JACOBS
UNIVERSITY

COOPERATION IN DYNAMICAL NETWORKS

by

Filippo Radicchi

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Physics

Approved, Thesis Committee

Prof. Hildegard Meyer-Ortmanns .....
Jacobs University Bremen, chair

Prof. Jürgen Fritz .....
Jacobs University Bremen

Prof. Herbert Jaeger .....
Jacobs University Bremen

Prof. Albert Díaz-Guilera .....
Universitat de Barcelona

Date of Defense: May 18, 2007

School of Engineering and Sciences

Abstract

I study various cooperation phenomena on different network topologies. To these belong synchronization of Kuramoto oscillators, entrained by pacemakers; evolutionary algorithms on spin systems, reinterpreted as an approach to social balance or as the solution of optimization problems in computer science; and the influence of enforced synchronization in the updating procedure of algorithms. In the case of synchronization of oscillators, my results allow the prediction of whether a pacemaker is able to entrain a given set of oscillators, depending on their number, their coupling strength and their connections. Furthermore, I use spin dynamics to establish a mapping between an approach of social balance and the solution of a special optimization problem, called the XOR-SAT problem in the context of computer science (XOR refers to the logical operator and SAT stands for satisfiability). In the language of spin dynamics, the equivalence of both problems becomes evident. Along with these cooperation phenomena and from the statistical-physics point of view, I observe phase transitions between active and absorbing states or different absorbing states in systems out-of-equilibrium. The time to reach these stationary states also turns out to depend on the system size in a characteristic way. Where these phase transitions are of the second order, I identify the universality class by measuring their critical exponents and checking the hyperscaling relations. For the spin dynamics on a triangular lattice, I find in this way a new universality class, while it turns out to be the known universality class of parity conservation for the transition between synchronous and asynchronous updating schemes. In general, my out-of-equilibrium processes show a sensitive dependence on the network topology that is varied between various regular, diluted random, and fully connected versions.

Preface

This thesis studies cooperative phenomena in dynamical models defined on various topologies of complex networks. The contents of the thesis is based on six articles published in international journals of physics:

Paper I S.H. Yook, **F. Radicchi** and H. Meyer-Ortmanns
Self-similar scale-free networks and disassortativity
Physical Review E **72**, 045105(R) (2005)

Paper II **F. Radicchi** and H. Meyer-Ortmanns
Entrainment of coupled oscillators on regular networks by pacemakers
Physical Review E **73**, 036217 (2006)

Paper III **F. Radicchi** and H. Meyer-Ortmanns
Reentrant synchronization and pattern formation in pacemaker-entrained Kuramoto oscillators
Physical Review E **74**, 026203 (2006)

Paper IV **F. Radicchi**, D. Vilone, S. Yoon and H. Meyer-Ortmanns
Social balance as a satisfiability problem of computer science
Physical Review E **75**, 026106 (2007)

Paper V **F. Radicchi**, D. Vilone and H. Meyer-Ortmanns
Universality class of triad dynamics on a triangular lattice
Physical Review E **75**, 021118 (2007)

Paper VI **F. Radicchi**, D. Vilone and H. Meyer-Ortmanns
Phase transition between synchronous and asynchronous updating algorithms
Journal of Statistical Physics, in press (2007)

In this thesis I will refer to these papers by their Roman numerals. They will be reported in chronological order according to their date of publication.

This thesis represents my own contribution to the study of the novel field of complex networks in connection with phase transitions, a topic with a long tradition in statistical physics. As with any thesis in natural science, the present one is also specialized to a narrow subject. This is unavoidable in order to reach results of a certain significance, but makes it also difficult for non-specialists to enter the topic and evaluate the results. Nevertheless, I tried to make the content of this thesis understandable to scientists as much as possible, giving concise and intuitive introductions to every topic and leaving more specialized explanations to the papers.

Filippo Radicchi
Bremen, May 2007

Contents

List of Figures

List of Tables

Introduction

Much interest of the scientific community in the analysis of complex networks recent years revealed. The concept of networks as sets of many interacting units is used in many different disciplines, ranging from physics to mathematics, from biology to psychology, from social science to computer science. Natural and artificial systems in fact offer a large variety of networks: technological networks (like the Internet), information networks (like the World Wide Web), biological networks (like protein interaction networks and genetic networks), social networks, economic networks, etc. Statistical physics provides the glue among all these different disciplines. An analysis with tools proper of statistical physics reveals that many real networks have several common topological features: this result is amazing since all these systems, coming from different disciplines and therefore apparently not related, have unexpected analogies.

The study of networks began in the middle of the 1960s. Social scientists were the first to be interested in networks. At roughly the same time, mathematicians started to study networks in a more formal way, considering networks as discrete geometrical objects. Nevertheless, the study of networks had a big increment since the end of the 1990s. The "boom" in the use of computers was essential for two main reasons. First, it was possible to collect a great deal of data describing the topologies of real systems and store them in public and freely accessible databases. Second, the high computational speed of modern computers allowed a fast and fruitful analysis of these datasets. In this context, statistical physics revealed its utility for understanding the properties of real networks. Statistical physics is in fact the discipline which studies many particles systems and therefore all the mathematical methods and tools developed in the context of statistical physics are also extremely appropriate for the study of networks. Soon the attribute "complex" was assigned to networks, understood in the sense that non-trivial topological and often dynamical features may arise from simple local properties or rules.

The study of statistical physicists at the beginning focussed on the topological properties of networks. Networks were studied as purely geometrical objects without paying attention to their particular functions or dynamics. More recently, the interest of scientists moved to the

study of dynamical properties of networks. Natural and artificial systems are not just static objects: every unit has its own dynamics that is affected by the underlying complex network of interactions. The study of the dynamical properties of a network is therefore needed for a more complete knowledge of the system associated with the network itself. Often, the global dynamics of the whole system strongly depends on the underlying network of interactions: the same local dynamical rules defined on two different networks of interactions may lead to two completely different global features; even different local dynamical rules defined on the same network of interactions can lead to similar behaviors.
The present thesis is a contribution to this field of research. I followed the same historical path of the scientific community, in the sense that I first studied some topological properties of real networks and then specialized my attention on the study of dynamical systems defined on networks. All the results obtained are new and interesting. All of them have been accepted for publication in international journal of physics. In this thesis, I decided to report the results in their original formulation, therefore the thesis is essentially a collection of papers.

The thesis consists of two main parts. The first part, chapters 1 and 2, is a general introduction to specific topics of statistical physics on which the arguments of the second part are based, which are reported in chapters 3, 4, 5 and 6.
In chapter 1, I introduce complex networks, illustrated with concrete examples of real networks with complex features, describing the mathematical background needed for the study of networks. In chapter 2 I give a short and basic description of phase transitions. Phase transitions are a frequently discussed topic of statistical physics with a huge range of applications. The results of the present thesis mainly concern phase transitions on networks and phase transitions for dynamical systems on networks. Therefore, the reader needs the general introduction to this topic of chapter 2 to understand the rest of the thesis.
The second and original part of the thesis starts from chapter 3 on. In this part, each chapter is organized in the same way, with a short preamble to the topic of the chapter itself followed by my original papers on the same topic. The aim of the introductions to each topic is not to give full and self-contained knowledge about all the background of the topic itself, since this would go beyond the scope of the thesis, but should be considered as guidelines for the reader in order to understand my papers. In particular, in chapter 3 I study a novel topological property of real networks: the fact that some of them have fractal dimensions and are self-similar. I therefore talk about fractals in general and then specialize to the definition of fractal networks. I conclude the chapter by reporting **Paper I** [165] . In chapter 4 I consider dynamical systems on networks. Each unit of the system is an oscillator and I study the conditions which allow a coherent behavior of the whole system. The first part of chapter 4 is therefore an introduction to synchronization, in particular to models representative of the

phenomenon of collective synchronization. **Paper II** [120] and **Paper III** [121] are my own contributions to this field of research. In chapter 5 I study another type of dynamical system on networks. I present a model that retains some basics features of networks of social relationships. The chapter starts with an introduction in which I show how the interest in this model is shared by different disciplines like social science, mathematics, physics and computer science. I then report in **Paper IV** [124] and **Paper V** [123] my own contributions. Finally in chapter 6, I conclude with the last topic analyzed during this thesis. I investigate the role of the updating scheme in numerical simulations for dynamical systems. I approach the problem in its simplest formulation, showing nevertheless how even in a very simple dynamics, non-trivial features can arise. I close the chapter with **Paper VI** [122] .
I conclude the thesis with two appendices, A and B. In appendix A, I perform some analytical calculations not reported in detail in any of the papers cumulated in this thesis. In appendix B, I give exact prescriptions regarding the numerical implementations of all the models studied in my papers. This appendix may be of interest to those who would like to reproduce my numerical measurements and for further applications or extensions of my results.

Chapter 1

Complex Networks

Introduction

The structural properties of a huge class of real systems can be studied by describing them as networks. One uses the word "network" to describe an ensemble of many interacting units: the most natural way to visualize a network is to draw the units as vertices and the interaction between two points as an edge or a line that connects them [see for example Figure 1.1]. But, in spite of the simplicity of this kind of description, these systems are complex when their many elementary interactions between units generates non-trivial topological and dynamical features. For this reason one refers to them as **complex networks**.
One of the main reasons for the recent increased interest of the scientific community in networks is motivated by the universality of this kind of description. One can speak about networks in physics, biology, social science , also in information science , as well in chemistry and technology. A very interesting result of this intensive study is the fact that many of these real networks have common topological properties, despite their completely different range of applications.

This chapter is organized as follows. I start with a short section [section 1.1] in which I describe concrete examples of natural systems that can be easily described as networks. I then start with the mathematical description of networks in terms of graphs [section 1.2]. The possibility of viewing a network as a graph allows to introduce in section 1.3 some basic quantities which allow to describe statistically the topological properties of real networks as reported in section 1.4. In section 1.5 I briefly describe the most prominent models of networks proposed by the mathematical and physical communities in order to explain the topological features of real networks. Finally, in section 1.6 I introduce networks as dynamical systems.

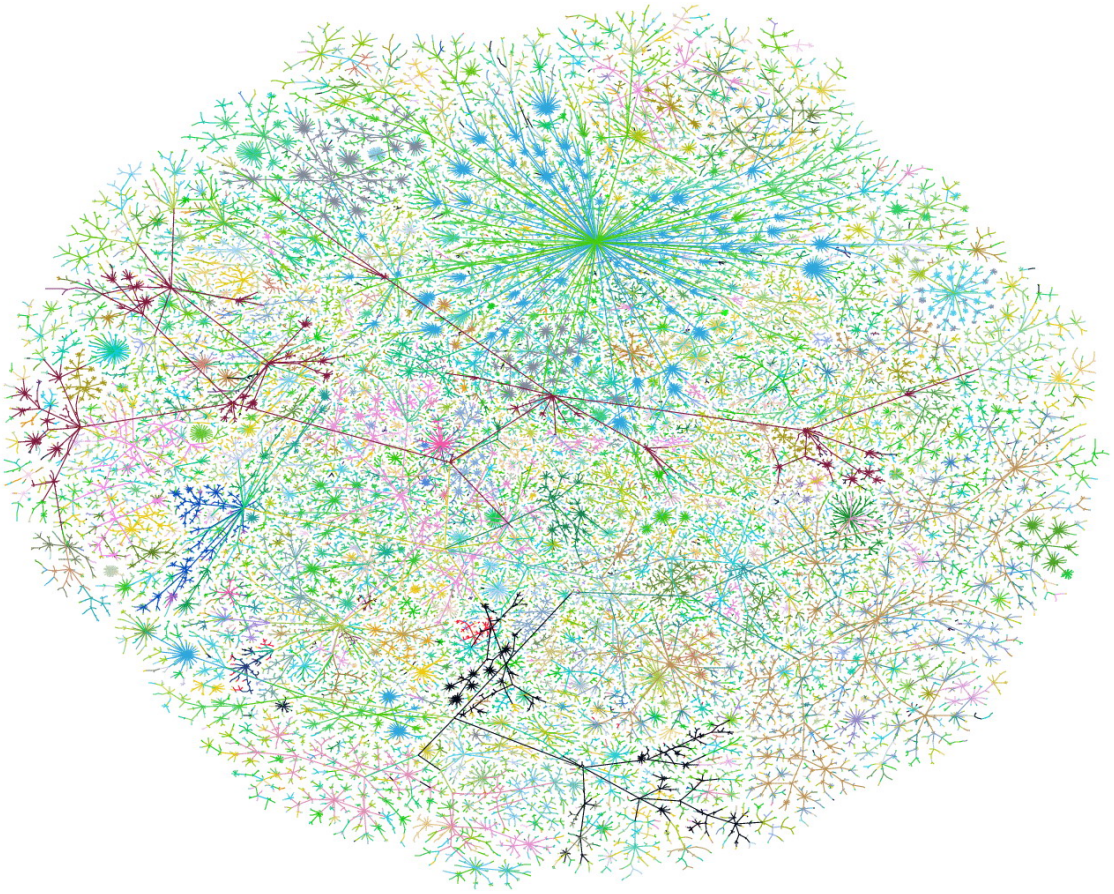


Figure 1.1: Internet Map, colored by IP address, of 29 June 1999 (By William R. Cheswick , http://www.cs.bell-labs.com/ ches/map/index.html).

1.1 Real Complex Networks

Despite the fact that the study of complex networks is quite a recent topic, the topological properties of a large amount of natural systems have already been studied describing them as networks. Examples of real networks can be found in several different scientific disciplines. In this section I will give a short list of these real networks. The examples of real networks reported here are organized in different subsections referring to their origins: among the real networks I will distinguish between social networks [subsection 1.1.1], information networks [subsection 1.1.2], technological networks [subsection 1.1.3] and biological networks [subsection 1.1.4]. The aim of this list is just to motivate the importance of the study of complex networks; therefore, the following will not be a complete list. An interested reader can find

more complete lists of analyzed real-world networks in reviews such as [3, 40, 105].

1.1.1 Social Networks

Chronologically, social science was the first discipline interested in networks. A social network is a set of people with some set of contacts or social relationships between them. Several examples can be taken out of the context of social networks. Historically important, since it was the first social network of a large dimension to be studied, is the social network analyzed by S.Milgram and J.Travers at the end of the 1960s [94, 152]: a network based on archives of the American mail delivery system, where two persons are connected if they have communicated via at least one letter. Other social networks have been studied recently in the context of statistical physics. For example, the **actor network** [8], based on the "Internet Movie Database[1]", where each vertex of the network is an actor and two actors are connected if they have played together at least during one movie. Another well studied type of social network is the so-called **collaboration network** [37, 52, 102, 103]: these networks use the information stored in the electronic databases on scientific publications, where each author is a vertex and two authors are connected if they have written a paper together.

1.1.2 Information Networks

Another category of real networks is given by the information networks (sometimes also called "knowledge networks"). Networks belonging to this category are represented by the **citations networks** [128]. These networks are again based on scientific publications but now each vertex of the network is given by a paper, while two papers are directly connected if one of the two papers cites the other. To this category also belongs the most famous example of real complex network: the **World Wide Web**, a network of web pages directly connected by hyperlinks. Due to the huge dimensions of the network, only small portions or domains of the World Wide Web have been studied in [6, 7, 25]. **Semantic networks** also belong to the category of information networks [39, 67]. Human language can be represented as a network in the following way: each word is a vertex of the network, while two words are connected either because they are neighbors in a text or because they have a similar meaning.

1.1.3 Technological Networks

A big category is represented by the technological networks, that is by the networks physically made by humans. To this category belong **power grid networks** [159, 158, 8], **road net-**

[1] http://imdb.com

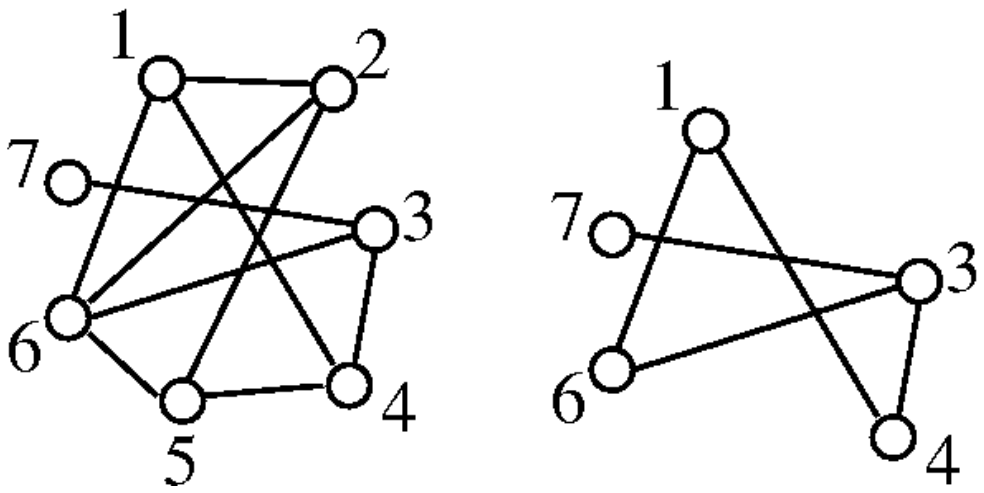
works [72], **water-pipe networks**, **airline** [8, 12] and **railway transportation networks** [136], etc. The most studied technological network is the **Internet**, the network of computers physically connected by wires and wireless connections. The extremely large number of computers belonging to the Internet prohibits a complete study of this network. At the moment, only the highest level of representation of the Internet, the **autonomous system network**, has been studied [44]. Each autonomous system is a provider and stands as a local and independent group of computers connected to the Internet.

1.1.4 Biological Networks

Several biological systems can be represented as networks. One example is the **metabolic pathways**, where each vertex is a substrate or a product of a metabolic reaction. Of course, in these networks the edges have a proper direction (i.e., substrates $\rightarrow$ product) [69]. Other examples of biological networks are the **protein folding networks** [68]. Each vertex is a specific folding state of a protein and two vertices are connected if the protein can pass between the two folding ways via simple elementary movements. Also, **genetic regulatory networks** belong to the category of biological networks [45, 53, 82]. The expression of a gene, or better the production of the protein for which the gene codes, can be controlled by the presence of other proteins, both activators and inhibitors. Therefore, the genome itself forms a switching network with vertices representing proteins and directed connections standing for the dependence of protein production on the proteins at each vertex. The **food webs** are another type of biological network [65]. An ecosystem can be represented as a network where each organism (animal or plant) can be predator or prey. Directed connections of the type predator $\rightarrow$ prey are present in food webs. Finally, **neural networks** clearly belong to the category of biological networks [160, 159]. Here the neurons are assigned to the vertices and the synapses to the edges connecting the neurons.

1.2 The Graph

From a mathematical point of view, a network is described as a graph. A graph G is a pair of sets $G = (E, V)$. E is a set of **vertices**, often called **nodes** or **units**. V is a set of pairs of nodes belonging to G. The elements of E are called **edges**, but often **arcs** or **links**. A graph has a simple representation [see Figure 1.2(a)]. Each node of the graph can be drawn as a dot, while a link connecting two nodes as a segment between the two respective dots. If in a graph G there are two nodes $i \in V$ and $j \in V$ and a link $(i, j) \in E$, the nodes i and j are called **adjacent** or simply **connected**. In particular, the nodes i and j are said to be **incident** on the

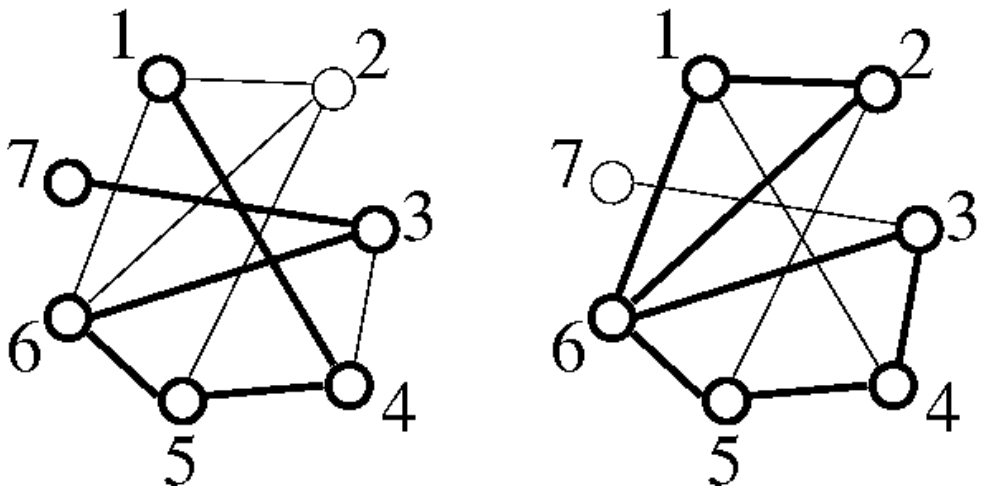


(a) Graph (b) Subgraph

Figure 1.2: **(a)** A graph composed of $N = 7$ nodes and $L = 10$ edges. **(b)** One possible subgraph of the graph in (a). This subgraph is composed of $N = 5$ vertices and $L = 5$ links.

arc (i, j). Moreover, one can say that the nodes i and j are next neighbors since there exists a connection (i, j) between them. This kind of neighbor relation can be simply extended: for example, if in the graph G there are three nodes $i, j, k \in V$ and the two links $(i, j) \in E$ and $(j, k) \in E$, but not the link $(i, k) \notin E$, the nodes i and k are said next-to-next neighbors.
The total number of nodes present in a graph G is called the **size** of the graph G and it is usually denoted by N. The total number of edges in the graph G is usually denoted by L.
All the operations generally defined for sets, such as inclusion ($\subset$), union ($\cup$), intersection ($\cap$), etc., are also valid for graphs. For example, if G and G' are two graphs, their intersection is given by $G \cap G' = (V \cap V', E \cap E')$, that is, by the simultaneous intersections between the sets of nodes and the set of edges. Again for example, given a graph $G = (V, E)$, a graph $G' = (V', E')$ is a subgraph of G and is denoted by $G' \subset G$ if both the relations $V' \subset V$ and $E' \subset E$ are valid. All the nodes of V' belong also to V and all the links in E' are also inside E: one can simply deduce in this case $G \cap G' = G'$. In Figure 1.2(b) I report one possible subgraph of the graph reported in Figure 1.2(a).

It is possible to define **paths** in a graph G [see Figure 1.3(a)]. A path G_P represents a way to go from a node $i_1 \in G$ to a node $i_{N_P} \in G$, passing over other different nodes $i_2, \ldots, i_{N_P-1} \in G$. A path G_P is again a subset of the graph G, $G_P \subset G$. G_P is given by a pair of sets: $G_P = (V_P, E_P)$. In particular $V_P = \{i_1, i_2, \ldots, i_{N_P-1}, i_{N_P}\}$ and $E_P = \{(i_1, i_2), \ldots, (i_{N_P-1}, i_{N_P})\}$, with $i_1 \neq i_j \neq i_{N_P}$, $\forall\, j = 2, \ldots, N_P - 1$, since all the nodes in the path should be different. In



(a) Path (b) Cycle

Figure 1.3: **(a)** A path in the graph of Figure 1.2(a). The path considered here involves the nodes $\{1, 3, 4, 5, 6, 7\}$, denoted as bold circles, and the edges $\{(1, 4), (4, 5), (5, 6), (6, 3), (3, 7)\}$, represented by bold lines. The length of the path is 5, since five links are involved. **(b)** Two cycles are shown, both represented by bold circles and lines. The first has length 3 and is composed of the nodes $\{1, 2, 6\}$ and the edges $\{(1, 2), (2, 6), (6, 1)\}$. The second one has order 4 and is composed of the nodes $\{3, 4, 5, 6\}$ and the links $\{(3, 4), (4, 5), (5, 6), (6, 3)\}$.

the special case for which $i_1 \equiv i_{N_P}$, the path is closed and it is called **cycle** [see Figure 1.3(b)]. The number of links $L_P \equiv N_P$ belonging to the path G_P is called the **order** or the **length** of the path. In a natural way, it is possible to introduce a metrics inside a graph G. Having assigned a "real length" to each link of the graph G, the real length assigned to each path is given by the simple sum of the real lengths of all the links belonging to the specific path considered.

When it is possible starting from every node i of the graph G to arrive at every other node j of the graph G, the graph G is called **connected**. Sometimes it happens that the graph G is not connected since the former condition is not valid. In these cases, the graph G is effectively composed by m independent connected subgraphs $G_1 \subset G$, $G_2 \subset G$, ..., $G_m \subset G$, with $G_q \cap G_r = \emptyset$, $\forall\ q \neq r$, and $G = \cup_{q=1}^{m} G_q$. Each subgraph G_q is called a **component** of the graph G, and the component G_{q^*} with the largest size is called the **giant component** of G.

The simplest and most often used way to describe the entire topology of a graph G is to use its **adjacency matrix** A. A is a square matrix where the number of columns equals the

number of rows and both are equal to the size N of the graph G. In the simplest case of unweighted graphs, the generic element $A_{i,j}$ of the adjacency matrix A is by definition

$$A_{i,j} = \begin{cases} 1 & \text{, if i and j are connected} \\ 0 & \text{, if i and j are not connected} \end{cases} . \tag{1.1}$$

If the graph G is undirected, A is a symmetric matrix ($A_{i,j} = A_{j,i}$, $\forall\ i, j \in G$). For example, the graph reported in Figure 1.2(a) has the adjacency matrix

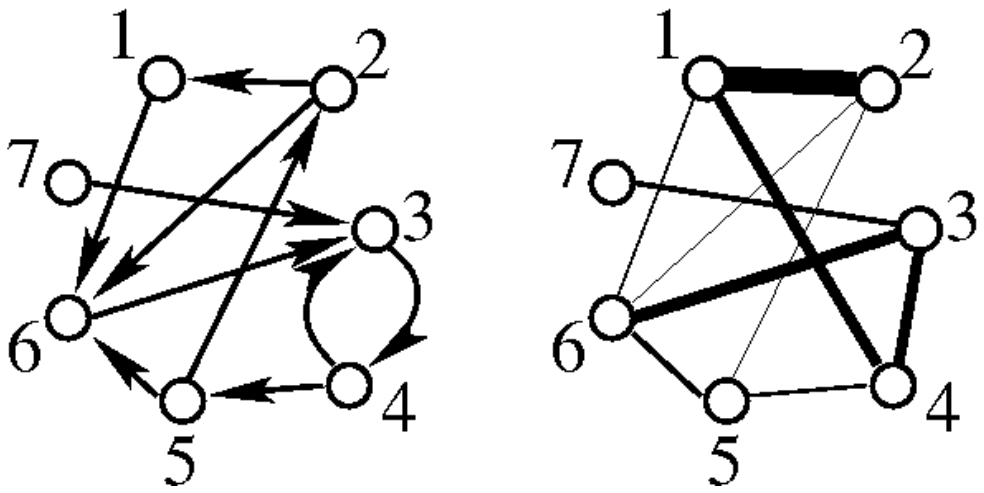
	1	2	3	4	5	6	7
1	0	1	0	1	0	1	0
2	1	0	0	0	1	1	0
3	0	0	0	1	0	1	1
4	1	0	1	0	1	0	0
5	0	1	0	1	0	1	0
6	1	1	1	0	1	0	0
7	0	0	1	0	0	0	0

.

The adjacency matrix represents the most compact way of storing all the topological features of a graph. It should be noted that there are in principle $N!$ different adjacency matrices for each graph, since there are $N!$ different ordering for the indices of the nodes in a graph of size N. All the $N!$ different adjacency matrices are equivalent because they always represent the same geometrical object.
As shown in this section, a network is fully equivalent to a graph. Therefore in the following I will use both the words "network" and graph synonymously for indicating the same geometrical object.

1.2.1 Other Types of Graphs

Up to now I considered for simplicity only undirected and unweighted graphs. By undirected graphs I mean that there is no distinction between the edges (i, j) and (j, i), since they stand for the same edge of the graph. By unweighted I mean that all the edges in the graph have the same weight or "importance." In real networks or in some specific models of graph, it makes sense to consider the **directions** and the **weights** of the links. In the case of directed networks, the presence of the link (i, j) does not imply the presence of the reverse link (j, i). For example, in citation networks [see section 1.1.2] an older paper cannot cite a newer one. This means that sometimes it is possible to go from i to j, but not from j to i. The adjacency matrix is therefore no longer symmetric since in general $A_{i,j} \neq A_{j,i}$. All



(a) Directed Graph (b) Weighted Graph

Figure 1.4: **(a)** The arrows assigned to the edges stand for the right direction assigned to the respective arc. **(b)** The thickness of one edge is proportional to the weight of the respective link.

the former definitions, such as paths, cycles, etc., are also valid for directed graphs. For example, a path in a directed graph should be a directed walk along the links of the graph. In the graphical representation, the lines standing now for directed links should also have an arrow which prescribes the direction that should be followed [see Figure 1.4(a)]. Similar considerations can be made in the case of weighted graphs. In this case, the generic entry $A_{i,j}$ of the adjacency matrix cannot be just one or zero, but should be a real number $w_{i,j}$ standing for the weight of the link (i, j). An example can be given by a network of pipes for the transportation of water, where not all the pipes (i.e., links) have the same section (i.e., weight). All the former definitions such as paths, cycles, etc., do not need to be changed in the case of weighted graphs. In a graphical representation, a way to distinguish the weight of each link is to draw the respective link as a line with a thickness proportional to its weight [see Figure 1.4(b)]. Weighted graphs will be used in chapter 5 and in **Paper IV** [124] and **Paper V** [123] . In this particular case, I quantify a social relationship between two social agents by assigning to it a weight or a "sign." In this framework, a positive relationship stands for the social agents being "friends," while the sign of the relationship is negative when the social agents are "enemies."

1.3 Measures on Complex Networks

As stated before, the adjacency matrix completely describes all the topological features (including the eventual direction or weight of the links) of a network. The adjacency matrix does not represent an object that is easy to handle, since real networks are in general composed of a large number of nodes and links. For the same reason, the human eye is not able to understand the topological features of large networks by simply having a look at the graphical representation of the network. Therefore, several observables that allow more manageable knowledge of the topology of the networks are introduced. Here I report a list of the quantities that are used mostly in the literature and that tell something about the topological properties of networks.

1.3.1 Degree of Connections

The simplest quantity that can be measured in a network is the degree of connections. Given a graph G, it is defined as the degree of connections or simply **degree** of the node $i \in G$, the total number of connections between the node i and all the other nodes $j \in G$. In terms of the adjacency matrix A, the degree of the node i, usually denoted as k_i, is simply calculated as

$$k_i = \sum_{j=1}^{N} A_{i,j} \quad , \tag{1.2}$$

if the network considered is undirected. Given a network, it is therefore possible to define a degree distribution $P(k)$. $P(k)$ represents the probability that a node, randomly chosen out of all the possible nodes of the network, has a degree equal to k. In practice, given a network, one can compute the degree of each node. Then one constructs a histogram: each column of the histogram refers to a value of the degree and the height of the k-th column is linearly proportional to the number of nodes with degree k present in the network. After a proper normalization, the histogram represents the actually measured degree distribution $P(k)$ of the network. Moreover, one can define the **joint probability** or **degree correlation** $P(k, k')$, standing for the probability that two nodes with degree respectively k and k' are connected. In other words, $P(k, k')$ represents the probability that one randomly selected edge of the network has at its ends two nodes with degrees k and k' respectively. Later I will make use of the joint probability in chapter 3 and in **Paper I** [165] .
In the case of a directed network, one can distinguish between the in-coming degree of the node i, ${}^{in}k_i$, and the out-going degree of the node i, ${}^{out}k_i$. Clearly ${}^{in}k_i = \sum_{j=1}^{N} A_{j,i}$ is the total number of connections that from the other nodes of the network come to i, while ${}^{out}k_i = \sum_{j=1}^{N} A_{i,j}$ is the total number of connections that go from i to other nodes of the

network. Therefore, in case of a directed network one can define in-coming and out-going degree distributions indicated by $P\left({}^{in}k\right)$ and $P\left({}^{out}k\right)$, respectively. Also in this case it is appropriate to consider degree correlations. There are several choices: $P\left({}^{in}k,\ {}^{out}k\right)$, the probability that a node with an in-coming degree ${}^{in}k$ has an out-going degree ${}^{out}k$; $P\left({}^{in}k,\ {}^{in}k'\right)$, the probability that a node with an in-coming degree ${}^{in}k$ is connected to a node with an in-coming degree ${}^{in}k'$; etc. In the case of weighted networks, one should talk not only about the degree of connections of the node i, but also about the **strength** of connections of the node i : $s_i = \sum_{j=1}^{N} w_{i,j}$. Again in this case it makes sense to define the strength distribution $P(s)$ in analogy to the degree distribution so far defined. I do not repeat the definitions, but it should be clear that for the strengths one can also define strength correlations.

