sqrt{N}$.

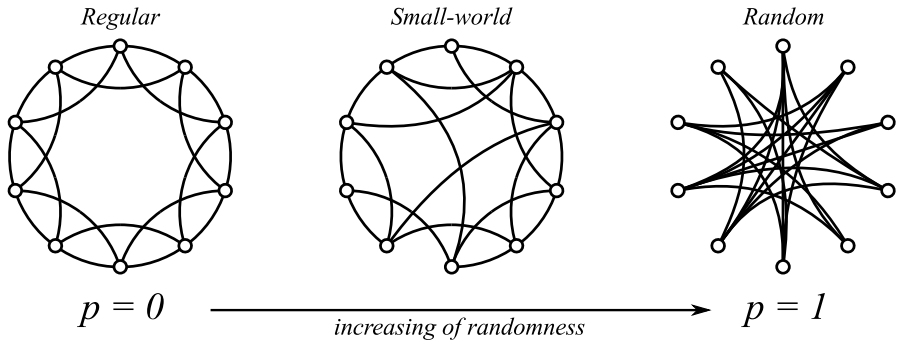


Figure 2.5: Rewiring of regular links. The number of links is unchanged. We started with a regular ring with $N = 10$ and $\langle k \rangle = 4$. For $p = 0$, the original network remains, but with the increase of p, there is an increasing of the disorder, and when $p = 1$, it becomes a random network.

connected with their K first neighbors to the left and to the right, leading to $\langle k \rangle = 2K$. Starting from this circular networks, one performs the rewiring process[8], as depicted on Fig. (2.5). On the WS model self-loops and redundant links are forbidden.

To better understand the effect of the addition of shortcuts, one could analyze the behavior of the clustering coefficient, $C(p)$, and of the average shortest path length, $\langle \ell \rangle$ as a function of p. When $p = 0$, the network is a regular ring. Since $C(p)$ is the probability to find a triangle in the network, we can start by computing the number of triangles in the network. Observing that a walk around any triangle, on such a network, is composed of two steps in the same direction on the ring, say a clockwise, followed by a third step to close the triangle. The third step can take at most $k/2$ network spacements, since it is the longest possible link on the network. Thus, the number of the two first steps is the number of different ways of choosing the node for those steps from the $k/2$ possibilities, which is, $k(k/2-1)/4$. Therefore, the number of triangles on the ring is $Nk(k/2-1)/4$. The number of possible triangles involving a particular node belongs to is simply $k(k-1)/2$, which leads, for the whole network, to $Nk(k-1)/2$. As a result, one can write the clustering coefficient as

$$C = \frac{\frac{1}{4}Nk(\frac{1}{2}k-1)\times 3}{\frac{1}{2}Nk(k-1)} = \frac{3(k-2)}{4(k-1)}. \tag{2.12}$$

The factor 3 arises at the numerator since a triangle gets counted three times when counting the number of possible triangles of the network. So, for $p = 0$, $C = 0$ for $k = 2$ (which is the case of a ring where each node is connected to its two first neighbors), and to the limit of $k \to \infty$, $C = 3/4$. In this regime ($p = 0$) the shortest path length is linearly dependent on N. At the

[8] A simple algorithm to perform the rewiring process consists in choosing two existing links, namely $\{(A,B),(C,D)\}$, and replace them by two nonexisting links, $\{(A,D),(C,B)\}$.

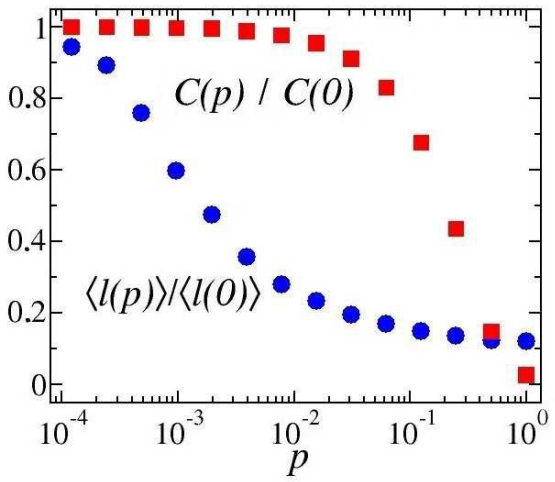


Figure 2.6: Average shortest path length $\langle \ell(p) \rangle$ and the clustering coefficient $C(p)$ of the WS model, as a function of p. The results were renormalized by $\langle \ell(0) \rangle$ and $C(0)$. The comparison between $\langle \ell(p) \rangle$ and $C(p)$ highlights a significant range of p where the network presents the small-world phenomenon. Here, $N = 1000$ and $k = 10$, and the simulations were performed over 20 realizations

other extreme limit, $p \to 1$, the network becomes completely random, $\langle \ell \rangle \simeq \ln(N)/\ln(k)$, and $C \simeq 2k/N$. Figure 2.6 show results of the computations of $C(p)$ and $\langle \ell(p) \rangle$ for the WS model. One can notice that, there is a considerable range of p where there is a small characteristic path length, as in a random network, but $C(p) \gg C_{ER}$, i.e., highly clustered like a regular lattice.

One can notice that for small values of p the presence of these "defects", from the rewiring of few links, reflects drastically as a non-linear behavior of $\langle \ell(p) \rangle$. On the other hand, only when the randomness is predominant, there is a considerable change in the value of the clustering coefficient, which leads $C(p)$ to be practically constant in the small p regime. The logarithm scale of Fig. 2.6 shows how fast it decreases, and that, during this intense variation, $C(p)$ remains virtually unchanged, indicating a transition between a regular topology and a small-world topology, almost imperceptible at the local level.

3 *The navigation problem*

The navigation problem consists in sending a message, data or some piece of information to a particular node of the network. The Milgram's experiment is just an example of this task, where participants were asked to delivery a message to a target person passing it through chains of acquaintances. The Milgram's experiment brought to light evidence of how close connected people are, coining the "six degrees of separation" concept. Many years latter, Jon Kleinberg pointed out another, but not less important, characteristic of the experiment: short paths not only exist on the network, but people are good at finding them. The main result found by Kleinberg explained how the underlying topology, and the particular structural characteristics of network can be determinant on how easy and efficient the transport of message, data, or information throughout the network can be done.

In this chapter we introduce the problem of navigation, the Kleinberg's model for a spatially embedded small-world network, and the consequences of considering the transport of information on such an underlying geographical scenario.

3.1 The Kleinberg's network model

As we presented earlier, social networks exhibit the small-world property. Since the results presented by Milgram, many models have been proposed in order to reproduce such a phenomenon. The WS model was the first to be successful being considered as a paradigm to other models. However, other models were able to show why two arbitrary nodes on real networks are capable to find effective short chains of links which connected them. Kleinberg showed how these chains are found as a consequence of the constrained space where the network is placed. Thus, he proposed a family of random networks, which naturally generalize the WS model, showing that a model of this family is efficient to reproduce the Milgram's Experiment.

The Kleinberg's model has been inspired on the paradigm presented by Watts and Strogatz, in other words, a network that is rich in local links, and has few long-range links. However, in-

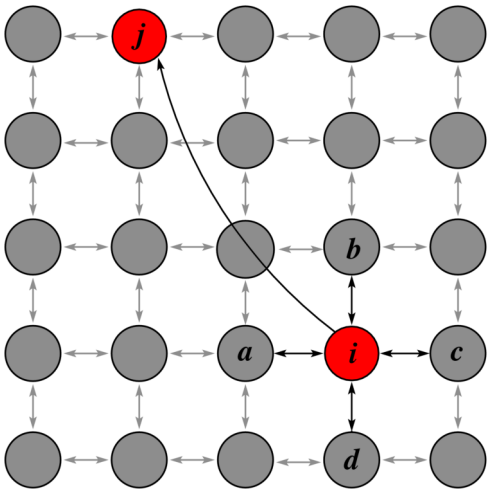


Figure 3.1: Connections of a node i. Each node i receives a link with its first neighbors within a distance $l = 1$ (a, b, c, and d), and a link with a randomly chosen node j. The node j is chosen with probability $P(r_{i,j}) \sim {r_{i,j}}^{-\alpha}$.

stead of using an one-dimensional ring as the underlying network, Kleinberg opted to use a two-dimensional square lattice. With this, a set of nodes, the individuals of the social network, are mapped by the nodes of an $L \times L$ square lattice, $\{i = (u,v) : u \in \{1,2,\ldots,L\}, v \in \{1,2,\ldots,L\}\}$. Thus, the distance between two distinct nodes $i = (u,v)$ and $j = (m,n)$ is

$$r_{i,j} = |m-u| + |n-v|. \tag{3.1}$$

Notice that, geographically, the distance $r_{i,j}$ is the "Manhattan distance", meaning the number of links that connect the nodes i and j. On the Kleinberg's model, there are three universal parameters:

- $l \geqslant 1$ - Each node i has a directed link with any other node within a Manhattan distance l. These are the local neighbors of node i.
- $q \geqslant 0$ e $\alpha \geqslant 0$ - Each node i receives q long-range links randomly with a node j, with probability $P(r_{i,j}) \sim {r_{i,j}}^{-\alpha}$.

The Kleinberg's model has a simple geographical interpretation: Individuals on a social network know their neighbors within a certain distance l, and other people distributed throughout the network. The power law probability distribution, $P(r) \sim r^{-\alpha}$, indicates that the probability of two different people to know each other is higher if they "live" close geographically to each other. Taking l and q as constants, the Kleinberg's model has only one parameter, α, which defines an entire family of random networks. For $\alpha = 0$, the Kleinberg's model is equivalent to the WS model, where the choice of a new link is independent on the lattice position of the new

acquaintance of i. As α increases, the long-range acquaintances of u become nearer to i.

3.2 Kleinberg's local navigation

3.2.1 The decentralized algorithm

The procedure defined in Milgram's experiment is a basic algorithm, remarkably a particular one from the class that computer scientists call *decentralized algorithms*, called the *greedy algorithm*. The algorithm that constitutes the Kleinberg's model to navigation is clearly inspired by the process stated by Milgram's experiment. The algorithm proposed by Kleinberg aims to compute the number of steps necessary to deliver a message from a source node s to a target node t. Both nodes are randomly chosen. At each step of the algorithm, the holder of the message u should choose among its connections, whether they are local or long-range connections, a node v to be the next message's holder. This choice takes into account the distance $r_{v,t}$ between node v and the target one t. This characteristic of the Kleinberg's decentralized algorithm reproduces the main ingredients of the Milgram's experiment: The message holder should choose a connection that he/she believes will speed up the message delivery.

During this process, the holder u has only local knowledge of the network, in other words he/she knows:

(i) The underlying geography of the network (the square lattice).

(ii) The network position of t.

(iii) The network position of its connections, the local and long-range ones, as well as the network position of any node that, once, was the message holder.

The main property of the algorithm is that u does not know the long-range connections of any other node on the network. The *expected delivery time* τ is the expected value of the number of steps necessary to the decentralized algorithm to deliver the message throughout the network from a node s to a node t, making use or not of the long-range connections added to the geographical lattice in terms of the power law distribution $P(r)$.

In the original work of Kleinberg [11], he presented simulations for $d = 2$, obtaining that the effective exponent is $\alpha = 2$ (when the delivery time is minimum to this scenario). With the analysis of the results, Kleinberg conjectured that the effective exponent is $\alpha = d$. Roberson and ben-Avrahan [26] showed that this is valid also for fractal networks, where $\ell \sim N^{1/d_f}$, in other words, the effective exponent $\alpha = d_f$ holds for any real value of d.

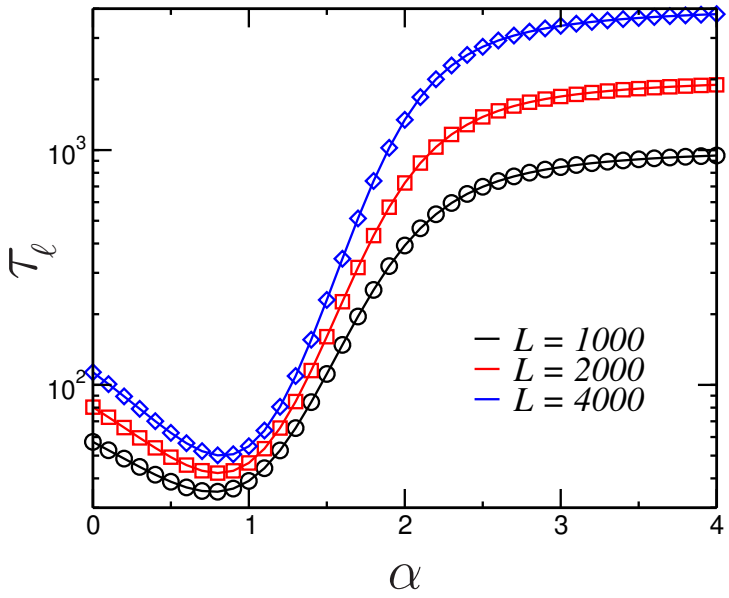


Figure 3.2: The expected mean delivery time τ_ℓ, namely the number of steps needed for a message starting from a source node $s = 0$ to reach the target node $t = L$ in a $4L$ length ring, as a function of α. The most efficient navigation is reached for $\alpha = 1$.

The Kleinberg's local navigation was solved analytically [27, 28] for the 1-dimensional case. For convenience, one can take $P(r_{ij})$ independent of i [27] by adopting periodic boundary conditions. In order to do it, let us consider case of a ring of length $4L$, where the source node $s = 0$, the target node $t = L$, and with the length of long-range connections limited to $2L-1$. Consequently

$$P(r_{i,i+b}) = Ab^{-\alpha}; \tag{3.2}$$

where A is a normalization factor given by

$$A = \frac{1}{2\sum_{b=1}^{2L-1} b^{-\alpha}}. \tag{3.3}$$

One can consider the probability $\Pi(n;\ell)$, which is the probability that a message, at a distance ℓ from t, takes n additional steps to reach t. Clearly, $\Pi(n;\ell)$ has two "boundary conditions": $\Pi(n,0) = \delta_{n,0}$ and $\Pi(0,\ell) = \delta_{0,\ell}$, where $\delta_{i,j}$ is the Kronecker delta function, defined as $\delta_{i,j} = 0$, if $i \neq j$, and $\delta_{i,j} = 1$, if $i = j$. These two conditions are due to the fact that, if the message holder is the target, it takes no additional step to reach t.

The probability $\Pi(n;\ell)$ should satisfy the following equation:

$$\Pi(n;\ell) = A\sum_{b=1}^{2\ell-1} b^{-\alpha}\Pi(n-1;|\ell-b|) + \left(1 - A\sum_{b=1}^{2\ell-1} b^{-\alpha}\right)\Pi(n-1;\ell-1). \tag{3.4}$$

The first term on the right hand side represents the probability that the first of the additional

n steps to be a long-range connection of length b, with which one may leave the message to a distance of $|\ell - b|$. The second, represents the probability for a step take place on the regular ring, where the message will advance only a single lattice spacement.

Considering the first moment of $\Pi(n;\ell)$, $\langle n \rangle \equiv \tau_\ell$, i.e., the expected delivery time from a node at a distance ℓ away from the target node. So,

$$\tau_\ell = \sum_n n\Pi(n;\ell). \tag{3.5}$$

As a consequence, we have that,

$$\sum_n n\Pi(n-1;\ell) = \sum_n (n+1-1)\Pi(n-1;\ell) = \sum_n \Pi(n-1;\ell) + \sum_m m\Pi(m;\ell) = 1 + \tau_\ell. \tag{3.6}$$

Here, $m = n-1$. Multiplying Eq. (3.4) by n and summing over n, we obtain the master equation for τ_ℓ, given by

$$\tau_\ell = A \sum_{b=1}^{2\ell-1} b^{-\alpha} \left(1 + \tau_{|\ell-b|}\right) + \left(1 - A \sum_{b=1}^{2\ell-1} b^{-\alpha}\right) (1 + \tau_{\ell-1}), \tag{3.7}$$

for $\ell = 1, 2, \ldots, L$. Figure (3.2) shows simulations results (symbols) for the Kleinberg's greeding algorithm in $d = 1$ yielding a perfect agreement with the numerical integration of Eq. (3.7)[1]. One can notice that, with the increase of L, the minimum value of τ_ℓ approaches the effective value of $\alpha = 1$.

In his original study [11], Kleinberg conjectured that the correlations between the local network structure and the long-range connections in small-world networks provide fundamental "clues" to the nodes, so that they are able to find effective chains of acquaintances that connect them. This interplay between the local topology and the long-range connection structure generates a "gradient" which permits the nodes to find such chains. For larger values of α, this interplay vanishes, since this type of effective chain does not exist anymore. One can notice that, for $\alpha < d$, short chains exist, but, the decentralized algorithm, based on local information, makes the nodes unable to find these shortcuts.