1.3.2 Distance and Diameter

Given a graph G and two nodes i and j belonging to G, one indicates with $\ell_{i,j}$ the length of the shortest path or the **distance** between the nodes i and j. The definition of distance between two nodes in a graph is clearly identical to the definition of distance between two points in a standard Euclidean space, where the distance is effectively defined as the length of the segment (i.e., the shortest path) that connects the two points. In directed graphs in general one should have $\ell_{i,j} \neq \ell_{j,i}$; in a unweighted graph, as already stated, $\ell_{i,j} = 1$ for any existing link (i,j) in the network, so that the distance between two nodes corresponds to the number of links belonging to the shortest path connecting them; in a weighted graph, as stated before in section 1.2.1, one can define $\ell_{i,j} = 1/\left|w_{i,j}\right|$ for any link (i,j) and also calculate the distances according to this definition. In the literature, the distance between two nodes in a graph is usually called geodesic distance or simply **geodesic**. Two nodes that are at distance $\ell = 1$ are called next neighbors or simply neighbors, two nodes at distance $\ell = 2$ are called next-to-next neighbors, and so on. It should be noted that a graph can be composed of several disconnected components. In this case there are no paths between two nodes belonging to two different components by definition of the component [see section 1.2]. For convenience the distance between two unreachable nodes is set to be infinite.
Having defined the distance between two nodes in a graph, it is obvious how to define the **average distance** between all the pairs of nodes in the graph, namely

$$\langle \ell \rangle = \frac{2}{N(N-1)} \sum_{i>j} \ell_{i,j} \quad . \tag{1.3}$$

The inverse of the normalizing factor $N(N-1)/2$ appropriately counts the total number of different pairs of nodes that one can find in a graph of size N. Moreover, one can define

several other measures. For example, it is possible to mix the notions of degree and distance in a distance-dependent degree correlation $P[k, k'(\ell)]$, standing for the probability that a node with degree k has to be at distance ℓ from a node with degree k'. For $\ell = 1$, $P[k, k'(\ell)]$ clearly reduces to the former introduced degree correlation between nearest neighbors [see section 1.3.1].
The quantity

$$d = \max_{i,j \in G} \ell_{i,j} \quad , \tag{1.4}$$

that is the maximum value of all the distances between pairs of nodes in the graph G, is called the **diameter** of the graph G. Again, the definition of the diameter of a graph is analogous to the one in Euclidean space: the diameter of a circle is the maximal distance between all pairs of points lying on that circle.

1.3.3 Clustering Coefficient

Given a node i with a degree k_i in an undirected and unweighted network, there are exactly $\frac{k_i(k_i-1)}{2}$ possible couples of nodes directly connected to the node i. Two nodes j and q connected to the i-th node [$A_{i,j} = 1$ and $A_{i,q} = 1$] are not generally connected with each other [$A_{j,q}$ can be zero or one]. Call y_i the total number of these couples that are actually connected: $y_i = \sum_{j,q} A_{i,j} A_{i,q} A_{j,q}$. One defines as a **clustering coefficient** of the node i the ratio

$$C_i = \frac{y_i}{\frac{k_i(k_i-1)}{2}} \quad . \tag{1.5}$$

By definition one has $0 \leq C_i \leq 1$. C_i stands for the probability that two nodes, randomly chosen out of all the nodes connected to the node i, are connected with each other. Again, from the knowledge of the clustering coefficient of all the nodes of a network, it is possible to define a probability distribution $P(C)$ and to eventually construct clustering coefficient correlation functions. The average value of the clustering coefficients over all the nodes of the network,

$$\langle C \rangle \; = \; \frac{1}{N} \sum_{i=1}^{N} C_i \quad , \tag{1.6}$$

is in general said to be simply the clustering coefficient of the network.
Effectively, Eq.(1.5) represents the density of triangles [or triads as I call them in section 5.1.1] attached to the node i. A more appropriate way to indicate it would be $C_i^{(3)}$ instead of simply C_i, because Eq.(1.5) actually defines the clustering coefficient of the third order for the node i. There is nothing that prohibits considerations of different local cyclic structures instead of cycles of order 3 [17, 18]. In principle, one can define the clustering coefficient for every

cycle of order g in the same way as it was done so far for triangles. Therefore, one defines the clustering coefficient of order g of the node i as $C_i^{(g)} = \frac{r_i^{(g)}}{p_i^{(g)}}$, so that the ratio of the total number of cycles of order g actually formed in the network by the node i and the number of cycles of order g that a node with same degree as the node i could, in principle, form in the network [119]. The clustering coefficient of order 3 was chronologically the first clustering coefficient measured on networks: therefore when the order is not explicitly specified, one refers to the clustering coefficient of the third order by simply indicating it as the clustering coefficient.
In the case of directed networks, the former definition of Eq.(1.5) remains the same. One has just to consider the direction of the links in order to compute the clustering coefficient. In particular, in a directed network it also makes sense to consider cycles of length 2 and therefore to define a clustering coefficient of order 2: $C_i^{(2)}$. For weighted networks several possibilities of defining a weighted clustering coefficient were proposed [12] but I do not report on them here.

1.4 Topological Properties of Real Complex Networks

In the former section 1.3 I introduced several quantities that can be easily measured on a network. Numerical measurements of the former quantities on real networks have revealed that almost all of these systems possess similar topological statistical properties. In this section I summarize the main results obtained in this context. I give several definitions of characteristic properties of real complex networks. Moreover, I report in Table 1.1 some concrete examples.

1.4.1 Scale-free Property

The degree distribution [see section 1.3.1] of almost all of the real networks follows a power law [see Figure 1.5]

$$P(k) \sim k^{-\gamma} \quad . \tag{1.7}$$

If a network has degree distribution given by Eq.(1.7), one says that the network has a **scale-free**[2] or scale-invariant degree distribution, or simply that the network is scale-free. Actual values of the exponent γ for the degree distribution of real networks range between 1.5 and 3.5 [see Figure 1.5 and Table 1.1].

[2] A function $f(x)$ is called scale-free or scale-invariant if it satisfies the relation $f(x) = \lambda^{-\Delta} f(\lambda x)$ for some choice of the exponent Δ and for all values of the scaling factor λ. Power laws $f(x) = x^n$ are clearly scale-invariant for $\Delta = n$.

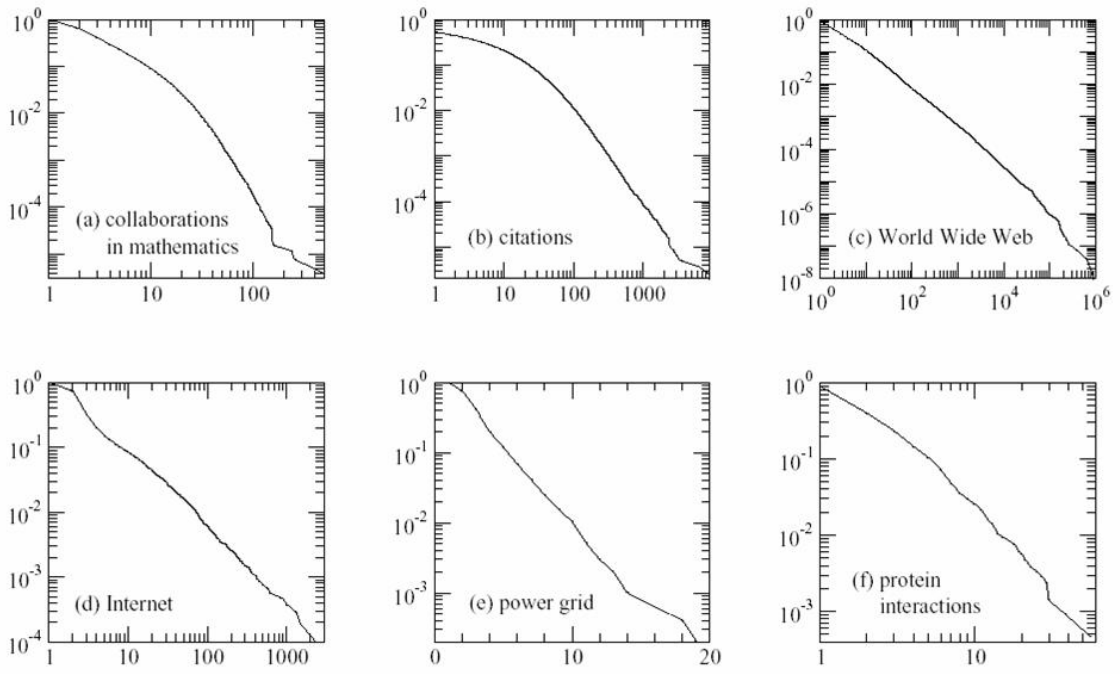


Figure 1.5: Degree distribution (power law) of several real world networks. In general, almost all of the real world networks have a degree distribution of the form $P(k) \sim k^{-\gamma}$, with $1.5 \leq \gamma \leq 3.5$. (from ref. [3])

1.4.2 Degree-assortative Mixing

A network is said to show degree assortative mixing if the nodes in the network that have many connections tend to be connected to other nodes with many connections. A typical way to measure this property is to look at the joint probability $P(k, k')$ defined in section 1.3.1. If $P(k, k')$ exhibits positive curvature this means that nodes with high degrees are, on the average, preferentially connected to nodes with high degrees. For practical reasons, it is better to introduce Pearson's coefficient [104, 108]

$$r = \frac{L^{-1}\sum_i j_i k_i - \left[L^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2}{L^{-1}\sum_i \frac{1}{2}(j_i^2 + k_i^2) - \left[L^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2} \quad , \tag{1.8}$$

where L stands for the total number of links in the network and all sums run over all the links (i.e., $i = 1, \ldots, L$). The Pearson's coefficient r effectively measures the average degree correlation of neighbor nodes (j_i and k_i are the degrees of nodes placed at the end of the i-th link of the network). When a network shows assortative mixing, one should have $r > 0$ and one says that the network is **assortative**. On the other hand, when $r < 0$ the network

is not assortative or **disassortative**. Table 1.1 reveals a strong difference between social networks that are in general assortative and technological and biological networks that are disassortative. Information networks, on the other hand, do not exhibit a clear tendency for the assortative mixing, since we can find positive and negative values of r for real networks belonging to this category. In chapter 3 and in **Paper I** [165] I shall report on my own studies of dissasortativity in connection with scale-free and self-similar networks.

1.4.3 Small-world Property

The average distance [Eq.(1.3)] of almost all the real networks is small when compared to their size N. In particular one can say that the average distance of real networks scales as

$$\langle \ell(N) \rangle \sim \log N \quad . \tag{1.9}$$

Typical values of $\langle \ell \rangle$ for real networks are shown in Figure 1.6 and reported in Table 1.1. One refers to this property of real networks, saying that they have **small-world** features (the expression *small-world* was introduced by S.Milgram [94] in the context of social networks).

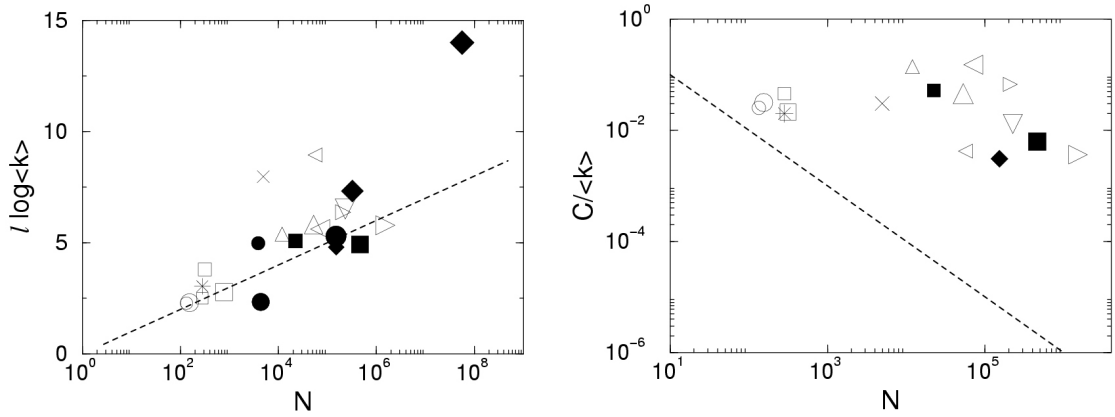


Figure 1.6: Average distance (left) and average clustering coefficient of order 3 (right) of several real networks (symbols) and those analytically calculated for the random graph associated with the network. (from ref. [3])

	Network	type	N	L	$\langle k \rangle$	$\langle \ell \rangle$	$\langle C^{(3)} \rangle$	γ	r	Ref(s).
social	film actors	und.	449 913	25 516 482	113.43	3.48	0.20	2.3	0.208	[8]
	physics coauthorship	und.	52 909	245 300	9.27	6.19	0.45	-	0.363	[102, 103]
	math coauthorship	und.	253 339	496 489	3.92	7.57	0.15	-	0.120	[37, 52]
information	citation network	dir.	1 022	5 103	4.99	4.87	0.13	3.0	0.157	[128]
	WWW nd.edu	dir.	269 504	1 497 135	5.55	11.27	0.11	2.1/2.4	-0.067	[6, 7]
	WWW Altavista	dir.	203 549 046	2 130 000 000	10.46	16.18	-	2.1/2.7	-	[25]
	word co-occurence	und.	460 902	17 000 000	70.13	-	-	2.7	-	[39, 67]
technol.	power grid	und.	4 941	6 594	2.67	18.99	0.10	-	-0.003	[159, 158, 8]
	train routes	und.	587	19 603	66.79	2.16	-	-	-0.033	[136]
	Internet	und.	10 697	31 992	5.98	3.31	0.035	2.5	-0.189	[44]
biological	metabolic network	und.	765	3 686	9.64	2.56	0.090	2.2	-0.240	[69]
	protein interactions	und.	2 115	2240	2.12	6.80	0.072	2.4	-0.156	[68]
	marine food web	dir.	135	598	4.43	2.05	0.16	-	-0.263	[65]
	neural network	dir.	307	2 359	7.68	3.97	0.18	-	-0.226	[160, 159]

Table 1.1: Statistical measures on real networks. In this table, indicated from left to right are: the name of the network, the type of the network [directed (dir.) or undirected (und.)], the size of the network N, the number of edges L, the average degree $\langle k \rangle$, the average distance $\langle \ell \rangle$, the average clustering coefficient of order 3 $\langle C^{(3)} \rangle$, the exponent γ of the power-law for the degree distribution, the Pearson coefficient r and the papers in which the respective network is studied. Moreover, the networks are organized according to the same classification as given in section 1.1 [see first column on the left].

1.4.4 Transitivity or Clustering Property

Almost all the real networks have large values of the average clustering coefficient [see Eq.(1.6)]. A real network is said to have a "large" clustering coefficient if this value is greater than the actual value calculated for the random model[3] associated with the network itself. Moreover, this means that the clustering coefficient has no explicit dependence on the size N of the network [see Figure 1.6 and Table 1.1]. In particular, high values are not only registered for the "standard" clustering coefficient of order 3 but for any order $g \geq 3$ of the clustering coefficient.

1.4.5 Other Properties

Other statistical properties have been studied for real complex networks. For example, it was also found that the distribution of the clustering coefficient [126] and the strength distribution in weighted networks [12] obey power laws. Moreover, one can describe one of the most-studied and interesting features of real networks: the so called **community structure** of complex networks [48, 119, 30, 105]. A real network, in general, has an internal organization in clusters or communities of nodes, and these communities are usually well organized in a hierarchical way. The interested reader can find more information and references in the review [105]. Finally, I should mention that recently a new topological property of real complex networks has been discovered [140]: some real networks are self-similar. Here I will not describe this property, despite the fact that it appears to be extremely interesting, because it will be the topic of chapter 3.

1.4.6 Regular Networks

In the previous sections, I summarized the basic topological features of real networks that can be determined by simple statistical measurements. Here I would like to remark on which are the main topological differences between real complex networks and the standard topologies usually considered in physics: the lattices. The reader can easily realize that a d-dimensional lattice can also be represented as a network: each site of the lattice is a node and each bond is an edge. The degree distribution for lattices is not scale-free: in the case of periodic boundary conditions it is just a delta[4] function centered on the actual value of the number of

[3] Given a network one defines as its random model or the random graph associated with the network a graph where the degree of each node of the original network is preserved, but all the edges of the original network are rewired randomly.

[4] $\delta(x - x_0)$ is called Kronecker's function or simply discrete delta function centered in x_0, if $\delta(0) = 1$ and $\delta(x - x_0) = 0 \ \ \forall x \neq x_0$.

its nearest neighbor in the lattice. The average distance $\langle \ell \rangle \sim N^{1/d}$, so that a d-dimensional lattice does not exhibit the small-world feature. The clustering coefficient is in general high, but this depends on the particular lattice considered and on the order g of the clustering coefficient calculated: for example, on two-dimensional square lattices one measures $C_i^{(3)} = 0$ and $C_i^{(4)} = 1$ for any node i.

1.5 Models for Complex Networks

In this section, I briefly review three main models used to describe the topological features of real complex networks [see section 1.2]. In fact, the main goal of statistical physicists is not only to characterize the statistical features of real complex networks by numerical measurements but also to construct models that can reproduce these features and explain the reasons for how these topological properties can arise.
The descriptions below refer to the random graph or Erdös-Rènyi model, the Watts-Strogatz model and the Albert-Barabási model. These three models refer to undirected and unweighted networks. I pay more attention here to the first two models because I will make use of them in chapter 5 and in chapter 4, respectively. The Albert-Barabási model is presented here just for educational purposes, since it represents the most cited model of complex networks so far proposed by physicists. Of course, the big attention given recently by physicists and mathematicians to complex networks led to the formulation of a huge number of different models. Good descriptions about these models can be found in the reviews [3, 40, 105]. Complex networks represent an intensively studied topic now days. This means that new models are continuously introduced, so that no complete review exists about all the models introduced up to now.

1.5.1 Random Graph (Erdös-Rènyi Model)

The random graph model was introduced in the middle of the 1950s by A.Rapoport and R.Solomonoff [137, 125] and independently at the beginning of the 1960s by P.Erdös and A.Rènyi [42, 43]. The main theoretical results of this model were obtained by P.Erdös and A.Rènyi, so that the random graph model is also known as Erdös-Rènyi model.

Construction of the Model

The random graph model is an extremely simple model. Consider a set of N nodes without any connections and fix a value for the parameter $0 \leq p \leq 1$. Then pick up every pair of nodes out of all the $\frac{N(N-1)}{2}$ couples of distinct nodes. For each pair of nodes, say i and j,

extract a number $0 \leq r \leq 1$ at random out of a uniform distribution: if $r < p$, then connect the nodes i and j so that $A_{i,j} = A_{j,i} = 1$; if $r \geq p$, do not connect the nodes i and j so that $A_{i,j} = A_{j,i} = 0$. Of course the diagonal elements of the adjacency matrix are all zero, $A_{i,i} = 0 \ \ \forall i$, since no self-connections are allowed.
By construction each link in the network is independent of the others. The probability that the total number of links L is equal to l is therefore

$$P\left(L = l\right) \ = \ B_l^{N(N-1)/2} \, p^l \ \left(1-p\right)^{N(N-1)/2-l} \quad , \tag{1.10}$$

where B_q^s stands as binomial coefficient[5]. The probability of Eq.(1.10) is therefore a binomial distribution. The peak of the distribution is located at the average value

$$\langle L \rangle \ = \ \frac{N(N-1)}{2} \, p \quad ,$$

with possible fluctuations characterized by the variance

$$\sigma_L \ = \ \sqrt{\frac{N(N-1)}{2} \, p \ \left(1-p\right)} \quad .$$

Degree Distribution

As it should by now be clear to the reader, other probabilities in the random graph model are expressed by binomial distributions, or eventually approximated by Gaussian or normal distributions and Poissonian distributions. In particular the probability that an arbitrary chosen node $i \in G$ has exactly k connections in the network is

$$P(k_i = k) \ = \ B_k^{N-1} \, p^k \ \left(1-p\right)^{N-1-k} \quad . \tag{1.11}$$

For sufficiently large values of N and a sufficiently small value of p, Eq.(1.11) is well approximated by the distribution of Poisson

$$P(k_i = k) \ = \ e^{-\langle k \rangle} \, \frac{\langle k \rangle^k}{k!} \quad , \tag{1.12}$$

where $\langle k \rangle = (N-1)p$ is the average degree of connections. If one wants to calculate the probability distribution over all the nodes of the graph G, one should say first of all that the average number of nodes with degree k is

$$\langle N_k \rangle \ = \ N \, P(k_i = k) \quad .$$

[5]The binomial coefficient, $B_q^s = \binom{s}{q} = \frac{s!}{(s-q)! \ q!}$, with $s \geq q \geq 0$ and $q, s \in \mathbb{N}$, counts the total number of possible ways in which one can select q elements out of s total elements in a set.

Then the probability to find exactly n nodes with degree k is

$$P(N_k = n) \;=\; e^{-\langle N_k \rangle}\, \frac{\langle N_k \rangle^n}{n!} \quad , \tag{1.13}$$

that is a Poissonian inside an other Poissonian. For sufficiently large values of N, Eq.(1.12) can be well approximated by its maximal value $P\,(k_i = \langle k \rangle)$. Therefore, one may rewrite Eq.(1.13) according to this approximation and evaluate the degree distribution as

$$P(k) \;=\; e^{-Np}\, \frac{(Np)^k}{k!} \quad . \tag{1.14}$$

The degree distribution of Eq.(1.14) can therefore be well represented by its average value $\langle k \rangle = pN$ and its variance $\sigma_k = \sqrt{Np(1-p)}$. The former result can be verified by numerical simulations such as the ones reported in Figure 1.7. As one can easily realize, the degree distribution of the random graph model [Eq.(1.14)] is not representative of the one typical of real complex networks [see section 1.3.1].

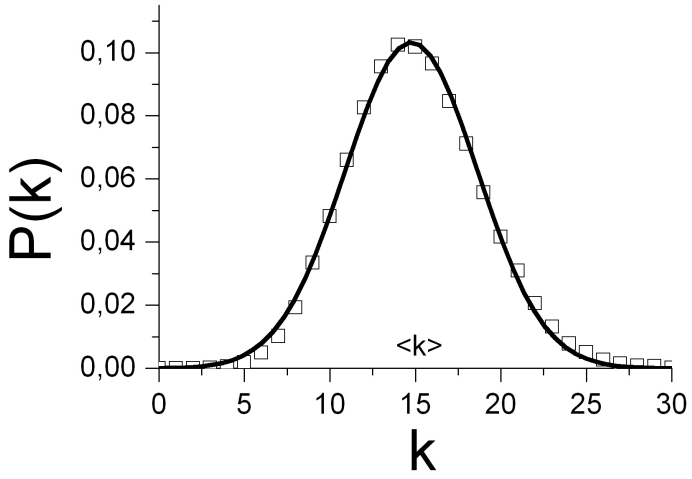


Figure 1.7: Degree distribution for a random graph with $N = 5000$ and $p = 0.003$. The numerical results (symbols) are obtained after averaging over 100 different realizations of the model. The full line stands for the Gaussian distribution with average value and variance equal to those of Eq.(1.14). (from ref. [118])

Diameter, Average Distance and Components

In general, the Erdös-Rényi model is not composed of only one connected component. Effectively there exists a threshold value for the probability of connection p, in general indicated by p_c, below which the graph is composed of several disconnected components and above which the graph is composed of just a single component. The actual value of the threshold is $p_c = 1/N$, where N stands again for the size of the graph.
When the graph is composed of several components, the diameter of the graph is by definition

equal to infinity. When the graph is composed of only one component, however it can be calculated easily. By construction, the number of nodes that should be at distance ℓ from a randomly selected node is $n_\ell \simeq \langle k \rangle^\ell$. Since the diameter [Eq.(1.4)] is the maximal value of distances inside the graph, all the nodes should be within the distance equal to the diameter from a randomly selected node. Therefore, one can write

$$N = n_d \simeq \langle k \rangle^d \quad \Rightarrow \quad d \simeq \frac{\log N}{\log \langle k \rangle} \quad . \tag{1.15}$$

Since the Erdös-Rényi model is a random model, Eq.(1.15) represents not only the diameter of the graph but also its average distance for sufficiently large values of N, since no node is distinguished and for large N most of the nodes have the same distance from the one selected, that is at distance of the order of the diameter. Therefore, the random graph model possesses the small-world feature typical of real networks [see section 1.3.3].

Clustering Coefficient

The clustering coefficient of order 3 [Eq.(1.5)] for the Erdös-Rényi model can be estimated easily. Given a node i inside the graph, the fact that two nodes connected to i are again connected with each other happens with probability p. Thus, it happens with the same probability that two randomly chosen nodes are connected. This implies that the clustering coefficient of order 3 of a random graph is just

$$C^{(3)}_{ran} \; = \; p \; = \frac{\langle k \rangle}{N} \quad . \tag{1.16}$$

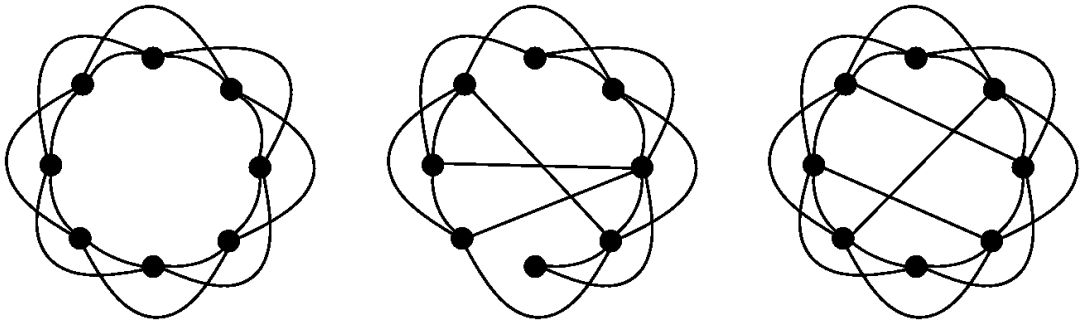
Again, it should be stressed that the random graph model is not representative of real complex networks [see section 1.4.4] since the average clustering coefficient depends on the size of the system [see Eq.(1.16)].

1.5.2 Watts-Strogatz Model

In 1998 the two American physicists D.J.Watts and S.H.Strogatz [159, 157, 158] introduced an interesting new model. Due to its topological features, the Watts-Strogatz model is well known in the literature as the **small-world** model.

Construction

One starts from a ring of N nodes [see Figure 1.8(a)]. Each of these nodes is connected to the first K neighbor nodes ($K/2$ to its right and $K/2$ to its left). With the aim of having



(a) Starting (b) Rewiring (c) Adding

Figure 1.8: **(a)** Starting configuration of the Watts-Strogatz model. Here the size of the lattice is $N = 8$ and the number of connections per node is $K = 4$. **(b)** Three links of the original lattice are rewired according to the prescriptions of the model. **(c)** Modified version of the Watts-Strogatz model: here the "short cuts" are simply added and do not result from the rewiring of formerly present connections.

a sufficiently sparse graph, one should consider $N \gg K$. In this way, one has constructed a one-dimensional lattice with periodic boundary conditions. A similar construction can be evidently performed also in higher dimensions and with different boundary conditions. In the following I restrict my attention only to the original formulation of Watts and Strogatz.
After one has constructed the ring of connections, the links of the network can be rewired with probability $0 \leq p \leq 1$ [see Figure 1.8(b)]. This means that one visits all the links of the network. For each of them, say (i, j), one extracts a number $0 \leq r \leq 1$ randomly out of a uniform distribution. If $r < p$, then one rewires the link: this means that one deletes the former link $A_{i,j} = 1 \rightarrow A_{i,j} = 0$, then one selects a new node q not yet connected to i and one draws a new connection $A_{i,q} = 0 \rightarrow A_{i,q} = 1$. However, if $r \geq p$, nothing happens. The first procedure allows an evolution of the original regular lattice towards a completely random structure. For $p = 0$ one evidently does not touch the original lattice, while for $p = 1$ every link is rewired and the resulting graph is quite similar to the random graph model proposed above in section 1.4.1. The rewired links are appropriately called "short cuts" since they graphically "cut" the ring. In a slightly different model [109] the short cuts are not induced by the rewiring of the preexisting links, but are simply added [see Figure 1.8(c)]. The resulting graph is therefore a superposition of the original lattice and of a random graph. The main features of the model are not changed at all.

Degree Distribution

For sufficiently large networks ($N \gg K$), the degree distribution $P(k)$ of the small-world model has a *cutoff* at K. This means that there is a null probability to find a node with a degree less than K. The general expression for the degree distribution can be calculated [13] and is given by

$$P(k) \;=\; \sum_{n=0}^{\min(k-K,K)} \binom{K}{n} (1-p)^n \, p^{K-n} \, \frac{(pK)^{k-K-n}}{(k-K-n)!} \, \exp(-pK) \quad , \tag{1.17}$$

for $k \geq K$ and $P(k) = 0$ for $k < K$. The degree distribution of the Watts-Strogatz model is not scale-free as the one measured for real complex networks [see section 1.3.1].

Average Distance

For the average distance $\langle \ell \rangle$ of the Watts-Strogatz model there is not yet any analytical estimation. Numerical simulations reveal a smooth transition of $\langle \ell \rangle$ as a function of p [see Figure 1.9]. At the moment, the only fact known is that the average distance obeys the scaling relation [109]

$$\langle \ell \rangle \;=\; f(NKp) \quad , \tag{1.18}$$

where

$$f(x) = \frac{1}{2\sqrt{x^2+2x}} \, \tanh^{-1}\left(\frac{x}{\sqrt{x^2+2x}}\right) \quad . \tag{1.19}$$

The former scaling relation of Eq.(1.18) tells that one can pass from the *large world* regime (i.e. large values of $\langle \ell \rangle$) to the *small-world* regime (i.e. small values of $\langle \ell \rangle$) by increasing p. This can be also understood from the asymptotic values of Eq.(1.18):

$$\lim_{p \to 0} \langle \ell(p) \rangle \sim \frac{N}{K} \quad , \quad \lim_{p \to 1} \langle \ell(p) \rangle \sim \frac{\log N}{\log K} \quad . \tag{1.20}$$

As expected when all links are rewired ($p = 1$) the small-world model behaves as a pure random graph [see Eq.(1.15)].

Clustering Coefficient

The analytic expression for the clustering coefficient of the small-world model is calculated in [109] with the result

$$C_{sm}^{(3)} \;=\; \frac{3(K-1)}{2(2K-1)} \, (1-p)^3 \quad . \tag{1.21}$$

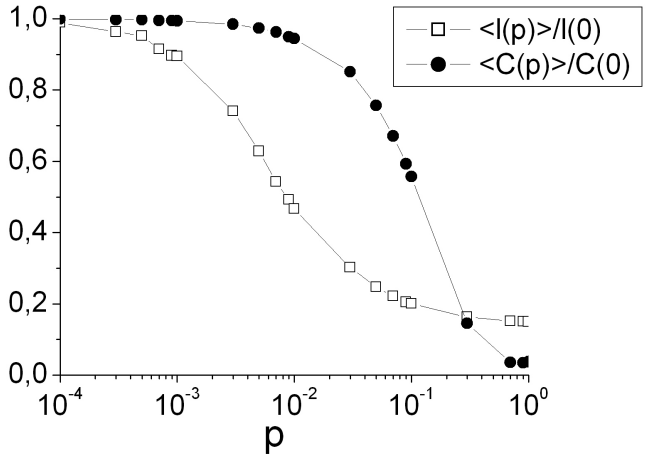


Figure 1.9: Average distance $\langle \ell(p) \rangle$ (squares) and average clustering coefficient $\langle C(p) \rangle$ (circles), both divided for the respective values at $p = 0$, for a Watts-Strogatz model with $N = 200$ and $K = 2$. (from ref. [118])

As one can see from Eq.(1.21) [see also Figure 1.9] the clustering coefficient of the Watts-Strogtaz model has large values up to $p \sim 0.1$ and it does not depend on the size of the system N.

Observations

Here I would like to summarize the results shown so far for the Watts-Strogatz model. This model qualitatively shows three different regimes as a function of the rewiring probability p [see Figure 1.9]:

- For $p \lesssim \frac{1}{NK}$, the topological features of the model are extremely close to those of a one-dimensional lattice, with a large clustering coefficient and large average distance.
- For $\frac{1}{NK} \lesssim p \lesssim 0.1$, the model is in the small-world regime, characterized by a large clustering coefficient and small average distance. It is this regime in which the model is representative of real networks since it possesses both a small average distance and high clustering coefficient [see sections 1.3.3 and 1.4.4 respectively].
- For $p \gtrsim 0.1$, the model exhibits properties similar to the random graph model [see section 1.4], with a small clustering coefficient and small average distance.