3.2.2 Scaling behavior for the expected delivery time

Kleinberg [11] studied how the underlying structure of the network may interfere on the decentralized algorithm which aims to find short chains of acquaintances between arbitrary nodes of a network. When the parameter $\alpha = 0$, the long-range connections are added to the network

[1] Here, we used $\tau_1 = 1$ as initial condition.

independently of its underlying structure. For this scenario, we can use random networks theory arguments to prove that the short chains between two nodes exist, and that they are limited by $O(\ln L)$, in other words, the average shortest path length $\langle \ell \rangle$ increases with the network size N logarithmically, at most, as in the WS model. This property is somehow obvious, since Kleinberg proposed his model inspired on the paradigm defined by Watts and Strogatz. However, as we will show in the following, this scenario is not efficient in finding effective chains of acquaintances. To analyze this problem in a general way, we consider a d-dimensional square lattice.

From the proposed model, we know that the probability of a node u having the node v connected through a long-range connection is given by the power law:

$$P(r_{u,v}) \sim [r_{u,v}]^{-\alpha}, \tag{3.8}$$

which should be normalized by a factor of $\sum_{v \neq u} {r_{u,v}}^{-\alpha}$, scaling as:

$$\begin{aligned} \sum_{v \neq u}^{\infty} {r_{u,v}}^{-\alpha} &\sim \int_1^L x^{-\alpha} x^{(d-1)}\, dx \\ &\sim \begin{cases} (\alpha - d)^{-1} & \text{, if } \alpha > d, \\ \ln L & \text{, if } \alpha = d, \\ L^{d-\alpha} & \text{, if } \alpha < d. \end{cases} \end{aligned} \tag{3.9}$$

For $0 < \alpha < d$, one can consider a region with an arbitrary radius R around the target node t, such that $R = L^{\delta}$, where $\delta < 1$. Equation 3.9 states that the probability for a node u to have a long-range connection with a node v is

$$P(r_{u,v}) \sim \frac{{r_{u,v}}^{-\alpha}}{L^{d-\alpha}} \leqslant \frac{1}{L^{d-\alpha}}. \tag{3.10}$$

Thus, the probability for a node to be chosen among one of the connections of node u within the shell defined by R must obey the relation:

$$\begin{aligned} P(r_{u,v} \mid r_{v,t} < R) &\leqslant \frac{R^d}{L^{d-\alpha}} \\ &\leqslant L^{\delta d - d + \alpha}. \end{aligned} \tag{3.11}$$

Since $\delta < 1$ and $R = L^{\delta}$, for large L, R increases more slowly than L if $\delta d > d - \alpha$, or it decreases if $\delta d < d - \alpha$, which indicates that the source node is likely to be outside the shell. This result has two consequences. First, any path between s and t should pass through, at least, one long-range connection that leaves the message into the shell. Second, the expected delivery time is limited to $\tau \geq R$.

The probability for a node with such a long-range connection to be reached in at least R steps is less than $R \times L^{\delta d - d + \alpha}$. If this probability goes to zero when $L \to \infty$, then the expected value for the number of steps is also limited by R, at least. For this probability to become zero, it is necessary that,

$$0 < R \times L^{\delta d - d + \alpha} < 1, \tag{3.12}$$

in other words,

$$0 < L^{\delta(d+1) - d + \alpha} < 1, \tag{3.13}$$

implying,

$$\begin{aligned} \delta(d+1) - d + \alpha &< 0, \\ \delta &< \tfrac{(d-\alpha)}{(1+d)}. \end{aligned} \tag{3.14}$$

Therefore, the expected number of steps can not be lower than $L^{(d-\alpha)/(1+d)}$ for $0 < \alpha < d$.

For $\alpha > d$, the probability of a long-range connection longer than L^{γ} ($0 < \gamma < 1$), according to Eq. (3.9) should scale with the network size as,

$$\int_{L^{\gamma}}^{\infty} \frac{x^{-(\alpha-d+1)}}{\alpha - d}\, dx \sim L^{\gamma(d-\alpha)}, \tag{3.15}$$

so, the probability of the message taking a step longer than L^{γ}, e.g. L^{β} ($\gamma < \beta < 1$), is lower than $L^{\beta} L^{\gamma(d-\alpha)}$. However, this probability vanishes when $L \to \infty$, which means that the total distance traveled by the message does not overcome the value $L^{\beta+\gamma}$. Nevertheless, since the distance that separates the source node s from the target node t should be proportional to L, it gives that,

$$\beta + \gamma = 1. \tag{3.16}$$

On the other hand, the probability that long-range connections longer than L^{γ} exist vanishes when $\beta + \gamma(d - \alpha) < 0$. These two conditions give that,

$$\beta < \frac{(\alpha - d)}{(\alpha - d + 1)}, \tag{3.17}$$

so that, τ should not be lower than $L^{\beta} \sim L^{(\alpha-d)/(\alpha-d+1)}$.

For the singular value $\alpha = d$, one can make the following analysis: the target node t is surrounded by a number y of shells with radius $e^{y-1} < R < e^{y}$, $y = 1, 2, \ldots$. Now, suppose that the message holder u is located on shell y, the probability that u has a connection with some

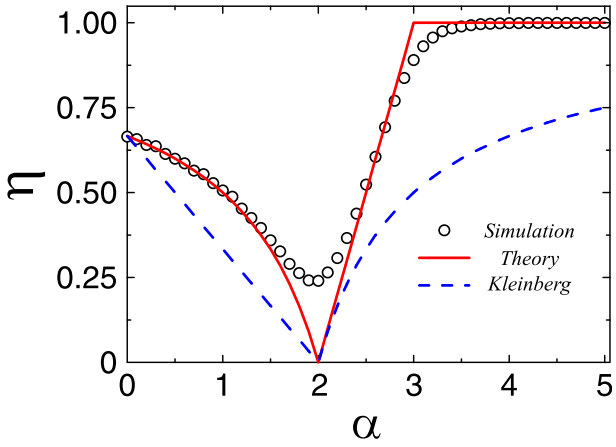


Figure 3.3: Delivery time exponent, η, as a function of α. Results from Eq. (3.20) (red solid line) are compared with the lower bounds proved by Kleinberg (blue dashed line) and with the simulation results from the delivery decentralized algorithm (black circles). One can notice that, for $\alpha = 2$, the behavior of τ does not depend of L as a power law. Indeed, its behavior is given by $O(\ln^2 N)$.

node v on the $(y-1)$ shell is, according with Eq. (3.9),

$$P(r_{u,v}) \sim \int_{e^{m-1}}^{e^m} \frac{y^{-1}}{\ln L}\, dy = \frac{1}{\ln L}. \tag{3.18}$$

The probability to reach the $y-1$-shell in more than x steps is $p(x) = (1 - 1/\ln L)^x$, and the mean number of steps needed to do it is

$$\langle x \rangle = \int_0^\infty p(x)\, dx \sim \ln L. \tag{3.19}$$

Since the larger shell is the one where $e^y = L$, the number of shells between s and t should be $O(\ln L)$, therefore, τ should not overcome $O\left((\ln L)^2\right)$.

Analytically, one can notice that τ is bounded by L^η, except for $\alpha = d$, where τ has the upper bound of $(\ln L)^2$, which increases more slowly than the powers of L presented above. Carmi *et al.* [27] showed that the asymptotic behavior of the exponent η is given by

$$\eta = \begin{cases} \frac{d-\alpha}{d+1-\alpha} & , 0 \le \alpha < d, \\ \alpha - d & , d < \alpha < d+1, \\ 1 & , \alpha > d+1. \end{cases} \tag{3.20}$$

Figure (3.3) shows the behavior of the exponent η as a function of α. The bounds calculated by Kleinberg are shown as the blue dashed lines. The asymptotic values of η are presented as the red solid lines. The symbols (black circles) are the results from the simulation. All three results are for the 2-dimensional case.

3.3 Global navigation on Kleinberg's network

The decentralized algorithm presented by Kleinberg approaches the navigation problem in a small-world network model [10, 11], in a procedure that makes use only of local information. As was presented in the last section, one can consider a d-dimensional square lattice with $N = L^d$ nodes. Kleinberg considered that each node u would be connected with its first neighbors within a perimeter $l \geq 1$, and in addition, would receive $q \geq 0$ long-range connections. The probability that u has a long-range connection with node v, randomly chosen, is proportional to $r_{u,v}^{-\alpha}$, where $r_{u,v}$ is the Manhattan distance between them, and α is a tunable parameter. One can notice that for $\alpha = 0$, $P(t_{u,v})$ is uniformly distributed, as in the WS model, where the long-range connections are added to u independently of the "geographical" characteristics of the network. When $\alpha > 0$, the long-range connections of u become even more clustered around its neighborhood.

In this section, we present the behavior of $\langle \ell \rangle$ for the Kleinberg's network, wherewith we can study the navigation throughout the network with global knowledge. In order to do it, we use a variant of Kleinberg's model, where $l = 1$, but only 10% of nodes receive a long-range connection. To compute $\langle \ell \rangle$ we use

$$\langle \ell \rangle = \frac{1}{N(N-1)} \sum_{u \neq v} \ell_{u,v}, \tag{3.21}$$

where $\ell_{u,v}$ is the shortest path length between the pair of nodes u and v.

The relation between navigation with global knowledge and $\langle \ell \rangle$ is justified since, with global access to the entire information of the network structure, the nodes "know" which is the shortest path between them and any other node on the network. Figure (3.4) show the computation of $\langle \ell \rangle$ for different network dimensions d, according to Eq. (3.21). For global navigation, efficient navigation occurs for $\alpha < d$, where $\langle \ell \rangle$ seems to be invariant. Since $\alpha = 0$ is a trivial solution, when $P(r_{u,v})$ is uniform, and permits high frequency of long-range connections, one can consider it as the optimal value for the exponent α, in sharp contrast with what is observed for Kleinberg's decentralized algorithm, where the message holder only has access to local information, and the efficient exponent is $\alpha = d$ for a d-dimensional square lattice. So, since the navigation is based on a global information algorithm, the message holder takes greater advantage of the power law distribution $P(r_{u,v}, \alpha = 0) \sim [r_{u,v}]^{-\alpha}$.

This result does not change when a constraint is imposed to the size of long-range connections, e.g., if long-range connection is restricted to be smaller than $b_\ell \times L$ ($0 \leq b_\ell \leqslant 1$). Figure (3.5) shows the behavior of $\langle \ell \rangle$ for different values of b_ℓ and $L = 10^3$, while in Fig. (3.6)

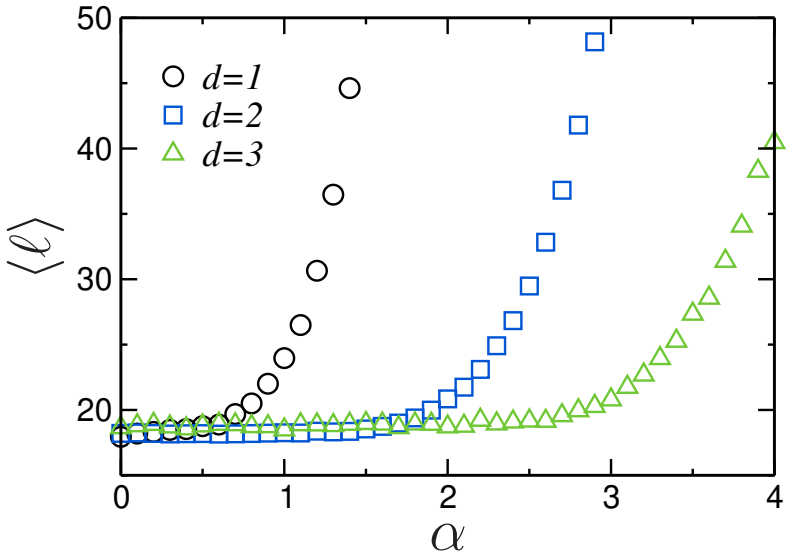


Figure 3.4: Shortest path length $\langle \ell \rangle$ as a function of α for different lattice dimensions d ($d = 1$, 2, and 3) over 4×10^3 realizations. 10% of nodes of the d-dimensional regular lattice, with linear size $L = 10^3$, are randomly selected to receive a long-range connection. The optimal $\langle \ell \rangle$ is achieved for $\alpha = 0$.

we fix $b_\ell = 0.04$ and compute $\langle \ell \rangle$ for different linear sizes L. In both cases, the effective navigation was achieved at $\alpha = 0$. So, one can conclude that the underlying geographical structure of the network does not affect the process with global knowledge of the network location of nodes.

Kosmidis, Havlin, and Bunde [17] showed equivalent results to the ones presented here for a network model similar to the Kleinberg's model and to the long-range percolation model [29]. Moreover, they showed that

$$\langle \ell \rangle \sim \begin{cases} log(N), & \alpha \leq d, \\ (log(N))^{\delta}, & d \leq \alpha < 2d, \\ N^{1/d}, 2d \leq \alpha. \end{cases} \tag{3.22}$$

The above equation indicates that for $\alpha \geq 2d$, the network behaves as a lattice, belonging to the "large-world" class of networks, but, more important, for $\alpha < 2d$ the Kleinberg's network model behaves as a small-world network. Their results suggest that the exponent δ of the scaling logarithm for $d \leq \alpha < 2d$ is approximately $\delta = 1/(\alpha(2-\alpha))$ for $d = 1$, and $\delta = 4/(\alpha(4-\alpha))$ for $d = 2$ [17].

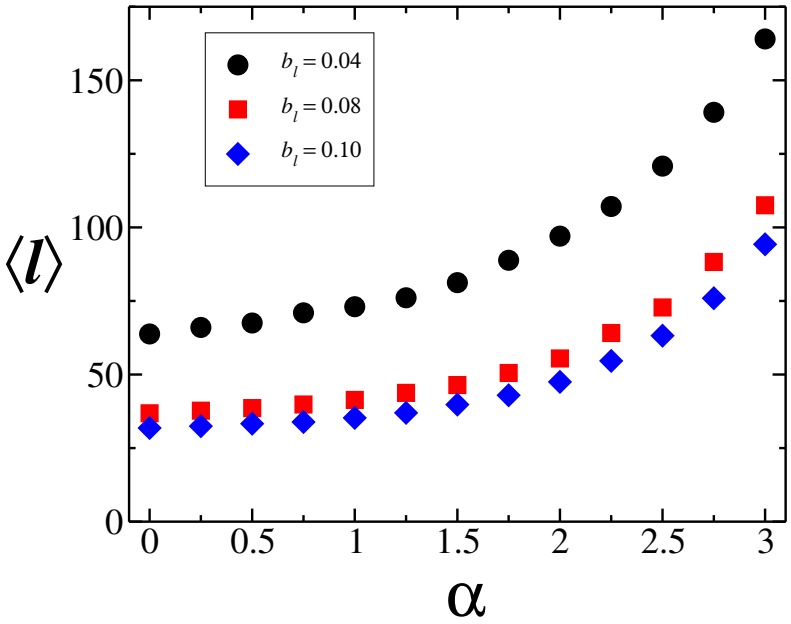


Figure 3.5: $\langle \ell \rangle$ as a function of α. When the long-range connections are constrained to be within a region $r_{u,v} \leqslant b_\ell \times L$, the effective navigation is achieved for $\alpha = 0$. The behavior of $\langle \ell \rangle$ is similar to the unconstrained case. Here, $L = 10^3$ is the linear size of the square lattice, and 20% of nodes were randomly selected to receive a long-range connection.

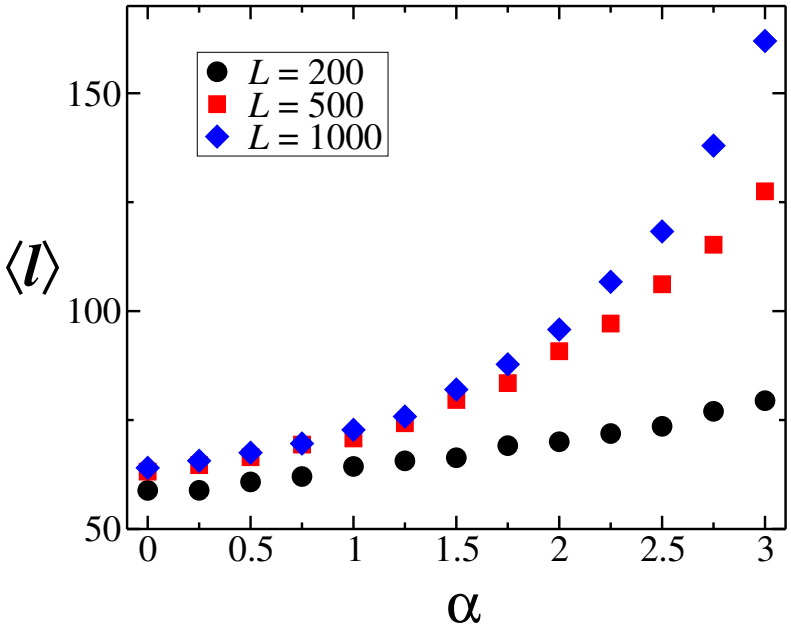


Figure 3.6: $\langle \ell \rangle$ as a function of α to different linear size L of the square lattice. The nodes u and v are constrained to a distance $b_\ell \times L$ again, but to a fixed $b(\ell) = 0.04$. One can notice that the more efficient navigation does not change, and is achieved for $\alpha = 0$. Again, 20% of nodes were randomly selected to receive a long-range connection.