1.5.3 Albert-Barabási Model

Both models introduced so far in sections 1.4.1 and 1.4.2 exhibit degree distributions that do not correspond to those typical of real complex networks. In general, a real complex network

has a degree distribution which is scale-free [see sections 1.3.1]. Moreover both the random graph and the Watts-Strogatz models are static models. In their construction all the nodes are introduced simultaneously. This fact does not agree with a large number of real complex networks. In the World Wide Web network, new pages are added continuously over time. In the network of scientific citations, new papers are written from time to time.
The model introduced by R.Albert and A.L.Barabási in 1999 [4, 5] includes both the former properties: it is an *evolving* network and it shows a scale-free degree distribution. Here I give for this model a description that is shorter than the ones performed before for the random graph and the small-world model, since I will not use the Albert-Barabási model in the present thesis. Still, the Albert-Barabási model has pedagogic relevance for any reader who is interested in complex networks.
In the Albert-Barabási model there are two main ingredients: regular growth of the network and **preferential attachment**. One starts at time $t = 0$ with an initial set of N_0 nodes and a set of L_0 connections between them. The initial network can be considered fully-connected, but the properties of the model, valid in the thermodynamical limit, are not affected by the initial configuration. At time $t = 1$, a new node is introduced in the network. The new node introduces $m \leq N_0$ new links. Each of these links is therefore created between the new node and a formerly present node, say i. In particular, the node i is chosen at random out of all the possible nodes of the network with the probability

$$\Pi(k_i) \;=\; \frac{k_i}{\sum_j k_j} \quad . \tag{1.22}$$

The numerator of Eq.(1.22) stands for the degree of the node i, while the denominator of Eq.(1.22) is the proper normalizing factor for the probability $\Pi(k_i)$. Eq.(1.22) represents the so-called preferential attachment rule. The larger the degree of the node i, the larger is the probability to increase its degree even more. "The Rich get richer." At each instant of time $t > 1$, a new node will be added to the network bringing again m new links. These links will then connect the new node to old nodes already present in the network at time $t - 1$, where the discriminant rule for the choice of the nodes will be given again by the preferential attachment of Eq.(1.22).
In the long time limit, $t \to \infty$, the degree distribution approaches a power law [Eq.(1.7)] telling that the network has a scale-free degree distribution. In particular, the exponent of the power law is $\gamma = 3$ for the Albert-Barabási model.

1.6 Dynamics on Complex Networks

Up to now, I have focussed just on the topological properties of real systems that can be described as networks. The topology represents in most cases only the backbone of real systems. After several years of research focusing on the topological properties of networks, scientific attention recently moved towards the more interesting relationship between topology and dynamics. Note that at the end almost all the real systems are dynamical: each unit, as well as each edge, may have its intrinsic dynamics, its time evolution being regulated by the interactions with some part of the system, and so by the network. The dynamical variables, which one is interested in, obviously differ from system to system. An example of a typical physical network is an electric circuit: on each edge one has an electric device (generator, resistor, capacitor or inductor) while the nodes represent the junctions between them. In this section, I will not give a complete review of all types of dynamics studied on complex networks. Dynamics on complex networks will be the main topic of the thesis and specific dynamical models will be discussed in detail in chapters 4, 5 and 6. Now I would like to stress the fact that it is in general useful to distinguish between two types of dynamical systems: discrete or continuous in time and/or internal space (that is, the state space on an individual unit). Moreover, I should remark that so far very simple dynamical models have been studied on networks since the complexity of the topologies is often sufficient to give rise to non-trivial dynamical properties. In the following subsections, I give some specific examples of both discrete and continuous dynamical systems. The following should not be understood as a complete review of all the dynamical systems studied on networks but just a simple list of examples for a better understanding of the topic. Here I only consider systems that are either discrete in both state space and time or continuous in space and time, but mixtures of both types also exist.

1.6.1 Discrete Dynamical Systems

In a discrete dynamical system, the state of one unit is characterized by a discrete set of numbers. In the simplest and most studied discrete dynamical models, the variable associated with each unit is a binary variable: this means that the unit can be in two states only. Depending on the context, these two states can be interpreted as up and down, as in the case of spin systems, true and false, as in the case of Boolean systems, friend or enemy, as in the case of social systems, etc. Often the intrinsic time of the dynamical system is also defined in discrete units. Typical numerical algorithms for the simulations of these systems can be asynchronous or synchronous: one randomly chosen unit is updated at each step in the former case, while all the units evolve synchronously in the latter. Generally different

updating schemes lead to different dynamical properties of the system. I will approach this problem in chapter 6. A concrete example of discrete dynamical systems studied on networks is represented by the so-called **Boolean networks** [74, 73, 75] used successfully for the dynamical description of genetic regulatory networks [see section 1.1.4]. Other kinds of discrete models have been studied on networks: **Ising** [59, 64, 60, 81, 138] and **Potts** [38, 129] models, the **voter model** [32, 143, 149], and **contact processes** [31, 90] are just a few examples. In chapter 5 and in my **Paper IV** [124] and **Paper V** [123] , I will study a discrete dynamical model for complex social networks.

1.6.2 Continuous Dynamical Systems

In a continuous dynamical system both the state variables of the units and the intrinsic time of the dynamics are continuous variables. Therefore the dynamical equations describing the system are coupled ordinary or partial differential equations. For this framework there are also many examples. **Neural networks** have been studied as networks of **pulse-coupled oscillators** [95, 150, 151], where the interaction itself happens in discrete steps but the state space and time evolve continuously. Networks of **chaotic oscillators** have been studied as well [22]. Networks of coupled **limit-cycle oscillators** is another well studied case of continuous dynamical processes applied to networks [62, 168, 78, 98, 99, 113, 63, 1, 114, 11]. In particular, the latter example will be discussed in detail in chapter 4 and in my **Paper II** [120] and **Paper III** [121] .

Chapter 2

Phase Transitions and Critical Phenomena

Introduction

One of the most important features of natural systems is the great diversity in size or length of their structure. For the description of a rigid body rolling down an inclined plane, there is no need to know anything about the internal structure of the rigid body. On the other hand, the atomic interactions inside the rigid body are the same if the rigid body is rolling down an inclined plane or if it is just at rest on a non-inclined plane. In other words, for understanding the dynamics of rigid bodies, one does not need to know anything about the atomic physics, and *vice versa*. In general, physical events differing with great disparity in order of magnitude have little influence on each other. They simply do not communicate, so that the phenomena associated with each scale can be analyzed independently.
There is a class of phenomena for which many length scales are equally important. These phenomena are called **critical phenomena**. The word "critical" is used in science with different meanings. In statistical physics, it is used in the context of **phase transitions**. In this framework, the word "critical" often describes a system at the border between order and disorder. At this border the system is characterized by an extremely large susceptibility to external factors and strong correlations between different parts of the systems itself. Both these properties are consequences of the fact that events at any scale contribute equally.
This chapter will be an introduction to phase transitions and critical phenomena. I will discuss this field of physics in a very generic and brief way, since it represents a broad topic of statistical physics and a huge number of books and papers have been written about it. I will start introducing the concept of phase transition [section 2.1]. I will describe phase transitions in general, introducing the most important observables and parameters able to characterize them

[sections 2.1.1 and 2.1.2]. Moreover, I will distinguish between discontinuous and continuous phase transitions [section 2.1.3]. In section 2.2 I will focus on continous phase transitions and in particular on the percolation transition [section 2.2.1]. Using the percolation model, I will briefly describe the main aspects of continuous phase transitions [section 2.2.2]. Moreover I will dedicate the section 2.2.3 to the introduction of the standard numerical tools used in order to characterize a particular continuous phase transition. I will conclude the chapter with a discussion of a special case of percolation, namely directed percolation, in section 2.2.4 and finally, paying attention to phase transitions occurring on networks [section 2.3]

2.1 Phase Transitions

Thermal phase transitions are matters of everyday experience. Think about H_2O, for example: ice turns into water and water turns into vapour as the temperature is increased. The former transitions can be read in the inverse ways and one can observe also transitions as ice into vapour and *vice versa*. Ice, water and vapour are the possible "phases" of H_2O and the transitions between different phases like the ones so far reported are called "phase transitions". There are many other possible phase transitions, where the word "phase" is not used with the meaning of solid, liquid or gas, but with a broader meaning of a particular state or property of the system analyzed. A physical example is given by the transition between the ferromagnetic and paramagnetic phases of magnetic materials at the Curie point. A non-physical example is represented by the transitions in intractable computational complexity problems such as k-satisfiability problems, where the transition from solvable to unsolvable instances exhibits threshold behavior depending on the ratio of number of clauses to number of variables. I will consider the latter example in detail in chapter 5 and in **Paper IV** [124] . Independently of the particular kind of system undergoing a phase transition, the description of the phase transition can be performed with a general procedure. In the following, I will introduce some basic quantities that are common to every system that shows a phase transition.

2.1.1 Order Parameter

The first observable, which should be defined on a system showing a phase transition, is the **order parameter**. Originally, the order parameter was introduced as a measure for the degree of order of the system. Therefore, it is able to show when the system undergoes a transition between an ordered and an disordered phase. Again, the notion of order parameter has origin in thermal phase transitions. For example, in the case of a system that changes

its phase from liquid to gas, a suitable order parameter is represented by its density ρ. In the case of the ferromagnetic-paramagnetic transitions, a good order parameter is given by the magnetization m. Meanwhile, the order parameter does not always represent a measure of the "order" of the system but it should indicate the phase. For example, in case the of a percolation transition, as the reader shall see in section 2.2.1, a good choice of the parameter is given by P_∞ that stands for the probability of having percolation.

2.1.2 Critical Point

The particular value of the external parameters for which the system undergoes a change of phase is called **critical point** or **critical threshold**. In the absence of an external magnetic field, one can determine whether iron is in the ferromagnetic or the paramagnetic phase by knowing whether the actual temperature T is smaller or larger than the Curie or critical temperature T_c, respectively. In case of $\mathrm{H_2O}$, the phase of this element is determined by the temperature T and the external pressure p. Therefore, in this case the critical point is not just a single point but a curve (T_c, p_c) in the plane (T, p). The plot of (T_c, p_c) is also called **phase diagram** and in the particular case of $\mathrm{H_2O}$, this plot shows interesting features about the presence of the "tricritical" point.
In experiments or numerical simulations, the determination of the critical values of the external parameters for which a system undergoes a phase transition is in general not easy. In case of models, simulated in computer programs, finite-size scaling analysis provides a proper tool for the estimation of the critical point, whose position depends on the size of the system, but converges to the critical point in the infinite-volume limit. I refer to section 2.2.3 for further details.

2.1.3 Classification of Phase Transitions

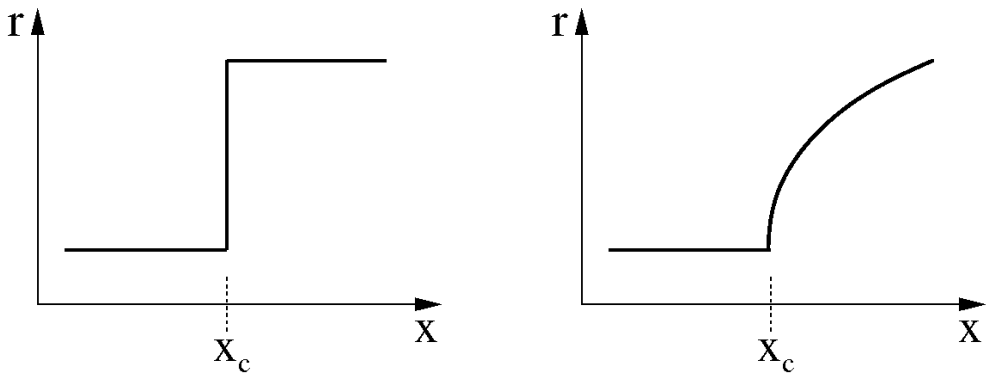
There are two main categories of phase transitions which depend on how the transition between the different phases takes place[1]. A schematic graphical representation is reported in Figure 2.1.
The order parameter at the critical point can have a sudden transition and in this case one will talk about a discontinuous phase transition or simply a **first-order phase transition** [see

[1]According to the original scheme of P.Ehrenfest, there are in principle an infinite number of categories or "orders" of phase transitions. Ehrenfest's scheme was to classify a phase change as n^{th}-order if the lowest-order derivative of the order parameter to be discontinuous or non-analytic across the transition is an n^{th} derivative. In the present thesis, I will follow the actual classification of phase transitions which prescribes the consideration of all the phase changes of orders larger than one as continuous phase transitions.

Figure 2.1(a)]. In the case of thermal physics, transitions of this type are phase transitions involving a finite latent heat: for example, phase transitions of H_2O involve a finite latent heat and the emission or absorption of latent heat justify the radical change of structure of H_2O at the critical temperature. I will study some transitions of this type during the present thesis. In chapter 4 and in particular in **Paper II** [120] and **Paper III** [121] , I will consider systems of identical limit-cycle oscillators in the presence of a pacemaker. Depending on the particular choice of the typical parameters of the system and of the pacemaker, I will observe a first-order phase transition between a fully synchronized phase and a non-synchronized phase. In chapter 5 and in **Paper IV** [124] , I will consider a model using the same description as that of *k*-satisfiability problems, which undergo a discontinuous transition between a satisfiable phase and an unsatisfiable phase [47].
Another class of systems exhibits smooth phase transitions. The order parameter changes from a finite value, denoting in general an ordered phase, continuously to vanishing values, standing for disordered phases [see Figure 2.1(b)]. This kind of phase transitions are called **second-order phase transitions** or **critical phenomena**. Transitions of the second-order are illustrated by the magnetic transition of iron as a function of the temperature. The Kuramoto model, introduced in section 4.1.1, also shows a second-order phase transition. In this thesis I will discuss several types of phase transitions belonging to this category. In chapter 5 and in **Paper IV** [124] , I will introduce a social model for which the stationary states show a continous phase transition. Moreover, in chapters 5 and 6 and in **Paper V** [123] and **Paper VI** [122] , I will discuss two models, similar to the one of percolation [introduced in section 2.2] and the one of directed percolation [introduced in section 2.2.4], respectively, both undergoing a second-order phase transition.

2.2 Application to Percolation

In this section I will focus on the percolation model. This model exhibits a second-order phase transition. Since the theory of continuous phase transitions predicts features which are independent of the specific model, one can consider the percolation model as a prototype of critical phenomena. This part does not serve as a complete introduction to percolation theory, for which I refer to more complete reviews [144, 28]. Moreover, in section 2.2.4 I will consider an "anisotropic" version of percolation, namely directed percolation. For directed percolation I will write just a short introduction and I refer the reader to the book [89] and to reviews [61, 112] for more details.

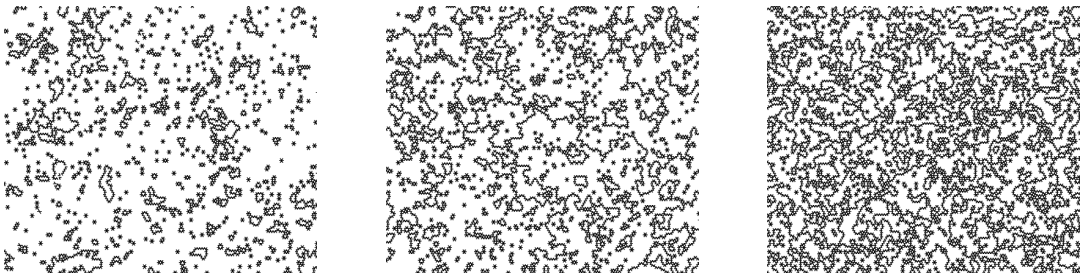


(a) First-Order (b) Second-Order

Figure 2.1: **(a)** In a first-order phase transition, the order parameter r jumps discontinuously at the critical threshold X_c. **(b)** In a second-order phase transition, there is no sudden change of the order parameter but above the critical point X_C the order parameter continuously changes its value as a function of the parameter X.

2.2.1 The Model

Percolation deals with a geometric phase transition. Consider a square lattice in which each site can be occupied with probability p or not with probability $1 - p$. The construction is made in a very simple way: one fixes *a priori* a value of the probability p, then visits all the sites. For each site, extracts a random number $0 \leq r \leq 1$ out of a uniform distribution: if $r < p$, then the site is marked as occupied, while if $r \geq p$ the site is marked as non-occupied or empty. Depending on the initial choice of p, the occupied sites can be said to percolate or not through the lattice. This means that a hypothetical walker can enter from one side of the lattice and exit to the opposite side by just walking on occupied sites and never jumping over non-occupied sites. The model is physically interesting if the occupied sites are interpreted as conductors and the empty sites as insulators. For small values of p, the lattice is composed mainly of insulators and only local clusters of conductors are present in the lattice. For sufficiently large values of p, however, many different conduction paths exist between opposite sides of the lattice, so that current can flow or percolate over the lattice that behaves as a conductive medium. Other physical phenomena can be easily interpreted as percolation phenomena. For example, percolation was used fruitfully for the description of porous and amorphous materials [169], disordered ionic conductors [27] and fragmentation [41]. In biology, percolation also provides a useful model for describing epidemic spreading [49, 50].



(a) Non percolation (b) Percolation threshold (c) Percolation

Figure 2.2: Example of bond percolation on a triangular lattice: in **(a)** the lattice does not percolate, while in **(c)** the lattice percolates. In **(b)** one is at the percolation threshold and one is at border between percolation and non percolation regimes. Figures taken from **Paper V** [123] .

In the former paragraph I presented the so-called **site percolation**, since I was considering the fact that the sites of the lattice can be occupied or not. Exactly in the same way but considering the bonds of the lattice instead of sites, it is possible to introduce the so-called **bond percolation**. Depending on the particular physical phenomenon that one wants to describe, it is useful to consider site or bond percolation. The main features of percolation, like critical exponents, do not depend on this choice.

2.2.2 Percolation as a Critical Phenomenon

Since the value of p can be tuned in a continuous way, a transition should exist between the non-percolating phase and the percolating phase. The value of p for which this transition takes place is the so-called **percolation threshold** and is generally denoted by p_c [see section 2.1.2 and see Figure 2.2(b)]. The percolation transition is a *geometrical* phase transition. The quantity which is usually used to describe the percolation transition is P_∞, the probability for a given site to belong to an infinite cluster of connected occupied sites. The word "connected" means that the occupied sites of the infinite cluster should be neighbors in the sense that one walker can move over all the sites of the cluster passing only over sites within the same cluster. The quantity P_∞ is a good **order parameter** [see section 2.1.1], able to describe the percolation transition: for $p < p_c$, $P_\infty = 0$, since only finite clusters exist and no percolation is possible in this phase; for $p > p_c$, $P_\infty > 0$, so that there is a finite probability that a given site belongs to an infinite cluster that percolates through the lattice. In particular, for $p = 0$, no finite and infinite clusters exist at all, while for $p = 1$ all the sites are occupied and all of them belong to the infinite cluster, so that $P_\infty = 1$. Above the critical threshold p_c, but in

its vicinity, the order parameter increases as a power of the distance from p_c,

$$P_\infty \sim (p - p_c)^\beta \quad . \tag{2.1}$$

The exponent β introduced in the Eq.(2.1) is called **critical exponent**. It is denoted by β in analogy with the standard notation of thermal critical phenomena where the critical exponent referring to the order parameter is always indicated by the Greek letter β. There are several other critical exponents depending on the quantity to which they refer: ν, γ, τ, σ, d_f, etc. Here in the introduction, I do not give an explicit definition of all of them, but I make use of them in my **Paper V** [123] . It should be noted that as in the case of β the Greek letters indicating the various critical exponents are taken from standard notation used for thermal phase transitions: therefore ν is the critical exponent of the divergence of the **typical linear size** ξ of a cluster of occupied sites, γ stands for the critical exponent of the divergence of the **susceptibility** χ around the critical point, and so on. The exponent d_f represents the **fractal dimension**[2] of the percolation cluster, since at the critical threshold the occupied sites (or bonds) of the lattice have a fractal geometry. All the critical exponents are well described in the reviews [144, 28] and reported in an understandable manner in my paper **Paper V** [123] . Here it is instructive to comment on several properties of the percolation transition and its critical exponents. The critical exponents of the percolation transition are the same for both site and bond percolation. They do not depend on the particular lattice considered as support for the model but only on the dimensionality of the lattice. This means that it does not make sense to distinguish critical exponents for square, triangular, honey-comb, or other lattices, but it makes sense to talk about critical exponents in one, two, three, etc. dimensions. The critical exponents of percolation in d-dimensions are the same for every type of lattice. Therefore, one can say that they are **universal**. The same observation can be made for the other continuous phase-transitions: for example, the critical exponents of the ferromagnetic-paramagnetic phase transitions at the Curie temperature are the same independently of whether one is considering iron, copper or zinc. Two phase transitions having the same critical exponents are said to belong to the same **universality class**. A number of universality classes have been discovered by statistical physicists during the recent years: starting from the ones typical of equilibrium systems (see for example refs. [85, 19]) to those

[2] A fractal is "a rough or fragmented geometric shape that can be subdivided in parts, each of which is (at least approximately) a reduced-size copy of the whole." The word fractal was coined by B.B.Mandelbrot in 1975 [87, 88] and was derived from the Latin word "*fractus*" meaning broken or fractured. The fractal dimension is the geometric dimension of a fractal and by definition is a non-integer number differently from standard geometric objects which always have an integer dimension. I refer to the books [88, 28] for a complete introduction to fractals. I will describe in chapter 3 two ways to estimate the dimension of a fractal. I will make use of them in **Paper I** [165] and in **Paper V** [123] , respectively.

of out-equilibrium systems (see for example refs. [61, 112]).
Finally, it should be noted that there exist well known algebraic equations between different critical exponents, so-called **hyper-scaling relations**. I will verify some of them in **Paper V** [123] . As more complete introductions to phase transitions and percolation transition, I recommend references [85, 19] and [144, 28], respectively, to the reader who is interested in getting a deeper knowledge of critical phenomena and a specific definition of critical exponents.

2.2.3 Finite-Size Scaling Analysis

Eq.(2.1) is valid only in the infinite-volume limit. This limit is taken in order to simplify the theoretical analysis that stands behind percolation theory and critical phenomena in general. In numerical calculations and in experiments one can handle only finite systems. Nevertheless, it is possible to determine the critical exponents of the transition just by looking at finite-size systems, since large enough but finite systems already exhibit characteristic properties of the phase transitions in the thermodynamic limit. I describe here the main idea of the so-called "finite-size scaling analysis" applied to percolation [130, 144] and I explain briefly the way in which this analysis constitutes a powerful tool for calculating the critical exponents of the percolation transition.
Consider a lattice with linear size L. Since one is considering finite-size lattices, L will be a large but finite number (L is large as compared to the typical correlation length of the system). In a finite lattice, the infinite cluster cannot exist and one should replace it with the largest cluster [see Figure 5.3(c)]. The ratio of sites belonging to the largest cluster and the total number of sites is denoted again as P_∞ and obeys the finite-size scaling relation

$$P_\infty = L^{-\beta/\nu}\ \tilde{P}\left[(p - p_c)\, L^{1/\nu}\right] \quad , \tag{2.2}$$

where ν is the critical exponent standing for the power-like behavior of the correlation length ξ [see Figure 5.3(c)], while $\tilde{P}(\cdot)$ is a suitable universal function. The mathematical theory which leads to Eq.(2.2) is known as **renormalization group** analysis[3]. It is beyond the scope of this thesis to enter a discussion of the renormalization group analysis. It should be only noted that this kind of analysis is extremely important for the theoretical treatment of critical phenomena and also for their numerical study, as the finite-size scaling analysis demonstrates.

[3]The renormalization group idea has its origin in the seminal work of L.P.Kadanoff in 1966 [70], who realized the possibility of reducing the degrees of freedom of a spin system by renormalizing the system itself in larger blocks of spins. The main and final formulation was made by K.G.Wilson [161], who was honored by the Nobel Prize in Physics in 1982.

In Eq.(2.2), the universal function $\tilde{P}(\cdot)$ is non-singular at zero [$\tilde{P}(0) \neq 0$] , and this fact is practically important, since at $p = p_c$, Eq.(2.2) reduces to

$$P_\infty \sim L^{-\beta/\nu} \quad . \tag{2.3}$$

Eq.(2.3) shows a simple way to measure the ratio $-\beta/\nu$. One just needs to measure P_∞ at p_c for different sizes L of the lattice and then measure the slope of this function on a double-logarithmic plot. Mathematical relations similar to Eq.(2.2) are valid for other observables of interest, which means almost all the critical exponents can be easily calculated numerically by varying the linear size of the lattice L at the critical point p_c.
The application of the finite-size scaling analysis to percolation and in particular the application of mathematical relations like Eq.(2.3) requires the knowledge of the critical threshold p_c. In the case of standard percolation, the actual value of p_c is known from analytical calculations for the cases of one- and two-dimensional lattices. For higher dimensions or in general for critical phenomena for which the critical point of the transition is not known, finite-size scaling analysis gives some specific tools for the determination of p_c. Here I illustrate just one of them and for more complete discussions I recommend the books [144, 28, 19]. Based on the finite-size scaling relation of Eq.(2.2), one can make a plot of the order parameter P_∞, rescaled by $L^{-\beta/\nu}$, as a function of the parameter p for which one would like to determine the critical value. By considering different sizes of the lattice L, one can tune the ratio β/ν appropriately until all curves for different L have a common intersection at a particular value of p, that actually represents the critical threshold p_c. In this way, one has determined simultaneously the ratio between the critical exponents β and ν and the critical threshold p_c. Its value can then be used for further numerical estimates of other critical exponents. This numerical tool is used in **Paper V** [123] for the determination of the critical threshold for a non-standard percolation model.

2.2.4 Directed Percolation

Among the critical phenomena, directed percolation is associated with an extraordinarily large variety of real phenomena. Directed percolation is defined in exactly the same way as percolation, discussed above, with only one additional condition: each bond of the lattice carries an arrow, which is determined by a preferred direction. Therefore, percolation is possible only when it is possible to go from one side of the lattice to the opposite side by walking on occupied sites and following the arrows on the bonds according to the prescribed direction. In other words, directed percolation represents the "anisotropic" version of percolation, which corresponds to an "isotropic" problem [144]. A good example can be taken from geology. Imagine a porous rock in which neighboring pores are connected by channels of different

permeability. Considering this system as a directed percolation model, one can pose the interesting question of how deep the water, which follows a direction prescribed by gravity, penetrates into the rock.

An alternative definition of directed percolation is given in terms of evolutionary processes. In this case the preferred direction is represented by the time of the evolution. Therefore if the process is defined on a lattice of d dimensions, the related evolutionary model is said to be $(d+1)$-dimensional. A standard example of these models is represented by the **contact process** [83]. This process is defined as a set of evolutionary rules on a population of particles living on a lattice of d dimensions. Every site of the lattice can be either occupied or empty. The evolutionary rules are simple. At each stage of the evolution, an occupied site is selected and a random number $0 \leq r \leq 1$ is taken out from a uniform distribution. If $r < p$, then the selected site becomes empty. If $r \geq p$, then one neighbor of the selected site is chosen at random: if this site is already occupied, nothing happens; if this site is empty, it becomes occupied. p is the spreading parameter of the model and effectively controls the behavior of the model itself. For $p < p_c$, the final configuration is composed of empty sites [see Figure 2.3(a)]; for $p > p_c$, the final configuration is composed of both empty and occupied sites [see Figure 2.3(c)]. By definition of the model, the configuration in which all the sites are empty is an absorbing configuration, since empty sites cannot create occupied sites. However, every configuration with a non-zero number of occupied sites is active because the evolutionary algorithm can go on. Therefore at p_c one has a transition between an inactive and an active phase [see Figure 2.3(b)]. This particular phase transition can be studied by using as **order parameter** the density of occupied sites ρ. Since the theory is again formulated in the thermodynamic limit (here this means in large-volume limit) and in the limit of infinite time, numerical studies can be performed by looking to finite-size systems and again performing a **finite-size scaling analysis**. In particular it can be shown that the contact process gives rise to a continuous phase transition and the critical exponents of the transition are the same as in the case of standard directed percolation in $(d+1)$-dimensions. Therefore, they both belong to same **universality class**.

In chapter 6 and particularly in **Paper VI** [122] , I will consider again a $(1+1)$-dimensional evolutionary process. In this case, I will see that the parity conservation of occupied sites leads to a classification of my model as belonging to the **parity conserving universality class** [61, 112].

(a) Non percolation (b) Percolation threshold (c) Percolation

Figure 2.3: Example of directed percolation in $(1+1)$-dimensions. The preferred direction is represented by the time, oriented from top to bottom in all figures. Occupied sites are denoted as full dots, while empty sites are not plotted. The initial configuration is given by one occupied site placed in the middle of the lattice. **(a)** If the spreading rate is smaller than the critical one, the system ends up in the inactive phases. **(b)** At the critical point one is at the boundary between inactive and active phase. The occupied sites are a non-zero number, but still small as compared to the total number of sites of the lattice. **(c)** Above the critical point, the occupied sites spread all over the lattice, meaning that the system is in the active phase. (from ref. [61])

2.3 Phase Transitions on Networks

As already shown in this chapter, examples of phase transitions can be found in a large variety of real systems. Moreover, the theory behind phase transitions is not restricted to a particular case of phase change, but all phase transitions can be analyzed with the help of the same mathematical and theoretical tools. In this sense the theory of phase transitions is "universal". The examples of phase transitions proposed so far are all taken from standard physics. Both percolation and directed percolation, described in section 2.2, are "usually" defined on regular lattices. The same underlying topology is used as support for the description of the ferromagnetic-paramagnetic phase transitions, well explained theoretically by the Ising model [see section 6.2]. In this thesis, moreover, I am primarily interested in the study of phase transitions occurring on complex networks, instead of phase transitions occurring on lattices. I have already anticipated some examples of phase transitions studied during this thesis and I will illustrate them in the following chapters. In this section, I would like to give a few more examples of phase transitions on networks which have been studied by scientists during the recent years.

Roughly, one can divide the phase transitions studied on networks into two main categories: purely topological phase transitions and phase transitions of dynamical models living on net-

works. Examples of phase transitions belonging to the first class can be illustrated by the models of complex networks proposed in chapter 1. The transition occurring at the critical probability of connection p_c in the Erdös-Renyi model [see section 1.5.1], between a phase in which the network is composed of several components ($p < p_c$) and a phase in which the network is composed of just a single component ($p > p_c$), can be easily described as percolation transition [3]. Similarly, the two transitions occurring in the Watts-Strogatz model, one for the average distance and one for the clustering coefficient [see section 1.5.2], as a function of the rewiring probability p are again purely topological phase transitions on networks. Moreover, in section 3.2 and in **Paper I** [165] I will consider fractal and self-similar networks. This can be viewed as a study of networks at "criticality", since many critical phenomena at their critical point can be described as fractal or self-similar geometrical objects. Moreover in **Paper IV** [124] , I will explore several phase transitions occurring in randomly diluted networks in the context of social balance [see chapter 5]. Examples of phase transitions, belonging to the second category of phase changes for dynamical models defined on networks, are given by every model that already undergoes a phase transition on regular lattices and that preserves the same feature on networks. Meanwhile, physicists studied the phase structure of the Ising model [59, 64, 60, 81, 138], the Potts model [38, 129], the contact process [32, 143], and synchronization transitions [62, 168, 78, 98, 99, 113, 63, 1, 114, 11], etc., also on generic network topologies.

Chapter 3

Self-similarity and Scale-free Degree Disassortativity on Complex Networks

Introduction

In chapter 1, I introduced complex networks. I illustrated several examples of real systems that can be described as networks [section 1.1] and I summarized the main topological properties emerging from the statistical analysis of their structures [section 1.4]. I stated that the majority of real networks are scale-free and small-world like [sections 1.4.1 and 1.4.2, respectively], and that they have a large clustering coefficient [section 1.4.3] and internal organization in communities [section 1.4.4]. In this chapter, I will describe in detail a novel topological feature of real complex networks, recently discovered by C.Song *et al.* [141]: the fact that the majority of real networks are self-similar so that they can be described as fractals. This property is of extreme importance for several reasons. The "fractality" of real networks is counter-intuitive since it is apparently in contrast to the small-world feature. Moreover, the possibility of describing networks as fractals allows a direct connection between networks and critical phenomena [introduced in chapter 2].
In order to make the concepts of self-similarity and fractality on networks understandable, I will first summarize some standard definitions about fractals [section 3.1]. Then I will see how it is possible to extend these standard definitions to complex networks [section 3.2]. Finally, I will report in **Paper I** [165] my own new results about the relation between self-similarity and degree disassortativity on scale-free networks [refer to the definitions given in sections 1.3.1 and 1.4.1, respectively].

3.1 Fractals

A fractal is a geometric object with a fractional dimension. This differs from standard geometric objects like lines and surfaces which have an integer dimension. The study of fractals was initiated in 1975 by the mathematician B.B.Mandelbrot [87, 88] who coined the word "fractal" in order to describe rough, broken and irregular characteristics of these objects ("*fractus*" is a Latin word meaning "broken"). Quite often, fractals are also **self-similar** objects in the sense that they can be subdivided into parts, each of which is a reduced-size copy of the whole.
Fractals are generally defined on regular topologies like Euclidean d-dimensional spaces or d-dimensional lattices. There are two main categories of fractals: *non-random* and *random* fractals. All the deterministic fractals generated by iterated functions or *growth rules* belong to the first category. An example is given by the Sierpiński gasket fractal. The construction of this fractal is extremely simple. One starts from a single triangle considered as occupied. This triangle is then divided into four smaller triangles generated by tracing three segments which connect the middle of the three sides of the starting triangle. Three of the new smaller triangles are considered as occupied and the one in the middle as empty. This procedure can then be iterated for each of the new three occupied triangles and so on for every occupied triangle at every stage. The result obtained after five stages of iteration is reported in Figure 3.1. By simple calculations one can determine that the fractal dimension of the Sierpiński gasket is $d_f = \frac{\log 3}{\log 2}$. On the other hand, random fractals are non-regular geometric objects, since

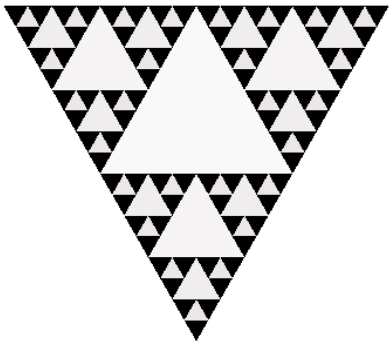


Figure 3.1: Sierpiński gasket fractal after five stages of iteration. [Figure generated by Xaos (http://xaos.sourceforge.net)]

they are not generated in a deterministic way. All the fractals found in nature are random. Examples can be found in a large variety of contexts: the corrugated coast of Britain [86], fractures in rocks [16, 15], the spatial distribution of epicenters of earthquakes [71], the spatial displacement of rivers and lakes [134], and so on. Moreover it should be emphasized that the majority of systems characterized by a second-order phase transition give rise to fractal

structures at their critical point. For example, the percolation picture at the critical threshold [see Figure 2.2(b)] is a fractal. In this case, the fractal dimension d_f is related to the other critical exponents β and ν by the hyper-scaling relation $d_f = d - \beta/\nu$ [refer to section 2.2.2 for the definitions of these quantities], where d stands for the dimension of the space in which the percolation model is defined.
There are two possible procedures for determining the fractal dimension d_f: the "**box counting**" and the "**cluster growing**" methods [28]. For simplicity, I describe them in the case of a percolation model, at its critical threshold p_c, defined on a square lattice [see section 2.2]. [Although these methods are valid independently of the specific model analyzed and the dimensionality of the Euclidean space or of the lattice on which the fractal is defined.] In the former method, one counts the minimal value of the total number of square boxes of side ℓ_B, namely $N_B(\ell_B)$, needed for tiling the entire fractal, that is for covering all occupied sites. Since $N_B(\ell_B) \sim \ell_B^{-d_f}$, one can simply calculate d_f in a double-logarithmic plot of $N_B(\ell_B)$ versus ℓ_B. In the latter method, sometimes also called the "window box scaling" method, one considers a square box of size ℓ_B centered on a randomly chosen site of the lattice and counts how many occupied sites are contained in the box. Averaging over many boxes, one can determine the average "mass" of a box of side ℓ_B, namely $\langle M(\ell_B)\rangle$. The scaling of the average mass with the linear size of the box is $\langle M(\ell_B)\rangle \sim \ell_B^{d_f}$, which allows the determination of the fractal dimension. Both these methods are based on the self-similarity of fractals, since a part of the system (i.e., the box) is representative for the whole system as long as one considers regular topologies like Euclidean spaces or lattices. Moreover, on regular topologies, both methods are equivalent in the sense that they predict a power-law scaling with the same exponent d_f. I will make use of both these methods during this thesis. In particular, I will use the box counting method on fractal networks in **Paper I** [165] where it is the method of choice on a scale-free network [defined in section 1.4.1], while I will determine the fractal dimension in a particular percolation transition using the cluster growing method in **Paper V** [123] .

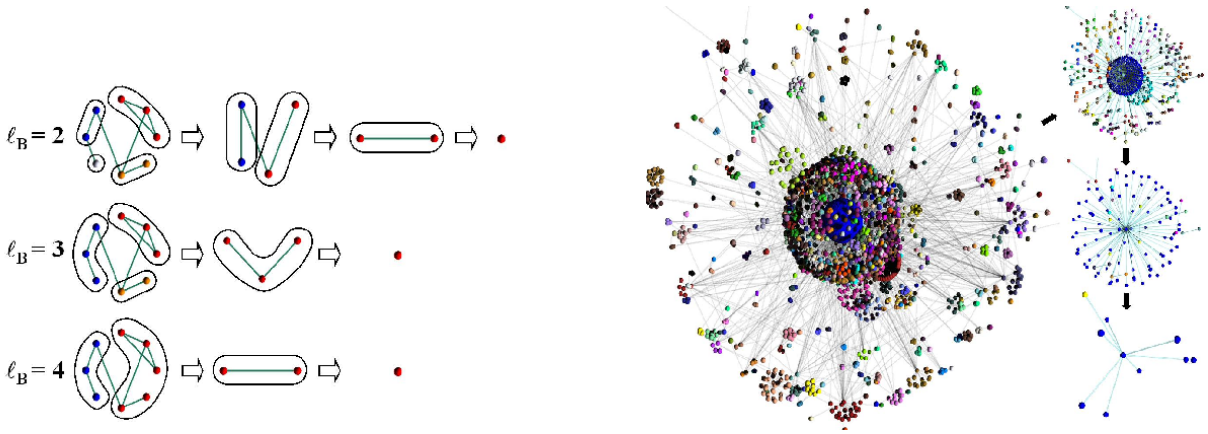
3.2 Self-similarity of Complex Networks

Up to now, I have considered fractals defined on regular topologies. Recently C.Song, S.Havlin and H.A.Makse [141] extended the definitions of fractality and self-similarity to topologies with inhomogeneous properties like networks. Both of these notions are apparently incompatible with networks since the small-world feature of real networks [i.e., Eq.(1.9] implies that the relation between the average distance and the size of the network is $\langle \ell \rangle \sim \log N$. Therefore the expected number of boxes needed to tile a real network should obey $N_B(\ell_B) \sim \exp(-\ell_B/\ell_0)$,

with ℓ_0 the characteristic length of the network.
C.Song *et al.* [141] solved this apparent incompatibility. They defined a box of size ℓ_B in a network as a subset of nodes with mutual distance less than ℓ_B. This means that the distance between all the pairs of nodes in a box is less than ℓ_B [refer to the definition of distance in a network given in section 1.3.2]. With this definition of box, the former methods, illustrated above for "standard" fractals and able to determine the fractal dimension d_f, are no longer equivalent. Measures performed on real networks revealed the following [141, 165]. The cluster growing method effectively leads to an exponential scaling $\langle M(\ell_B)\rangle \sim \exp(\ell_B/\ell_1)$, consistent with the small-world feature of real networks. Alternatively, the box counting method reveals that

$$N_B(\ell_B) \sim \ell_B^{-d_B} \quad . \tag{3.1}$$

One uses here d_B to indicate the fractal dimension of the network, since this exponent can be found only with the box counting method. In particular, the power-law scaling is not true for all real networks, but is strongly related to the degree disassortativity of the networks, as the reader will see in **Paper I** [165] .
There are different ways of implementing the tiling procedure on a network. They are well



(a) Tiling (b) Renormalization

Figure 3.2: **(a)** Examples of tiling procedure for the case of $\ell_B = 2, 3$ and 4. **(b)** Different stages of renormalization for a real network. (Both figures from ref. [140])

described in [139], while I describe in the appendix B.1 the one used in my **Paper I** [165] . The tiling procedure can be also used for renormalizing the network. Each box of the original network becomes a node in the renormalized network, and two boxes in the renormalized network are connected if two of their nodes are connected in the original network [see Figure 3.2].

This renormalization scheme can be iterated until it converges to a network composed by a single node. Interestingly, it is possible to relate the fractal dimension d_B to the exponent of the degree distribution γ of scale-free networks [by the definition given in section 1.4.1, scale-free networks have degree distribution which is a power-law: $P(k) \sim k^{-\gamma}$]. This is a consequence of the self-similarity of fractal scale-free networks, since the power-law governing their degree distributions is invariant under the tiling procedure. At each stage of the renormalization, the network has a degree distribution which is a power-law with the same exponent γ. This allows to relate γ to the fractal dimension d_B, by the relation

$$\gamma = 1 + d_B/d_k \quad . \tag{3.2}$$

d_k is another critical exponent and is defined by the relation $k' \sim \ell_B^{d_k} k$, where k' is the degree of the box (i.e. the degree of the new node in the renormalized network) and k stands for the maximal degree of the nodes of the original network contained in the box. Therefore, d_k is the exponent characteristic of the scaling of the degrees under the tiling procedure.

Paper I

Self-similar scale-free networks and disassortativity

Soon-Hyung Yook, Filippo Radicchi and Hildegard Meyer-Ortmanns

Physical Review E **72**, 045105(R) (2005)

The following paper is devoted to the study of self-similarity and fractality of real networks. The results found in this paper extend the validity of the conjecture made by C.Song *et al.* in [140]. Moreover, we found one of the most important ingredients which allows us to distinguish whether a network is fractal and self-similar or not. This criterion is related to the joint degree distribution [see section 1.3.1]. If a scale-free network has degree-disassortative mixing [definition given in section 1.4.2], then it shows both the features to be a fractal [it satisfies Eq.(3.1)] and to be self-similar [it satisfies Eq.(3.2)]. On the other hand, if the scale-free network is assortative, then it cannot satisfy Eq.(3.1), nor does it make sense to consider Eq.(3.2). A standard way to distinguish assortative and disassortative networks is to look at Pearson's coefficient defined in Eq.(1.8), which is negative for disassortative networks and positive for assortative ones. The intuitive idea behind the formulation of this criterion is the following. If a network has degree-assortative mixing this means that the "hubs" (i.e., the nodes with the largest degrees) of the network are directly connected. Therefore the number of boxes needed to cover the network decreases faster than a power-law as a function of the linear size of the boxes. This means that Eq.(3.1) is not satisfied and no fractal dimension d_B can be assigned to the network. All these ideas are well explained in the paper. Moreover, it should be noticed that the results of this paper inspired C.Song *et al.* to formulate a minimal model in which "hubs repulsion" (i.e., degree-disassortativity) is used as a basic ingredient for the generation of a fractal self-similar network [142].

Personal Contribution

For the realization of this paper, I performed all numerical simulations. In particular, since the paper is essentially a numerical study, the same numerical simulations were separately performed by me and Dr.S.Yook in order to check the consistency of our results. I discuss the algorithmic part of the simulations in the appendix B.1.

PHYSICAL REVIEW E **72**, 045105(R) (2005)

Self-similar scale-free networks and disassortativity

Soon-Hyung Yook,* Filippo Radicchi, and Hildegard Meyer-Ortmanns†
SES, International University Bremen, P.O.Box 750561, D-28725 Bremen, Germany
(Received 8 July 2005; published 26 October 2005)

Self-similar networks with scale-free degree distribution have recently attracted much attention, since these apparently incompatible properties were reconciled in [C. Song, S. Havlin, and H. A. Makse, Nature **433**, 392 (2005)] by an appropriate box-counting method that enters the measurement of the fractal dimension. We study two genetic regulatory networks (*Saccharomyces cerevisiae* [N. M. Luscombe, M. M. Babu, H. Yu, M. Snyder, S. Teichmann, and M. Gerstein, Nature **431**, 308 (2004)] and *Escherichia coli* [http://www.ccg.unam.mx/Computational_Genomics/regulondb/DataSets/RegulonNetDataSets.html and http://www.gbf.de/SystemsBiology]) and show their self-similar and scale-free features, in extension to the datasets studied by [C. Song, S. Havlin, and H. A. Makse, Nature **433**, 392 (2005)]. Moreover, by a number of numerical results we support the conjecture that self-similar scale-free networks are not assortative. From our simulations so far these networks seem to be disassortative instead. We also find that the qualitative feature of disassortativity is scale-invariant under renormalization, but it appears as an intrinsic feature of the renormalization prescription, as even assortative networks become disassortative after a sufficient number of renormalization steps.

DOI: 10.1103/PhysRevE.72.045105 PACS number(s): 89.75.Hc, 05.45.Df, 87.16.Yc

I. INTRODUCTION

Until very recently, the celebrated properties of a scale-free degree distribution seemed to be incompatible with self-similar features of networks, in which the number of boxes N_B of linear size ℓ_B scales with ℓ_B according to a power-law $N_B \sim \ell_B^{-d_B}$, with an exponent that is given by the fractal dimension d_B. Using the box-counting method, Song *et al.* showed that many scale-free (SF) networks observed in nature can have a fractal structure as well [1]. This result is striking, because the tiling and renormalization according to the linear size of the boxes, in which all pairs of nodes inside a box have mutual distance less than ℓ_B, appear to be physically relevant rather than being a formal procedure. Therefore, the essential quantity in the tiling is the linear size of the box, ℓ_B, defined by $\ell_B - 1$ being the maximal distance between the nodes of the box. In the renormalization procedure the boxes are contracted to the nodes of the renormalized network whose edges are the interconnecting edges between the boxes on the original network.

In this paper we study the genetic regulatory network of two well-known organisms, *Saccharomyces cerevisiae* [2] and *Escherichia coli* [3]. We first determine the degree-distribution $P(k)$, that is the probability for finding a node with degree k, i.e., with k edges, to read off the exponent γ according to $P(k) \sim k^{-\gamma}$ [4] in order to check that the networks are scale-free. The edges correspond to regulatory interactions between transcription factors and target genes (either between transcription factors and nontranscription factor targets, or between two transcription factors). Next we measure the ratio of the total number of boxes N_B of linear size ℓ_B over the total number of nodes N in the network, that is N_B/N, for various box-sizes ℓ_B, to obtain the fractal dimension d_B from $N_B/N \sim \ell_B^{-d_B}$. After renormalizing the networks according to the procedure proposed in [1], we measure the scaling behavior of the degree k' according to $k' = s(\ell_B)k$, where k' stands for the degree of a node in the renormalized network, k is the largest degree inside the box that was contracted to one node with degree k' in the renormalization process, and $s(\ell_B)$ is assumed to scale like $s(\ell_B) \sim \ell_B^{-d_k}$ with a new exponent d_k. The invariance of γ under renormalization and the transformation behavior of the degree itself imply the relation [1]

$$\gamma = 1 + d_B/d_k \tag{1}$$

between the exponents. Therefore, we check this relation by measuring d_B, d_k and comparing the values of γ from Eq. (1) with the measured γ from the degree distribution.

One of the important features of networks is their "degree" of assortativity. The notion of assortative mixing was known from epidemiology and ecology when it was introduced as a characteristic feature of generic networks by Newman [5]. Assortativity refers to correlations between properties of adjacent nodes. One particular property is the (in- or out-) degree of a node as the number of its incoming or outgoing links, respectively. Degree-degree correlations can be recorded as histograms; in order to facilitate the comparison between networks of different size, they can be also characterized by the Pearson coefficient. The Pearson coefficient is obtained from the connected degree-degree correlation function $\langle jk \rangle - \langle j \rangle \langle k \rangle$ after normalizing by its maximal value, which is achieved on a perfectly assortative network. Here, $\langle jk \rangle$ stands for the average of having vertex degrees j and k at the end of an arbitrary edge. The Pearson coefficient r takes values between $-1 \leq r \leq 1$, it is positive for assorta-

*Present address: Dept. of Physics and Research Institute of Basic Sciences, Kyung Hee University, Seoul, 130–701, Korea.
†Electronic address: h.ortmanns@iu-bremen.de

tive networks ($r=1$ for complete assortativity) and negative for disassortative ones. As it was shown in [5], the Pearson coefficient can be rewritten according to

$$r = \frac{M^{-1}\sum_i j_i k_i - \left[M^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2}{M^{-1}\sum_i \frac{1}{2}(j_i^2 + k_i^2) - \left[M^{-1}\sum_i \frac{1}{2}(j_i + k_i)\right]^2}, \qquad (2)$$

where M denotes the total number of edges i, $i=1,\ldots,M$, and j_i, k_i denote the degrees of the two vertices at the ends of edge i. Equation (2) has actually the form in which we measured the Pearson coefficient for a number of self-similar networks. The results are presented below. The reason why this feature is of interest in the present context is its relation to the power-law or exponential behavior of $N_B(\ell_B)$. In particular, we are interested in the question whether disassortativity is scale-invariant on a qualitative level under renormalization according to the prescription proposed in [1], and why these properties go along. Disassortative features in protein interaction networks were found and explained by Maslov and Sneppen [6] on the level of interacting proteins and genetic regulatory interactions. According to their results links between highly connected nodes are systematically suppressed, while those between highly connected and low-connected pairs of proteins are favored. In this way there is little cross-talk between different functional modules of the cell and protection against intentional attacks, since the failure of one module is less likely to spread to another one. Also in immunological networks one speaks of lock- and key-interactions between molecular receptors and antigenic determinants [7]. In general, complementarity is essential for pattern recognition interactions, underlying biological and biochemical processes as well as for symbiotic species in ecological networks. Of course, it is not at all obvious or necessary that complementarity in "internal" (functional) properties should be manifest in topological features like the degree-degree correlations. Therefore, we study the relation between self-similarity and degree-assortativity in this paper.

II. MEASUREMENTS AND RESULTS

For the genetic regulatory networks *Saccharomyces cerevisiae* [2] and *Escherichia coli* [3] we observe a power-law behavior of N_B/N for $\ell_B > 3$ with $d_B = 5.1 \pm 0.3$ for *S. cerevisiae* and $d_B = 3.4 \pm 0.2$, respectively. The obtained degree-distributions are scale-free and satisfy a power-law with exponent $\gamma \simeq 2.6 \pm 0.2$ for *S. cerevisiae* and $\gamma \simeq 2.8 \pm 0.3$ for *E.coli*, cf. Fig. 1. The scaling relation (1) between the exponents d_B and d_k is also satisfied within the error bars for both networks. From (1) we obtain $\gamma = 2.8 \pm 0.3$ for *Saccharomyces cerevisiae* and $\gamma = 2.7 \pm 0.5$ for *Escherichia coli*.

In Table I we summarize the results also for some additional networks, for which we list their properties of self-similarity and disassortativity. If we confirm the property of self-similarity it means not only the scaling behavior of N_B/N according to a power-law, but also the numerical verification of the scaling relation of Eq. (1) and the invariance of γ under renormalization [1]. This is more conclusive, because it is sometimes difficult to disentangle exponential from power-law behavior of N_B/N for networks with a small diameter (for example, see Fig. 2 for the regulatory network of *S.cerevisiae* with an inset that shows the same data points on a log-linear scale instead of the log-log scale), whereas the scaling relation only holds for a power-law of the decay, it is easier to prove or disprove. A confirmation of (dis)assortativity refers to histograms with (negative) positive slope of next-neighbor degree-degree correlations and/or a (negative) positive Pearson coefficient, respectively. In most cases we measured the degree-degree correlations also between nodes at distance $d=2,3$, as indicated in the figures, where the distance is measured in units of edges.

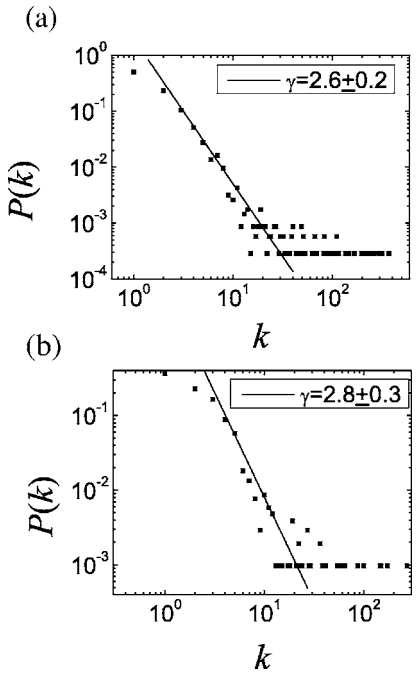


FIG. 1. Degree distribution $P(k)$ for (a) *S. cerevisiae* and (b) *E. coli* to read off γ.

The first four networks of Table I refer to the genetic regulatory networks of *S.cerevisia* and *E.coli*, the scientific collaboration network, and the internet on the autonomous systems level. For these networks the properties of columns 2 and 3 were examined by us, while for the last three networks (the biochemical pathway network of *E.coli*, the actor network and the world-wide-web), the self-similarity was es-

TABLE I. Networks from datasets characterized by properties of self-similarity and disassortativity.

Network	Self-similarity	Disassortativity
genetic reg.*S.cerevisia* [2]	yes Fig. 2	yes Fig. 5(a)
genetic reg.*E.coli* [3]	yes Fig. 4	yes Fig. 5(b)
scient.collab. [8]	no Fig. 7(a)	no Fig. 7(b)
internet aut.sys. [8]	yes [10]	yes [10]
biochem.pathway *E.coli* [8]	yes [1]	yes [10]
actor [8]	yes [1]	(yes, $k>1000$) Fig. 3
www [8]	yes [1]	yes [10]

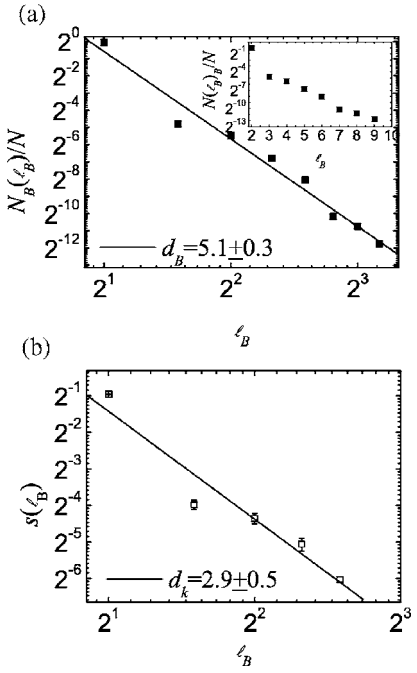


FIG. 2. *S.cerevisiae*: (a) Normalized number of boxes N_B as a function of linear box size ℓ_B to read off d_B. (b) Rescaling factor $s(\ell_B)$ as function of the box size to read off d_k.

tablished before [1], and we studied their property of disassortativity in addition. In particular the actor network deserves some further comments. The actor-network is self-similar [1], but its positive Pearson coefficient suggests that it is assortative, in contrast to all other self-similar networks we have studied so far. A closer look at its next neighbor-degree-degree correlation (Fig. 3) shows an assortative behavior for degrees up to the order of 1000, but slowly decays for larger degrees and becomes disassortative. The degree-degree correlation between nodes at distance larger than one is decreasing with degree k for all k.

Moreover, some comments are in order to the yeast-genetic regulatory network with 3456 nodes and 14 117 edges, (cf. Fig. 2). Since it has a diameter of 9, the largest ℓ_B value for the tiling is 10. Therefore, we have only 8 data points available for the fit. Each point corresponds to an

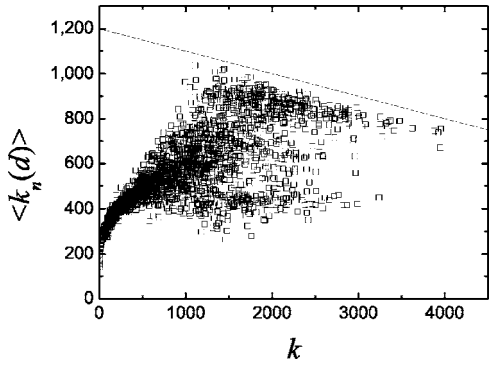


FIG. 3. Degree-degree correlation $\langle k_n(d)\rangle$ for $d=1$ against k for the actor network, showing initially assortative behavior for $k \leqslant 1000$.

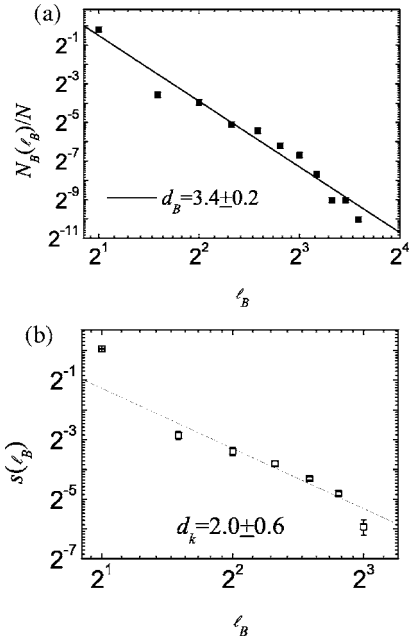


FIG. 4. Same as Fig. 2, but for *E.coli* to read off (a) d_B (b) d_k.

average over 100 tiling configurations. Different tiling configurations result from different starting seeds as well as the random selection of neighbors during the tiling process. The data point at $\ell_B=3$ in Fig. 2 lies clearly outside the fluctuations about the average over different tiling configurations, thus outside the error bars, which are at least two orders of magnitude smaller than the respective value of N_B/N, so that they are not visible on the scale of the figure. In Fig. 4, we find a similar behavior for *E.coli*. The deviation from the power-law behavior at $\ell_B=3$ goes along with an assortative degree-degree-correlation between nodes at distance $d=2$ as

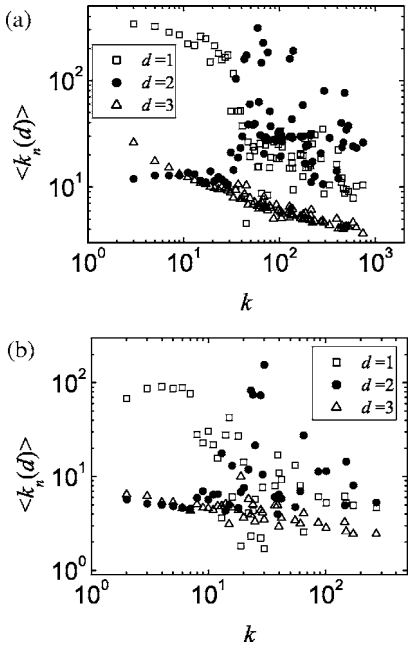


FIG. 5. Degree-degree correlation $\langle k_n(d)\rangle$ vs degree k for distances $d=1,2,3$, (a) *S.cerevisiae* and (b) *E.coli*.

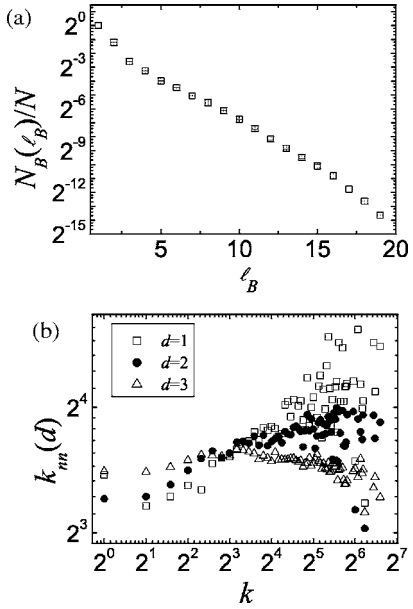


FIG. 6. Scientific collaboration network (a) exponential decay of $N_B/N(\ell_B)$ (b) degree-degree correlation $\langle k_n(d)\rangle$ for $d=1,2,3$ against k.

it is seen from Figs. 5(a) and 5(b), showing the degree-degree correlation of *S.cerevisiae* and *E.coli*, respectively, at distances $d=1,2,3$. The data in Fig. 5 explicitly show the disassortative behavior at $d=1$ and $d=3$ for both *S.cerevisiae* and *E.coli*. However, for $d=2$, we find that there is a certain value of $k, k=k^*$, at which $\langle k_n\rangle$ abruptly increases and slowly decreases for $k>k^*$. Here $k^*\sim 30$ for *S.cerevisiae* and $k^*\sim 10$ for *E.coli*. These mixed properties of assortativity and disassortativity seem to go along with the deviation from the power-law behavior of N_B/N. On a qualitative level, this is plausible if we focus on a hub that should be present in a scale-free network. In an assortative network (assortative say at distance d, for example, $d=2$), this hub is likely connected to another hub within the distance d. If this hub is chosen as a seed of a box in a tiling with linear box size $\ell_B>d$, we need much less boxes to cover the many nodes in the neighborhood of the hub than in a disassortative network.

In a network which is assortative not only for a certain range of k, but for all k and at distances $d\geqslant 1$, like the scientific collaboration network, N_B/N actually decays faster than power-like for all ℓ_B, as it is seen from the exponential fit of Fig. 6(a).

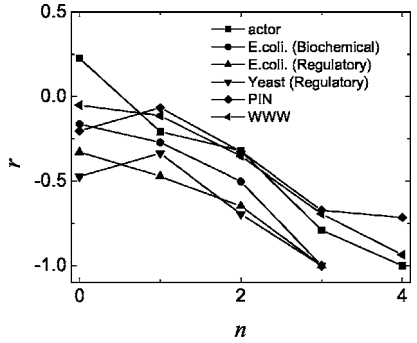


FIG. 7. Pearson coefficient r for several data sets, in particular the PIN-data taken from [9].

The scaling relation between the exponents γ, d_B, and d_k assumes the scale-invariance (under renormalization) of the degree distribution, that is the invariance of the exponent γ. Similarly, it is of interest how the disassortativity transforms under renormalization (as defined in [1]). As we see from Fig. 7, even networks like the scientific collaboration network (Fig. 6), which are originally assortative, transform to more and more disassortative ones under iterated renormalization. (The number of renormalization steps is determined by the size of the networks, in particular by its diameter. The final step is achieved when the reduced network consists of just one node.) Therefore the transformation behavior of disassortativity seems to be an effect of the renormalization procedure rather than an intrinsic self-similar property of the network. Similarly, we measured the transformation behavior of the clustering coefficient under renormalization of self-similar networks. As the data [10] show, it is an invariant property of scale-free networks, while it changes under renormalization for non-self-similar ones like the Barabási-Albert one [11].

To summarize, we find numerical evidence that self-similar scale-free networks are preferably disassortative in their degree-degree correlations. For biological networks this result may reflect the complementarity in interactions that is observed on various levels, as mentioned in the introduction, although it is still far from being obvious.

ACKNOWLEDGMENT

One of us (H.M.-O.) would like to thank S.Havlin for drawing her attention to self-similar networks at the COSIN-final meeting at Salou (Spain) March 2005.

[1] C. Song, S. Havlin, and H. A. Makse, Nature (London) **433**, 392 (2005).
[2] N. M. Luscombe, M. M. Babu, H. Yu, M. Snyder, S. Teichmann, and M. Gerstein, Nature (London) **431**, 308 (2004).
[3] http://www.ccg.unam.mx/Computational_Genomics/regulondb/DataSets/RegulonNetDataSets.html and http://www.gbf.de/SystemsBiology
[4] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[5] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002); M. E. J. Newman and J. Park, Phys. Rev. E **68**, 036122 (2003).
[6] S. Maslov and K. Sneppen, Science **296**, 910 (2002).
[7] M. Copelli, R. M. Zorzenon dos Santos, and J. S. Sa Martins, cond-mat/0110350 (2001).
[8] www.nd.edu/networks
[9] P. Uetz *et al.*, Nature (London) **403**, 623 (2000).
[10] The additional data are available on the web http://imperia.iu-bremen.de/ses/physics/ortmanns/36753/index.shtml
[11] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

Chapter 4

Synchronization on Complex Networks

Introduction

The word "synchronization" takes its root from the Greek words "$\sigma\upsilon\nu$" and "$\chi\rho o\nu o\varsigma$", which literally mean "common" and "time", respectively. Therefore the word synchronization is used with the meaning of "occurring at the same time" or "sharing the same time" in connection with a large variety of phenomena appearing in natural sciences, engineering and social life.
The first study about synchronization was made in the 17^{th} century by the physicst, mathematician and astronomer C.Huygens [66]. Huygens discovered that two pendulum clocks fastened at a common beam become synchronized, in the sense that their oscillations coincide perfectly and the pendula always move in opposite directions. Later, systems exhibiting synchronization were studied and observed in other fields of physics such as acoustics [127] and electrodynamics [154]. Nowdays, the attention of scientists is devoted to the phenomenon of collective synchronization where complex systems of many interacting units tend to become synchronized in the sense that they perform the same function at the same time. Examples of natural systems that exhibit mutual synchronization can be found everywhere: radio communication and electrical devices, simultaneous hand clapping in an audience [101], fireflies emitting sequences of light pulses [26], crickets chirping simultaneously [156], chemical systems exhibiting oscillatory variation of the concentration of reactans [167], coherent contraction of the human heart, and so on.
This chapter gives a short introduction to synchronization phenomena on complex networks. Since networks are complex many-bodies systems, I will focus my attention on the phenomenon of collective synchronization [see section 4.1]. I will start with the description of the Kuramoto model in section 4.1.1. This model is described in detail since it represents the core model of my research on the phenomenon of synchronization. Then I will list other

models proposed in the context of collective synchronization in section 4.1.2. Finally, I will report my **Paper II** [120] and **Paper III** [121] . Both papers will be supplemented by a short introduction with the aim of justifying the importance of the analysis performed in the papers themselves. The papers will explain and motivate the results I obtained in this context.