3.4 Local vs. global navigation

The algorithms presented by Kleinberg for local navigation (decentralized), and the algorithm for global navigation (basically, the breadth-first-search algorithm [22]), provide simple

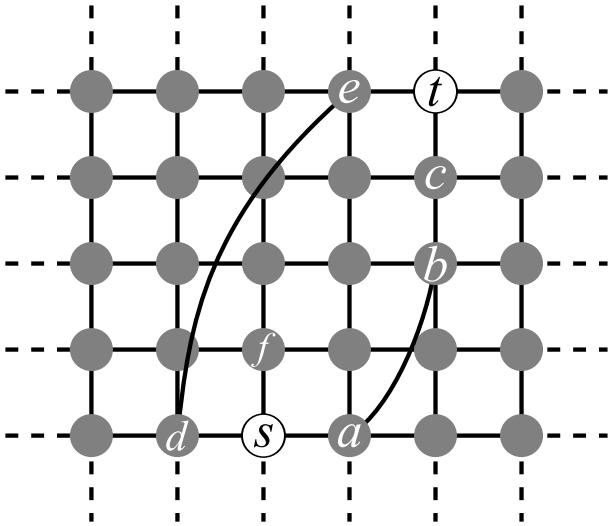


Figure 3.7: How the navigation algorithms work. If the goal is to reach the target (node t) starting from the source node s, Kleinberg's decentralized (greedy) algorithm gives two solutions. The first, passing through node a is the better one in this scenario and has length 4. The second, passing through node f has length 6. On the other hand, the global navigation algorithm (breadth-first-search) results in a path passing through node d, which is 3 steps long.

methods for the analysis of the navigation problem. The results presented in the previous sections imply a simple interpretation, which can be illustrated by the following example.

Figure 3.7 presents a limited region of a square lattice, which may be infinite. That region contains the source node s, and the target node t. There are two long-range connections, the first between nodes a and b, and a second, between nodes d and e. In this simple example, the source node s at the beginning of the Kleinberg's local navigation algorithm, is the the message holder. At the first step of the algorithm, s verifies that node a and f are closer to the target node t than its other neighbor d. If s chooses node f as the next message holder, the decentralized algorithm will result in many different chains of acquaintances of length 6. If, however, s chooses node a, the algorithm will result in only one possible chain of acquaintance of length 4. This better solution is the sequence of links $[(s,a);(a,b);(b,c);(c,f)]$. These dubious solutions are natural on Kleinberg's local navigation, since s does not "know" that node a has a long-range connection with node b. For the same reason, although this solution is the optimal one in the local navigation scenario, since s does not know that its neighbor d has a long-range connection with node e, in an extreme case, it is far from the optimal solution to this delivery task.

In the global information navigation scenario, any network node has knowledge of all the long-range connections of any other network node. Thus, for this case, the breadth-first-search algorithm provides s knowledge about the existence of the long-range connection between d and e. And after that, in the third step of the delivery task, e is the message holder and delivers

it to the target node t. The sequence $[(s,d);(d,e);(e,t)]$, which is 3 steps long, is the chain that solves the delivery task for the global information scenario.

Clearly, in a real situation it is not always possible to make use of global knowledge of the network to solve transport problems, but this simple comparison indicates that sometimes it is better to give a step back than to go further.

4 *Towards design principles for optimal transport networks*

The interplay between topology and dynamics in complex systems represents the focus of many studies in different fields of research with important scientific and technological applications. Due to their enormous potential to represent the intricate topology of numerous systems in nature, complex networks [30, 20, 31] have recently been used as substrates in combination with a plethora of dynamical models to describe the behavior of biological, social, chemical, physical and technological networks [32, 33, 34]. Much attention has been dedicated to the problem of navigation in complex network geometries [10, 26, 27, 28, 35, 36, 37, 38, 39, 40]. In most cases, the influence of the underlying network geography on the performance of the transport process is investigated assuming that only local information is available for navigation [10, 26, 27, 28, 41].

For many navigation problems of interest in science and technology, global rather than local information is required, i.e., any source node s possesses the knowledge of the entire network topology. In this situation, the average shortest path $\langle \ell \rangle$ from source to target becomes the relevant navigation variable to be optimized. For example, in a subway network, such as in Manhattan, the travel routes should be planned or changed in such a way as to minimize the travel time for a given limited reconstruction cost. This task is performed by considering the whole structure of the network in terms of its nodes and links, namely, by knowing the location of all subway stations, their connections and the shortest path between any two stations. If we now consider an underlying network of streets and avenues over which one has to plan or improve an existing subway network, and if the aim is to minimize the average travel time between its stations, the search for an optimal strategy to add new connections in the network for a given budget should therefore play a key role. In this chapter we show that the imposition of a cost constraint represents a crucial ingredient in the design and development of efficient navigation networks. Optimal navigation with the presence of long-range links in a lattice network without constraints was studied recently by Kleinberg [10]. Here we show that a rather different behavior can be observed for this problem when realistic constraints on total length

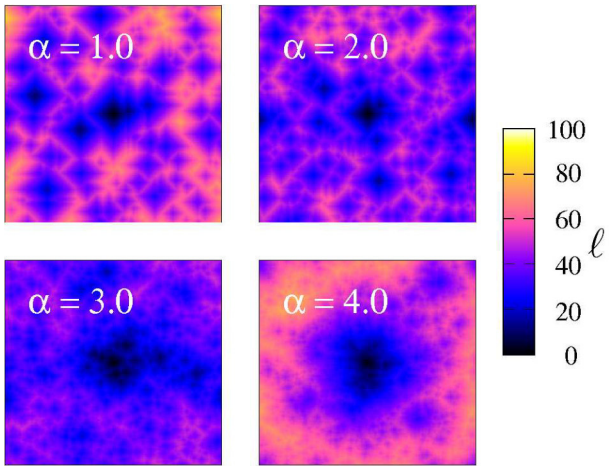


Figure 4.1: Shortest-path length, ℓ, from each node to the central node in the network for different values of α. In this case we impose a constraint in the length of the long-range connections. The sum of the length of these connections is limited, $\Lambda = \sum r_{ij} = N$, where N is the number of nodes in the underlying lattice. The network model is constructed from a square lattice with L^2 nodes, with $L = 256$. We can clearly observe that the best condition for shortest path length is obtained for $\alpha = 3$.

are imposed on the process of adding long-range connections.

4.1 Model for network navigation with cost

Consider the case of an existing subway network which needs improvements [41, 42]. The financial cost to build up a large number of *new direct connections* between distant stations (i.e., non-neighboring sites) can make it prohibitive, since only limited resources are normally available for this task. This problem can be modeled by the following system. In a 2-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 stations and the bonds represent the routes of the subway. In our model, pairs of sites ij are then randomly chosen to receive long-range connections with probability proportional to $r_{ij}^{-\alpha}$, where r_{ij} is the Manhattan distance between sites i and j, as in the Kleinberg's model. Finally, the addition of long-range connections to the system stops when their total length (cost), $\sum r_{ij}$, reaches a given value Λ. 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 connections added to the system.

To better demonstrate the competition between total length and number of links, we gener-

ate a single network realization ($L = 256$) for a given value of α and compute the shortest-path length, ℓ, from each node in the network to its central node. This calculation is performed as follows. Once we choose the root node (e.g., the central one), we visit all its neighbors, including the neighboring nodes connected by long-range connections. These visited nodes are classified as *shell one nodes*, meaning that they are only one time step away from the root node [1]. After that, we visit all the neighbors of these nodes not visited before and classify them as *shell two nodes*. Following this procedure for all network nodes, we obtain the ℓ values (time) for each node to be reached from the root node. Figure 4.1 shows the contour plots representation of the ℓ values performed for four different values of the parameter α. For $\alpha = 1$ and 2 the number of long-range connections is small. As a result, only a few little islands sparsely dispersed in these networks are really close to their central nodes (only a few short and/or long-range connections away). For $\alpha = 3$ the added long-range links are shorter, but more numerous, thus substantially decreasing the shortest path over the whole network. For $\alpha = 4$, due to the very short length size of the added connections, only a limited region surrounding the central node displays a reduced shortest path. Sites which are further away from the origin have significantly larger path length ℓ to the origin.

4.2 Global navigation in a small-world network with cost

We extract more quantitative information about this navigation problem by performing extensive simulations for different values of α and many realizations of different system sizes. In each case, the average shortest path $\langle \ell \rangle$ is calculated over all realizations, considering all the shortest distances between each pair of nodes. We assume that the total length (cost) is proportional to the total length of the links in the underlying network, i.e., $\Lambda = AL^2$, 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) [2].

The results presented in Fig. 4.2 clearly indicate the presence of a minimum $\langle \ell \rangle$ for different system sizes at the same value of the exponent $\alpha = 3$, where optimal navigation is achieved with (global) knowledge of the shortest paths and cost limitations. The way in which $\langle \ell \rangle$ scales with system size L, however, seems to follow rather different behaviors, depending on the value of α. We tested two possible forms for $\langle \ell \rangle$ vs N, a power law and a power of $\log N$. As shown in the main plot of Fig. 4.3a, our results for $\alpha \neq 3$ suggest that the shortest path $\langle \ell \rangle$ follows a

[1] We assume that the traverse time of a link is almost the same for both short-range and long-range links since most of the time is spent in station and for decelerating and accelerating

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

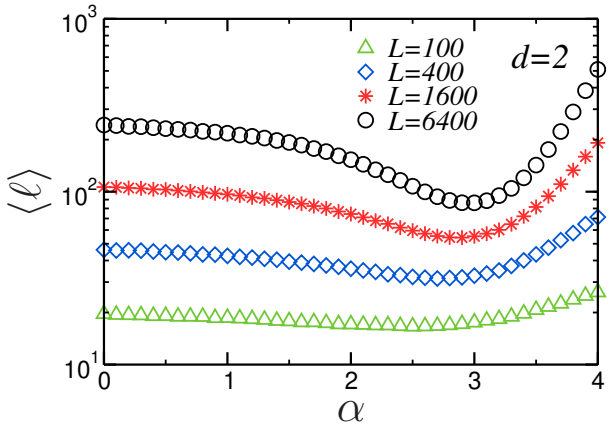


Figure 4.2: Average shortest path length $\langle \ell \rangle$ as a function of α. There is a constraint in the total length of the long-range connections, $\Lambda = \sum r_{ij} = L^2$, where L is the size of the underlying square lattice. We find that the optimal shortest-path is achieved for $\alpha = 3$. With the restriction on total length, the number of long-range connections are not fixed (e.g.. with $\alpha = 0$, large long-range connections become frequent, which reduces the total number of long-range connections.) To obtain these results, we simulated $10,000$ realizations for $L = 512$, 3500 realizations for $L = 1024$ and 2048, and 25 realizations for $L = 4096$.

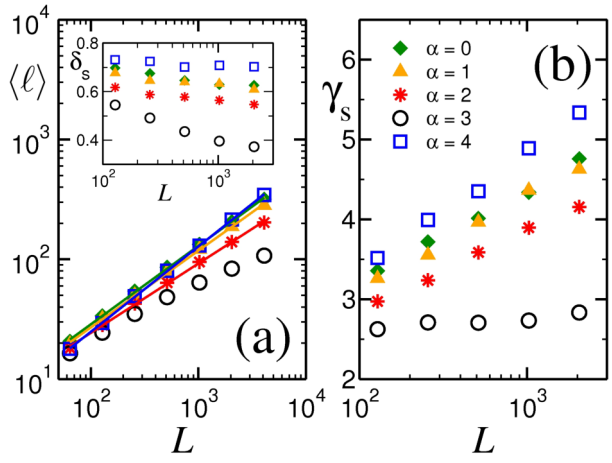


Figure 4.3: In (a) we show the average shortest path length $\langle \ell \rangle$ as a function of the lattice size L for the square lattice. The constraint in the total length of the long-range connections is $\Lambda = L^2$. The curve with $\alpha = 3$ increases more slowly with L compared to any other value of α. In the inset, the plot of the successive slopes δ_S obtained from $\log_{10}\langle \ell \rangle$ versus $\log_{10} L$ reinforces the display of power law behavior of $\langle \ell \rangle$ with L for $\alpha \neq 3$. The plot of the successive slopes γ_S obtained from $\log_{10}\langle \ell \rangle$ versus $\log_{10}(\log_{10} L)$ shown in (b) indicates that $\langle \ell \rangle$ increases as a power of the logarithm of L for the optimal condition $\alpha = 3$.

power law with system size L. This is supported by the plot of successive slopes δ_S obtained from $\log_{10}\langle \ell \rangle$ versus $\log_{10} L$, which are almost invariant, as shown in the inset of Fig. 4.3a. In contrast, in the case of $\alpha = 3$, the increase with L of $\langle \ell \rangle$ appears to be less rapid than a power law. Interestingly, the successive slopes γ_S obtained from $\log_{10}\langle \ell \rangle$ versus $\log_{10}(\log_{10} L)$, as

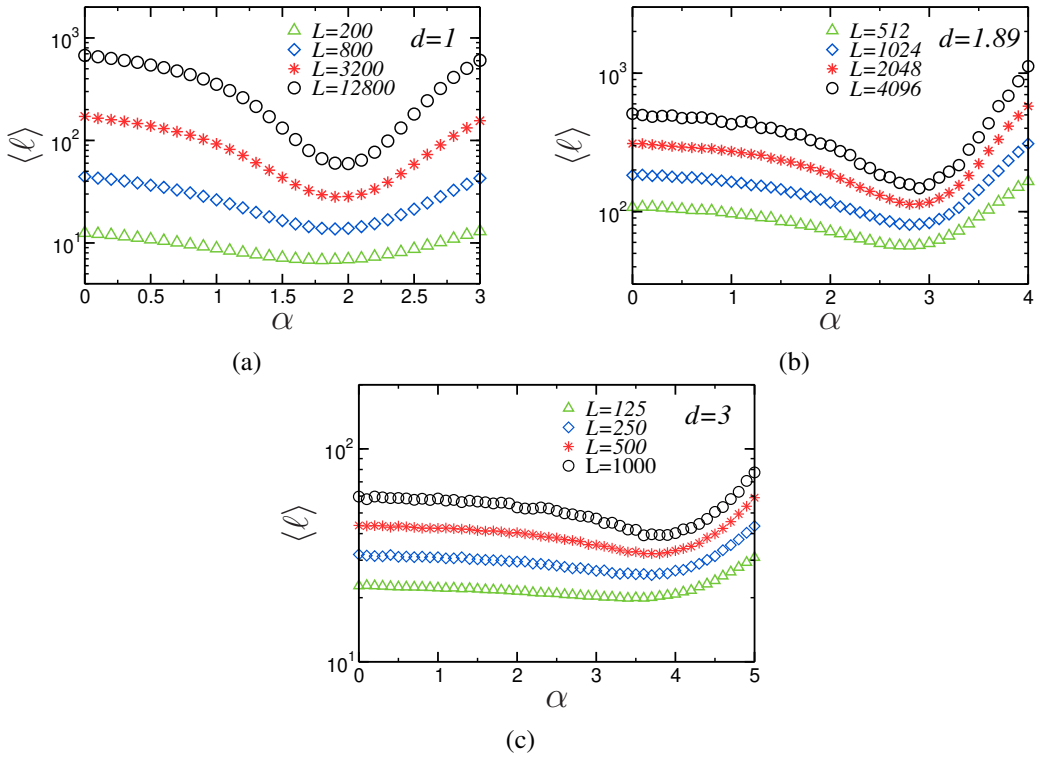


Figure 4.4: Average shortest path length $\langle \ell \rangle$ as a function of α for (a) a one dimensional lattice, for (b) a fractal of dimension $d_f = 1.89$, and (c) for a three dimensional lattice. (a) the total length Λ of long range connections is limited 10% of N. We find the optimal shortest path length at $\alpha = d + 1 = 2$. We simulated 4,000 realizations for each L. (b) The total length Λ is limited to N_f, namely the number of nodes that benlong to the largest percolation cluster. We find the optimal navigation is achieved at $\alpha = d_f + 1 = 2.9$. We simulated 4,000 realizations for $L = 512$ and 1024, 1,600 realizations for $L = 2048$, and 400 realizations for 4096. (c) Average shortest path length $\langle \ell \rangle$ as a function of α for three-dimensional lattice with additional long-range connections. The total length Λ of long-range connections is limited to $N = L^3$. We find the optimal shortest path length is achieved at $\alpha = d + 1 = 4$. We simulated 4000 realizations for $L = 125$ and 250, 400 realizations for $L = 500$ and 80 realization for $L = 1000$.

presented in Fig. 4.3b, indicate that $\langle \ell \rangle$ increases as a power of the logarithm of L, $\langle \ell \rangle \sim \log_{10}^{\gamma_S} L$, rather than a power of L, only for $\alpha = 3$. This provides clear support for the fact that in the optimal condition, $\alpha = 3$, the transport will improve even further as L increases, as suggested by Fig. 4.2.

We also studied our model for a one-dimensional lattice, for a three-dimensional simple cubic lattice, and for the a fractal network (the largest percolation cluster on a square lattice at the critical threshold obtained from the regular bond percolation process) and observed similar behavior. The optimal condition we obtained in these case at $\alpha = 2$ ($d = 1$), $\alpha = 4$ ($d = 3$), and $\alpha = 2.9$ (to the percolation cluster, where $d_f = 1.89$). These results, shown in Fig. 4.4, lead us to conjecture that the optimal value is obtained at $\alpha = d + 1$, where d is the dimension of the

underlying lattice.

In the following, we present analytical arguments showing that $\alpha = 3$ is indeed the only case where logarithmic scaling of $\langle \ell \rangle$ with L can occur for the square lattice, while for $\alpha \neq 3$ a power law with L should exist. By arbitrarily fixing the cost parameter to $\Lambda = AL^2$, we obtain that $\rho \sim \langle r \rangle^{-1}$, where ρ is the density and $\langle r \rangle$ is the average length of the added long-range connections. Since $\langle r \rangle \sim \int_1^L r^{2-\alpha} dr$, it follows that for $2 \leq \alpha < 3$, $\rho \sim L^{\alpha-3}$ and for $\alpha < 2$, $\langle r \rangle$ is limited by the network size leading to $\rho \sim L^{-1}$. Thus, for all values of $\alpha < 3$ the density of the long-range links added, due to the constraint, 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 see this we argue that $\langle \ell \rangle$ is bounded by the relation $\langle \ell \rangle > \rho^{-1/d}$. The right hand side, $\rho^{-1/d}$, appears for the case of the small world model, where $\alpha = 0$, with a fixed concentration of links, $\langle \ell \rangle \sim \rho^{-1/d} \ln L$ [31]. Since for the case $0 < \alpha < 3$, $\langle \ell \rangle$ decreases with increasing α, the bound $\langle \ell \rangle > L^{(3-\alpha)/d}$ is rigorous and $\langle \ell \rangle$ in this range must scale as a power of L. For $\alpha > 3$ and sufficiently large networks, $\langle r \rangle$ is finite and the density becomes independent of the system size, i.e., $\rho \sim L^0$. Thus, the effect of the constraint Λ on navigation should become negligible. However, the finite value of $\langle r \rangle$ suggests that long-range links can be neglected and therefore $\langle \ell \rangle$ should scale as a power of L. Thus, it follows that only for $\alpha = 3$, $\langle \ell \rangle$ can scale logarithmically with L, as suggested by our numerical simulations (see Fig. 4.3).

4.3 Local navigation in a small-world network with cost

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. [10], for example, the *greedy algorithm* is introduced as a paradigm based on local information, where the traveler, when leaving a 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 for $d = 2$ [10]. 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. 4.5, 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 $\Lambda = L^2$. 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

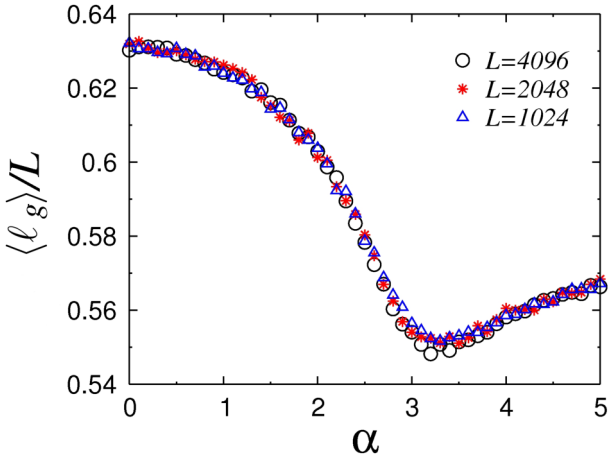


Figure 4.5: The characteristic average delivery time $\langle \ell_g \rangle / L$ as a function of α for navigation with the *greedy algorithm*. The cost Λ involved to add long-range connections changes the behavior of the density of long-range connections. As a result of that, a minimum is observed at $\alpha \approx 3$. Each data point is a result of 4000 simulations and the cost Λ is fixed at L^2.

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. 4.5) of the curves of $\langle \ell_g \rangle / L$ vs. α. The $\langle \ell_g \rangle$ dependence on L for different α for the Kleinberg model was recently derived analytically [27, 28].

4.4 Conclusion

In conclusion, we have investigated the effect of introducing a cost constraint on the optimal design of a transportation network. Our results show 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. The results recently reported by Bianconi *et al.* [43] on the US airport network provide striking support for the validity of our optimal exponent $\alpha = 3$. The fact that the probability of a flight connection within US decays as a power law with the distance between airports, $r^{-\alpha}$, where $\alpha = 3.0 \pm 0.2$, reveals the optimized aspect of the network under the conditions of uniform geographical availability (for customer satisfaction) and cost limitations (for airline companies profit). The result $\alpha = 3$ is in sharp contrast with the results obtained for unconstrained systems with global and local information, where the optimal conditions are

$\alpha = 0$ [35] and $\alpha = d$ [10, 26], 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 on the navigation problem offers a different and more realistic theoretical framework to understand the evolving topologies of other important complex network structures in nature, such as subways, trains, or the Internet.

5 *The nonlocal explosive percolation process*

The percolation paradigm represents a formidable example where a simple geometrical construction leads to profound concepts in statistical physics, with special emphasis on phase transitions, and real applications in science and technology [44, 45, 46, 47, 48, 49]. Standard percolation processes are based on local rules, since they are accomplished through random allocation of sites or bonds, therefore disregarding any spatial correlation or global information involved in the occupation of other elements on the lattice. However, in the case of long-range spatially-correlated percolation [50, 51, 52, 53, 54], the probability for a site to be occupied depends on the occupancy of other sites. Moreover, it has been shown that spatial long-range correlations in site occupancy can give rise to important changes on the structural characteristics of the spanning cluster as well as its corresponding conducting backbone [54]. These changes are strong enough to modify the scaling exponents of traditional (local) percolation.

Recently, a new percolation model has been proposed, based on the so-called Achlioptas process, in terms of a bond occupation process that is essentially nonlocal [12]. In this model, at each step, two unoccupied bonds are randomly chosen and associated with weights related to the cluster sizes they would potentially connect. By comparison with the traditional percolation model [44], the Achlioptas process presents a more abrupt transition when applied to different network topologies [55, 56, 57, 19, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69]. As potential applications, the Achlioptas process has been recently associated to the growth dynamics of Protein Homology Networks [70] as well as to the formation of bundles of single-walled nanotubes [71]. In this chapter, we shortly present the classical bond percolation, we discuss its locality and briefly analyze its evolution on a square lattice. In the following, we present the explosive percolation model proposed by Achlioptas *et al.* [12] and its implementation on a square lattice.

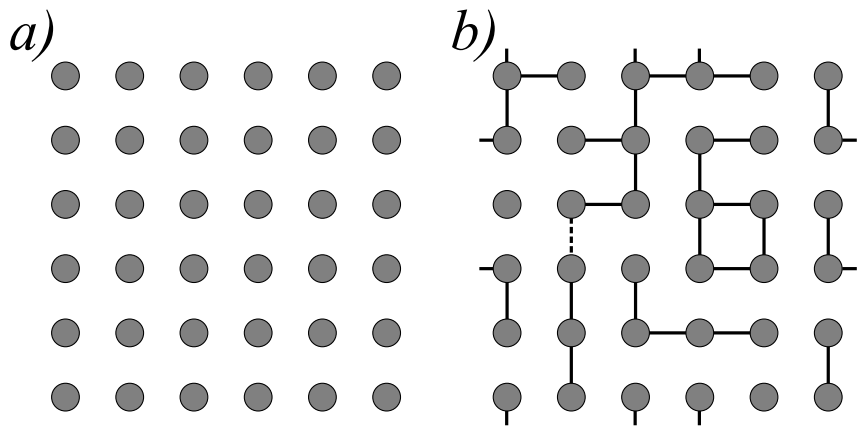


Figure 5.1: Example of a percolation. (a) The lattice nodes are occupied, but the absence of links renders the network complete disconnected. (b) As the links are gradually occupied, the global connectivity of the network can be achieved with the insertion of a single link, for example, the link connecting the nodes at the center of the second column (dashed).

5.1 What's percolation?

Imagine a sheet of paper where small holes were pierced in such a way that they look like a square lattice (Figure 5.1(a)). One can successively choose randomly pairs of neighboring holes and connect them by cutting the paper. After some steps of this "algorithm" one of the possible configurations is depicted in Fig. 5.1(b), before it breaks apart. If in the next step the pair of nodes (holes) placed at the center of the second column is chosen, this sheet of paper will break into two, losing its global connectivity. There are various ways to cut the sheet of paper with very few or very large number of cuts, but configurations like the one in Fig. 5.1 appear with much higher probability.

Percolation theory deals, specifically, with problems like the one posed in the preceding paragraph, the problem of long-range connectivity in macroscopic systems. In a formal sense, percolation corresponds to an equilibrium phase transition, which became a crucial paradigm of criticality of continuous transitions, due to its simplicity and broad range of application.

In section 1.2, we presented a percolation problem when we discussed the structural evolution of the Erdő-Rényi random network. Due to its generality the percolation problem can be stated in many ways and in different network models. In the following, in order to illustrate the beauty and general features of the percolation problem, we present the bond percolation problem, which is the type of percolation problem treated in the following sections and chapters, on a square lattice.

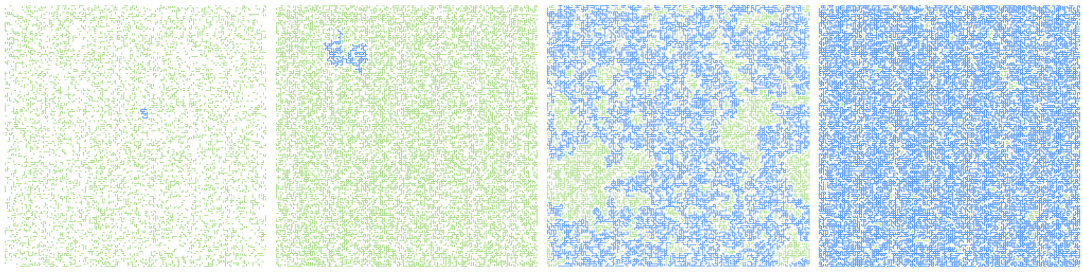


Figure 5.2: An ensemble's member of the bond percolation process on a square lattice of $L = 124$. Suddenly, a spanning cluster appears at the vicinity of the percolation threshold $p_c = 0.5$. Before this critical point the clusters are disconnected. After its emergence, the spanning cluster rapidly becomes predominant ($p > p_c$).

5.2 The traditional percolation process

In the percolation example presented for the evolution of the ER network all the nodes of the network exist, but they are not connected. At each step of the evolution of the ER network one unoccupied link is chosen and the pair of nodes that define this link becomes connected. In this situation, for a fully disconnected network, there are no constraints to the spatial position of the nodes, and one node can be connected to any other node of the network. This is not true for lower dimensional cases, e.g., when the percolation occurs on a 2-dimensional square lattice. Here we present the simplest case, in which a node can be connected only to its first neighbors. One can define a control parameter p as the fraction of occupied links. At the beginning of the process, $p = 0$, any link is empty and there is no global connectivity in the network. For $p = 1$, all the links exist and any path between any pair of node is possible. It is possible to define the order parameter for this percolation process as the fraction of nodes belonging to the largest cluster, P_∞. For small values of p, where the lattice is sparsely occupied, P_∞ must be zero if compared with sufficiently large lattice sizes, but with the increase of p, there is a critical value of $p = p_c$, called the percolation threshold, for which $P_\infty > 0$. Figure 5.2 shows the evolution of the bond percolation on a square lattice. The links belonging to the largest cluster are highlighted in blue, and any other occupied link is highlighted in green. When $p < p_c$ the largest cluster size is small when compared with the system size. At the neighborhood of the transition, the largest cluster expands throughout the lattice, providing global connectivity for the lattice. When $p > p_c$, there is almost no link that does not belong to the largest cluster.

For bond percolation on a square lattice it is easy to show that the percolation threshold is $p_c = 1/2$ using a simple argument, which makes use of a *dual lattice*. The position of the nodes

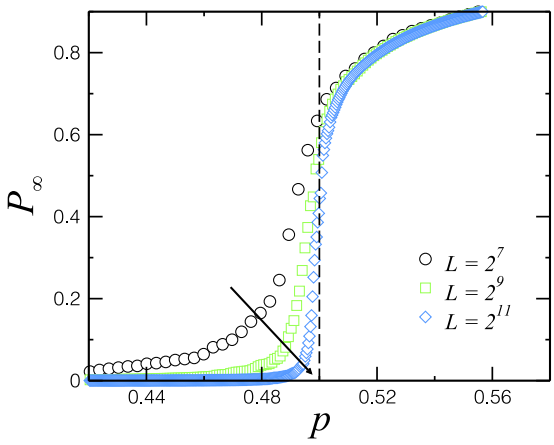


Figure 5.3: Behavior of the order parameter P_∞. With the increment for the lattice size L, the value of p_c becomes close to 0.5, the thermodynamic limit of the transition. For sufficient large systems, Eq. 5.1 is valid, and the microscopic details of the system do not affect the behavior of macroscopic quantities.

of a dual lattice are placed at the centers of the squares defined by the nodes of the "real" lattice. Therefore it is easy to argue that the links of the dual lattice cross the links of the original lattice. The argument is the following: The links of the dual lattice are occupied if, and only if, the links of the original lattice are not occupied, thus the control parameter p^* of the dual lattice is $1-p$. Thus, the dual lattice percolates only if the original lattice does not percolate. Consequently, because of the symmetry between the original and the dual lattice, the percolation threshold is $p_c = 1/2$.

This is easier to see graphically, either by Fig. 5.2, or by analyzing the characteristic behavior of the order parameter P_∞. The behavior of P_∞ near the percolation at the thermodynamic limit ($L \to \infty$) is given by

$$P_\infty \sim (p - p_c)^{-\beta}, \tag{5.1}$$

which is an analogous to the equation that describes the behavior of the order parameter for a temperature-induced transition near the critical temperature T_c. Here $(p - p_c)$ is equivalently to $(T - T_c)$, and $\beta = 5/36$ for the 2-dimensional case. However, for our case of a finite 2 - dimensional lattice, the behavior of P_∞ can not be so simple. For a finite lattice with L^2 nodes, a spanning cluster can appear with only $2L$ links (two straight lines connecting the opposite sides of the lattice). So, even for small values of p, the percolation probability becomes larger than zero. Figure 5.3 shows the behavior of P_∞ versus p for different lattice sizes. Remarkably, for small finite sizes, the order parameter P_∞ increases with p far from $p_c = 1/2$ (highlighted by the vertical dashed line), which does not agree with Eq. 5.1. However, this finite size effect disappears for large L, when the value of the order parameter P_∞ remains zero for a broader

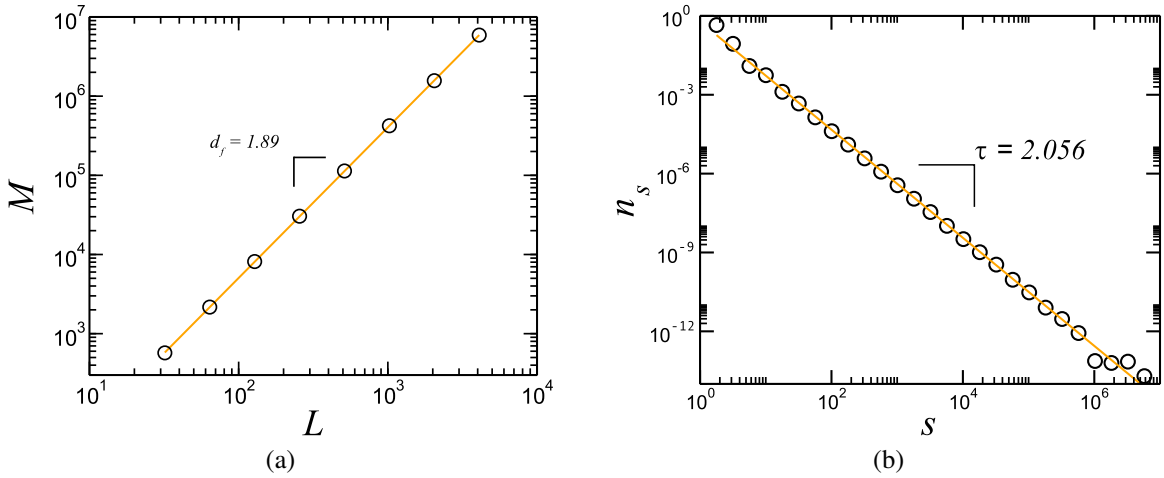


Figure 5.4: (a) Mass of the largest cluster M as a function of L. One can notice that the order parameter scales with a power law of L with an exponent of $d_f \approx 1.89$,the fractal dimension of the spanning cluster, revealing a scale invariance with the system sizes. (b) The distribution of cluster sizes n_s, which also follows a power law, but with s. At criticality, all the scales of cluster size are present, another consequence of the scale invariance of a typical second-order phase transition.

range of p, and the behavior described by Eq. 5.1 emerges near the percolation threshold. A natural question arises from the analysis of Fig. 5.3: How is the finite size behavior of the largest cluster at the percolation threshold? Figure 5.4(a) shows the behavior of the mass of the largest cluster M as a function of L. One can notice that

$$M \sim L^{d_f}, \tag{5.2}$$

where $d_f \approx 1.89$ is the fractal dimension of the spanning cluster. For this random displacement of bonds, clusters of different sizes s are formed at criticality, and the spanning cluster can present a complex shape. Figure 5.4(b) shows the cluster size distribution, which also follows a power law

$$n_s \sim s^{-\tau}, \tag{5.3}$$

with $\tau \approx 2.05$. A pictorial example of this last property is depicted on Fig. 5.5, where several clusters of different sizes are shown (in black we highlight the largest cluster).