4.1 Models for Collective Synchronization

During the last forty years the attention of physicists, in particular those studying non-linear dynamics, was directed to the study of mathematical models able to describe the phenomenon of synchronization. One of the aims behind the proposal and the analysis of a model is often to give a qualitative description of a phenomenon, rather than a quantitative one, in order to describe the phenomenon in its minimal formulation and to understand its basic ingredients. In the framework of synchronization several models were proposed. In this section, I will introduce just some of them. In particular I focus my attention on the Kuramoto model [see section 4.1.1], since this model will be the core of both of my papers about synchronization (see **Paper II** [120] and **Paper III** [121]). Moreover, I will mention in section 4.1.2 other models of particular interest that show synchronized patterns as well.

4.1.1 The Kuramoto Model

In 1975, Y.Kuramoto introduced one of most fruitful models to describe the phenomenon of collective synchronization. This model was originally designed for biological systems in which a large number of oscillators, despite the difference in their natural frequencies, lock their motion to a common frequency. This model was used as a description of different real systems: arrays of lasers [79, 155], microwave oscillators [166], superconducting Josephson junctions [34, 153]. In biological applications, the oscillating quantity is often the concentration of substrates described by an amplitude and a phase φ.
In his book [80], Kuramoto studied the dynamics of these phases in detail. He started from a system of differential equations. Using a perturbative method in the coupling strength and averaging over time, he showed that for any system of weakly coupled and nearly identical limit-cycle oscillators[1] the long-term dynamics is given by the equation for the i-th oscillator

[1] *Limit-cycle*, *phase* or *self-sustained* oscillator simply means that this oscillator has a stable periodic motion when left to itself. The notion of limit-cycle was introduced by H.Poincaré and then expressed in rigorous mathematical and physical terms by A.A.Adronov and A.A.Vitt [2].

phase of the following form

$$\dot{\varphi}_i = \omega_i + \sum_{j=1}^{N} \Gamma_{ij}(\varphi_j - \varphi_i) \quad , \tag{4.1}$$

where the i-th oscillator itself follows the differential equation

$$\dot{\vec{x}}_i \;=\; F(\vec{x}_i) \;+\; \epsilon \sum_{j=1\,,j\neq i}^{N} V_{ij}(\vec{x}_i, \vec{x}_j) \quad . \tag{4.2}$$

$F(\vec{x}_i)$ is a force term for the independent dynamics of $\vec{x}_i$, ϵ is the weak coupling strength and $V_{ij}(\vec{x}_i, \vec{x}_j)$ indicates the coupling between the i-th and the j-th oscillators. The degrees of freedom of the system are reduced: the space variable $\vec{x}_i \in \mathbb{R}^n$ of the i-th oscillator is mapped into the phase $\varphi_i \in [0, 2\pi]$ along the unitary circle and the interaction between two different oscillators depends only on the difference of the respective phases via the coupling function $\Gamma_{ij}(\varphi_j - \varphi_i)$. This approximation is valid only in the case of small coupling (ϵ is supposed to be small in order to apply the perturbative method) and in the case of very similar oscillators (their natural periods are supposed to be nearly identical during the averaging procedure).
In spite of the simple form of the system of differential equations (4.1), its general treatment is difficult. Kuramoto solved it only for a very particular case: an all-to-all coupled system, where the interaction function is identical for all pairs of oscillators

$$\Gamma_{ij}(\varphi_j - \varphi_i) = \frac{K}{N} \sin(\varphi_j - \varphi_i) \quad , \tag{4.3}$$

and the natural frequencies are chosen at random from a symmetric distribution $g(\omega)$, with a center in ${\omega_0}^2$.
The case of an all-to-all coupled system is rather special and corresponds to an unrealistic choice but shows an interesting behavior. It allows a simple solution using the mean field approach. On the other hand it is sufficiently complex to have a non-trivial solution.
One introduces first the time-dependent **order parameter** [see Figure 4.1]:

$$R e^{i\psi} = \frac{1}{N} \sum_{j=1}^{N} e^{i\varphi_j} \; . \tag{4.4}$$

The order parameter is a macroscopic quantity that can be interpreted as the collective rhythm produced by the whole population. The radius $R(t)$ measures the phase coherence

[2]System (4.1) is clearly invariant under simultaneous transformations $\varphi_i \to \varphi_i + \omega_0 t$, $\forall i$, so one can assume without loss of generality that $\omega_0 \equiv 0$.

(for complete coherence, when all the phases are the same, it equals one; for complete incoherence, when the phases are equally distributed over the circle, it equals zero), while $\psi(t)$ is the average phase. Rewriting the system of equations, in the particular case of (4.3), using the definition (4.4) and finally equating the imaginary parts leads to

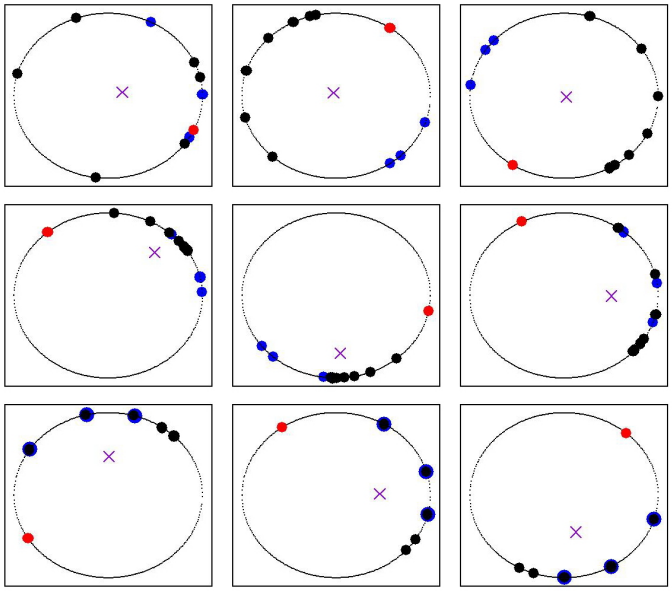


Figure 4.1: Phase representation on the unitary circle that represents the limit-cycle in phase space. The dots correspond to the actual phases of the oscillators φ_i, while the cross to the phase ψ of the order parameter. Moreover the distance of the cross from the center is equal to the radius R of the order parameter.

$$\dot{\varphi}_i = \omega_i + KR \ \sin(\psi - \varphi_i) \ . \tag{4.5}$$

In this form, the i-th oscillator appears as decoupled[3] from all the others, but coupled only to the mean field. From equation (4.5) one can see that the i-th oscillator's phase φ_i is pulled towards the mean phase ψ[4]. The coupling strength is proportional to the coherence of the system: this means that there is effectively a feedback between coupling and coherence: as the population becomes more coherent, R grows towards one, so the effective coupling KR increases, which tends to recruit even more oscillators into the synchronized state.
The original work of Kuramoto studied primarily the steady solution of (4.5), for which $R(t)$ is a constant and $\psi(t)$ rotates with a constant velocity Ω. Shifting (4.5) into the mean field rotating frame (where $\psi \equiv 0$ as well as $\Omega \equiv 0$), one can rewrite (4.5) as

$$\dot{\varphi}_i = \omega_i - KR \sin \varphi_i \ \ . \tag{4.6}$$

[3]Of course this is not strictly true because the mean field is the effect of the whole system.

[4]for $K > 0$, $\sin(\psi - \varphi_i)$ is a "ferromagnetic" interaction because it tends to align the two phases (i.e., tends to reduce the phase lag between the phases φ_i and ψ).

So the oscillators can exhibit two different types of long-term behavior. If $|\omega_i| \leq KR$, the i-th oscillator falls into a stable fixed point ($\dot{\varphi}_i = 0$), defined by

$$\omega_i = KR \sin \varphi_i \ .$$

This oscillator belongs to the locked population, because its phase is locked at frequency Ω in the original frame. The i-th oscillator with $|\omega_i| > KR$ is drifting (or "wild") because it moves, in the original frame, with a non-uniform velocity due to the interaction with the other oscillators, accelerating near some phases and decelerating near others. The frequency of the locked population corresponds to the center of the distribution $g(\omega)$, while the wild oscillator frequencies correspond to the tail; in general, this result can be extended to non-unimodal distribution [see Figure 4.2], where oscillators with natural frequencies close to the peaks of the distribution form synchronized clusters.

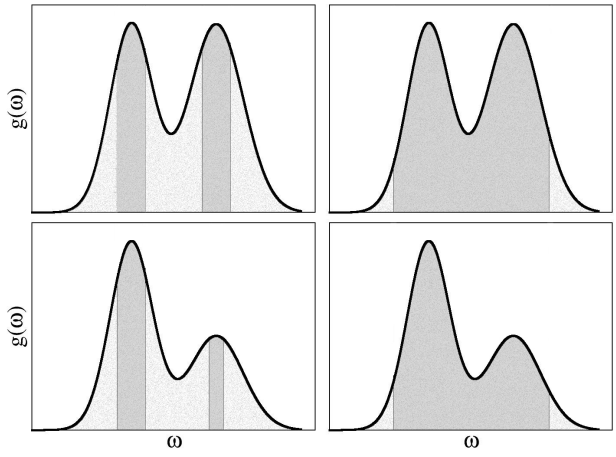


Figure 4.2: Formation of cluster of synchronized oscillators and their mutual synchronization in the cases of symmetric (upper row) and asymmetric (lower row) distribution of natural frequencies. When the coupling strength is weak (left column) one sees the formation of distinct clusters of coherent oscillators (dark grey) around the peaks of the distribution, while the rest of the population is wild (light grey); for larger coupling strength the two different clusters collapse into only one coherent (right column).

It is also possible, in the limit of large N, to estimate the fraction of the oscillators belonging to the two different populations. When the system size goes to infinity, one can replace all discrete quantities by continuous ones and sums by integrals. This allows the problem to be solved explicitly, using a self-consistent approach, and leads to the most important result

$$K_c = \frac{2}{\pi g(\omega_0)} \ , \tag{4.7}$$

if the distribution for the eigenfrequencies $g(\omega)$ is a Lorentzian distribution[5]. K_c represents the critical value for the coupling strength under which one does not have synchronization as

[5]The Lorenztian distribution for the eigenfrequencies is given by $g(\omega) = \frac{\gamma}{\pi} \frac{1}{(\omega-\omega_0)^2+\gamma^2}$. Kuramoto found K_c also in the case of a Gaussian distribution.

long-term behavior at all and so $R = 0$ [see Figure 4.3]. For values of the coupling strength $K > K_c$, the system starts to synchronize. This is reflected in values of $R > 0$. The module of the order parameter increases continuously towards one as the coupling strength increases. This means that the system has a second-order phase transition at $K = K_c$ [see the definition given in section 2.1.3].

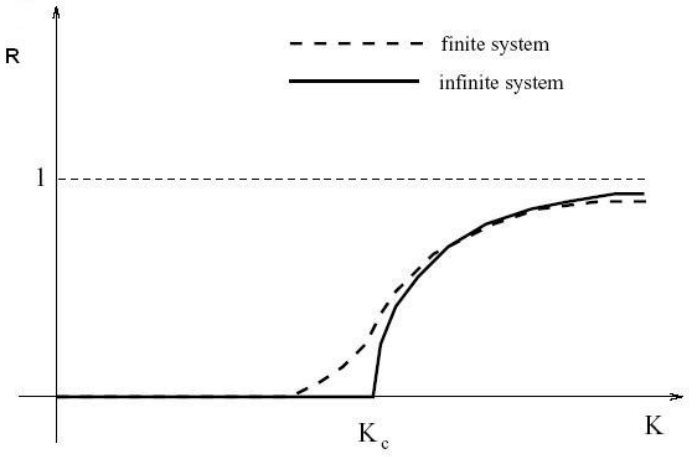


Figure 4.3: Typical second-order transition in the order parameter R of the Kuramoto model [see the definition and classification of phase transitions given in section 2.1.3].

The seminal work of Kuramoto inspired many successive papers (see [145] for a short review and [1] for a more complete one). In some of them, particular mathematical aspects were solved (that were not considered in the original analysis of Kuramoto); in others, different mathematical tools were applied.
On the other side, many successive papers consider simple modifications of the original model. The most common one is to consider different kinds of coupling functions. The sine function represents only one particular and the simplest choice (which can be solved analytically). Periodic coupling functions [36], containing more harmonics, were also proposed. Moreover, there are oscillator models described by two angles (so-called top models [132]), instead of just one.
In other papers [164, 110] a more realistic model with time-delayed interactions is proposed:

$$\Gamma_{ij}(\varphi_j - \varphi_i, t) = \frac{K}{N} \sin\left(\varphi_j(t-\tau) - \varphi_i(t)\right) , \tag{4.8}$$

where τ stands for the time-delay constant. Time delay implements the finite velocity of signal propagation. It must be considered when it has the same order of magnitude as the intrinsic time of the system ($\tau \sim \frac{1}{\Omega}$). The introduction of the delay τ between stimulus and effect gives some surprising results: multiple stable solutions with more than one synchronization frequency and different basins of attraction.

Other papers study the effect of different topologies. The consequence of a change in the backbone is that the mean-field approach is no longer valid. In the simplest case, the interaction network is built on a regular d-dimensional lattice [35, 148, 147, 163], where at each site is a Kuramoto oscillator, coupled via short-range interactions with its nearest neighbors. In this case, the problem can be treated using the renormalization group approach, but with some difficulties to include the heterogeneity of the oscillator's frequencies. The general feature is that the system tends to synchronize locally, and to organize into clusters of coherent oscillators. At the moment, the analysis, which was performed by numerical simulations, still misses some important points: to predict the lower critical dimension below which no kind of synchronization can occur (if the dimension d of a lattice is smaller than the lower critical dimension, a system of heterogeneous oscillators living on this lattice cannot longer synchronize), the dynamical correlation length for the cluster formation and the upper critical dimension above which the synchronization transition is mean-field like. Other authors [62, 78, 98, 99, 113, 63, 1, 114, 11] also studied the Kuramoto model with short-range interactions on complex topologies: random, small-world and scale-free [in models as the ones described in sections 1.5.1, 1.5.2 and 1.5.3, respectively], but only numerically. In general, these systems also tend to organize locally, and partial synchronization is found especially between oscillators belonging to the same local ordered structures (communities and cycles).

4.1.2 Other Models

The Kuramoto model and its variations represent a "paradigm for synchronization" as the large number of papers shows. It turned out to be useful for describing the phenomenon of collective synchronization. But there are other models to tackle this problem as well (and maybe in a more appropriate way). Historically, the Kuramoto model was not the first model proposed to describe synchronization as a collective phenomenon, but the work of Kuramoto was anticipated by A.T.Winfree [162].
Another model which was considered by physicists was a system of **pulse-coupled** oscillators [95]. Each unit is described by a variable that evolves continuously with its own frequency but when it reaches a certain threshold, it fires and launches a pulse to its neighbors. The period of the oscillator that receives the pulse becomes shorter (in the case of excitatory interactions) or longer (in case inhibitory interactions). This model is reasonable for biological oscillators, which often exhibit oscillations based on the buildup and sudden discharge of the membrane voltage or other variables. For this reason it is also called the **integrate-and-fire** model. It is likely that the activity of the neurons in the brain is regulated by this kind of dynamical interaction, rather than a continuous interaction as in the case of Kuramoto oscillators.

The phenomenon of synchronization can also be understood and studied in a system of coupled chaotic oscillators [22]. Despite the fact that a chaotic oscillator exhibits a non periodic motion when it does not interact with other oscillators, it is nevertheless possible to define for a chaotic oscillator the average (in time) phase and frequency. In these terms, a system of coupled chaotic oscillators can be studied using the same formalism as in the case of regularly oscillating systems and can be analyzed in the context of synchronization. Studies of this type of system were performed, for example, in [117, 115, 168].
I note that the present section does not want to give a complete review of models for synchronization. Therefore, I invite the interested reader to a have a look at books like [116, 146], for more examples and more detailed explanations.

In the following **Paper II** [120] and **Paper III** [121] we present a study of the Kuramoto model [see section 4.1.1] posed to the topology of complex networks. Previous papers like [62, 78, 113, 98, 99, 1, 114] and papers published after our work like [11, 10], together with our results revealed that the behavior of the Kuramoto model strongly depends on the underlying topology. Within these papers we specialize to a particular formulation of the Kuramoto model, where the system consists of a homogeneous oscillating medium (i.e., a population of Kuramoto oscillators with the same natural frequency) and a defect (i.e., a Kuramoto oscillator with a natural frequency which differs from that of all the other oscillators). We call this "defect" **pacemaker**.

Paper II

Entrainment of coupled oscillators on regular networks by pacemakers

Filippo Radicchi and Hildegard Meyer-Ortmanns

Physical Review E **73**, 036218 (2006)

As a function of the difference between the natural frequency of the medium and the frequency of the pacemaker, as well as a function of the coupling strength, we observe a first-order phase transition [see section 2.1.3] between a phase in which all the oscillators move together with the pacemaker (i.e. full synchronization) and a phase in which the system loses synchronization. We investigate these transitions on d-dimensional regular lattices. Interestingly, taking the limit of infinite dimension ($d \to \infty$) in the analytical expressions on regular lattices, we reproduce the functional dependence of the critical threshold numerically measured in networks [78]. This function depends only on the fact that networks are infinite-dimensional geometric objects.

Personal Contribution

I performed the analytical calculations and the numerical simulations reported in the following paper. Moreover in the appendix A.1, I report some analytical results about the explicit dependence of the depth of a lattice on the linear size of the lattice. The results obtained in the following paper were based upon these calculations. In the appendix B.2, I briefly describe the methods adopted for the numerical simulations performed in the paper. The same pseudo-code is also valid for the simulations performed in **Paper III** [121] .

Entrainment of coupled oscillators on regular networks by pacemakers

Filippo Radicchi* and Hildegard Meyer-Ortmanns†
School of Engineering and Science, International University Bremen, P. O. Box 750561, D-28725 Bremen, Germany
(Received 12 December 2005; published 24 March 2006)

We study Kuramoto oscillators, driven by one pacemaker, on d-dimensional regular topologies with nearest neighbor interactions. We derive the analytical expressions for the common frequency in the case of phase-locked motion and for the critical frequency of the pacemaker, placed at an arbitrary position in the lattice, so that above the critical frequency no phase-locked motion is possible. We show that the mere change in topology from an open chain to a ring induces synchronization for a certain range of pacemaker frequencies and couplings, while keeping the other parameters fixed. Moreover, we demonstrate numerically that the critical frequency of the pacemaker decreases as a power of the linear size of the lattice with an exponent equal to the dimension of the system. This leads in particular to the conclusion that for infinite-dimensional topologies the critical frequency for having entrainment decreases exponentially with increasing size of the system, or, more generally, with increasing depth of the network, that is, the average distance of the oscillators from the pacemaker.

DOI: 10.1103/PhysRevE.73.036218 PACS number(s): 05.45.Xt, 05.70.Fh

I. INTRODUCTION

Synchronization is a ubiquitous phenomenon, found in a variety of natural systems like fireflies [1], chirping crickets [2], and neural systems, but it is also utilized for artificial systems of information science in order to enable a well-coordinated behavior in time [3]. As an important special case, the coordinated behavior refers to similar or even identical units like oscillators that are individually characterized by their phases and amplitudes. A further reduction in the description was proposed by Kuramoto for an ensemble of oscillators [4], after Winfree had started with a first model of coupled oscillators [5]. Within a perturbative approach Kuramoto showed that for any system of weakly coupled and nearly identical limit-cycle oscillators, the long-term dynamics is described by differential equations just for the phases (not for the amplitudes) with mutual interactions, depending on the phase differences in a bounded form. It is this model, later named after him, that nowadays plays the role of a paradigm for weakly and continuously interacting oscillators. In a large number of succeeding publications the original Kuramoto model was generalized in various directions; for a recent review see [6]. In particular the natural frequencies were specialized in such a way that one oscillator plays the role of a pacemaker with frequency higher than the natural frequencies of all other oscillators [7,8]. Pacemakers play an important role for the formation of patterns in Belousov-Zhabotinsky system [9]. Special families of wave solutions of the phases arise as a consequence of dynamically created pacemakers [4,10]. Moreover pacemakers are important for the functioning of the heart [5] and for the collective behavior of *Dictyostelium discoideum* [11], as well as for large-scale ecosystems [12]. In [13] it is shown that a single periodic pendulum oscillator can entrain or at least drastically influence the dynamics of all chaotic pendula on two-dimensional lattices. Furthermore in [8] the role of pacemakers on complex topologies was analyzed in order to understand the functioning of the neural network at the basis of the circadian rhythm in mammals.

In this paper we consider a system of Kuramoto oscillators, coupled with their nearest neighbors on various regular lattice topologies, and driven by a pacemaker, placed at an arbitrary site of the lattice (Sec. II). In particular, we analytically derive the common frequency of phase-locked motion in case of generic networks (in particular for d-dimensional regular lattices with open or periodic boundaries) in Sec. III. Locked phases will be also called phase entrainment throughout this paper. We also analytically derive the upper bound on the absolute value of the ratio of the pacemaker's frequency to the coupling strength in case of one-dimensional regular lattices (Sec. IV). In Sec. V we consider higher-dimensional lattices and extend the results obtained for $d=1$ to any dimension $d \geqslant 2$ of the lattice. We find that the range of pacemaker frequencies for which one obtains synchronization, the so-called entrainment window, decreases as an inverse power of the linear size N of the lattice with an exponent given by the dimension d of the lattice. This leads to the conclusion that the entrainment window of an infinite-dimensional network decreases exponentially with increasing linear size N if the pacemaker is asymmetrically coupled to the other oscillators. This conclusion is supported by our analysis of coupled oscillators on a Cayley tree, a topology that amounts to an infinite-dimensional regular lattice. These results confirm the results recently obtained by Kori and Mikhailov [8] for random network topologies. For random topologies the entrainment window decays exponentially with increasing so-called depth of the network, that is, the average distance of all other oscillators from the pacemaker. For our regular topologies, the linear size of the networks for hypercubic lattices and the radius of the Cayley tree are proportional to the network depth.

*Electronic address: f.radicchi@iu-bremen.de
†Electronic address: h.ortmanns@iu-bremen.de

II. THE MODEL

The system is defined on a regular network. To each node i, $i=0,\ldots,N$, we assign a limit-cycle oscillator, characterized by its phase φ_i that follows the dynamics

$$\dot{\varphi}_i = \omega + \delta_{i,s}\Delta\omega + (1+\delta_{i,s}\epsilon)\frac{K}{k_i}\sum_j A_{j,i}\sin(\varphi_j - \varphi_i). \quad (1)$$

A is the adjacency matrix of the system ($A_{i,j}=A_{j,i}=1$ if the nodes i and j are connected and $A_{i,j}=0$ otherwise); it reflects the underlying topology of the network. Here only nearest neighbors are coupled. Moreover, $k_i=\Sigma_j A_{j,i}$ is the degree of the ith node; it gives the total number of connections of this node in the network. $\delta_{i,j}$ denotes the Kronecker delta ($\delta_{i,j}=1$ if $i=j$ and $\delta_{i,j}=0$ if $i\neq j$). The oscillator at position s represents the pacemaker. Its natural frequency differs by $\Delta\omega$ with respect to the natural frequency ω of all other oscillators. Without loss of generality we set $\omega=0$, because system (1) is invariant under the transformation $\varphi_i \rightarrow \varphi_i+\omega t\ \forall\ i$. Moreover, the interaction of the pacemaker with the other oscillators can be linearly tuned by the parameter $-1\leqslant\epsilon\leqslant 0$. For $\epsilon=0$ the pacemaker is on the same footing as the other oscillators. For $\epsilon=-1$ its interaction is asymmetric in the sense that the pacemaker influences the other oscillators, but not vice versa (the pacemaker acts like an external force). In natural systems both extreme cases as well as intermediate couplings can be realized. The constant $K>0$ parametrizes the coupling strength. The phases of the ith and jth oscillators interact via the sine function of their difference, as originally proposed by Kuramoto.

III. PHASE-LOCKED MOTION

We consider the conditions for having phase-locked motion, in which the phase differences between any pair of oscillators remain constant over time, after an initial short transient time. First we calculate the frequency Ω in common to all oscillators in the phase-locked state. Imposing the phase-locked condition $\dot{\varphi}_i\equiv\Omega\ \forall\ i=0,\ldots,N$ to system (1) and using the fact that the sine is an odd function (see Appendix A for details), we obtain

$$\Omega = \Delta\omega\frac{k_s}{(1+\epsilon)\sum_{i\neq s} k_i + k_s}. \quad (2)$$

As long as Ω depends on the degree k_s we see that the common frequency Ω increases with the degree of the pacemaker. On the other hand, when the network has a homogeneous degree of connections, $k_i\equiv k\ \forall\ i=0,\ldots,N$, Eq. (2) takes the form

$$\Omega = \frac{\Delta\omega}{(1+\epsilon)N+1}. \quad (3)$$

It should be noticed that in this case the common frequency does not depend on the common degree k of the network.

In terms of the original parametrization of the model (1), the common frequency after synchronization is $\Omega+\omega$. For the derivation of Eqs. (2) and (3) we made use only of the odd parity of the coupling function. The former results are still valid for any other odd coupling function $f(\varphi_j-\varphi_j)$ which is 2π periodic and bounded.

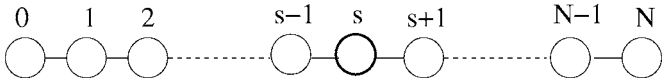


FIG. 1. One-dimensional lattice with $N+1$ sites labeled by their coordinates.

IV. ONE-DIMENSIONAL LATTICE

A. Linear chain

Let us consider first the case of $N+1$ Kuramoto oscillators coupled along a chain (Fig. 1) with free boundary conditions at the positions 0 and N ($k_i=1$ if $i=0$ or $i=N$, $k_i=2$ otherwise), the pacemaker located at position s, $0\leqslant s\leqslant N$. In this case it is convenient to introduce the phase lag between nearest neighbors $\theta_i=\varphi_i-\varphi_{i-1}$. Consider first the pacemaker placed at position $0<s<N$. We start with the nearest oscillators to the right of the pacemaker, placed at position $i=s+1$,

$$\Omega = K/2[\sin(-\theta_{s+1}) + \sin(\theta_{s+2})]$$

$$\Rightarrow \sin(\theta_{s+2}) = 2\Omega/K + \sin(\theta_{s+1}).$$

Moving again to the right, the equation of the $(s+2)$th oscillator reads

$$\begin{aligned}\Omega &= K/2[\sin(-\theta_{s+2}) + \sin(\theta_{s+3})] \Rightarrow \sin(\theta_{s+3}) \\ &= 2\Omega/K + \sin(\theta_{s+2}) \Rightarrow \sin(\theta_{s+3}) \\ &= 4\Omega/K + \sin(\theta_{s+1}).\end{aligned}$$

Iteratively, we can write for each $1\leqslant j\leqslant N-s$

$$\sin(\theta_{s+j}) = 2(j-1)\frac{\Omega}{K} + \sin(\theta_{s+1}). \quad (4)$$

In particular when $j=N-s$ we have

$$\sin(\theta_N) = 2(N-s-1)\frac{\Omega}{K} + \sin(\theta_{s+1}),$$

but also at the boundary

$$\Omega = K\sin(-\theta_N).$$

From the last two equations we can simply determine the value of $\sin(\theta_{s+1})$ as a function of s, N, and Ω. Substituting this value into Eq. (4), we obtain

$$\sin(\theta_{s+j}) = 2(j-1)\frac{\Omega}{K} + (2s-2N+1)\frac{\Omega}{K}. \quad (5)$$

When $\Omega>0$ ($\Omega<0$), Eq. (5) is always negative (positive) and has its minimum (maximum) value for $j=1$. This means that when the pacemaker succeeds in "convincing" its nearest neighbor to the right to adopt its frequency, all the others to the right do the same. Now, the absolute value of the critical threshold can be calculated by using only the fact

that the sine function is bounded ($|\sin(\theta)| \leq 1$) and using the expression for Ω as a function of ϵ and N, as it is derived in the Appendix A in Eq. (A2):

$$^{R}\left|\frac{\Delta\omega}{K}\right|_{C} = \frac{(1+\epsilon)N-\epsilon}{2N-2s-1}. \tag{6}$$

Equation (6) yields the bound for oscillators to the right (R) of s to approach a phase-locked state. Following the same procedure, but moving to the left of the pacemaker, we find

$$^{L}\left|\frac{\Delta\omega}{K}\right|_{C} = \frac{(1+\epsilon)N-\epsilon}{2s-1} \tag{7}$$

as bound for oscillators to the left (L) of the pacemaker to synchronize in a phase-locked motion. Since we are interested in a state with all oscillators of the chain being phase entrained, we need the stronger condition given by

$$\left|\frac{\Delta\omega}{K}\right|_{C} = \min\left[^{R}\left|\frac{\Delta\omega}{K}\right|_{C}, ^{L}\left|\frac{\Delta\omega}{K}\right|_{C}\right]. \tag{8}$$

For the pacemaker placed at the boundaries $s=0$ or N, using Eqs. (2) and (A2), we obtain

$$\left|\frac{\Delta\omega}{K}\right|_{C} = \frac{(1+\epsilon)(2N-1)+1}{2N-1}. \tag{9}$$

B. Ring topology

If we close the chain of $N+1$ oscillators to a ring, $k_i \equiv 2 \ \forall \ i$, the derivation of the upper bound on the pacemaker's frequency $|\Delta\omega/K|_C$ proceeds in analogy to that of Eq. (9). Using Eq. (3), the final result is then given as

$$\left|\frac{\Delta\omega}{K}\right|_{C} = \frac{1}{N} + (1+\epsilon). \tag{10}$$

In all former cases, for $N\to\infty$, the critical threshold $|\Delta\omega/K|_C$ goes to $(1+\epsilon)$ [or values proportional to $(1+\epsilon)$]. Differently, the common frequency Ω goes to ω for $0 \geq \epsilon > -1$ and $N\to\infty$, while Ω goes to $\omega+\Delta\omega$ for $\epsilon=-1$ and $N\to\infty$. Therefore the symmetric coupling of the pacemaker to the rest of the system favors the synchronizability of the system, while it can no longer synchronize for a completely asymmetric coupling of the pacemaker ($\epsilon=-1$). This is plausible as it must be easier for N oscillators to convince one pacemaker to follow them (case $\epsilon>-1$) than the opposite case, in which the pacemaker must convince N oscillators to follow it (case $\epsilon=-1$). For increasing system size, the latter case becomes impossible, while the former is still possible.

C. Topological switch to synchronization

The numerical results of this paper are obtained by integrating the set of Eqs. (1) with the Runge-Kutta method of fourth order ($dt=0.1$). The numerical value of the critical threshold $|\Delta\omega/K|_C$ is evaluated with an accuracy of 5×10^{-5}. As one knows from [4], in case of regular topologies, stable solutions with different winding numbers are possible. To avoid different winding numbers, we always choose homogeneous initial conditions (also the distribution of *phases* would be different, in particular, in the synchronized state).

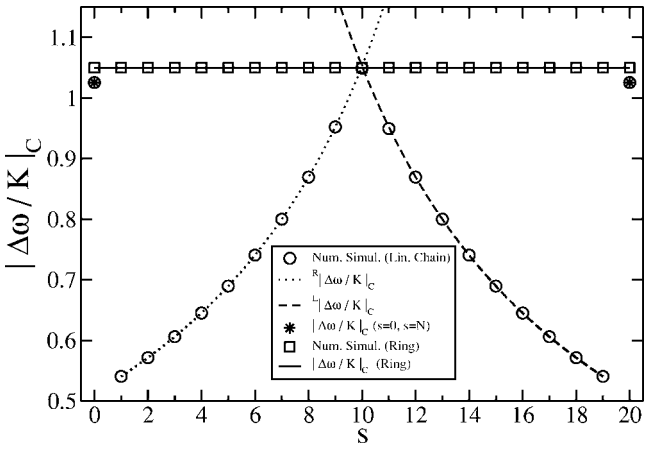


FIG. 2. Critical threshold $|\Delta\omega/K|_C$ as function of the position s of the pacemaker on a one-dimensional lattice. For further details see the text.