Briefly, we presented three critical exponents. The first, β, which describes the divergence of the order parameter near the transition, the fractal dimension d_f of the spanning cluster, and finally the *Fisher exponent* τ for the cluster size distribution. They are just a fraction of the several critical exponents that describe the behavior of many macroscopic (thermodynamic) parameters of percolation near the transition, usually defined by relations similar to Eq. 5.1. Theses relations, as stated earlier by Eq. 5.1, emerge only in the thermodynamic limit, which

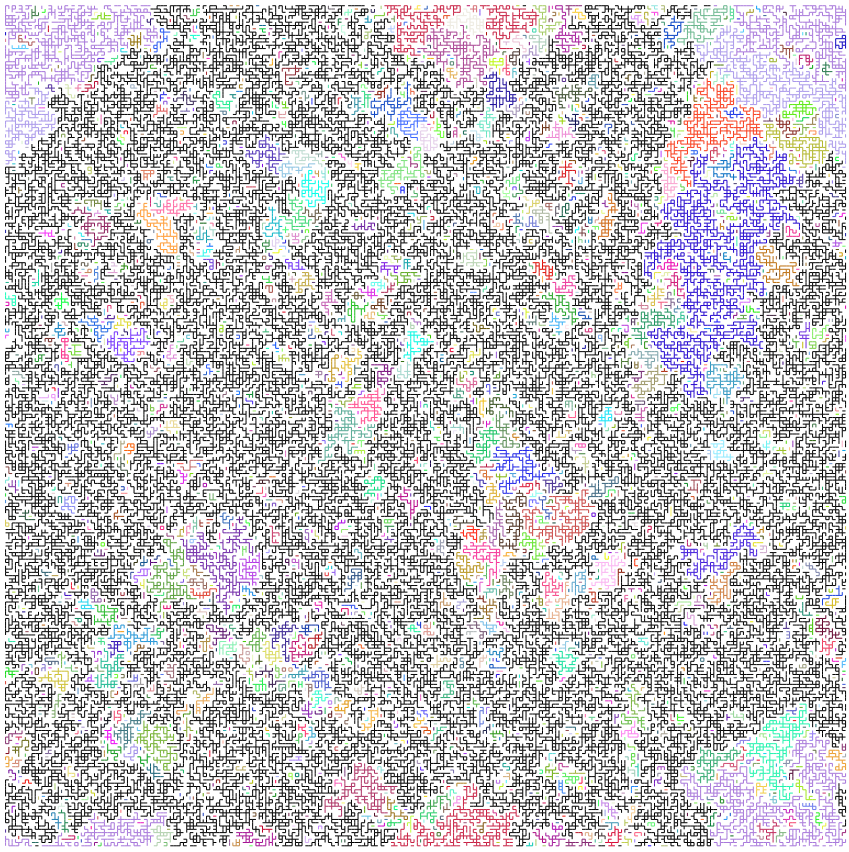


Figure 5.5: Example of bond percolation process at the critical limit ($p = p_c = 0.5$), in a 256×256 square lattice. The black bonds belong to the largest cluster.

is natural since we are dealing with macroscopic parameters. The emergence of power laws describing the behavior of macroscopic quantities is possible since, for sufficiently large systems, the microscopic details of the system are negligible and only the long length scales are important. A consequence of losing microscopic degrees of freedom is that distinct systems present similar macroscopic behavior, characterized by the same critical exponents. This property is know as *universality* [1]. Systems that present the same set of critical exponents belong to the same *universality class*, for example the bond and site percolation. If instead of random occupation of empty links in the lattice, we occupy the nodes of an empty lattice, and at the transition we measure macroscopic quantities, we will find the same critical exponents, e.g., $d_f = 1.89$. These two percolation models are different at the small scales, however, they seem the same at large scales. The other remarkable feature of continuous phase transitions is the ubiquitous presence of power laws related with the scales of the system. By analyzing Eq. 5.2 and Eq. 5.3 one can say that the bond percolation presents *scale invariance*. This roughness and fractal-looking of the spanning cluster, the presence of all cluster sizes at criticality, both

[1] This notion becomes clear when considered under the light of the renormalization group theory [72].

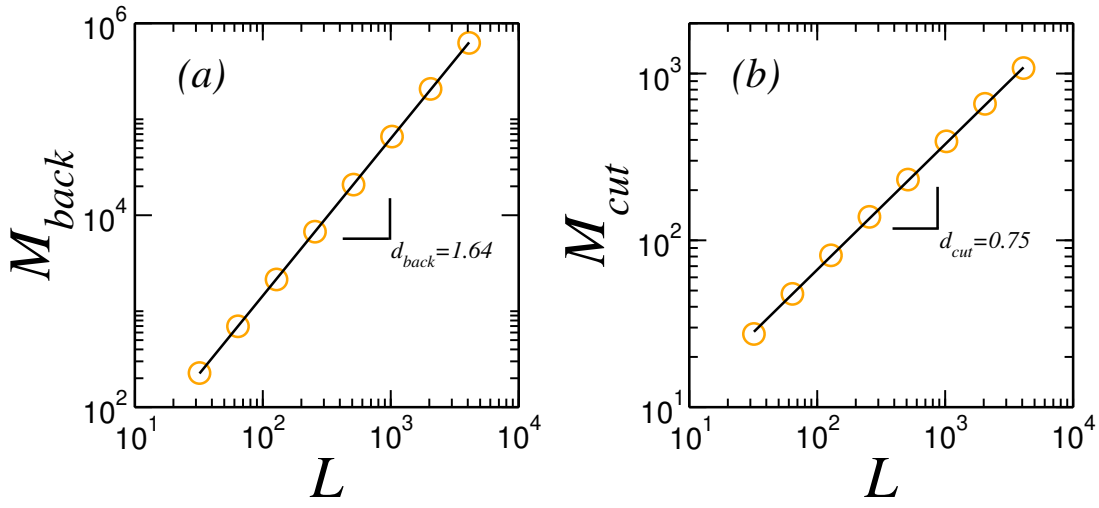


Figure 5.6: Scale dependence of the mass of the backbone (a) and the mass of cutting bonds (b). Both follow power laws with L, defining the fractal dimension of the backbone ($d_f = 1.64$) and of the cutting bonds ($d_f = 0.75$).

come from a hidden symmetry of the problem, namely, percolation is invariant under a change in *length scale*.

When we look at the internal structure of the spanning cluster, this self-similarity plays an important role. The spanning cluster is responsible for the global connectivity of the lattice. If one applies an electrical potential difference between two opposite sides of the lattice, the transport of charge will occur through the spanning cluster. The number of bonds that conduct current in the cluster constitutes the so-called *conducting backbone* of the spanning cluster. One can imagine that, at the thermodynamic limit, the spanning cluster appears precisely at $p_c = 1/2$, thus there are some bonds, occupied just before the percolation threshold that, if removed, will brake the global connectivity of the spanning cluster. These bonds are called *cutting bonds* [2]. Interestingly, the mass of the conducting backbone M_{back} and the mass of cutting bonds M_{cut} scales with the linear system size L as power laws, $M_{back} \sim L^{d_{back}}$ and $M_{cut} \sim L^{d_{cut}}$ (see Fig. 5.6).

5.3 Explosive percolation

5.3.1 The product rule for explosive percolation

The percolation transition is a type of second-order transition, where the order parameter changes continuously beyond the critical threshold. Many thermal induced transitions on mag-

[2] The cutting bonds are also called red bonds.

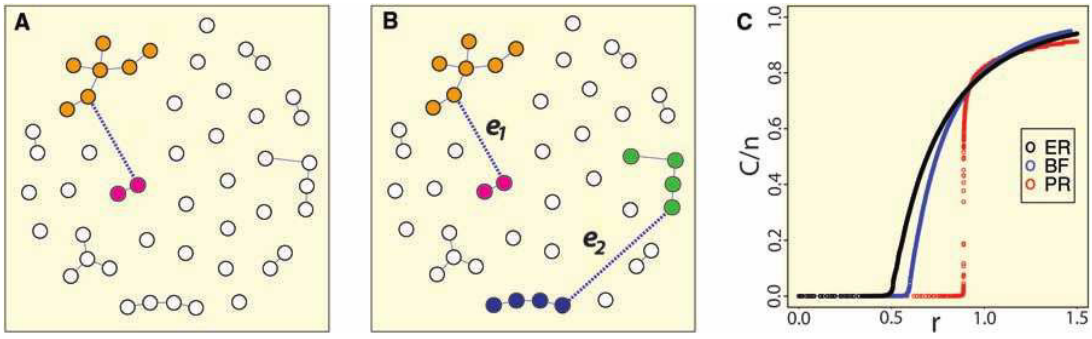


Figure 5.7: Comparison between the traditional percolation (a) and the explosive percolation (b) on a fully-connected network (Erdös-Rényi). In (c) the evolution of the order parameter is presented. One can notice the different transition of the PR (red) if compared with the traditional (black).

netic systems are also within this class of transitions. However, the transition of an amount of liquid water into a block of solid ice at $0°C$, perhaps the most familiar phase transition, does not belong to this family of phase transitions. The liquid-solid transition of water, which involves latent heat, is a first order phase transition, where the system (water) either absorbs or releases a fixed amount of energy, but during this process the temperature $T = T_c$, the control parameter, remains constant.

Due to the simplicity of the percolation transition, the question of whether it could present a first order phase transition was an open question until Achlioptas and his collaborators [12] presented a modification to the rule of the classical percolation process on the ER model of random network, which led to an abrupt transition called explosive percolation. At their seminal paper, it is proposed the product rule (PR) for percolation. Figures 5.7(a-b) depict the difference between the traditional percolation and explosive percolation. At the traditional percolation process, one unoccupied bond is chosen one by one connecting two nodes of a cluster, or connecting two different clusters per step [3]. For the PR process, at each step, two different unoccupied bonds $\{e_1, e_2\}$ are chosen, but only one is selected and inserted in the network, namely the one that minimizes the product of the sizes of the clusters it connects [4]. Figure 5.7(c) shows that the PR postpone the critical threshold, leading to an abrupt transition. Interestingly, the traditional percolation process is a local process, since the insertion of a bond is independent of the occupation of the others. In the case of explosive percolation this is not true. At each step it is necessary to know the distribution of cluster sizes for the entire network.

[3] An isolated node is a cluster of unitary size.

[4] If one bond is internal to a cluster, the squared value of this cluster size is considered by the PR.

5.3.2 Explosive percolation on a square lattice

The PR process on an $L \times L$ square lattice was first considered by Robert Ziff [67, 68], where the links were added randomly and one at a time, tracking the internal cluster structure of clusters during the process [73]. Using finite size scaling analysis, Ziff presented a set of critical exponents. In particular, he showed that the scaling of M can be described in terms of the fractal dimension $d_f \approx 1.95$ [68]. This result was confirmed later by Andrade *et al.* [69]. Moreover, they studied the transport behavior of the exploding percolation cluster on a square lattice, using the best-of-m generalization of the PR model [64]. They showed that for diverse values of m the mass of the spanning cluster M_{clus}, and the mass of the conducting backbone M_{back}, and the mass of the cutting bonds M_{cut} are all described by power laws of the system size L. The behavior of the fractal dimension of the spanning cluster d_{clus} increases monotonically with m due to the delay of the percolation transition, since the transition occurs with a higher fraction of bond occupation. Interestingly, the scaling exponents d_{back} and d_{cut} displays a non-monotonic behavior with m. In particular, they showed that for the original PR process, the fractal dimensions are given by $d_{clus} \approx 1.95$, $d_{back} \approx 1.52$, and $d_{cut} \approx 1.02$ respectively, which correspond to different exponents when compared with the ordinary percolation.

5.4 Local vs. global percolation process

The remarkable difference between the ordinary percolation process and the explosive percolation resides on the nonlocal feature of the product rule. For the traditional percolation, the only information accessible is if the bond chosen to be inserted in the lattice already exists. On the other hand, the abrupt of the PR process is possible only by controlling the largest cluster size [58]. Since at each step the information about the cluster size distribution of the entire network must be available to apply the PR, in the thermodynamic limit there is no loss of information, and the macroscopic parameters are affected by the microscopic details of the network. Therefore, the PR breaks the universality class of classical percolation.

Regardless of initial claims of a first order transition in the Achlioptas process, however, recent analytical and numerical works [74, 75, 76, 77] have demonstrated that the alleged "Explosive Percolation" process actually displays a continuous, i.e., a second-order phase transition. This apparent drawback of the PR model has been somehow overstated, in the sense that the model proposed by Achlioptas and his colleagues [12] certainly represents an original and interesting contribution to the field. For instance, much less importance has been given to the nonlocal attributes of the Achlioptas process. This will be investigated in the next chapters.

6 A Hamiltonian approach for explosive percolation process

In this chapter we investigate what are the basic principles that lead to the abrupt phase transition observed in the explosive percolation model. First we name *merging bonds* those edges that connect nodes in distinct clusters, while *redundant bonds* are edges connecting nodes in the same cluster. We show that two conditions are necessary for obtaining a first-order transition in a growth process where bonds are included one by one, namely, the process has to favor the inclusion of bonds that keep all the clusters at about the same size, and the process has to preclude the introduction of redundant bonds, at least below the critical point. More precisely, merging bonds must be introduced with much higher probability than redundant bonds. In Fig. 6.1 we show a pictorial description of these two ingredients.

6.1 Hamiltonian model for explosive percolation

In order to validate our hypothesis, we propose an extension of the percolation model that describes a general growth process in the space of graphs. For this, we define a Hamiltonian that depends on the graph G describing the network. The probability of finding the system in a certain state G will be given by $P(G) = Z^{-1}\exp(-\beta H(G))$, where $Z = \sum_G \exp(-\beta H(G))$. A simple form for a Hamiltonian that includes the two ingredients is

$$H(G) = \sum_{i\in\mathbf{C}} s_i^2 + \ell_i s_i^\alpha, \tag{6.1}$$

where the sum is over the entire set of clusters $\mathbf{C}$, s_i is the number of nodes in cluster i, and ℓ_i is the number of redundant bonds added to this cluster. If the number of bonds in the cluster is b_i, we have $\ell_i = 1 + b_i - s_i$. Note that each time one includes a redundant bond, one also closes a new loop in the cluster, thus ℓ_i is also a measure for the number of loops in the cluster.

We can now simulate a process of cluster growth controlled by the Hamiltonian of Eq. (6.1). This is performed by starting with a network of N nodes without bonds, so that each vertex

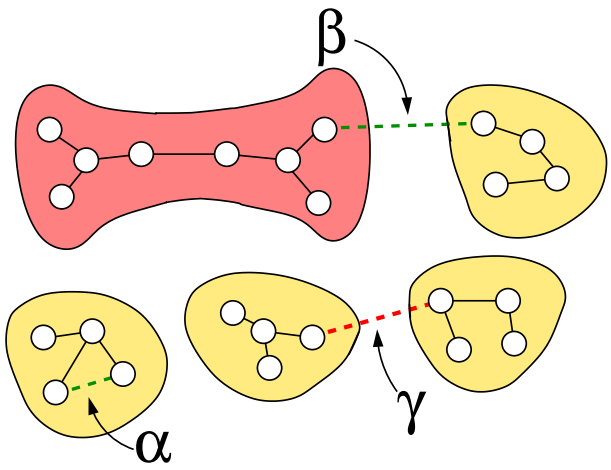


Figure 6.1: Two ingredients for explosive percolation. Here we show a possible configuration for a growth process where, at each step, any unoccupied bond can be introduced in the graph. For instance, in this figure we show three bonds that could be added in the next step, namely, α, β, and γ. The two ingredients for obtaining a sharp transition are the following: (*i*) bonds that keep the clusters approximately at the same size are favored over bonds that result in larger size discrepancies; and (*ii*) bonds that connect nodes in the distinct clusters (*merging bonds*) are favored over bonds that connect nodes in the same cluster (*redundant bonds*). Thus, among the bonds indicated, α has the smallest probability due to condition (*ii*), β is not accepted due to condition (*i*), and the most probable is the γ bond.

initially belongs to a different cluster. At each step, a new bond can be placed between any pair of nodes not yet connected. The probability of including a particular bond b is given by $\Pi_b \sim \exp(-\beta \Delta H_b)$, where ΔH_b is the energy change after including this bond. Such a growth model emulates equilibrium configurations of graphs following the Eq. (6.1) and having a given number of bonds N_b. However, since the removal/rewiring of bonds is not considered during growth, this corresponds to an out-of-equilibrium process. Consequently, some differences should be expected between the observed results and the actual thermal equilibrium.