We summarize the results obtained so far in Fig. 2. The analytical results for $|\Delta\omega/K|_C$ are represented by lines for the open chain [dotted line from Eq. (6) and dashed line from Eq. (7)] and for the ring [full line from Eq. (10)], and by crosses [Eq. (9)] in the case of an open chain with the pacemaker at the boundaries, while the circles (open chain) and squares (ring) represent numerical data that reproduce the analytical predictions within the numerical accuracy. All results are obtained for $N=20$ and $\epsilon=0$; they are plotted as a function of the pacemaker's position s which matters only in the case of the open chain. The horizontal line obviously refers to the ring; the two branches (left and right), obtained for the chain, cross this line when the pacemaker is placed at $s=10$ in the middle of the chain. When the pacemaker is located at the boundaries $s=0$ and 20, we obtain two isolated data points close to the horizontal line.

Let us imagine that for given N and ϵ the absolute value of the pacemaker's frequency $|\Delta\omega|$ and the coupling K are specified out of range, such that the ratio is too large to allow for phase-locked motion on a chain, but small enough to allow the phase entrainment on a ring. It is then the mere closure of the open chain to a ring that leads from nonsynchronized to synchronized oscillators with phase-locked motion. Therefore, for a whole range of ratios $|\Delta\omega/K|$, no fine tuning is needed to switch to a synchronized state, but just a simple change in topology, the closure of a chain to a ring. Because this closure may be much more feasible in real systems than a fine tuning of parameters to achieve synchronization, we believe that this mechanism is realized in natural systems and should be utilized in artificial ones. In our numerical integration we simulated such a switch, and plot the phase portrait in Fig. 3. The phases φ_i as a function of time are always projected to the interval $[0,2\pi)$: we use a thick black line for the phase of the pacemaker and thin dark gray lines for the other oscillators. In this concrete numerical simulation with $T\,dt=4000$ integration steps altogether, we

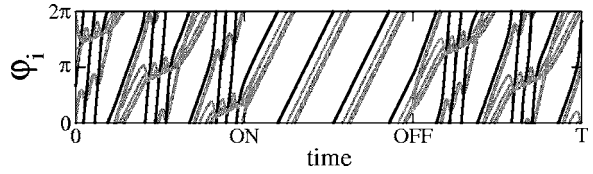


FIG. 3. Phase portrait of a system of Kuramoto oscillators on a one-dimensional lattice. This figure shows how it is possible to switch "ON" and "OFF" synchronization via a simple topological change of the system, passing from a linear chain to a ring, and vice versa. A more detailed description is given in the text.

analyzed a one-dimensional lattice of Kuramoto oscillators with $N=6$, $s=2$, $\Delta\omega/K=1$, and $\epsilon=0$. In the time interval from 0 to "ON" $(T/3)$ we see a phase evolution with different slopes and varying with time. Moreover, the pacemaker and the left part of the system (oscillators $i=0,1$) have larger frequencies than the right part of the system (oscillators $i=3,4,5,6$). At the instant "ON" we close the chain, passing to a ring topology; the system almost instantaneously reaches a phase-locked motion (all phases moving with the same and constant frequency). At time "OFF" $(2T/3)$ we open the ring; again, the system then behaves similarly to the first phase, i.e., for $t\in[0,ON)$.Furthermore it should be noticed from Fig. 2 that it is also favorable to put the pacemaker at the boundaries of an open chain to facilitate synchronization. For $d=1$ the pacemaker then has to entrain only one rather than two nearest neighbors so that the range of allowed frequencies $|\Delta\omega|$ increases.

V. HIGHER DIMENSIONS

All of our results obtained so far extend qualitatively to higher dimensions d, when system (1) is placed on a hypercubic lattice with (N_j+1) oscillators in each direction j, so that we have an ensemble of $\Pi_{j=1}^{d}(N_j+1)$ Kuramoto oscillators, where the ith oscillator's position is labeled by a d-dimensional vector $\vec{i}$, with $0\leqslant i_j\leqslant N_j\ \forall\ j=1,\ldots,d$. If the condition for having a phase-locked motion is satisfied, the system of oscillators reaches a common frequency still given by Eq. (2). This condition now is satisfied for $|\Delta\omega/K|\leqslant|\Delta\omega/K|_C$, the critical ratio for the pacemaker's frequency at position $\vec{s}=(s_1,\ldots,s_d)$ (see Fig. 4). In the simplest case when the lattice has $(N+1)$ sites in each direction, we can qualitatively extend all the previous results, obtained so far for $d=1$, to any dimension $d\geqslant 2$. For example, the "closure" of the open boundaries of the lattice to a torus in d dimensions favors synchronization of the system. We checked this numerically for $d=2$ and $N=10$ (Fig. 4). Except for the central node at $\vec{s}=(5,5)$, the critical threshold in case of open boundary conditions always lies below that for periodic boundary conditions. Moreover, it is natural to assume that the former results, obtained for one-dimensional lattices, in the case of periodic boundary conditions extend to d-dimensional lattices by replacing N in Eq. (10) by $(N+1)^d-1$

$$\left|\frac{\Delta\omega}{K}\right|_C=\frac{1}{(N+1)^d-1}+(1+\epsilon). \tag{11}$$

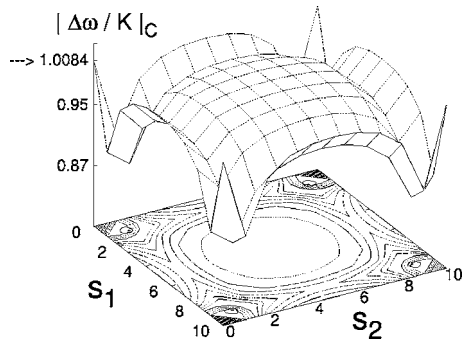


FIG. 4. Same as Fig. 2, but in $d=2$ dimensions. The linear size of the system is $N=10$. The arrow indicates the value of the critical threshold for periodic boundary conditions.

For $\epsilon=-1$ and fixed K, the entrainment window $\Delta\omega_c$ decreases exponentially with increasing dimension d of the lattice and as a power of the linear size of the lattice with exponent d. This conjecture is supported by the numerical results as documented in Fig. 5. For this reason we expect that for infinite-dimensional systems $(d\to\infty)$ the entrainment window decreases exponentially with increasing linear size of the system if the pacemaker is asymmetrically coupled with $\epsilon=-1$.

In order to verify this conjecture, we study a system of limit-cycle oscillators placed on a Cayley tree (Fig. 6). Cayley trees have z branches to nearest neighbors at each node, apart from nodes in the outermost shell, where the number of nearest neighbors is $z-1$. Cayley trees are infinite-dimensional objects in the sense that the surface of the system (i.e., the number of nodes in the outermost shell at distance R from the center) is proportional to the volume of the system (i.e., the total number of nodes in the tree). The set of evolution equations is still given by Eq. (1), with the adjacency matrix A specialized to the underlying Cayley tree

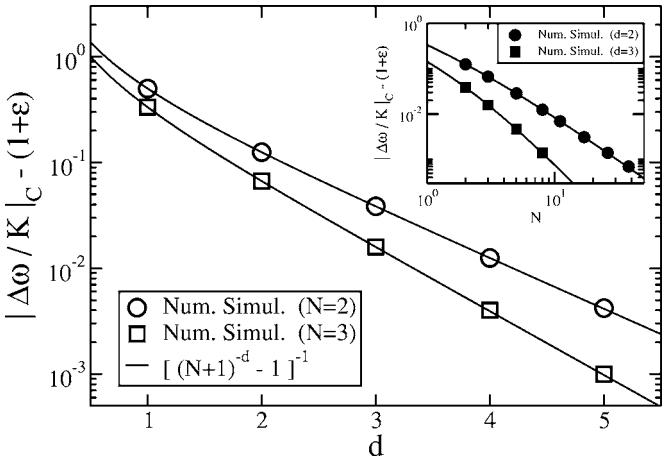


FIG. 5. Entrainment window for higher-dimensional lattices. We plot the critical threshold $|\Delta\omega/K|_C-(1+\epsilon)$ which is independent of ϵ as is seen from Eq. (11). The main plot shows the dependence of the entrainment window on the dimension d, while the inset shows its dependence on the linear size N of the lattice. Numerical results are plotted as symbols. They fit perfectly with our predictions according to Eq. (11).

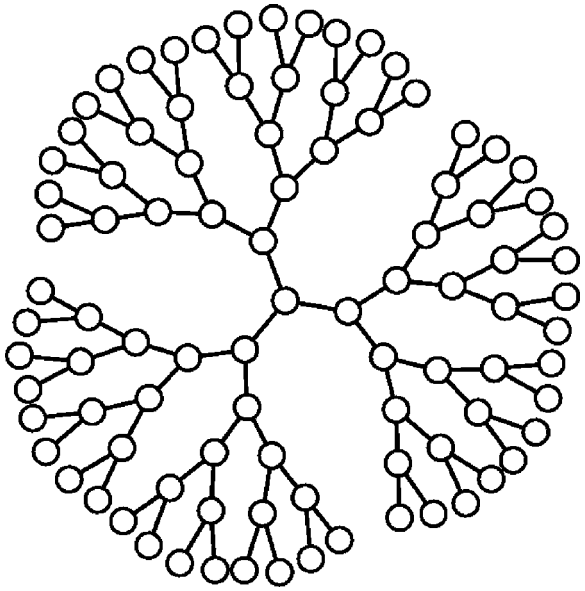


FIG. 6. Cayley tree with $z=3$ branches and radius $R=5$.

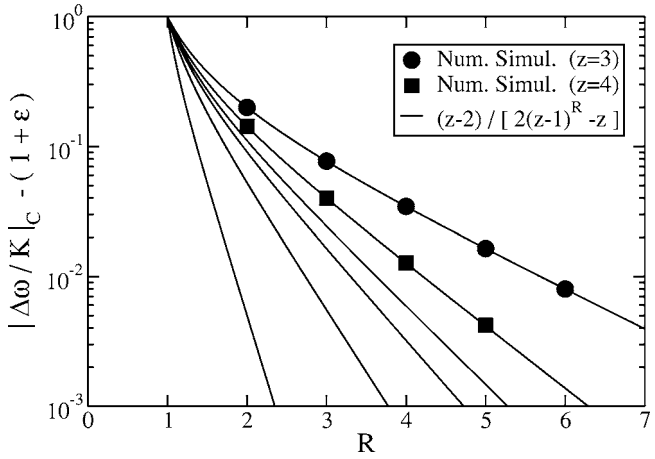


FIG. 7. Entrainment window for Cayley trees as a function of the radius R. Numerical results (full dots and squares) refer to $z=3$ and 4. The full lines correspond to plots of $|\Delta\omega/K|_C-(1+\epsilon)$, as defined in Eq. (13), for different coordination numbers z; from top to bottom $z=3$, 4, 5, 6, 10, and 100, respectively.

topology. For simplicity, we consider only the case of the pacemaker placed at the center of the tree ($s=0$).

When the number of branches per nodes z is 2, the Cayley tree becomes a linear chain, with $N=2R$ and the pacemaker placed at position $s=N/2=R$. Obviously for $z=2$ we find the same results as given in the previous section (see Appendix B).

The infinite dimensionality of the tree shows up for $z>2$. We obtain

$$\Omega = \Delta\omega \frac{z-2}{z-2+(1+\epsilon)[2(z-1)^R - z]} \tag{12}$$

as the common frequency for the phase-locked motion and

$$\left|\frac{\Delta\omega}{K}\right|_C = \frac{z-2}{2(z-1)^R - z} + (1+\epsilon) \tag{13}$$

as critical ratio for obtaining the phase-locked motion. Both Eqs. (12) and (13) are valid for $R \geqslant 2$. In particular Eq. (13) tell us that for $\epsilon=-1$ the entrainment window *decreases exponentially* with increasing radius R of the Cayley tree (see Fig. 7). For large $z \geqslant 2$ it is easily seen that the radius R becomes proportional to the depth D of the network (in the limit $z \gg 2$ one finds $D \simeq R$). These results nicely confirm those recently obtained by Kori and Mikhailov [8], who found that in infinite-dimensional systems such as random networks and small-world networks with a high number of rewired edges, the entrainment window decreases exponentially with increasing depth of the network. In order to connect their results to ours, we note that not only the radius R of the Cayley tree but also the linear size N of the hypercubic lattices is proportional to the depth of these regular networks.

VI. SUMMARY AND CONCLUSIONS

Entrainment of Kuramoto oscillators coupled on regular lattices via a pacemaker is possible for large system sizes ($N \to \infty$) only in the case of symmetric ($\epsilon=0$) couplings of the pacemaker. As the pacemaker coupling becomes asymmetric ($-1 \leqslant \epsilon < 0$), synchronization becomes more and more difficult, and impossible for large system sizes and $\epsilon=-1$. However, we are not only interested in the "thermodynamic" limit. For finite N and $-1 \leqslant \epsilon \leqslant 0$, we find that for a whole range of ratios $|\Delta\omega/K|$ it is possible to induce synchronization by a mere closure of a chain to a ring. The sensitive dependence of synchronization on the topology in a certain range of parameters may be exploited in artificial networks and is—very likely—already utilized in natural systems, in which a switch to a synchronized state should be easily feasible (although we are currently not aware of a concrete example from biological systems). If the pacemaker is coupled symmetrically to the other oscillators ($\epsilon=0$), the entrainment window stays finite in the large-N limit, but the common frequency approaches zero for $N \to \infty$. In the other extreme case, if the pacemaker is coupled asymmetrically to the rest of the system ($\epsilon=-1$), our main result is that the entrainment window decreases as a power of the depth of the network with the dimension d in the exponent. Extrapolating this behavior to arbitrary dimension d, we see that one of the reasons for the exponentially fast "closure" of the entrainment window in complex network topologies [8] is their effective infinite dimensionality.

APPENDIX A: COMMON FREQUENCY

Imposing the phase-locked condition, setting $\omega=0$, and dividing by $K/k_i \ \forall \ i=0,\ldots,N$, system (1) takes the form

$$\frac{k_i}{K}\Omega = \delta_{i,s}\frac{k_i}{K}\Delta\omega + (1+\delta_{i,s}\epsilon)\Omega_i, \tag{A1}$$

where

$$\Omega_i := \sum_{j\neq i} \sin(\varphi_j - \varphi_i).$$

From the odd parity of the sine function it follows that

$$\sum_i \Omega_i = 0 \quad \text{and} \quad \sum_{i \neq g} \Omega_i = -\Omega_g.$$

Summing (A1) over all i except $i=s$ we obtain

$$\sum_{i \neq s} \Omega \frac{k_i}{K} = \sum_{i \neq s} \delta_{i,s} \frac{k_i \Delta\omega}{K} + \sum_{i \neq s} (1 + \delta_{i,s}\epsilon)\Omega_i$$

$$\Rightarrow \frac{\Omega}{K} \sum_{i \neq s} k_i = -\Omega_s$$

while summing (A1) over all i except $i=j$, where $j \neq s$, we obtain

$$\sum_{i \neq j} \Omega \frac{k_i}{K} = \sum_{i \neq j} \delta_{i,s} \frac{k_i \Delta\omega}{K} + \sum_{i \neq j} (1 + \delta_{i,s}\epsilon)\Omega_i$$

$$\Rightarrow \frac{\Omega}{K} \sum_{i \neq j} k_i = \frac{k_s \Delta\omega}{K} - \Omega_j + \epsilon\Omega_s$$

$$\Rightarrow \frac{\Omega}{K} \sum_{j \neq s} \sum_{i \neq j} k_i = \sum_{j \neq s} \frac{k_s \Delta\omega}{K} - \sum_{j \neq s} \Omega_j + \sum_{j \neq s} \epsilon\Omega_s$$

$$\Rightarrow \frac{\Omega}{K} \sum_{j \neq s} \sum_{i \neq j} k_i = \frac{N k_s \Delta\omega}{K} + (1 + \epsilon N)\Omega_s.$$

It should be noticed that

$$\sum_{j \neq s} \sum_{i \neq j} k_i = N k_s + (N-1) \sum_{i \neq s} k_i,$$

so that we can write

$$\Omega \left((1+\epsilon) \sum_{i \neq s} k_i + k_s \right) = k_s \Delta\omega,$$

from which Eq. (2) is implied.

In the case of a one-dimensional lattice with open boundary conditions, we have $k_0 = 1 = k_N$ and $k_j = 2 \ \forall \ j \neq 0, N$, so that

$$\sum_{i \neq s} k_i = \begin{cases} 2N-1 & \text{if } s = 0 \text{ or } s = N, \\ 2(N-1) & \text{otherwise.} \end{cases}$$

The common frequency of Eq. (2) can be written as

$$\Omega = \Delta\omega \begin{cases} \dfrac{1}{(1+\epsilon)(2N-1)+1} & \text{if } s = 0 \text{ or } s = N, \\ \dfrac{1}{(1+\epsilon)(N-1)+1} & \text{otherwise.} \end{cases} \tag{A2}$$

APPENDIX B: COMMON FREQUENCY AND CRITICAL THRESHOLD FOR CAYLEY TREES

Consider a node in the outermost shell, i.e., at distance R from the pacemaker, reached along the path $\vec{s}_R = (s_1, s_2, s_3, \ldots, s_{R-1}, s_R)$ from the central node. Here $s_1 = 1, \ldots, z$, while $s_i = 1, \ldots, z-1 \ \forall \ i = 2, \ldots, R$ (see Fig. 6) label the choice of branch along the path. For this node we can write Eqs. (1) as

$$\dot{\varphi}_{\vec{s}_R} = \Omega = K \sin(\varphi_{\vec{s}_{R-1}} - \varphi_{\vec{s}_R}) \Rightarrow \sin(\varphi_{\vec{s}_{R-1}} - \varphi_{\vec{s}_R}) = \frac{\Omega}{K},$$

from which we can see that all oscillators at distance R, independently of the path along which they are reached from the center, satisfy the same equation.

Proceeding in an analogous way as before, we find for the shell $R-1$

$$\dot{\varphi}_{\vec{s}_{R-1}} = \Omega = \frac{K}{z} \left[\sum_{s_R=1}^{z-1} \sin(\varphi_{\vec{s}_R} - \varphi_{\vec{s}_{R-1}}) + \sin(\varphi_{\vec{s}_{R-2}} - \varphi_{\vec{s}_{R-1}}) \right]$$
$$= \frac{K}{z} \left[-(z-1)\frac{\Omega}{K} + \sin(\varphi_{\vec{s}_{R-2}} - \varphi_{\vec{s}_{R-1}}) \right]$$
$$\Rightarrow \sin(\varphi_{\vec{s}_{R-2}} - \varphi_{\vec{s}_{R-1}}) = \frac{\Omega}{K}[z + (z-1)].$$

Furthermore, for the shell $R-2$ we obtain

$$\dot{\varphi}_{\vec{s}_{R-2}} = \Omega$$
$$= \frac{K}{z} \left[\sum_{s_{R-1}=1}^{z-1} \sin(\varphi_{\vec{s}_{R-1}} - \varphi_{\vec{s}_{R-2}}) + \sin(\varphi_{\vec{s}_{R-3}} - \varphi_{\vec{s}_{R-2}}) \right]$$
$$= \frac{K}{z} \left[-(z-1)\frac{\Omega}{K}[z + (z-1)] + \sin(\varphi_{\vec{s}_{R-3}} - \varphi_{\vec{s}_{R-2}}) \right]$$
$$\Rightarrow \sin(\varphi_{\vec{s}_{R-3}} - \varphi_{\vec{s}_{R-2}})$$
$$= \frac{\Omega}{K}[z + z(z-1) + (z-1)^2]$$

until we arrive for $R-(t+1)$ at

$$\sin(\varphi_{\vec{s}_{R-(t+1)}} - \varphi_{\vec{s}_{R-t}}) = \frac{\Omega}{K} \left(z \sum_{q=0}^{t-1} (z-1)^q + (z-1)^t \right).$$

For $t = R-1$, we have

$$\sin(\varphi_{\vec{s}_0} - \varphi_{\vec{s}_1}) = \frac{\Omega}{K} \left(z \sum_{q=0}^{R-2} (z-1)^q + (z-1)^{R-1} \right),$$

but also, at the center of the Cayley tree,

$$\Omega = \dot{\varphi}_{\vec{s}_0} = \Delta\omega + (1+\epsilon)\frac{K}{z} \sum_{s_1=1}^{z} \sin(\varphi_{\vec{s}_1} - \varphi_{\vec{s}_0})$$
$$= \Delta\omega - (1+\epsilon)\Omega \left(z \sum_{q=0}^{R-2} (z-1)^q + (z-1)^{R-1} \right)$$

from which

$$\Omega = \Delta\omega\left[1 + (1+\epsilon)\left(z\sum_{q=0}^{R-2}(z-1)^q + (z-1)^{R-1}\right)\right]^{-1}. \tag{B1}$$

For $z=2$ the Cayley tree reduces to a linear chain with $N+1$ oscillators ($N=2R$), open boundary conditions, and the pacemaker placed at position $s=N/2$. The common frequency given in Eq. (B1) becomes the same as in Eq. (A2).

For $z>2$ we can rewrite the truncated geometric series as

$$\sum_{q=0}^{R-2}(z-1)^q = \frac{(z-1)^{R-1}-1}{z-2} \tag{B2}$$

and obtain after some algebra Eq. (12).

As we have seen, all oscillators at the same distance r from the pacemaker satisfy the same equation. We can write

$$\sin(\varphi_r - \varphi_{r-1}) = -\frac{\Omega}{K}\left(z\sum_{q=0}^{R-r-1}(z-1)^q + (z-1)^{R-r}\right), \tag{B3}$$

suppressing the path that was followed to reach the node. It is easily seen that when $\Omega>0$ ($\Omega<0$), Eq. (B3) is always negative (positive), monotone and increasing (decreasing), and takes its minimum (maximum) value for $r=1$. Imposing the bound of the sine function [$|\sin(\varphi)| \leqslant 1$] to Eq. (B3), with $r=1$, and inserting Eq. (B1) we obtain

$$\left|\frac{\Delta\omega}{K}\right|_C = \frac{1}{z\sum_{q=0}^{R-2}(z-1)^q + (z-1)^{R-1}} + (1+\epsilon).$$

For $z=2$ again we obtain Eqs. (6) and (7) with $s=N/2$. For $z>2$ we can again use the truncated geometric series of Eq. (B2) and obtain Eq. (13).

[1] J. Buck, Nature (London) **211**, 562 (1966).
[2] T. J. Walker, Science **166**, 891 (1969).
[3] I. Kanter, W. Kinzel, and E. Kanter, Europhys. Lett. **57**, 141 (2002).
[4] Y. Kuramoto, *Chemical Oscillators, Waves, and Turbulence* (Springer, New York, 1984).
[5] A. T. Winfree, *The Geometry of Biological Time* (Springer, New York, 1980).
[6] J. A. Acebrón, L. L. Bonilla, C. J. Pérez-Vicente, F. Ritort, and R. Spigler, Rev. Mod. Phys. **77**, 137 (2005).
[7] H. Yamada, Prog. Theor. Phys. **108**, 13 (2002).
[8] H. Kori and A. S. Mikhailov, Phys. Rev. Lett. **93**, 254101 (2004).
[9] A. N. Zaikin and A. M. Zhabotinsky, Nature (London) **225**, 535 (1970).
[10] B. Blasius and R. Tönjes, Phys. Rev. Lett. **95**, 084101 (2005); Y. Kuramoto and T. Yamada, Prog. Theor. Phys. **56**, 724 (1976).
[11] K. J. Lee, E. C. Cox, and R. E. Goldstein, Phys. Rev. Lett. **76**, 1174 (1996).
[12] B. Blasius, A. Huppert, and L. Stone, Nature (London) **399**, 354 (1999).
[13] M. Weiss, T. Kottos, and T. Geisel, Phys. Rev. E **63**, 056211 (2001).

Paper III

Reentrant synchronization and pattern formation in pacemaker-entrained Kuramoto oscillators

Filippo Radicchi and Hildegard Meyer-Ortmanns

Physical Review E **74**, 026203 (2006)

In this paper, we continue our analysis of the Kuramoto model on complex topologies and in the presence of a pacemaker. We determine that the behavior of this dynamical system on a small-world model [see section 1.5.2] can be effectively described by introducing a range-dependent factor in the interaction term. Moreover, the critical point of the transition between the synchronized and the non-synchronized phases is a non-monotonic function of the parameter which characterizes the range of the interaction, leading therefore to a so-called reentrant phase transition. In the second part of the paper we explain the presence of pacemakers in real systems, since they can be "dynamically" created [21] by local gradients in the natural frequencies of a quasi-identical population of oscillators.

Personal Contribution

I performed the analytical calculations and the numerical simulations reported in the following paper.

Reentrant synchronization and pattern formation in pacemaker-entrained Kuramoto oscillators

Filippo Radicchi* and Hildegard Meyer-Ortmanns†
School of Engineering and Science, International University Bremen, P.O. Box 750561, D-28725 Bremen, Germany
(Received 13 March 2006; published 3 August 2006)

We study phase entrainment of Kuramoto oscillators under different conditions on the interaction range and the natural frequencies. In the first part the oscillators are entrained by a pacemaker acting like an impurity or a defect. We analytically derive the entrainment frequency for arbitrary interaction range and the entrainment threshold for all-to-all couplings. For intermediate couplings our numerical results show a reentrance of the synchronization transition as a function of the coupling range. The origin of this reentrance can be traced back to the normalization of the coupling strength. In the second part we consider a system of oscillators with an initial gradient in their natural frequencies, extended over a one-dimensional chain or a two-dimensional lattice. Here it is the oscillator with the highest natural frequency that becomes the pacemaker of the ensemble, sending out circular waves in oscillator-phase space. No asymmetric coupling between the oscillators is needed for this dynamical induction of the pacemaker property nor need it be distinguished by a gap in the natural frequency.

DOI: 10.1103/PhysRevE.74.026203 PACS number(s): 05.45.Xt, 05.70.Fh, 89.75.Kd

I. INTRODUCTION

Synchronization in the sense of coordinated behavior in time is essential for any efficient organization of systems, natural as well as artificial ones. In artificial systems like factories the sequence of production processes should be synchronized in a way that time- and space-consuming storage of input or intermediate products is avoided. The same applies to natural systems like networks of cells which obviously perform very efficiently in fulfilling a variety of functions and tasks. In a more special sense, the synchronized behavior refers to oscillators with almost identical individual units, in particular to phase oscillators with continuous interactions as described by the Kuramoto model [1]. These sets of limit-cycle oscillators describe synchronization phenomena in a wide range of applications [2]. One of these applications is pattern formation in chemical oscillatory systems [3], described by reaction-diffusion systems. A well-known example for such a chemical system is the Belousov-Zabotinsky reaction, a mixture of bromate, bromomalonic acid, and ferroin, which periodically changes its color corresponding to oscillating concentrations of the red, reduced state and the blue, oxidized state. In these systems, expanding targetlike waves and rotating spiral waves are typical patterns [4]. As Kuramoto showed (as, e.g., in [1]), these dynamical systems can be well-approximated by phase oscillators if the interaction is weak. For reaction-diffusion equations with a nonlinear interaction term he predicted circular waves with strong similarities to the experimental observation of target patterns. In these systems, defects or impurities seem to play the role of pacemakers, driving the system into a synchronized state. Therefore he treated pacemakers as local "defect" terms leading to heterogeneities in the reaction-diffusion equations.

In this paper we consider two types of pacemakers. In the first part they are introduced as "defects" in the sense that they have a different natural frequency from the rest of the system, whose oscillators have either exactly the same frequency or small random fluctuations about some common average by assumption, where in the second case the pacemaker's frequency is clearly different from the values of a typical fluctuation. Actually, we will not consider the latter case, since small fluctuations do not lead to any qualitative change. The *ad hoc* distinction of the pacemaker may reflect natural and artificial systems with built-in impurities like those in the Belousov-Zhabotinsky system. We have recently studied these systems for nearest-neighbor interactions [5] and regular networks in d dimensions for periodic and open boundary conditions. For such systems we have shown that it is the mean distance of all other nodes from the pacemaker (the so-called depth of the network) that determines its synchronizability. Periodic boundary conditions facilitate synchronization as compared to open boundary conditions. Here we extend these results to a larger interaction range. The effect of an extended interaction range on oscillatory systems without a pacemaker were considered in [6]. We analytically derive the entrainment frequency for arbitrary interaction range between next-neighbor and all-to-all interactions. We derive the entrainment window for all-to-all couplings analytically, for intermediate couplings numerically. The entrainment window depends nonmonotonically on the interaction range, so that the synchronization transition is reentrant. In our system the reentrance can be explained in terms of the weight of the interaction term and the normalization of the coupling strength.

In the second part we consider a system of oscillators without a "defect," all oscillators on the same footing up to the difference in their natural frequencies. An obvious choice for the natural frequencies would be a Gaussian distribution or another random distribution to describe fluctuations in natural frequencies in otherwise homogeneous systems, without impurities. Such systems have been studied by Blasius and Tönjes [7] for a Gaussian distribution. The authors showed that in this case an asymmetric interaction term is needed to have synchronization, this time driven by a dynamically established pacemaker, acting as a source of con-

*Email address: f.radicchi@iu-bremen.de
†Email address: h.ortmanns@iu-bremen.de

centric waves. This result shows that conditions exist under which a system establishes its own pacemaker in a "self-organized" way. In contrast to the random distribution of natural frequencies we are interested in a (deterministic) gradient distribution, in which the natural frequencies linearly decrease over a certain region in space. Such a deterministic gradient alone is probably not realistic for natural systems, but considered as a subsystem of a larger set of oscillators with natural frequency fluctuations it may be realized over a certain region in space. Here we suppress the fluctuations and focus on the effect of the gradient alone. As it turns out, the asymmetry in the natural frequency distribution is sufficient for creating a pacemaker as the center of circular waves in oscillator-phase space, without the need for an asymmetric term in the interaction. The oscillator with the highest natural frequency becomes the source of circular waves in the synchronized system. We call it dynamically induced as its only inherent difference is its local maximum at the boundary in an otherwise "smooth" frequency distribution. The type of pattern created by the pacemaker has the familiar form of circular waves. Beyond a critical slope in the natural frequencies, full synchronization is lost. It is first replaced by partial synchronization patterns with bifurcation in the frequencies of synchronized clusters, before it gets completely lost for too steep slopes.

The outline of the paper is as follows. In the first part we treat the pacemaker as defect and derive the common entrainment frequency for arbitrary interaction range and topology (Sec. II). Next we determine the entrainment window, analytically for all-to all coupling and numerically for intermediate interaction range (Sec. II). Here we see the reentrance of the transition as a function of the interaction range (Sec. II). In the second part (Sec. III) we consider dynamically induced pacemakers, first without the asymmetric interaction term, for which we analytically derive the synchronization transition as a function of the gradient in the frequencies (Sec. III A). In Sec. III B we add an asymmetric term, so that the pattern formation is no longer surprising due to the results of [7], but the intermediate patterns of partial synchronization are different. In Sec. IV we summarize our results and conclusions.

II. PACEMAKER AS DEFECT IN THE SYSTEM

The system is defined on a network, regular or small-world like. To each node i, $i=0,\ldots,N$, we assign a limit-cycle oscillator, characterized merely by its phase φ_i, which follows the dynamics

$$\dot{\varphi}_i = \omega + \delta_{i,s}\Delta\omega + \frac{K}{k_i}\sum_{j\neq i} A_{j,i}\sin(\varphi_j - \varphi_i) \tag{1}$$

with the following notations. The frequency ω denotes the natural frequencies of the system. In this system we treat the pacemaker as a defect. It is labeled by s and has a natural frequency that differs by $\Delta\omega$ from the frequency of the other oscillators having all the same frequency ω. $\delta_{i,j}$ denotes the Kronecker delta. The constant $K>0$ parametrizes the coupling strength. We consider regular networks and choose

$$A_{j,i} = r_{j,i}^{-\alpha},$$

$r_{j,i}$ is the distance between nodes j and i, that is

$$r_{j,i} = \min([|j-i|,(N+1)-|j-i|]) \tag{2}$$

on a one-dimensional lattice with periodic boundary conditions. In two or higher dimensions it is the shortest distance in lattice links. The parameter $0\leqslant\alpha\leqslant\infty$ tunes the interaction range. Alternatively we consider $A_{j,i}$ as the adjacency matrix on a small-world topology: $A_{j,i}=1$ if the nodes j and i are connected and $A_{j,i}=0$ otherwise). Moreover, $k_i=\Sigma_j A_{j,i}$ is the degree of the ith node, it gives the total number of connections of this node in the network. This system was considered before in [5] for nearest-neighbor interactions ($\alpha\to\infty$) on a d-dimensional hypercubic lattice and on a Cayley tree. Here we extend the results to long-range interactions via $0\leqslant\alpha<\infty$.

Entrainment frequency and entrainment window

In the Appendix of [5] we derived the common entrainment frequency Ω in the phase-locked state, for which $\dot{\varphi}_i\equiv\Omega$ for all i, to be given as

$$\Omega = \Delta\omega\frac{k_s}{\sum_i k_i} \tag{3}$$

in the rotated frame, in which the natural frequency ω is zero. In particular such a result is obtained directly from system (1), applying the transformation to the rotating frame $\varphi_i\to\varphi_i+\omega t\ \forall i$, multiplying the ith equation by k_i, summing over all equations, and then using the fact that $\Sigma_i\Sigma_j A_{j,i}\sin(\varphi_j-\varphi_i)=0$ because the symmetry of the adjacency matrix A and the antisymmetry of the sin function. Whenever $k_i\equiv k$ is independent of i as it happens for periodic boundary conditions and homogeneous degree distributions, in particular for $\alpha=0$ (all-to-all coupling), the common frequency (in the nonrotated frame) is given by

$$\Omega = \frac{\Delta\omega}{N+1} + \omega,$$

independent of α. In the limit $N\to\infty$ the system synchronizes at ω.

Analytical results for the entrainment window were derived for open and periodic boundary conditions and $\alpha\to\infty$ in [5]. For $\alpha=0$, the result is the same as for $\alpha\to\infty$, periodic boundary conditions and $d\geqslant 1$

$$\left|\frac{\Delta\omega}{K}\right|_c = \frac{1}{N} + 1. \tag{4}$$

This is seen as follows. Since for all-to-all couplings the mean-field approximation becomes exact, it is natural to rewrite the dynamical equations (1) in terms of the order parameter $Re^{i\psi}$ defined according to

$$Re^{i\psi} = \frac{1}{N}\sum_{j\neq s} e^{i\varphi_j}.$$

This quantity differs from the usual order parameter just by the fact that the sum runs over $j \neq s$. Equation (1) then takes the form

$$\Omega = \Delta\omega + KR\sin(\psi - \varphi_s) \quad \text{for } i = s,$$

$$\Omega = \frac{K}{N}\sin(\varphi_s - \varphi_i) + KR\sin(\psi - \varphi_i), \quad \forall\, i \neq s. \qquad (5)$$

Using $\varphi_i = \psi$ for all $i \neq s$ as one of the possible solutions and the fact that the sin function is bounded, these equations imply the entrainment window as given by Eq. (4) with N the total number of nodes minus 1. As we have shown in [5], the same dependence in terms of N holds for a d-dimensional hypercubic lattice with N' nodes per side and nearest-neighbor couplings, for which $N=(N'+1)^d-1$, so that the entrainment window for nearest-neighbor interactions and on a regular lattice decays exponentially with the depth of the network in the limit of infinite dimensions d; for random and small-world networks the same type of decay was derived in [8].

The intermediate interaction range $0<\alpha<\infty$ cannot be treated analytically. Here the numerical integration of Eq. (1) (via a fourth-order Runge-Kutta algorithm with time step $dt=0.05$ and homogeneous initial conditions) shows that the entrainment window is even smaller than in the limiting cases $\alpha=0$ or $\alpha=\infty$ for otherwise unchanged parameters (that is fixed size N, dimension d, coupling K). For these cases formula (4) gives an upper bound on the entrainment window, as one can easily see from Eq. (1) for $i=s$ because $\Sigma_{j\neq s}\sin(\varphi_j-\varphi_s)/r_{j,s}^{\alpha} \leq \Sigma_{j\neq s} 1/r_{j,s}^{\alpha} = k_s$, $\forall\alpha$. Actually the entrainment window becomes smaller for intermediate interaction range, a feature leading to the reentrance of the synchronization transition as a function of α, as we shall see in Figs. 1 and 2 below.

Reentrance of the phase transition

Reentrance phenomena are observed in a variety of phase transitions, ranging from superconductor-insulator transitions as a function of temperature [9] and noise-induced transitions [10] to chaotic coupled-map lattices as function of the coupling [11]. In [12] a reentrance phenomenon is discussed as an artifact of the approximation. Reentrance of phase transitions is challenging as long as it appears counterintuitive. For example, it is counterintuitive when synchronization is first facilitated for increasing coupling and later gets blocked when the coupling exceeds a certain threshold, so that synchronization depends nonmonotonically on the coupling. As we see in Fig. 1 for a one-dimensional system with periodic boundary conditions and for various system sizes N, the entrainment window $|\Delta\omega/K|_c$ depends nonmonotonically on the interaction range, parametrized by α. This dependence is easily explained by looking at the total weight coming from the sum of interactions. This contribution depends nonmonotonically on α, as $\Sigma_j \sin(\phi_j-\phi_i)/r_{j,i}^{\alpha}$ decreases for increasing α, while the coupling strength increases under the same

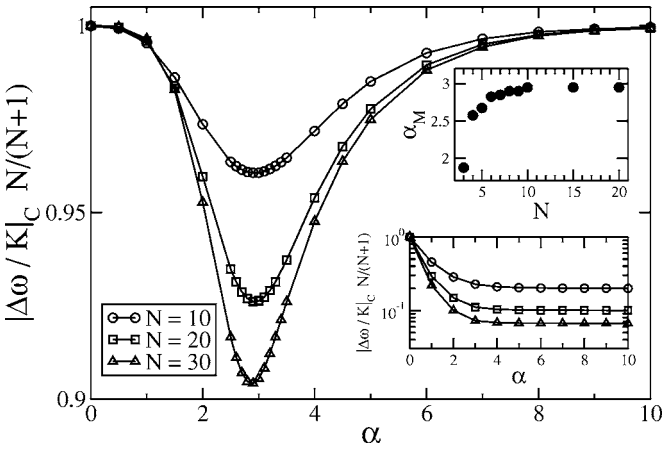


FIG. 1. Normalized entrainment window as a function of the interaction range parametrized via α, for a one-dimensional lattice with $N+1$ sites, periodic boundary conditions, and for interaction strength between the oscillators i and j suppressed by the power α of their distance $r_{j,i}$. The main plot indicates the reentrance of the synchronization transition. The upper inset shows the size dependence of α_m, for which the entrainment window becomes minimal, it saturates at $\alpha\sim 3$. The lower inset shows the monotonic dependence on α for constant coupling strength K/N. The figure was obtained for homogeneous initial conditions. The qualitative features remain the same for random initial conditions.

change of α, so that there is a competition between the two factors leading to a minimum of the weight for a certain value of α. In contrast, if we scale the coupling strength by a constant factor K/N, independently of the degree of the node i, the overall weight of the sum decreases for increasing α and along with this, the drive to the synchronized state. As it is shown in the inset of Fig. 1, the threshold monotonically decreases with α if the coupling strength is normalized by $1/N$ independently of the actual degree, as expected from our arguments above. The second inset shows the N-dependence of the value of α for which synchronization is most difficult to achieve, it converges to $\alpha=3$ for large N. Similar nonmonotonic behavior is observed if the interaction range is tuned via the number k of nearest neighbors on a ring rather than via α, see Fig. 2(a) and for a small-world topology [13], see Fig. 2(b) Here the small world is constructed by adding random shortcuts with probability p from a regular ring with nearest-neighbor interactions ($p=0$) to an all-to-all topology ($p=1$) [14].

III. DYNAMICALLY INDUCED PACEMAKERS

In view of dynamically induced pacemakers we study the dynamical system

$$\dot{\varphi}_i = \omega_i + \frac{K}{k_i}\sum_j A_{j,i}\Gamma(\varphi_j - \varphi_i) \qquad (6)$$

without a pacemaker. The interaction term Γ is given by

$$\Gamma(\varphi_j - \varphi_i) = \sin(\varphi_j - \varphi_i) + \gamma[1 - \cos(\varphi_j - \varphi_i)], \qquad (7)$$

where $\gamma=0$ corresponds to antisymmetric interaction, and $\gamma>0$ to asymmetric interaction. In the following we consider

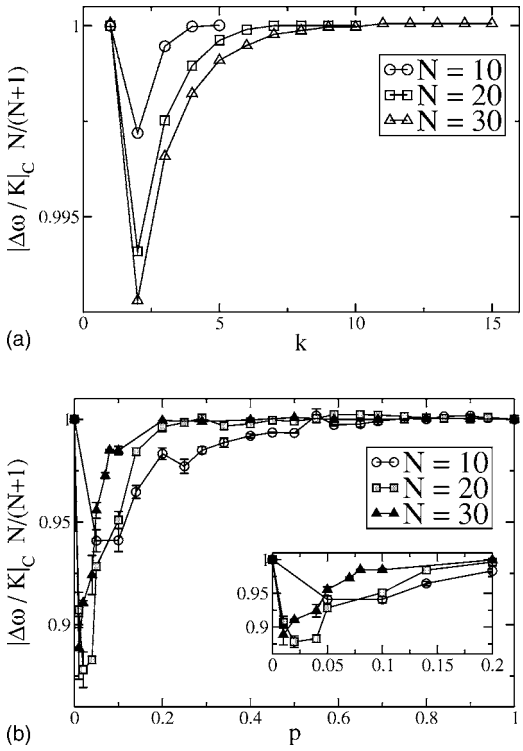


FIG. 2. Normalized entrainment window as function of the number of neighbors k (a) and of the probability p for adding shortcuts in a small-world topology (b) for a one-dimensional lattice with periodic boundary conditions and $N+1$ sites. Average values in (b) are taken over 50 realizations, error bars correspond to the standard deviation. Again the nonmonotonic behavior indicates reentrance of the synchronization transition.

the cases $\gamma=0$ and $\gamma=2$. The value $\gamma=2$ is chosen for convenience in order to compare our results with those of Blasius and Tönjes [7], but different values of $\gamma>0$ lead to qualitatively the same features.

We focus on lattices with periodic boundary conditions in one and two dimensions. Moreover, we consider only nearest-neighbor interactions, so that in $d=1$ we have $k_i=2\ \forall i$ and $k_i=4\ \forall i$ in $d=2$. The natural frequencies ω_i are chosen from a gradient distribution

$$\omega_i = \Delta\omega - r_{i,s}\frac{\Delta\omega}{\max_i r_{i,s}}, \tag{8}$$

where $r_{i,j}$ is the network distance between the oscillators i and j, given by Eq. (2) in one dimension and by the shortest path along the edges of the square lattice in two dimensions. The oscillator s is the oscillator with the highest natural frequency. For a one-dimensional lattice with periodic boundary conditions (ring) and $N+1$ oscillators, choosing $s=0$ for simplicity, $\max_i r_{i,0}=N'$, with $N'=N/2$ for N even and $N'=(N+1)/2$ for N odd. In the following we choose N as even; the case of N odd is very similar.

A. Pattern formation for $\gamma=0$

1. Entrainment frequency

Next we derive the common frequency in the case of phase-locking $\dot{\varphi}_i \equiv \Omega$, $\forall i$. Summing the equations of (6) and using the antisymmetry of Γ for $\gamma=0$, we obtain

$$\begin{aligned}(N+1)\Omega &= \sum_{i=0}^{N}\omega_i = \Delta\omega + 2\Delta\omega\sum_{i=1}^{N/2}(1-2i/N)\\ &= \Delta\omega\left[1+N-4/N\frac{N/2(N/2+1)}{2}\right]\\ &= \Delta\omega\frac{N}{2},\end{aligned}$$

leading to

$$\Omega = \Delta\omega\frac{N}{2(N+1)}, \tag{9}$$

so that $\Omega\to\Delta\omega/2$ for $N\to\infty$.

2. Entrainment window

It is convenient to rewrite the differential equations (6) in terms of phase lags $\theta_i=\varphi_i-\varphi_{i-1}$, so that the equation for the oscillator at position 0 reads

$$\Omega = \omega_0 + K\sin(\theta_1) \Rightarrow \sin(\theta_1) = \frac{\Omega-\omega_0}{K}.$$

For the other oscillators we obtain

$$\begin{aligned}\Omega &= \omega_1 + K/2[\sin(\theta_2)-\sin(\theta_1)]\\ &\Rightarrow \sin(\theta_2) = 2\frac{\Omega-\omega_1}{K}+\sin(\theta_1),\end{aligned}$$

$$\begin{aligned}\Omega &= \omega_2 + K/2[\sin(\theta_3)-\sin(\theta_2)]\\ &\Rightarrow \sin(\theta_3) = 2\frac{\Omega-\omega_2}{K}+\sin(\theta_2),\end{aligned}$$

so that in general

$$\begin{aligned}\Omega &= \omega_{i-1} + K/2[\sin(\theta_i)-\sin(\theta_{i-1})]\\ &\Rightarrow \sin(\theta_i) = 2\frac{\Omega-\omega_{i-1}}{K}+\sin(\theta_{i-1})\end{aligned}$$

or

$$\sin(\theta_i) = \sum_{j=1}^{i-1} 2\frac{\Omega-\omega_j}{K}+\sin(\theta_1).$$

The former sum can be performed as

$$\begin{aligned}\sum_{j=1}^{i-1} 2\frac{\Omega-\omega_j}{K} &= \frac{2}{K}\sum_{j=1}^{i-1}[\Omega-\Delta\omega(1-2j/N)]\\ &= \frac{2}{K}(\Omega-\Delta\omega)(i-1)+\Delta\omega\frac{2}{K}\frac{i(i-1)}{N},\end{aligned}$$

from which

$$\sin(\theta_i) = \frac{\Omega-\Delta\omega}{K}(2i-1)+\frac{\Delta\omega}{K}\frac{2i^2-2i}{N},$$

or, using Eq. (9),

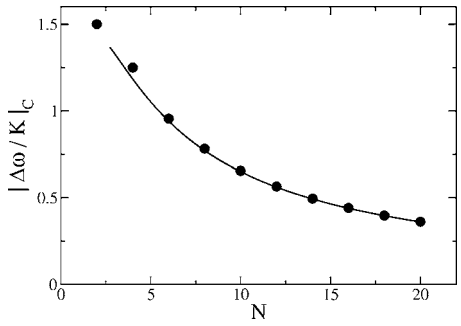


FIG. 3. Critical threshold for a one-dimensional lattice as function of its linear size N, full line corresponds to the analytic prediction, full dots to numerical results for $N \geqslant 2$, the symmetry breaking parameter is chosen as $\gamma=0$.

$$\sin(\theta_i) = \frac{\Delta\omega}{K}(2i-1)\left[\frac{N}{2(N+1)} - 1\right] + \frac{\Delta\omega}{K}\frac{2i(i-1)}{N}$$
$$= \frac{\Delta\omega}{K}\left[(1-2i)\frac{N+2}{2(N+1)} + \frac{2i(i-1)}{N}\right]. \quad (10)$$

For given N, we first determine the position i_m at which the right-hand side of Eq. (10) takes its minimum (maximum) value as a function of i. This value is constrained by the left-hand side of Eq. (10) to be out of $[-1,1]$. For i treated as a continuous variable the derivative of Eq. (10) with respect to i yields

$$i_m = \frac{N(N+2)}{4(N+1)} + \frac{1}{2}. \quad (11)$$

On the other hand

$$i_m \geqslant 1 \Leftrightarrow N \geqslant \sqrt{2} \quad (12)$$

and

$$i_m \leqslant \frac{N}{2} \Leftrightarrow N \geqslant 1 + \sqrt{3}, \quad (13)$$

so that for $N \gg 1$, $1 \leqslant i_m \leqslant N/2$ is certainly satisfied. Upon inserting i_m from Eq. (11) we have

$$\left|\frac{\Delta\omega}{K}\right|_C = \left|(1-2i_m)\frac{N+2}{2(N+1)} + \frac{2i_m(i_m-1)}{N}\right|^{-1}$$
$$= \left(\frac{N^2(N+2)^2 + 4(N+1)^2}{8N(N+1)^2}\right)^{-1}.$$

The critical threshold is then given by

$$\left|\frac{\Delta\omega}{K}\right|_C = \frac{8N(N+1)^2}{N^2(N+2)^2 + 4(N+1)^2}. \quad (14)$$

This relation is plotted in Fig. 3, where the full line represents the theoretical prediction, while the full dots correspond to the threshold values, numerically determined by integrating the system of differential Equations (6), starting with $N=2$. (The case of $N=2$ violates the constraint (13) in agreement with the fact that for $N=2$ the system degenerates to the former case of having a pacemaker as a defect and two oscillators with natural frequencies $\omega_1=\omega_2=0$.) Furthermore,

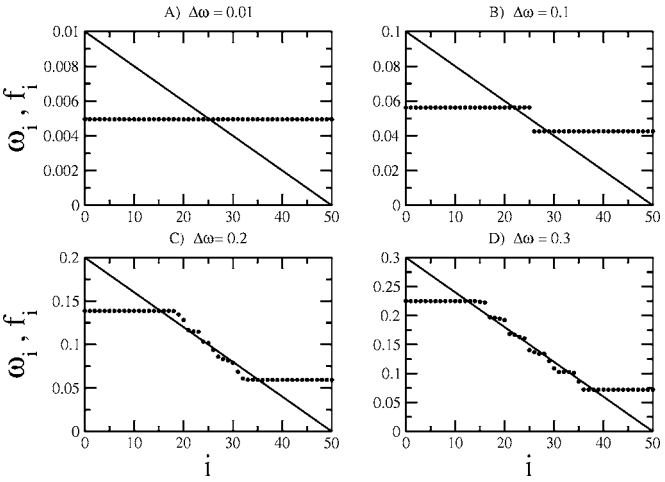


FIG. 4. Natural frequencies ω_i (full lines) and measured, stationary frequencies f_i (full dots) as a function of the position i of the oscillators on a one-dimensional lattice, for $N=100$, periodic boundary conditions, and four values of the maximal natural frequency $\Delta\omega$. The oscillator with maximal natural frequency $\Delta\omega$ is at position $s=0$. Only half of the oscillators up to index 50 are shown because of the symmetric arrangement. The symmetry breaking parameter is chosen as $\gamma=0$.

it should be noticed that Eq. (14) scales as $1/N$ for large values of N, the large-N behavior is, however, not visible for the range of N ($N \leqslant 20$) plotted in Fig. 3.

For a ring with maximal distance $N'=N/2$ and given N and K, the slope $\Delta\omega/N'$ in the natural frequencies determines whether synchronization is possible or not. The oscillators synchronize if the values of N lead to a ratio $|\Delta\omega/K|$ below the bound $|\Delta\omega/K|_C$ of Eq. (14), named as in Sec. II, Eq. (4), but with $\Delta\omega$ now standing for the maximal difference between the highest and the lowest natural frequencies. If for given N the slope, parametrized by $\Delta\omega$, exceeds this threshold, no entrainment is possible. For $N \to \infty$ the allowed slope goes to zero, the system does no longer synchronize. To further characterize the synchronization patterns, we measure the stationary frequency f_i of the individual oscillators i, defined according to

$$f_i = \lim_{t\to\infty} \frac{\varphi_i(t+t_0) - \varphi_i(t_0)}{t - t_0}, \quad (15)$$

as well as its average $f = \frac{1}{N+1}\Sigma_{i=0}^{N} f_i$ and variance $\sigma = \left[\frac{1}{N+1}\Sigma_{i=0}^{N}(f_i - f)^2\right]^{1/2}$. The average frequency and variance are introduced to characterize the intermediate patterns of partial synchronization and to distinguish it from the case $\gamma=2$ in Sec. III B. In all numerical calculations we set from now on $K=1$. Figure 4 shows the natural frequencies (ω_i) (represented as a full line) and the stationary (f_i) frequencies (shown as dots) for four slopes $\Delta\omega=0.01$, 0.1, 0.2, and 0.3 respectively, as function of the position on a linear chain of 100 oscillators, of which we show only one-half due to the symmetric arrangement. The case of $\Delta\omega=0.01$ corresponds to full synchronization, while the other three figures represent partial synchronization of two ($\Delta\omega=0.1$) and more clusters. It should be noticed that the two clusters for

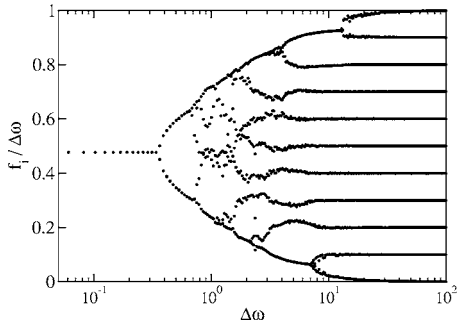


FIG. 5. Bifurcation in cluster-frequency space above a critical slope in the natural frequencies, parametrized by $\Delta\omega$, the maximal natural frequency. Numerically measured frequencies f_i (normalized such that $f_i/\Delta\omega \in [0,1]$) as a function of $\Delta\omega$ for a one-dimensional lattice with $N=20$ and periodic boundary conditions. The symmetry breaking parameter is chosen as $\gamma=0$.

($\Delta\omega=0.1$) would violate the condition (14) if they were isolated clusters of 50 oscillators each with a slope determined by $\Delta\omega=0.1$, but due to the nonlinear coupling to the oscillators of the second cluster, the partial synchronization is a stationary pattern. Fig. 5 displays the frequencies f_i as function of the slope, again parametrized by $\Delta\omega$. Here we have chosen $N=20$. We clearly see the bifurcation in frequency space starting at the critical value of $\Delta\omega \simeq 0.36$ and ending with complete desynchronization, in which the stationary frequencies equal the natural frequencies. For frequency synchronization the phases are locked, but their differences to the phase of say the oscillator $s=0$ ($\psi_i=\varphi_i-\varphi_0$) increases nonlinearly with their distance from s, as seen in Fig. 6. Therefore the distance between points of the same ψ decreases with i.

In two dimensions we simulated 100×100 oscillators on a square lattice with periodic boundary conditions and the oscillator s with the highest natural frequency placed at the center of the square lattice. The natural frequencies of the oscillators are still given by Eq. (8). Their spatial distribution looks like a square pyramid centered at the middle of the square lattice. Figure 7 displays $\sin\psi_i$ on this grid. It exhibits stationary patterns after 2×10^4 steps of integration, for six choices of $\Delta\omega$, all above the synchronization threshold. We see some remnants of synchronization, most pronounced in Fig. 7(A), in which a circular wave is created at the center at s, and coexists with waves absorbed by sinks at the four corners of the lattice. The projection of Fig. 7(A) on one dimension corresponds to Fig. 4(B) with a bifurcation into two cluster frequencies. Note that it is again the oscillator with the highest natural frequency that becomes the center of the outgoing wave, while the corners with the lowest natural frequencies ($\omega_i=0$) become sinks. Although Fig. 7(F) shows almost no remnant of synchronization, it is interesting to follow the time evolution towards this "disordered" pattern via a number of snapshots, as displayed in Fig. 8 after 9[22.5] (A), 41 [102.5] (B), 150 [375] (C), 393 [982.5] (D), 490 [1225] (E), and 1995 [4987.5] (F) integration steps T [T $\Delta\omega$ time steps in units of natural periods ($1/\Delta\omega$) of Eq. (8)]. While the pattern of Fig. 8(A) would be stationary in case of full synchronization, here it evolves after iterated reflections

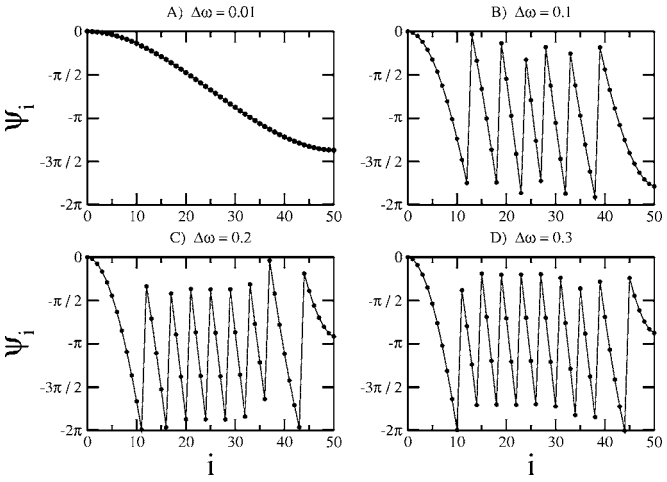


FIG. 6. Phase differences $\psi_i=\varphi_i-\varphi_0$ (from the oscillator at $i=0$) as a function of the oscillator index i for a one-dimensional lattice, with $N=100$ and periodic boundary conditions, for different values of the maximal natural frequency $\Delta\omega$ and $\gamma=0$. The difference depends nonlinearly on the distance from the oscillator at $i=0$. The "steps" in these plots are due to the projection of the variables ψ_i onto the interval $[0,2\pi)$.

to that of Fig. 8(F). The pattern of Fig. 8(F) is stationary in the sense that it stays "disordered," on a "microscopic" scale it shows fluctuations in the phases. The evolution takes a number of time steps, since the interaction is mediated only via nearest neighbors and not of the mean-field type (all-to-all coupling).

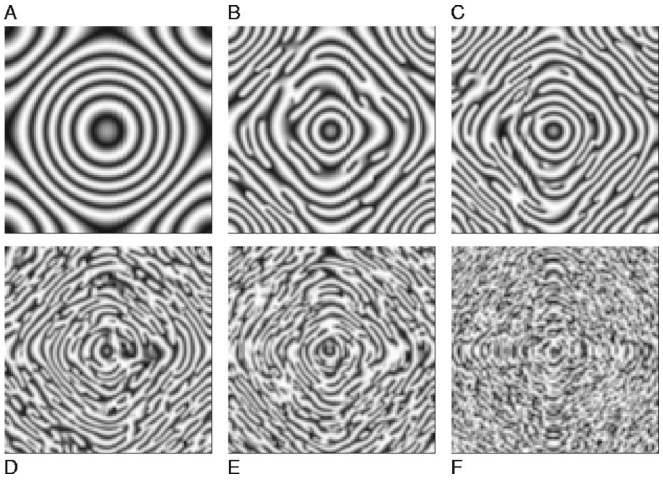


FIG. 7. Pattern formation for symmetry breaking parameter $\gamma=0$ on a two-dimensional lattice with $N=100$ oscillators per side. The oscillator s with the maximal natural frequency is placed at the center of the lattice. We plot $\sin\psi_i$ for every oscillator i. The colors vary between white, corresponding to -1, and black, corresponding to 1. The plots are realized after 2×10^4 integration steps and for different values of $\Delta\omega$ [therefore after different time in natural units of $1/\Delta\omega$]: (A) 0.1 [10^2], (B) 0.5 [5×10^2], (C) 1.0 [10^3], (D) 5.0 [5×10^3], (e) 10 [10^4], and (F) 50 [5×10^4]. A cross section through the pattern of (a) would reveal a frequency distribution as in Fig. 4(B).

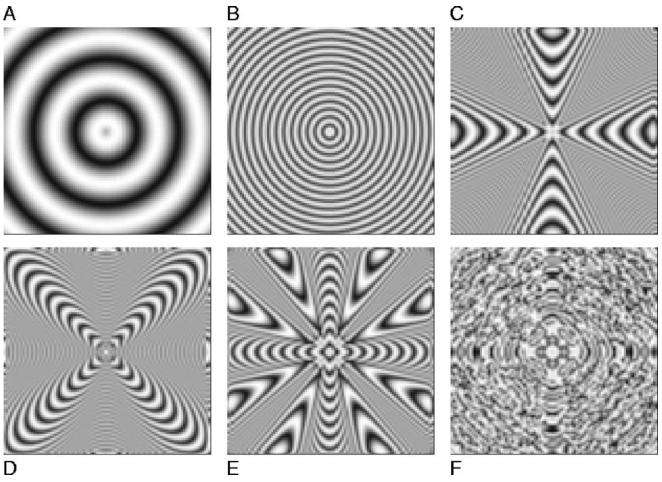


FIG. 8. Time evolution of the "disorder" in Fig. 7(f) for symmetry breaking parameter $\gamma=0$. All plots are realized for $\Delta\omega=50$, after a different number of integration steps [time in natural units of $1/\Delta\omega$]: (A) 9 [22.5], (B) 41 [102.5], (C) 150 [375], (D) 393 [982.5], (E) 490 [1225], and (F) 1995 [4987.5]. Only the pattern of (F) is stationary on a global scale, i.e., up to local fluctuations in the phases.

B. Pattern formation for $\gamma=2$

Next let us consider system (6) for $\gamma=2$, for which we expect pattern formation from the results of [7] due to the broken antisymmetry of sin via a cosine term in the interaction. The main difference shows up in the partial synchronization patterns above the synchronization threshold. As it is evident from Figs. 9 and 10 the size of the one synchronized cluster, filling the whole lattice below the threshold, shrinks with increasing $\Delta\omega$, since the oscillators with the highest natural frequencies decouple from the cluster, but the remaining oscillators do not organize in other synchronized clusters, the cluster-frequency does no longer bifurcate as before. Along with this, the two-dimensional stationary patterns change to those of Fig. 11, whereas the evolution towards Fig. 11(F) is now displayed in Fig. 12 with snapshots taken after 9[22.5] (A), 48[120] (B), 160 [400] (C), 277 [692.5] (D), 386 [965] (E), and 2926 [7315] (F) integration steps [T $1/\Delta\omega$ as above].

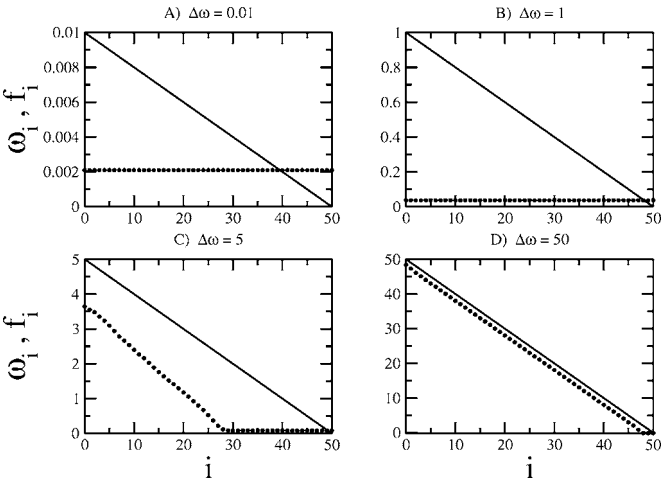


FIG. 9. Natural and measured frequencies for symmetry breaking parameter $\gamma=2$. The frequency of the synchronized cluster decreases and the cluster size shrinks, until all oscillators keep their natural frequency in a completely desynchronized state.

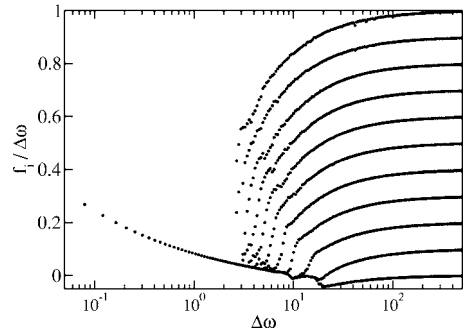


FIG. 10. Bifurcation in cluster-frequency for symmetry breaking parameter $\gamma=2$. Above the synchronization threshold the size of the original synchronized cluster shrinks with $\Delta\omega$. Oscillators outside this cluster stay isolated in contrast to Fig. 5.

A further manifestation of the difference in the partial synchronization patterns is seen in the average frequency f and the variance σ as a function of $\Delta\omega$ (Fig. 13). Above the synchronization transition the average frequency remains constant for $\gamma=0$, but increases for $\gamma=2$ in agreement with Figs. 9 and 10, the variance increases in both cases ($\gamma=0$ and $\gamma=2$) above the transition, so that it may serve as the order parameter. Therefore we can tune the synchronization features via the slope of the gradient.

The main qualitative feature, however, is in common to both systems with and without antisymmetric coupling: the oscillator with the highest natural frequency becomes the center of outgoing circular waves, it is dynamically established as the pacemaker. The patterns, seen here in the case of full and partial synchronization, are quite similar to those predicted by Kuramoto [1] for reaction-diffusion systems

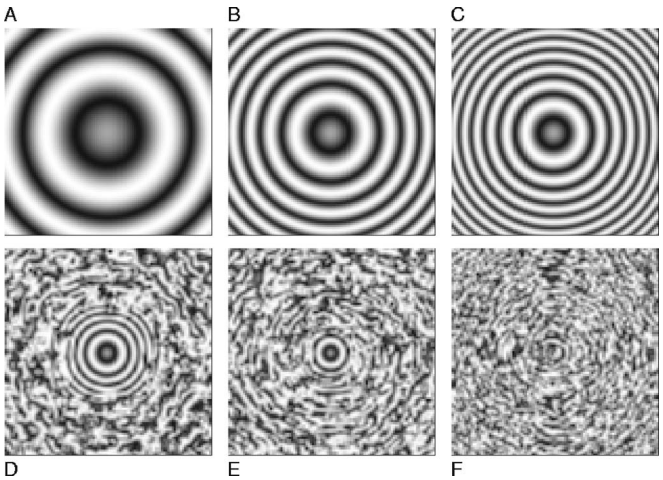


FIG. 11. Pattern formation for symmetry breaking parameter $\gamma=2$. We use the same parameter choice as in Fig. 7. Now (A), (B), and (C) correspond to full synchronization, while one-dimensional cross sections through (D), (E), and (F) would reveal distributions of f_i as in Figs. 9(C) and 9(D).