6.2 Asymptotic behavior of the hamiltonian model

For small values of α, redundant bonds are favored over merging bonds, while for large values of α the opposite takes place. Let us investigate the asymptotic behavior in the two different scenarios. If redundant bonds are favored, one might expect that a new merging bond will be included only after the addition of all possible redundant bonds. Since clusters of equal size minimize Eq. (6.1), we can assume that, for low temperatures, all clusters have about the same size S, so that fully connected sub-graphs with $S(S-1)/2$ bonds are formed with $\ell = (S-1)(S-2)/2$. After adding the next bond, two of these clusters shall merge to form a

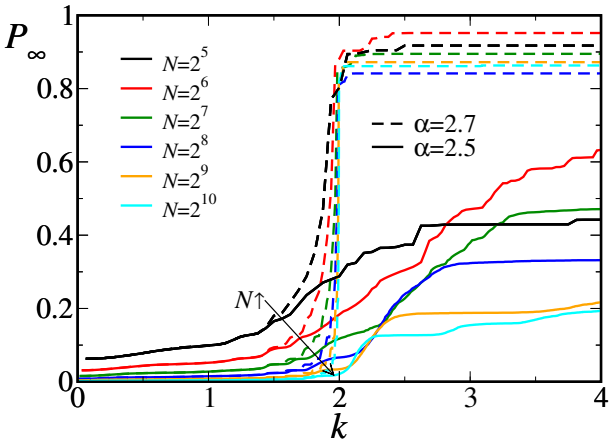


Figure 6.2: Transition to explosive percolation. When the process favors redundant bonds, $\alpha = 2.5$, the largest cluster follows a slow continuous growth. When the merging bonds are favored, $\alpha = 2.7$, the system displays an abrupt transition around a critical connectivity $k = 2$. In this case the transition becomes sharper as the system size increases, suggesting a first-order transition type of behavior. In all simulations, we use $\beta = 1.0$ and take an average over 1000 realizations of the growth process.

new largest cluster, into which redundant bonds can be included. At this point we can calculate the energy variation for a redundant bond, $\Delta H_r = (2S)^\alpha$, and for a merging bond between pair of clusters, $\Delta H_m = 2S^2 + (S-1)(S-2)(2^\alpha - 1)S^\alpha$. Surprisingly, for any value of α, in the asymptotic limit of very large clusters, $S \to \infty$, merging bonds have higher energy variation than redundant bonds, and the growth process with fully connected clusters is stable.

The situation becomes quite different when the presence of merging bonds is favored. As before, we use that all clusters have approximately the same size S. However, without redundant bonds, the clusters are all tree-like with exactly $S-1$ bonds, and $\ell = 0$. At this point, we have $\Delta H_r = S^\alpha$, and $\Delta H_m = 2S^2$. Thus, for large S, the inclusion of merging bonds will lead to smaller energy variations, as long as $\alpha > 2$. We then conclude that, in the large cluster limit, $S \gg 1$, both scenarios are stable for $\alpha > 2$.

6.3 Simulation results

The evolution of the system towards tree-like or fully connected clusters is determined at the beginning of the growth process. Considering that $S = 3$ represents the minimal size necessary for the inclusion of a redundant bond, we obtain $\Delta H_r = 3^\alpha$ and $\Delta H_m = 2 \times 3^2 = 18$. Thus, merging bonds become more probable when $\alpha > \ln(18)/\ln(3) = 2 + \ln(2)/\ln(3) \approx 2.63$, which corresponds to a threshold condition above which the system exhibits an abrupt transition. One should note that this is an approximate result, since we do not account for

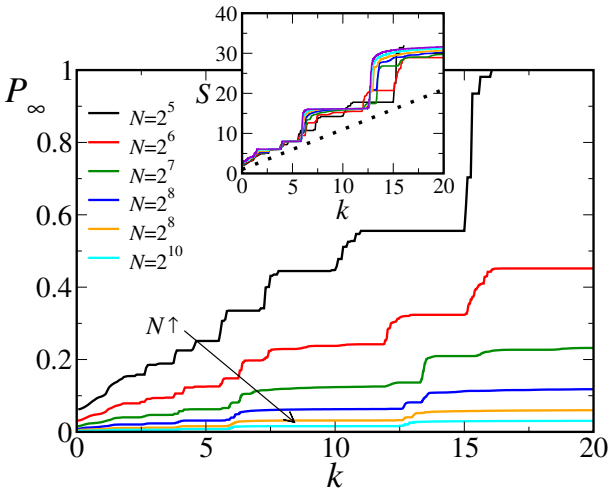


Figure 6.3: Growth process when redundant bonds are favored. Here we show results for $\beta = 1.0$ and $\alpha = 2.0$. Since in this situation merging bonds are less likely to be included, the graph has to reach states where it splits in several fully connected sub-graphs, before a new merging bond is introduced. When the merging bond is included, a new and larger cluster is created. This explains the presence of discontinuous jumps in the size of the largest cluster. Assuming that the system consists of only fully connected clusters of the same size, we obtain the dotted line shown in the inset, $S = k+1$. This condition corresponds to the minimum bound for the simulation results, that approximately follows this theoretical prediction. Since the largest cluster S is finite for any finite connectivity k, the system does not display a percolation transition.

fluctuations in the cluster size distribution. However, as shown in Fig. 6.2, the results for $\alpha = 2.5$ and 2.7 indeed confirm the change in behavior from a sharp transition for the larger value of α to a slow continuous growth for the smaller value. Note also that the threshold value for α is not universal and could be readily changed by adding a multiplicative constant to any of the two terms constituting the Hamiltonian of Eq. (6.1).

Let us examine in more detail the scenario for a small value of $\alpha = 2.0$. In Fig. 6.3 we show that the fraction occupied by the largest cluster P_∞ systematically increases with the average connectivity k, with a growth rate that decreases with system size N. The inset of Fig. 6.3 shows the same results, but for the size of the largest cluster $S = NP_\infty$. One can see that S follows approximately a linear growth with the connectivity k. In this scenario, a merging bond is expected to be placed only when all clusters become saturated with redundant bonds. If we now use that all clusters have about the same size S, we obtain $k = S - 1$, which corresponds to the dotted line in the inset. The deviations of the numerical results from this prediction should be expected. In the growth model, the merging of two clusters can only double the value of S, so that the values of S at the plateaus observed in the curves are approximately powers of two. However, we see that the curves always approach the dotted line before doubling S. This

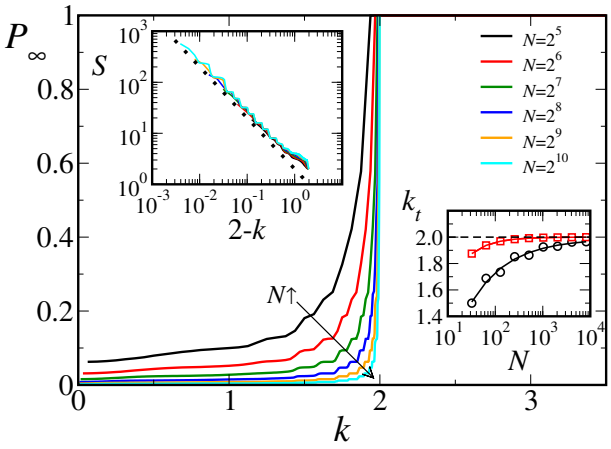


Figure 6.4: Growth process when merging bonds are favored. For $\beta = 1.0$ and $\alpha = 3.0$, the system does not include redundant bonds and all clusters remain tree-like until the critical point $k_c = 2$ is reached. Supposing that the system comprises only clusters of the same size S, we have that $S = 2/(2-k)$ for the case of trees. This relation works as a minimum bound for the simulation results, as shown in the inset on the left. Thus we have the critical condition $k_c = 2$, where the size of the largest cluster diverges to occupy the whole network. For $k < k_c$ the largest cluster remains finite and its occupation fraction P_∞ goes to zero as the system size grows, characterizing a typical first-order transition. The inset on the right shows the threshold connectivity k_t to obtain a largest cluster greater than the square root of the system size, $S > N^{1/2}$ (black circles), and greater than half the system size, $S > N/2$ (red squares). The red line follows $k = 2-4/N$, the expected behavior for the connectivity where $S = N/2$. The black line is a fit of the form $k = p_1 + p_2 \times N^{-p_3}$, with $p_1 = 1.99 \pm 0.03$, $p_2 = 3.14 \pm 0.02$, and $p_3 = 0.53 \pm 0.05$. In the limit $N \to \infty$ both curves converge to $k \approx 2$, that is in the thermodynamic limit we observe at $k = 2$ a discontinuous transition in the order parameter from a vanishing fraction, $P_\infty \sim N^{-1/2}$, to a finite fraction, $P_\infty = 1/2$, confirming the approach to a first-order transition.

linear growth for P_∞, with a slope that decays with the system size N, indicates that, in the thermodynamic limit, this system does not undergo a percolation transition, namely, $P_\infty = 0$ for any finite value of k.

Figure 6.4 shows results for $\alpha = 3.0$. Here we are in the scenario where the clusters grow as loopless trees. In this case, the system undergoes a transition that becomes sharper as the the number of nodes N increases. Again, if we assume that the system is divided in trees of the same size S, we obtain $S = 2/(2-k)$, as indicated by the dotted line in the inset. As before, the size S increases in steps due the out-of-equilibrium nature of the growth process. Strikingly, the theoretical relation for the equilibrium state still provides a consistent prediction for the lower bound of the largest cluster size. Since S remains finite for any $k < 2$, and at the critical point $k = 2$ a tree that spans all the system is formed, it follows that the order parameter P_∞ displays a first-order transition in the thermodynamic limit.

In addition, we performed additional simulations with our model implemented on a two-

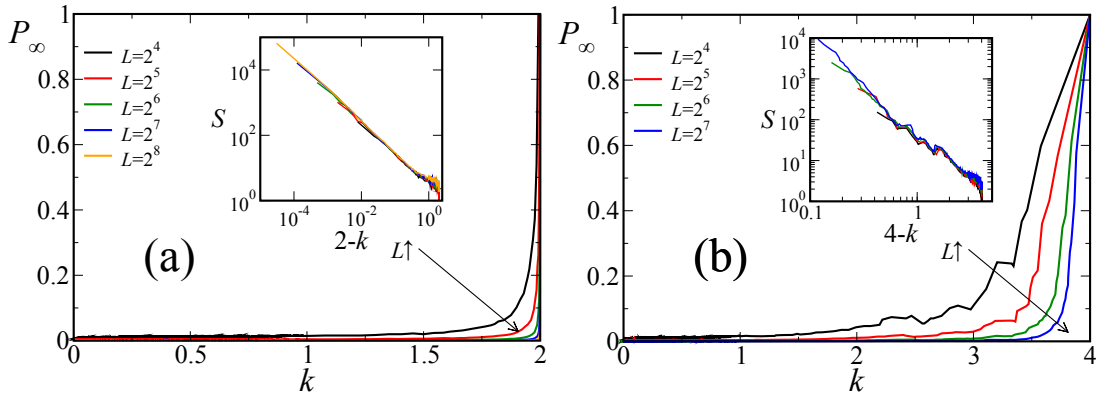


Figure 6.5: Results of the hamiltonian model on square lattice of size L. The main panels show the average fraction accupied by the largest cluster versus the average connectivity k, while the insets show log-log plots of the average size of the largest clusters versus the distance from the transition point. We use $\beta = 1.0$ in both cases, with $\alpha = 4.0$ on panel (a) (clusters grow as trees), and $\alpha = 2.0$ on panel (b) (clusters are saturated with internal bonds). The results in (a) are essentially identical to those of Fig. 6.4. As shown in (b), the finite dimensionality has some effect on the saturated case, $\alpha = 2.0$. Since there is a maximum possible connectivity $k_{max} = 4$ on the square lattice, the system has to form an infinite cluster when it approaches this limit. This effect is not relevant in the case of fully connected graph, since there is no upper limit for the connectivity in the thermodynamic limit.

dimensional square lattice. As shown in Fig. 6.5(a), for large values of α the results are essentially identical to those of Fig. 6.4, implying that the results are independent on the dimensionality of the underlying lattice on which the process is applied, that is, an abrupt transition is observed at $k_c = 2$, and the size of the largest cluster diverges as $(k_c - k)^{-1}$. For small values of α (see Fig. 6.5(b)), similar results are also obtained. However, differently from the fully-connected graph, there is an upper limit for the connectivity of the nodes in lattices of low dimensionality. On the square lattice, for instance, the size of the largest cluster remains finite until the saturation is reached at $k = 4$, where all bonds are included in the lattice.

6.4 Equilibrium solution

As already mentioned, this growth process bears some differences with a thermal equilibrium state of graphs with the proposed Hamiltonian Eq. (6.1) at low temperatures. In fact, for $k \to 2$ there is always an energy gain in breaking large trees into smaller highly connected graphs. One may then ask whether the sharp transition observed in the simulations is just a feature of the irreversible growth process or could be reproduced in an equilibrium statistical framework. We now show that in fact in the limit of large α we can obtain an exact equilibrium solution that exhibits a first-order transition.

If we impose that all clusters in the system are loopless trees, $\alpha \to \infty$, it is possible to enumerate all possible ways in which the network can be divided into a set of clusters of a given size. Let Ω represent the number of ways that a fully connected graph can be divided into trees with n_i trees of size $i = 1, 2, 3...$ We then have

$$\Omega = N! \prod_s \left(\frac{T_s}{s!} \right)^{n_s} \frac{1}{n_s!}, \tag{6.2}$$

where N is the total number of nodes in the network, and T_s is the number of trees that span a fully connected graph of size s, given by Cayley's formula, $T_s = s^{s-2}$ [78]. Since all clusters are trees, we can relate the number of clusters N_c with the system size N and the average connectivity k as $N_c = N(1-k/2)$. Therefore, given a fixed value of k, the values of n_s obey the following two constraints: $\sum n_s = N_c$ and $\sum s n_s = N$, where the sum is over all possible cluster sizes s. In our generalized percolation model, we still need to impose a fixed energy value, $\sum E_s n_s = E$, where $E_s = s^2$ is the energy of a tree with size s. Using Lagrange multipliers, η, λ, and β to deal with each of these constraints, we can find the cluster size distribution that maximizes Ω,

$$n_s = e^{\eta - \beta s^2 - \lambda s} \frac{s^{s-2}}{s!}. \tag{6.3}$$

The critical condition takes place when the distribution Eq. (6.3) diverges. One can verify that, for $\beta = 0$, this happens when $\lambda = \lambda_c(\beta = 0) = 1$. The critical connectivity can then be determined as

$$k = 2\left(1 - \frac{N_c}{N_s}\right) = 2\left(1 - \frac{\sum n_s}{\sum s n_s}\right), \tag{6.4}$$

yielding $k_c = k(\lambda_c = 1, \beta = 0, \eta \to \infty) = 1$. Note that, at the critical condition $n_s \sim s^{-5/2}$, the fraction occupied by the largest cluster follows $P_\infty = k - k_c$, thus reproducing the known critical properties of the standard Erdös-Rényi model [13].

For $\beta > 0$, the distribution always converges unless $\lambda \to -\infty$. From Eq. (6.4), we obtain that $k_c = k(\lambda \to -\infty, \beta, \eta \to \infty) = 2$. For $k < k_c$ all clusters are finite trees, therefore occupying a vanishing fraction of the network. At $k = k_c = 2$, a giant tree spans the entire network, characterizing a first-order transition. Of course, this simple approach to the problem is only possible due to the imposition of tree-like clusters. The general enumeration of connected graphs with any number of redundant bonds is not a simple task [79], and the cluster size distribution in this generalized condition might be quite different. However, at least in the situation where redundant bonds are not present, explosive percolation can be duly explained within the framework of equilibrium statistical mechanics.

6.5 conclusion

In conclusion, we have shown that two simple conditions, namely, the absence of loops and the imposition of clusters of similar sizes, are the only necessary ingredients for a percolation process to display first-order transition in the size of the infinite cluster as function of the average degree of the network. We argue that both conditions are implicitly present in the explosive percolation model proposed in Ref. [12]. We emphasize that these conditions are essentially non-local, namely, the probability of adding a particular bond depends on the global structure of the graph. Moreover, our model provides a simple connection between explosive percolation and equilibrium statistical physics, leading to a clear interpretation of the mechanisms behind this growth process. Finally, other possibilities for the energy function can also be investigated in different contexts, revealing a whole new family of percolation-like models.

7 *Nonlocal Product Rules for Percolation*

In this chapter we introduce a generalization of the PR model in which the range of its nonlocal features can be systematically controlled. This is carried out by imposing that pairs of bonds for selection are randomly chosen according to a probability that decays as a power law of their Manhattan distance, namely the distance measured as the number of connections separating the sites in a regular lattice. This physically plausible assumption is inspired on a geographical model for complex networks where long-ranged shortcuts are incorporated to regular lattices. Such a conceptual construction has been extensively used as a way to explain the emergence of optimal navigation and efficient transport in small-world systems [11, 80, 18, 81, 82, 83]. As a consequence of the selection rule adopted here, we show that the scaling properties of the system become dependent on the specific value of the corresponding power law exponent. A transition is then revealed from the traditional to the PR percolation behavior. Moreover, the results of our extensive numerical simulations provide strong evidence for the fact that the AER model [84], when applied to regular lattices in two-dimensions, falls in the same universality class as ordinary percolation.

7.1 The model

As shown in Fig. 7.1, our bond percolation process takes place on a square lattice of size L. At each step, two sites i and j are randomly selected with probability $P(r_{ij}) \sim r_{ij}^{-\alpha}$, where r_{ij} is the Manhattan distance between i and j, measured as the number of connections separating these sites in the underlying regular lattice [1]. From each site i and j, one bond is then selected among its four adjacent edges, namely e_i and e_j, respectively. If at least one of these two bonds is already occupied, the entire process of selection is restarted. If not, following the product

[1] In fact, to account for the number n_r of neighbors that are at the Manhattan distance r from a particular site in a square lattice ($N_r = 4r$), the random distance r mast be generated following a power law distribution proportional to $r^{(1-\alpha)}$

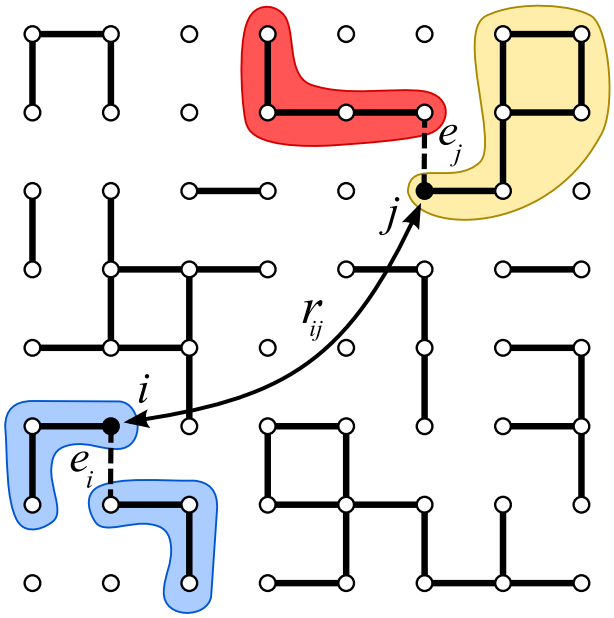


Figure 7.1: Pair of unoccupied bonds e_i and e_j (dashed lines) randomly selected for the application of the product rule, according to the probability $P(r_i) \sim r_j^{-\alpha}$, where r_{ij} is the Manhattan distance between sites i and j (black circles), and α is a variable exponent. Following the PR model, the bond e_i, merging the two clusters in blue (with 3 sites each), becomes occupied. The bond e_j would merge the the clusters in red (4 sites) and yellow (6 sites), but remains unoccupied.

rule, weights are assigned to each of these bonds, in proportionality to the product of the size (number of sites) of the clusters they would potentially connect. In the case that a bond connects two sites in the same cluster, the weight is equal to the square of the cluster size. The bond associated with the smallest weight becomes occupied, while the other stays unoccupied, but can be selected again in later steps.The percolation process stops when one among all clusters, namely the spanning cluster, connects the lattice from top to bottom [44]. At that point, the fraction p of occupied bonds corresponds to the percolation threshold p_c.

Our model displays two distinct limiting behaviors, depending on the exponent α. For $\alpha = 0$, we recover the usual PR, for which the preliminary random selection of the bonds e_1 and e_2 constitutes a highly nonlocal process [63, 74]. In the limit of $\alpha \to \infty$, the bonds e_1 and e_2 are always adjacent, which corresponds to the AER process proposed in Ref. [84], but applied here to regular lattices. Although more spatially restricted than the PR process, the AER is still nonlocal, since it requires information about the masses of the joining clusters [74]. As we show later, the finite low-dimensionality of the square lattice employed here attenuates even further the already weaker nonlocal features of the AER process.

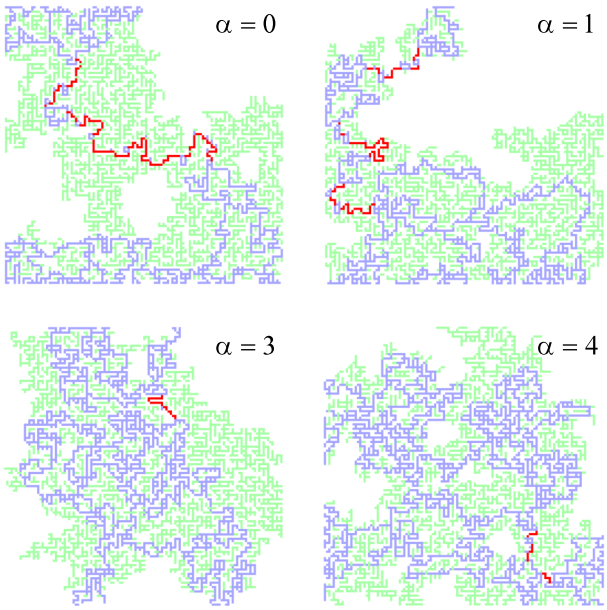


Figure 7.2: Snapshots of the largest cluster at p_c for different values of the exponent α, and a lattice size $L = 64$. The bonds forming the conducting backbone are in blue, the cutting bones are in red, and the remaining bonds of the largest cluster are presented in green. Although no major difference can be observed on the mass M_{clus} of the largest cluster, one can notice that the conducting backbone occupies a larger fraction of the largest cluster as α increases, leading to a substantial decrease on the number of cutting bones M_{cut}.

7.2 Results and discussion

Figure 7.2 shows typical realizations of the largest cluster at the critical point p_c for different values of α. For $\alpha = 0$, we obtain $p_c = 0.527 \pm 0.001$, which is in good agreement with previous simulation results of the PR on the square lattice [66, 68]. Moreover, we find that p_c decreases smoothly and monotonically with α from this value to 0.522 ± 0.001 at $\alpha = 4$ (not shown). Next we apply the burning algorithm [85] to compute the mass of the spanning cluster M_{clus}, the mass of its conducting backbone M_{back}, and the mass (number) M_{cut} of cutting bonds. The last ones, also called red bonds, if removed, would break the spanning cluster into two, therefore destroying the global connectivity of the system.

Our results show that, regardless of the value of α, all these critical quantities scale with the system size L as typical power laws, $M_{back} \sim L^{d_{back}}$, $M_{cut} \sim L^{d_{cut}}$ (see Fig. 7.3), and $M_{clus} \sim L^{d_{clus}}$, where d_{back}, d_{cut} and d_{clus} are the fractal dimensions of the conducting backbone, the cutting bonds, and the largest cluster, respectively. In Figs. 7.4(a)-(c) we show that all these exponents exhibit a monotonic variation with α, going from a saturation regime of (nonlocal) explosive percolation at $\alpha = 0$ to another compatible with ordinary bond percolation (BP) at sufficiently large values of α. Accordingly, for $\alpha = 0$, we recover the previously numerically

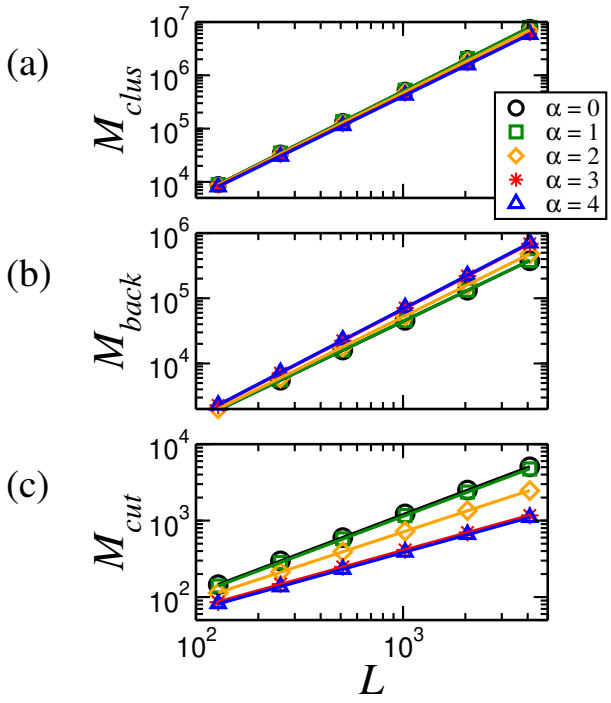


Figure 7.3: (a) Log-log dependence of the mass of the spanning cluster M_{clus} on the system size for different values of the exponent α. (b) and (c) the same as in (a), but for the mass of the conducting backbone M_{back} and the number of cutting bonds M_{cut}, respectively. In all cases and for all values of α, the evidence of scaling behavior substantiates the calculation of the fractal dimensions d_{clus}, d_{back}, and d_{cut} as the slopes of the corresponding straight lines that are best-fitted to the simulation data. All quantities are averaged over at least 2500 realizations precisely at the point in which the largest cluster appears.

calculated values of $d_{clus} = 1.96 \pm 0.01$ [65, 67], $d_{back} = 1.52 \pm 0.03$, and $d_{cut} = 1.02$ [69]. In all three cases, by increasing α, a crossover from PR to BP takes place in the interval $1 < \alpha < 3$. More precisely, d_{clus} decreases in this interval and starts fluctuating around 1.89 for $\alpha > 3$ (see Fig. 7.4a), in agreement with the classical $2D$ value of $91/48$ [44]. After increasing in the interval $1 < \alpha < 3$, the exponents d_{back} and d_{cut} remain practically constant for $\alpha > 3$, around 1.64 [45] (see Fig. 7.4b) and 0.75 [44] (see Fig. 7.4c), respectively. These values are fully compatible with previously reported numerical calculations for ordinary (local) percolation in $2D$. The variations of the exponents d_{clus} and d_{back} within $1 < \alpha < 3$ reflect relevant changes in the compactness of the spanning cluster and its conducting backbone. Although the spanning cluster becomes less compact as α increases (d_{clus} decreases), the mass of the backbone M_{back} tends to occupy a larger fraction of M_{clus}, since the dimension d_{back} increases in the same interval of α values, as shown in Fig 7.3. In these circumstances, a more compact conducting backbone implies a smaller number of cutting bones (see Figure 7.3), therefore explaining the decrease in the exponent d_{cut}.

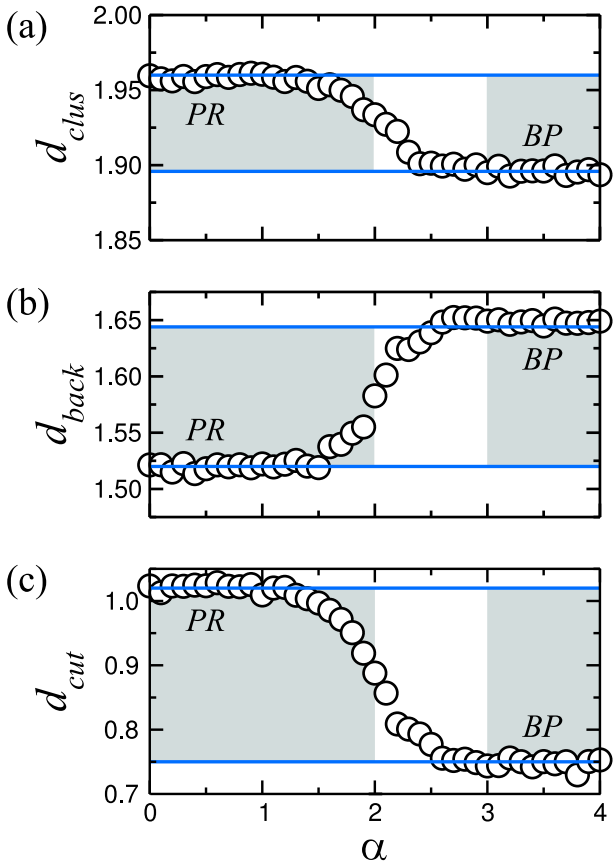


Figure 7.4: Dependence on the exponent α of the size-scaling exponents for (a) the mass of the spanning cluster d_{clus}, (b) the mass of the conducting backbone d_{back}, and (c) the number of cutting bonds d_{cut}. In all cases, a crossover can be observed in the interval $1 < \alpha < 3$ from a regime of non-local explosive percolation at $\alpha = 0$, to a regime that is compatible with ordinary bond percolation (BP), at sufficiently large values of α. The dashed red lines correspond to $d_{clus} = 1.96$ and $91/48$ in (a), $d_{back} = 1.52$ and 1.64 in (b), and $d_{cut} = 1.02$ and 0.75 in (c).

Next we provide some analytical arguments that indicate how nonlocal features are introduced in our percolation model through the power law probability $P(r_{ij})$. We first consider the average distance between all pairs of sites in a empty lattice, $\langle r \rangle = \sum_{r=1}^{L} r N_r r^{-\alpha} / \sum_{r=1}^{L} N_r r^{-\alpha}$, where $N_r = 4r$ is the number of sites that are at a Manhattan distance r from a given site in the square lattice. Approximating the sum by an integral, we obtain that $\langle r \rangle \sim \int_1^L r^{2-\alpha} dr$. It follows that, for $\alpha < 2$, $\langle r \rangle$ is limited by the network size, leading to $\langle r \rangle \sim L$, while, for $2 \leq \alpha < 3$, this distance scales as $\langle r \rangle \sim L^{3-\alpha}$. For $\alpha \geq 3$ and sufficiently large lattice sizes, $\langle r \rangle$ is always finite. As a consequence, the effect of nonlocality on the scaling properties of the system would only play a role for $\alpha < 3$. In addition, distinct non-local behaviors should be expected for the intervals $\alpha < 2$ and $2 \leq \alpha < 3$. These characteristics are consistent with the results displayed in Fig. 7.4. The observed mismatch between expected and numerically calculated crossover values of the exponent α is a consequence of finite-size scaling effects as well as the fact that the

sequential bond allocation leads to the presence of spatial correlations in the percolation process. These correlations make the assumption of an ever empty lattice, as adopted to compute $\langle r \rangle$, is no longer strictly valid.

7.3 Limits of the model on hyper-cubic lattices

D	$p_{c,AP}$	$p_{c,AER}$	d_{clus} - AP	d_{clus} - AER	d_{clus} - Classical [44]
2	0.52655(0)	0.5200(7)	1.955 ± 0.002	1.899 ± 0.001	$91/48$
3	0.32209(6)	0.28536(0)	2.788 ± 0.003	2.530 ± 0.003	2.53
4	0.23416(0)	0.20216(3)	3.665 ± 0.009	3.079 ± 0.005	3.06
5	0.18465(6)	0.16045(4)	4.61 ± 0.01	3.59 ± 0.04	3.54
6	0.15264(2)	0.13411(3)	5.558 ± 0.005	4.46 ± 0.01	4

Table 7.1: Estimated values of the percolation threshold p_c and the scaling exponent d_{clus} for hyper-cubic lattices of dimension D calculated using the jump method. The presented values correspond to averages over a minimum of 2500 realizations of systems with sizes up to $L =$ 4096 ($D = 2$), 256 ($D = 3$), 64 ($D = 4$), 28 ($D = 5$), and 16 ($D = 6$).

In order to better confirm the role of nonlocality on the PR process, we perform additional simulations in the two limits of the model at higher dimensions, namely for $\alpha = 0$, which corresponds to the original PR process, and for the AER process, $\alpha \to \infty$. In these cases, improved performance can be achieved by adopting the so-called jump method to analyze the behavior of the order parameter M_{clus} [64, 86, 87]. For each realization, we compute the average fraction p of occupied bonds at which a jump takes place, defined as the maximum change on M_{clus} from the occupation of a single bond. This value of p corresponds to the percolation threshold p_c. The results for p_c and d_{clus} in both limits and different dimensions are summarized in Table 7.1. Interestingly, the discrepancy between the fractal dimensions calculated for PR and AER models increases substantially with lattice dimensionality. Moreover, our calculations suggest that the resemblance between regular BP and the limiting case $\alpha \to \infty$ stands only up to five dimensions, when compared with previous results from the literature [44, 45].

7.4 Conclusion

In summary, we have proposed a generalization of the PR model where the range of nonlocality in the percolation process can be explicitly tuned. Our results show that this model displays a rich variety of scaling behaviors, going from ordinary to nonlocal explosive percolation. We expect our PR model, since it is based on a geographical choice of bond pairs,

to provide relevant physical insights into the role of nonlocality on the critical properties of percolation.

8 *Conclusions and perspectives*

In the first part this thesis, the navigation in small-world networks, under a cost constraint, is studied using computer simulation results. We have introduced a modification in Kleinberg's network model to characterize how this cost constraint change the navigability in spatially embedded networks. Under unconstrained conditions for a d-dimensional network, the effective exponents are $\alpha = d$ and $\alpha = 0$, for global and local navigation, respectively. However, in real situations, it is difficult to find systems where unlimited resources are available to introduce shortcuts in order to improve transportation in real networks. For the network model presented here, the cost constraint breaks this dichotomy between local and global navigation, leading to a value of the effective exponent $\alpha = d + 1$ that is unique for both local and global navigation, established here for $d = 1$, 2, 3 for regular lattices and for the spanning cluster of percolation with $d_f = 1.89$.

In our work, the long-range connections are equivalent to each other, i.e., these connections have no weight. In the future, the influence of long-range connections with weight, where this weight could be related to the length of the long-range connections, would provide new insights for a more complete understanding of the navigation problem in spatially embedded networks. Another open problem is the analytical solution for dimensions higher than one for the navigation problem with local knowledge with a cost constraint in Kleinberg's network.

In the second half of this thesis, we have addressed two different problems related with explosive percolation. To better understand the ingredients that produce the abrupt transition of explosive percolation, we propose a Hamiltonian approach which revealed that the absence of loops and an ensemble of networks composed by disconnected clusters of similar sizes are the two simple and sufficient conditions, which are essentially nonlocal, for a percolation process to display an abrupt transition.

The second percolation study was the generalization of the product rule (PR) for explosive percolation, in which the nonlocal features of the PR can be controlled by a parameter α. We showed that the scaling behavior of the mass of the spanning cluster, the mass of the conducting

backbone, and the mass of the cutting bonds are described by power-laws. Furthermore, we showed that the related scaling exponents present different values according to the parameter α, going from (local) ordinary to (nonlocal) explosive percolation. In addition, our results suggest that the similarity between the limiting case of $\alpha \to \infty$ and the traditional percolation stands only up to five dimensions.

The complete characterization of the critical behavior of the PR for explosive percolation in finite dimension lattices is an open question. The computation of all critical exponents which describe second order phase transitions of the PR for explosive percolation is crucial for a better understanding of the effects of such nonlocal percolation process.

Bibliography

[1] LAWRENCE, S.; GILES, C. L. Accessibility of information on the web. *Nature*, v. 400, p. 107–109, 1999.

[2] GILES, C. L.; LAWRENCE, S.; KROVETZ, B. Access to information on the web. *Science*, v. 280, p. 1815, 1998.

[3] FALOUTSOS, M.; FALOUTSOS, P.; FALOUTSOS, C. On power law relationships of the internet topology. *Computer Communications Review*, v. 29, 1999.

[4] JEONG, H. et al. The large-scale organization of metabolic networks. *Nature*, v. 407, p. 651–654, 2000.

[5] SCOTT, J. *Social Network Analysis: A Handbook*. London: Sage Publications, 2000.

[6] WASSERMANN, S.; FAUST, K. *Social Network Analysis*. Cambridge: Cambridge University Press, 1994.

[7] ALBERT, R.; BARABASI, A.-L. Statistical mechanics of complex networks. *Reviews of Modern Physics*, v. 74, p. 47–97, 2002.

[8] NEWMAN, M. E. J. The structure and function of complex networks. *SIAM REVIEW*, v. 45, p. 167–256, 2003.

[9] STROGATZ, S. H. Exploring complex networks. *Nature*, v. 410, p. 268–276, 2001.

[10] KLEINBERG, J. M. The small-world phenomenon: An algorithm perspective. *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, p. 163–170, 2000.

[11] KLEINBERG, J. M. Navigation on small-wolrd. *Nature*, v. 406, p. 845, 2000.

[12] ACHLIOPTAS, D. et al. Explosive percolation in random networks. *Science*, v. 323, p. 1453, 2009.

[13] ERDŐS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae*, v. 6, p. 290–297, 1959.

[14] MILGRAM, S. The small-world problem. *Psychology Today*, v. 1, 1967.

[15] DODDS, P.; MUHAMAD, R.; WATTS, D. J. An experimental study of search in global social networks. *Science*, v. 301, p. 827–829, 2003.

[16] WATTS, D. J. *Small Worlds*. Princeton, NJ: Princeton Press, 1999.

[17] KOSMIDIS, K.; HAVLIN, S.; BUNDE, A. Structural properties of spatially embedded networks. *Europhys. Lett.*, v. 82, p. 48005, 2008.

[18] LI, G. et al. Towards design principles for optimal transport networks. *Phys. Rev. Let.*, v. 104, p. 018701, 2010.

[19] MOREIRA, A. A. et al. Hamiltonian approach for explosive percolation. *Phys. Rev. E*, v. 81, p. R040101, 2010.

[20] WATTS, D.; STROGATZ, S. Collective dynamics of 'small-world' networks. *Nature*, v. 393, p. 440, 1998.

[21] MORENO, J. L. *Who Shall Survive?* Beacon, NY: Beacon House, 1934.

[22] CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L. *Introduction to Algorithms*. Cambridge & New York: McGraw-Hill Book Company & The MIT press, 1990.

[23] PRICE, D. J. S. Networks of scientific papers. *Science*, v. 149, p. 510–515, 1965.

[24] BOCCALETTI, S. et al. Complex networks: Structure and dynamics. *Physics Reports - Review Section of Physics Letters*, v. 424, p. 175–308, 2006.

[25] DOROGOVTSEV, S. N. *Lectures on Complex Networks*. New York, USA: Oxford University Press, 2010.

[26] ROBERSON, M. R.; BEN-AVRAHAN, D. Kleinberg navigation in fractal small-world networks. *Phys. Rev. E*, v. 74, p. 017101, 2006.

[27] CARMI, S. et al. Asymptotic behavior of the kleinberg model. *Phys. Rev. Lett.*, v. 102, p. 238702, 2009.

[28] CARTOZO, C. C.; RIOS, P. D. L. Extended navigability of small world networks: Exact results and new insights. *Phys. Rev. Lett.*, v. 102, p. 238703, 2009.

[29] MOUKARZEL, C. F.; MENEZES, M. A. de. Shortest paths on systems with power-law distributed long-range connections. *Phys. Rev. E*, v. 65, p. 056709, 2002.

[30] ALBERT, R.; BARABÁSI, A.-L. Internet - diameter of the world-wide web. *Nature*, v. 401, p. 130, 1999.

[31] BARTHÉLÉMY, M.; AMARAL, L. A. N. Small-world networks: Evidence for a crossover picture. *Phys. Rev. Lett.*, v. 82, p. 3180, 1999.

[32] JEONG, H. et al. The large-scale organization of metabolic networks. *Nature*, v. 407, p. 651, 2000.

[33] LAWRENCE, S.; GILES, C. L. Accessibility of information on the web. *Nature*, v. 400, p. 107, 1999.

[34] GILES, C. L.; LAWRENCE, S.; KROVRETZ, B. Access to information on the web. *Science*, v. 280, p. 1815, 1998.

[35] KOSMIDIS, K.; HAVLIN, S.; BUNDE, A. Structural properties of spatially embedded networks. *Europhys. Lett.*, v. 82, p. 48005, 2008.

[36] MOUKARZEL, C.; MENEZES, M. A. de. Shortest paths on systems with power-law distributed long-range connections. *Phys. Rev. E*, v. 65, p. 056709, 2002.

[37] GUIMERÀ, R. et al. Optimal network topologies for local search with congestion. *Phys. Rev. Lett.*, v. 89, 2002.

[38] DANON, L.; ARENAS, A.; DÍAZ-GUILERA, A. Impact of community structure on information transfer. *Phys. Rev. E*, v. 77, 2008.

[39] SANTOS, M. C. et al. Optimization of random searches on defective lattice networks. *Phys. Rev. E*, v. 77, 2008.

[40] BOGUNÁ, M.; KRIOUKOV, D. Navigating ultrasmall worlds in ultrashort time. *Phys. Rev. Lett.*, v. 102, 2009.

[41] LATORA, V.; MARCHIORI, M. Is the boston subway a small-world network? *Physica A*, v. 314A, p. 109, 2002.

[42] GASTNER, M.; JEONG, H. Price of anarchy in transportation networks: Efficiency and optimality control. *Phys. Rev. Lett.*, v. 101, p. 128701, 2008.

[43] BIANCONI, P. P. G.; MARSILLI, M. Assessing the relevance of node features for network structure. *Proc. Natl. Acad. Sci. U.S.A.*, v. 106, p. 11433, 2009.

[44] STAUFFER, D.; AHARONY, A. *Introduction to Percolation Theory*. London: Taylor & Francis, 1992.

[45] SAHIMI, M. *Applications of Percolation Theory*. London: Taylor & Francis, 1994.

[46] SAHIMI, M.; ARBABI, S. Percolation and fracture in disordered solids and granular media: Approach to a fixed point. *Phys. Rev. Lett.*, v. 68, p. 608, 1992.

[47] GRASSBERGER, P.; ZHANG, Y. C. "self-organized" formulation of standard percolation phenomena. *Physica A*, v. 224, p. 169, 1996.

[48] ANDRADE, J. S. et al. Self-organization in growth of branched polymers. *Physica A*, v. 238, p. 163, 1997.

[49] STANLEY, H. E. et al. Percolation phenomena: a broad-brush introduction with some recent applications to porous media, liquid water, and city growth. *Physica A*, v. 266, p. 5, 1999.

[50] HAVLIN, S. et al. Random multiplicative processes and transport in structures with correlated spatial disorder. *Phys. Rev. Lett.*, v. 61, p. 1438, 1988.

[51] HAVLIN, S. et al. Universality classes for diffusion in the presence of correlated spatial disorder. *Phys. Rev. A*, v. 40, p. 1717, 1989.

[52] MAKSE, H. A. et al. Method of generating long-range correlations for large systems. *Phys. Rev. E*, v. 53, p. 5445, 1996.

[53] MAKSE, H. A. et al. Long-range correlations in permeability fluctuations in porous rock. *Phys. Rev. E*, v. 54, p. 3129, 1996.

[54] PRAKASH, S. et al. Structural and dynamical properties of long-range correlated percolation. *Phys. Rev. A*, v. 46, p. R1724, 1992.

[55] CHO, Y. S. et al. Percolation transitions in scale-free networks under the achlioptas process. *Phys. Rev. Lett.*, v. 103, p. 135702, 2009.

[56] CHO, S.; KAHNG, B.; KIM, D. Cluster aggregation model for discontinuous percolation transitions. *Phys. Rev. E*, v. 81, p. R030103, 2010.

[57] CHO, Y. S. et al. Finite-size scaling theory for explosive percolation transitions. *Phys Rev. E*, v. 82, p. 042102, 2010.

[58] ARAÚJO, N. A. M.; HERRMANN, H. J. Explosive percolation via control of the largest cluster. *Phys. Rev. Lett.*, v. 105, p. 035701, 2010.

[59] ARAÚJO, N. A. M. et al. Tricritical point in explosive percolation. *Phys. Rev. Lett.*, v. 106, p. 095703, 2011.

[60] CHEN, W.; D'SOUZA, R. M. Explosive percolation with multiple giant components. *Phys. Rev. Lett.*, v. 106, p. 115701, 2011.

[61] GÓMES-GARDENES, J. et al. Explosive synchronization transitions in scale-free networks. *Phys. Rev. Lett.*, v. 106, p. 128701, 2011.

[62] HOOYBERGHS, H.; SCHAEYBROECK, B. V. Criterion for explosive percolation transitions on complex networks. *Phys. Rev. E*, v. 83, p. 032101, 2011.

[63] FRIEDMAN, E. J.; LANDSBERG, A. S. Construction and analysis of random networks with explosive percolation. *Phys. Rev. Lett*, v. 103, p. 255701, 2009.

[64] MANNA, S. S.; CHARTTERJEE, A. A new route to explosive percolation. *Physica A*, v. 390, p. 177, 2011.

[65] RADICCHI, F.; FORTUNATO, S. Explosive percolation in scale-free networks. *Phys. Rev. Lett.*, v. 103, p. 168701, 2009.

[66] RADICCHI, F.; FORTUNATO, S. Explosive percolation: A numerical analysis. *Phys. Rev. E*, v. 81, p. 036110, 2010.

[67] ZIFF, R. M. Explosive growth in biased dynamic percolation on two-dimensional regular lattice networks. *Phys. Rev. Lett.*, v. 103, p. 045701, 2009.

[68] ZIFF, R. M. Scaling behavior of explosive percolation on the square lattice. *Phys. Rev. E*, v. 82, p. 051105, 2010.

[69] ANDRADE, J. S. et al. Transport on exploding percolation clusters. *Phys. Rev. E*, v. 83, p. 031133, 2011.

[70] ROZENFELD, H. D. et al. Explosive percolation in the human protein homology network. *Eur. Phys. J. B*, v. 75, p. 305, 2010.

[71] KIM, Y.; YUN, Y. K.; YOOK, S. H. Explosive percolation in a nanotube-based system. *Phys. Rev. E*, v. 82, p. 061105, 2010.

[72] WILSON, K. G. Problems in physics with many scales of length. *Sci. Am.*, v. 241, p. 158, 1979.

[73] NEWMAN, M. E. J.; ZIFF, R. M. Efficient monte carlo algorithm and high-precision results for percolation. *Phys. Rev. Lett.*, v. 85, p. 4104, 2000.

[74] GRASSBERGER, P. et al. Explosive percolation is continuous, but with unusual finite size behavior. *Phys. Rev. Lett.*, v. 106, p. 225701, 2011.

[75] COSTA, R. A. da et al. Explosive percolation transition is actually continuous. *Phys. Rev. Lett*, v. 105, p. 255701, 2010.

[76] LEE, H. K.; KIM, B. J.; PARK, H. Continuity of the explosive percolation transition. *Phys. Rev. E*, v. 84, p. R020101, 2011.

[77] RIORDAN, O.; WARNKE, L. Explosive percolation is continuous. *Science*, v. 333, p. 322, 2011.

[78] CALEY, A. A theorem on trees. *Quart. J. Math.*, v. 23, p. 376, 1889.

[79] WRIGHT, E. M. The number of connected sparsely edged graphs. *J. Graph Theory*, v. 1, p. 317, 1977.

[80] BARTHELEMY, M. Spatial networks. *Phys. Rep.*, v. 499, p. 1, 2011.

[81] LI, Y. et al. Exact solution for optimal navigation with total cost restriction. *Europhys. Lett.*, v. 92, p. 58002, 2010.

[82] VIANA, M. P.; COSTA, L. da F. Fast long-range connections in transportation networks. *Phys. Lett. A*, v. 375, p. 1626, 2011.

[83] GALLOS, L. K.; MAKSE, H. A.; SIGMAN, M. A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proc. Natl. Acad. Sci. USA*, v. 109, p. 2825, 2012.

[84] D'SOUZA, R. M.; MITZENMACHER, M. Local cluster aggregation models of explosive percolation. *Phys. Rev. Lett.*, v. 104, p. 195702, 2010.

[85] HERRMANN, H. J.; HONG, D. C.; STANLEY, H. E. Backbone and elastic backbone of percolation clusters obtained by the new method of "burning". *J. Phys. A*, v. 17, p. L261, 1984.

[86] NAGLER, J. et al. Impact of single links in competitive percolation. *Nat. Phys.*, v. 7, p. 265, 2011.

[87] SCHRENK, K. J.; ARAÚJO, N. A. M.; HERRMANN, H. J. Gaussian model of explosive percolation in three and higher dimensions. *Phys. Rev. E*, v. 84, p. 041136, 2011.

APPENDIX A – Publications

A.1 List of publications related with this thesis

- 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 Re. Lett. **104**, 018701 (2010).
- A. A. Moreira, E. A. Oliveira, S. D. S. Reis, H. J. Herrmann, and J. S. Andrade, Jr., *Hamiltonian approach for explosive percolation*, Phys. Rev. E **81**, 040101(R) (2010).
- S. D. S. Reis, A. A. Moreira, and J. S. Andrade, Jr., *Nonlocal product rules for percolation*, Phys. Rev. E **85**, 041112 (2012).
- 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.

A.2 List of publications not related with this thesis

- J. S. Andrade, Jr., S. D. S. Reis, E. A. Oliveira, E. Fehr, and H. J. Herrmann, *Ubiquitous fractal dimension of optimal paths*, Comput. Sci. Eng. **13**, 74 (2011).
- S. D. S. Reis, N. A. M. Araújo, J. S. Andrade, Jr. and H. J. Herrmann, *How dense can one pack spheres of arbitrary size distribution?*, Europhys. Lett. **97**, 18004 (2012).
- H. F. Credidio, E. N. Teixeira, S. D. S. Reis, A. A. Moreira, and J. S. Andrade, Jr., *Statistical patterns of visual search for hidden objects*, Sci. Rep. **2**, 920 (2012).