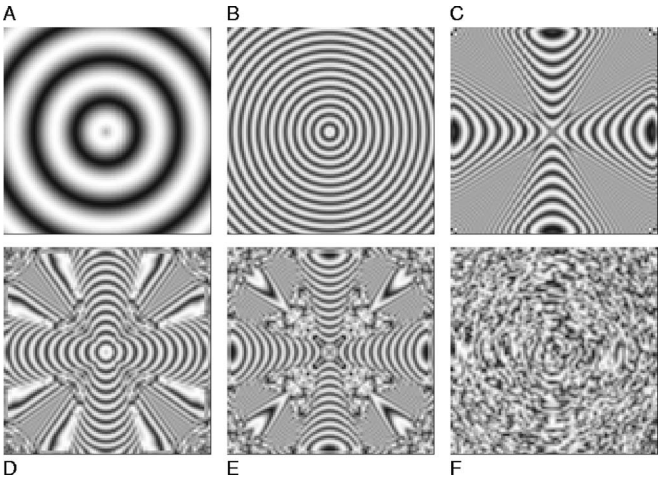


FIG. 12. Time evolution for symmetry breaking parameter $\gamma=2$, otherwise the same as in Fig. 8. This time the plots are realized after 9 [22.5] (A), 48 [120] (B), 160 [400] (C), 277 [692.5] (D), 386 [965] (E), and 2926 [7315] (F) time steps [time in units of $1/\Delta\omega$]. Only the disordered pattern of (F) is stationary up to local fluctuations in the phases.

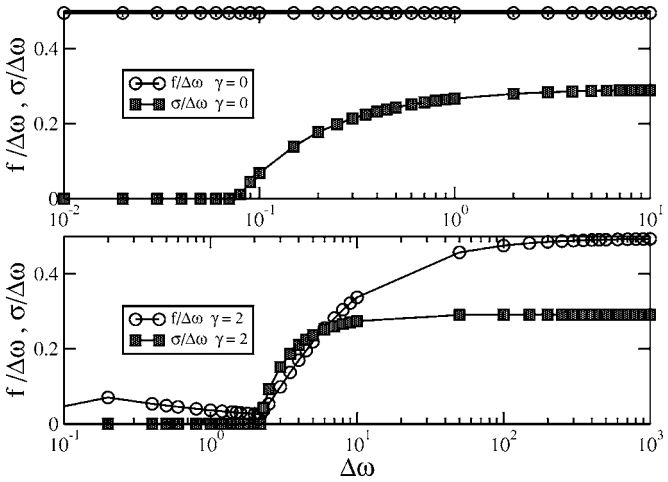


FIG. 13. Average frequency and variance as defined in the text (both divided by $\Delta\omega$) as a function of the maximal natural frequency $\Delta\omega$, in the case of a one-dimensional lattice with $N=100$ oscillators and for values of the symmetry breaking parameter $\gamma=0$ and $\gamma=2$. The raise in the variance from zero indicates the transition to the desynchronized phase.

with continuous diffusion terms, and to those experimentally observed in chemical oscillatory systems [4].

IV. SUMMARY AND CONCLUSIONS

For pacemakers implemented as defect we have extended former results on the entrainment frequency and the entrainment window to arbitrary interaction range, analytically for $\alpha=0$, numerically for intermediate $0<\alpha<\infty$. For large dimensions the entrainment window decays exponentially with the average distance of nodes from the pacemaker so that only shallow networks allow entrainment. The synchronization transition is reentrant as a function of α (or k, the number of neighbors on a ring, or p, the probability to add a random shortcut to the ring topology). The entrainment gets most difficult for $\alpha\sim 3$ and large N, while it is most easily achieved for next-neighbor and all-to-all couplings. This reentrance is easily explained in terms of the normalization of the coupling strength. For the same system without a pacemaker, but with a gradient in the initial natural frequency distribution, the oscillator with the highest natural frequency becomes the center of circular waves, its role as a pacemaker is dynamically induced without the need for an asymmetric term in the interaction. Here we analytically determined the synchronization transition as a function of the gradient slope. Above some threshold, full synchronization on one- or two-dimensional lattices is lost and replaced by partial synchronization patterns. These patterns depend on the asymmetry parameter γ. For $\gamma=0$ we observe a bifurcation in frequency space, for $\gamma=2$ the one synchronized cluster shrinks in its size, before for even steeper slopes synchronization is completely lost. For artificial networks these results may be used to optimize the placement and the number of pacemakers if full synchronization is needed, or to control synchronization by tuning the slope of natural frequency gradients.

[1] Y. Kuramoto, *Chemical Oscillators, Waves, and Turbulence* (Springer, New York, 1984).
[2] A. T. Winfree, *The Geometry of Biological Time* (Springer-Verlag, New York, 1980); J. Buck, Nature (London) **211**, 562 (1966); T. J. Walker, Science **166**, 891 (1969); I. Kanter, W. Kinzel, and E. Kanter, Europhys. Lett. **57**, 141 (2002); B. Blasius, A. Huppert, and L. Stone, Nature (London) **399**, 354 (1999); J. A. Acebrón, L. L. Bonilla, C. J. Pérez-Vicente, F. Ritort, and R. Spigler, Rev. Mod. Phys. **77**, 137 (2005).
[3] Y. Kuramoto and H. Nakao, Physica D **103**, 294 (1997); Y. Kuramoto, D. Battogtokh, and H. Nakao, Phys. Rev. Lett. **81**, 3543 (1998).
[4] A. N. Zaikin and A. M. Zhabotinsky, Nature (London) **225**, 535 (1970).
[5] F. Radicchi and H. Meyer-Ortmanns, Phys. Rev. E **73**, 036218 (2006).
[6] J. Rogers and L. T. Wille, Phys. Rev. E **54**, R2193 (1996); M. Marodi, F. d'Ovidio, and T. Vicsek, *ibid.* **66**, 011109 (2002).
[7] B. Blasius and R. Tönjes, Phys. Rev. Lett. **95**, 084101 (2005); Y. Kuramoto and T. Yamada, Prog. Theor. Phys. **56**, 724 (1976).
[8] H. Kori and A. S. Mikhailov, Phys. Rev. Lett. **93**, 254101 (2004).
[9] A. van Otterlo, K. H. Wagenblast, R. Fazio, and G. Schön, Phys. Rev. B **48**, 3316 (1993).
[10] F. Castro, A. D. Sanchez, and H. S. Wio, Phys. Rev. Lett. **75**,

1691 (1995).

[11] C. Anteneodo, S. E. de S. Pinto, A. M. Batista, and R. L. Viana, Phys. Rev. E **68**, 045202 (2003); C. Anteneodo, S. E. de S. Pinto, A. M. Batista, and R. L. Viana, Phys. Rev. E **69**, 029904(R) (2004).

[12] A. P. Kampf and G. Schön, Physica B **152**, 239 (1988); K. B. Efetov, Zh. Eksp. Teor. Fiz. **78**, 2017 (1980).

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

[14] R. Monasson, Eur. Phys. J. B **12**, 555 (1999); M. E. J. Newman and D. J. Watts, Phys. Lett. A **263**, 341 (1999).

Chapter 5

Evolutionary Algorithms on Social Complex Networks

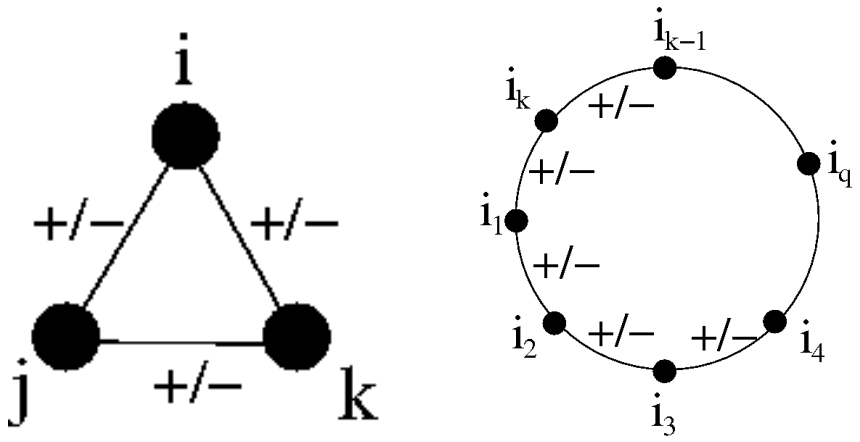
Introduction

During the recent years, statistical physicists extended their interest to problems apparently far from the standard topics of physics. Physicists have a propensity to generate models that are able to capture the basics ingredients of phenomena, independently of the "nature" of these phenomena, and statistical physics provides the proper mathematical tools for an accurate analysis of these models. In this context, the research about social phenomena gave some very interesting results. A multitude of recent papers shows how much statistical physics can be useful for understanding typical social phenomena: opinion formation, cultural assimilation and dissemination, social segregation, language formation and grammar construction are just a few examples.
In this chapter, I will discuss a particular social problem: the problem of social balance. The reader will see how social balance involves not only social scientists [section 5.1.1], but also mathematicians [section 5.1.2] and physicists [section 5.1.3]. Moreover, the reader will see in section 5.2 that social balance can be considered in an analogous way to a particular formulation of satisfiability problems of computer science. In this very interdisciplinary spirit, I conclude the chapter by reporting my own contributions in **Paper IV** [124] and **Paper V** [123] .

5.1 A Common Problem in Social Science, Mathematics and Physics

5.1.1 Social Balance in Social Science

The notion of **social balance** was introduced during the 1940s by the sociologist F.Heider [57, 58]. In its basic formulation, a social unit (also called triad) is composed of three persons or agents [see Figure 5.1(a)]. Each person knows the other two persons. Moreover there is a binary social relationship between every couple of persons: they can like or dislike each other. In social terms, this binary interpersonal relationship between two agents can easily be interpreted in several meaningful ways: the agents are friends or enemies, the agents have the same acquaintance or not, and so on. In a triad there are three binary relationships. Based on a well established principle in social science [58], Heider formulates its simple definition of social balance: a unit of three agents is called *balanced* if all three persons are friends or if two persons having a common enemy are friends, while the triad is called *imbalanced* otherwise.



(a) Heider's triad (b) k-cycle

Figure 5.1: **(a)** Basic unit in the social balance theory of F. Heider [57, 58]. The agents (i, j and k) are represented as dots, while the binary (here for simplicity $+/-$) interpersonal relationships as lines connecting pairs of agents. **(b)** Generic representation of a k-cycle composed of the nodes (persons) $i_1, i_2, i_3, \ldots, i_q, \ldots, i_{k-1}, i_k$.

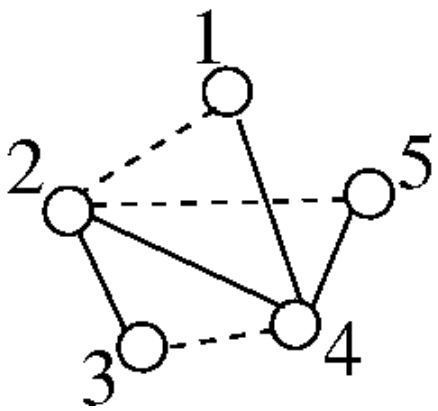
While the definition of balanced and imbalanced triads appear to be reasonable, the binary representation of a social relationship is evidently an oversimplification of reality. In real social

interactions, one person may not simply like or dislike another person. In general, one person should consider several aspects and characteristics of the other person: he likes some of them, while he dislikes others. Therefore, one can not simply say that two persons like or dislike each other, but one should introduce a sort of degree of liking between them. Moreover, in Heider's theory a relationship between two persons is completely symmetric, while sometimes in reality a social interaction can be asymmetric: a first person likes a second person, but the second person can dislike the first person or can be just indifferent to the first person. Despite the oversimplification (so far described by its most evident features of simplification) of the model, the Heider's triad will be the starting point in the following. In fact, this simple model applied to more complex social networks than a simple triad nevertheless leads to interesting and non-trivial aspects.

5.1.2 Social Balance in Mathematics

Further extensions followed after the seminal work of Heider, mostly due to the studies of mathematicians interested in social problems. The main mathematical results about social balance constitute a small branch of graph theory called signed graph theory and they are produced mainly by the works of D.Cartwright, F.Harary and R.Z.Norman [29, 56, 55] and F.S.Roberts [133]. From a mathematical point of view, a social network can be described as a graph [see section 1.2]: a set of nodes (or vertices), each representing a person, and a set of unordered pairs of nodes (also called as links or edges), each representing a social relationship. The triad of Heider [see Figure 5.1(a)] can therefore be viewed as a graph with three nodes (i,j and k) and three edges [(i,j), (j,k) and (k,i)]. As already stated in chapter 1, the simplest way of describing a graph is to use its adjacency matrix A. This is a square matrix, where the number of columns equals the rows and both are equal to the total number of nodes of the graph, and the generic element $A_{i,j} = 1$ if there exists a connection between the nodes i and j, while $A_{i,j} = 0$ in case this connection does not exist. In this framework, it is convenient to store the binary representation of social relationships in the adjacency matrix as well [see Figure 5.2], so that considering the social network as a weighted graph where only two weights, namely the signs, are allowed [see section 1.2.1]. The generic element $A_{i,j}$ standing for a social relationship between the agents i and j can be: $A_{i,j} = 1$ if i and j are friends; $A_{i,j} = -1$ if i and j are enemies; or $A_{i,j} = 0$ if i and j are just indifferent or if they do not know each other. The use of the adjacency matrix A leads to a more formal definition of social balance for an Heider's triad. Defining the sign $S_{i,j,k}$ of a triad, composed of the nodes i, j and k, as product

$$S_{i,j,k} = A_{i,j}A_{j,k}A_{k,i}$$



(a) Social network

	1	2	3	4	5
1	0	−1	0	1	0
2	−1	0	1	1	−1
3	0	1	0	−1	0
4	1	1	−1	0	1
5	0	−1	0	1	0

(b) Adjacency matrix

Figure 5.2: **(a)**Sample social network and **(b)** its adjacency matrix. The negative links are represented as dashed lines, while positive links are full lines. In (b) the bold small indices refer to the same indices of the nodes used in (a).

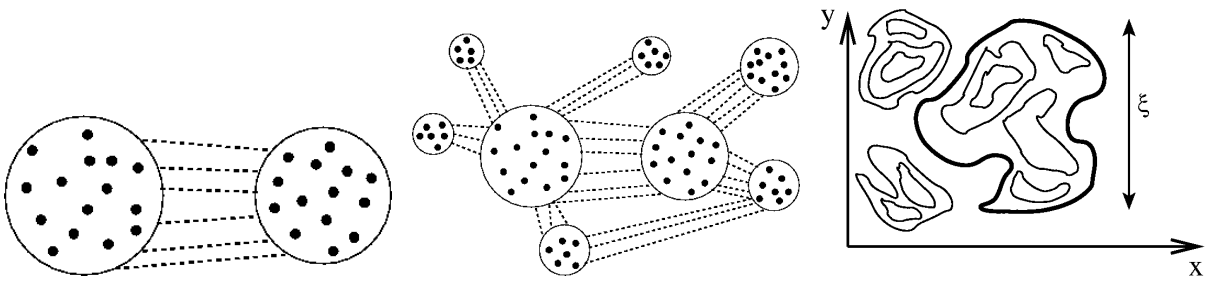
one can say: if $S_{i,j,k} = 1$ the triad is balanced, otherwise if $S_{i,j,k} = -1$ the triad is imbalanced. The mathematical description of social networks as graphs leads to a simple extension of Social Balance not only restricted to a simple triad, but extended to general k-cycles [29]. A k-cycle is defined as a close path between k nodes i_1, i_2, ..., i_q, ..., i_k, where the existence of the path is determined by the conditions $A_{i_1,i_2} \neq 0$, ..., $A_{i_{q-1},i_q} \neq 0$, $A_{i_q,i_{q+1}} \neq 0$ [refer to definition given in section 1.2 and see Figure 5.1(b)]. The sign of a k-cycle $i_1, \ldots, i_k$ is defined as

$$S_{i_1,\ldots,i_k} = \prod_{v=1}^{k} A_{i_v,i_{v+1}} \quad , \tag{5.1}$$

with $i_{k+1} \equiv i_1$. Therefore the k-cycle $i_1, \ldots, i_k$ is balanced if $S_{i_1,\ldots,i_k} = 1$, while it is imbalanced if $S_{i_1,\ldots,i_k} = -1$.

A generic network will be composed of several k-cycles and one will say that the network itself is balanced if and only if all the k-cycles belonging to the network are balanced [29]. Therefore if only one imbalanced k-cycle belongs to the network, the network is called imbalanced. Since every k-cycle of a network should be balanced in a balanced configuration of the network itself, the number of negative links in each k-cycle should be an even number. Depending on the network topology, this observation leads naturally to the formulation of the possible balanced configuration of a network. In the case of fully connected topology [Figure 5.3(a)], the only possible balanced configurations are given by bipartitions of the network in two "cliques": all the nodes belonging to the same clique are friends, while two nodes belonging to two different

cliques are enemies (*Structure theorem* [29, 133]). In the case of diluted random networks [Figure 5.3(b)], the possibility that some pairs of nodes are not connected leads to more elastic balanced configurations. A balanced configuration can be formed by several cliques such that only hyper-cycles[1] of even length are allowed between cliques. On two-dimensional regular topologies [Figure 5.3(c)], the space should be decomposed in different regions of local friendships bounded by closed paths of unfriendly relationships: a k-cycle is by definition a closed path and evidently it should cross an even number of times a boundary of negative links in order to be balanced. Since the former feature should be satisfied by every k-cycle in every region of the space, it naturally follows that in a balanced configuration each negative link should belong to only one closed path of negative links: this means that several closed paths of negative links can simultaneously be present, but they cannot have any intersections. In particular, in Figure 5.3(c) I draw with the thickest line the biggest cluster of connected negative links and I remark on the meaning of its typical length ξ [see section 2.2.2]: both are fundamental quantities in the numerical measurements perfomed in my **Paper V** [123] .



(a) All-to-all network (b) Diluted network (c) Two-dimensional lattice

Figure 5.3: Possible balanced configurations for different network topologies. For more detailed descriptions see the text.

5.1.3 Social Balance in Physics

Physicists have recently focused their attention on the dynamical features of social balance. The interest of physicists in social balance is more recent and, at the moment, restricted to the work of T.Antal *et al.* [9] and F.Radicchi *et al.* (**Paper IV** [124] and **Paper V** [123]). The study of social balance is not only devoted to the identification of the possible static balanced configurations but also to the dynamical aspects of social balance. Physicists stress

[1] One indicates with "hyper-cycle" a cycle between different cliques, or "hyper-nodes", of the social network.

their attention to the way in which a network can dynamically reach a balanced configuration, since real networks of social interactions evolve in time and they are not only static systems. In [9] a dynamical algorithm is introduced as a typical dynamical process towards social balance. This algorithm belongs to the class of Monte-Carlo algorithms, as they are quite often used in different physical contexts [106]. This algorithm is then generalized in **Paper IV** [124] . A propensity parameter $0 \leq p \leq 1$ is introduced with the meaning of the "positive propensity" that two persons in the network make friendship. The value of p as well the value of k, for the length of cycles considered in the dynamical process, are fixed *a priori*. At each step of the algorithm a k-cycle is selected randomly. If it is balanced nothing happens. If the selected k-cycle is imbalanced, then one of its links changes its value and automatically the selected k-cycle becomes balanced as well as the other imbalanced k-cycles, sharing the same link with the selected k-cycle. The other balanced k-cycles, however, sharing the same flipped link, become simultaneously imbalanced. The way in which one selects the link that should be flipped depends on the type of the selected imbalanced k-cycle and/or on the propensity parameter p. In particular, if the selected k-cycle contains both negative and positive links, then a random number $0 \leq r \leq 1$ is chosen out of a uniform distribution: if $r \leq p$, one of the negative links belonging to the selected k-cycle is selected randomly and flipped to a positive value; if $r > p$, one of the positive links belonging to the selected k-cycle is selected randomly and flipped to a negative value. If the selected imbalanced k-cycle contains only negative links (and this can happen only for odd values of k), one of its negative links is selected randomly and flipped to a positive value. At each selection, the time of the simulation is increased by $1/L$, where L is the total number of links of the network, so that L selections are performed in one time unit. It should be remarked that the fact that a k-cycle is selected does not correspond automatically to a flip of the sign of one link, since this happens only when the selected k-cycle is imbalanced and not when it is balanced.

The former dynamical algorithm allows the study of the social balance in a familiar way for physicists. The introduction of simple order parameters such as the density of positive links ρ (used in [9] and in **Paper IV** [124]) or the probability of having an infinity cluster of connected negative links P_∞ (used in **Paper V** [123]) become natural quantities for the analysis of the phase-structure of social balance [see section 2.1]. Of course the phase structure is studied as a function of the propensity parameter p and this analysis reveals a critical behavior of the system at a critical value p_c. As the reader will see in the following papers the critical value p_c as well as the features of the critical phenomenon itself are dependent on the particular topology of the network analyzed. Moreover, the introduction of a dynamics, and therefore of a time unit, in the problem of social balance also gives the chance to study the system from a dynamical point of view and to measure interesting dynamical observables: the relaxation time needed by the network to reach a balanced configuration or the dynamical critical exponents

of the transition are just a few examples.
In physical terms the notion of social balance becomes equivalent to the one of frustration in spin systems. The sign of each link can be interpreted as a spin value. Moreover, it is quite natural to introduce a cost function

$$H = -\frac{1}{M}\sum_{s=1}^{M} H_s \quad , \tag{5.2}$$

where

$$H_s = \prod_{v=1}^{k} A_{{}^s i_v,\, {}^s i_{v+1}} \quad , \tag{5.3}$$

with ${}^s i_{v+1} \equiv {}^s i_1$, is the local cost function associated with the s-th k-cycle. Effectively Eq.(5.3) is the same as Eq.(5.1) and represents the sign of the s-th k-cycle. The cost function of Eq.(5.2) is then given by the normalized sum of the signs of all the M k-cycles present in the system and represents effectively the degree of *frustration* of the system. It is obvious that balanced (or unfrustrated) configurations represent the absolute minima of the cost function. Moreover, it is natural to consider the algorithm, described so far, as equivalent to a Monte Carlo algorithm. Monte Carlo algorithms are typically used in physical problems in order to explore a cost function's landscape. They are able to find the global or local minima of this cost function. In this connection, it should be noticed that the ability of the algorithm so far proposed to find a global minima resides in the fact that it works effectively as a Monte Carlo algorithm with simulated annealing and so it is able to escape from eventual local (and non-global) minima of the cost function of Eq.(5.2). As already stated, the single flip of the spin associated with a link belonging to a unbalanced k-cycle can increase the degree of frustration of the network: since all the other k-cycles, sharing the same link, automatically change their state, then after the single flip there is the possibility that the total number of frustrated k-cycles of the system increases by this single spin flip. In a different algorithm proposed in [9], the constraint that the degree of frustration of the system can never be increased by a single spin flip leads to a completely different approach to social balance: during the dynamics, the network can also fall into and no longer exit from jamming frustrated configurations that actually represent local minima of the cost function of Eq.(5.2).

5.2 Satisfiability Problems in Computer Science

Satisfiability problems constitute an interesting theoretical clue in computer science. Let me first discuss in an informal way what I mean by a satisfiability problem in computer science

and then restrict my attention to a special class of satisfiability problems known as random k-satisfiability problems.
Every problem on a computer is effectively represented by a main Boolean function or formula. The main Boolean formula can then be composed of several independent Boolean functions. These secondary functions can then be decomposed in several other functions and so on. All the formulas can therefore be represented by a hierarchical structure of formulas nested in other formulas. At the lowest level of the hierarchy there should be the Boolean variables involved in the problem. The possible choices of the variables on the lowest level represent the space of the configurations. A particular configuration is called a solution of the problem if the actual value, dependent on the configuration of variables, of the formulas at all the levels of the hierarchy are such that the main formula is satisfied. General problems can have just one solution or several different solutions or they can have no solutions at all. In the former two cases, the problem is said to be satisfiable while when no solution exists, the problem is clearly said to be unsatisfiable. From a theoretical and practical point of view, computer scientists are first of all interested in understanding when the main problem is satisfiable or not. It should be stressed that the satisfiability of one problem implies the existence of at least one algorithm able to solve the problem. Therefore, a complete analysis of the satisfiability of one problem also includes the study of the computational speed or the so-called *complexity* of the possible algorithms able to find the solution, since in practical problems fast algorithms are clearly preferred to slow ones at parity of ability to find a solution (when a solution exists). Moreover, it should be stressed that the hierarchical description of the problem is made just for convenience. Effectively, the functions presented at the intermediate levels are not needed for knowing whether the problem is satisfied or not for a given configuration of the variables. In principle, one could apply a brute algorithm trying all the possible configurations to see whether the problem is satisfied or not for any given configuration. Of course, this is computationally the most expensive approach for finding a solution, but this approach is in principle applicable. The knowledge of the functions at the intermediate levels of the problem is convenient because one can use it in order to drive an algorithm towards the right direction for finding the solution of the problem, when this solution really exists. A good example of this last observation can be made in terms of one of the most famous problems in computer science: the problem of sorting a determined set of numbers [77]. I do not describe the possible ways to translate this problem in terms of Boolean functions: it should be clear to the reader that every problem is translated by a computer in terms of zeroes and ones and elementary operations between them, that is, in terms of Boolean variables and Boolean functions. In the case of sorting, the problem is always satisfiable, and there is only one sequence with the right order so that there exists only one solution to the problem. A brute approach here means that one has to try all the possible sequences of numbers and

one has to verify for each of them whether the numbers are really sorted or not. In this brute approach one has to try $N!$ different sequences if N is the total number of elements in the initial set. In contrast, using the "intermediate" knowledge that the numbers should be sorted locally in couples, one can construct an algorithm, called *bubble sort*, that takes N^2 elementary operations for finding the solution to the problem. Moreover, using some more tricks, one can realize the possibility of finding the solution using other kind of algorithms (*merge sort*, *quick sort*, etc.) and reduce the complexity of the algorithm to $N \log N$.
From a theoretical point of view it can be interesting to study an apparently simple situation. The attention is restricted to a problem with only one intermediate level of Q functions C_q, with $q = 1, \ldots, Q$. With each clause are associated k Boolean variables, randomly chosen out of B total variables. The former sentence explains why this model is called the k random satisfiability problem (kS). Moreover, in each clause C_q, the assigned variables $x_{i_1}^q$, $x_{i_2}^q$, ..., $x_{i_k}^q$ are connected by simple OR operations ($\bigvee$) and some of these variables can be eventually negated. All the clauses C_q are then connected via logical AND operations ($\bigwedge$) in the main formula F, that actually represents the problem. The problem will be satisfied when a properly chosen configuration of the variables simultaneously satisfies all the clauses C_q and therefore the formula F. The algorithm proposed for finding the solution to this theoretical model is the so-called random walk satisfiability (RWS) algorithm. At each update, one unsatisfied clause C_q is selected at random out of all the possible unsatisfied clauses present in the problem. Then one variable $x_{i^*}^q$ belonging to the selected C_q is chosen at random out of all k variables belonging to C_q. The value of $x_{i^*}^q$ is inverted. At each update the time of the simulation is increased by $1/B$. The algorithm runs until unsatisfied clauses are no longer present in the problem, which means until a solution to the problem is found. The algorithm is based on the knowledge of the effective structure of the problem since one tries to find a solution satisfying one clause per time. The k random satisfiability problem represents an interesting model for physicists since it can be directly mapped to a spin glass model [92, 96, 97, 33].

5.2.1 Connection between Social Balance and Satisfiability Problems

One can draw an interesting connection between satisfiability problems in computer science and social balance if considers a subclass of the kS: the k random XOR satisfiability problem (kXS) [20, 131, 14, 33, 93, 84]. In this case the connection between Boolean variables x_i^q is via the logical operator XOR so that the clauses C_q amount to parity checks of the sum of their k Boolean variables $x_{i_1}^q$,...,$x_{i_k}^q$, such that

$$C_q = x_{i_1}^q + x_{i_2}^q + \ldots + x_{i_k}^q \mod 2 \quad . \tag{5.4}$$

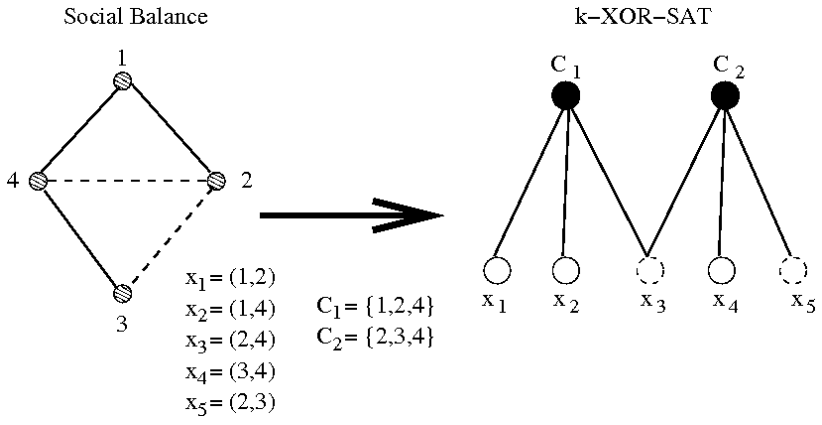


Figure 5.4: Translation of social balance into terms of the k-XOR satisfiability problem. The translation is represented by a one-to-one mapping.

As Eq.(5.4) says, if the number of true variables is an odd number, then $C_q = 1$, otherwise $C_q = 0$. It is simple to realize an analogy between socially balanced networks and the kXS [see Figure 5.4]. Each link in the social network is a variable in the corresponding kXS. Of course a negative value for one link corresponds to a false value for the respective variable, while a positive value for one link corresponds to a true value for the respective variable. In this way, each k-cycle in the social network corresponds to a clause in the kXS. If the sign of one k-cycle [defined in Eq.(5.1)] is positive then the respective clause is satisfied, while a negative sign for a k-cycle means that the corresponding clause is false. The "inverse" translation from a kXS to a social network is not directly true [see Figure 5.5]. When $k-1$ variables simultaneously belong to $k-1$ different clauses, it is not possible to consider the kXS in terms of a social network. In this case, the translation is possible only if multiple edges between nodes in the social network are allowed, otherwise there is no other way to draw the network. Nevertheless, this possibility can be neglected for kXS with a sufficiently large number of Boolean variables B.

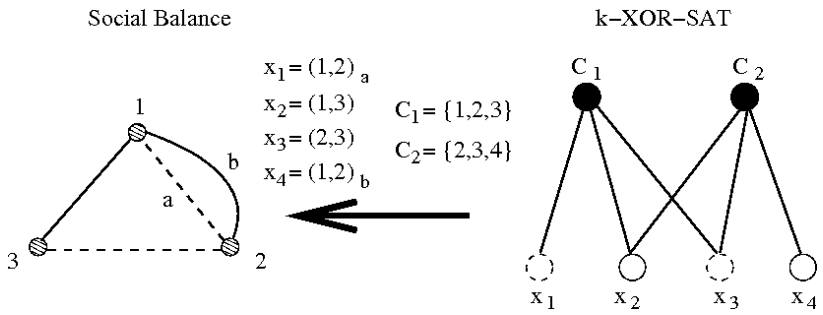


Figure 5.5: Translation of the k-XOR satisfiability problem into terms of social balance. In general, the translation is possible only if multiple edges are allowed in the social network.

The following papers, **Paper IV** [124] and **Paper V** [123] , represent our own contributions to the topic of social balance. We generalize the recent work of T.Antal *et al.* [9] in two main aspects. First we extend their triad dynamics to any cycle of length k and we fully solve the mean-field case. Then we extend the notion of social balance to more general and realistic topologies than the mere all-to-all network.

Paper IV

Social balance as a satisfiability problem of computer science

Filippo Radicchi, Daniele Vilone, Sooyeon Yoon and Hildegard Meyer-Ortmanns

Physical Review E **75**, 026106 (2007)

In this paper we first study social balance and k-cycle dynamics in cases of fully connected networks. Since the mean-field approximation becomes exact for this topology, several analytical predictions can be made. All these predictions are supported by numerical results obtained by computer simulations. In the second part of the paper, the social network is not fully connected anymore but diluted more realistically. This allows us to describe in detail the connection between social balance and k-XOR satisfiability problems, performing a meaningful qualitative and quantitative analysis.

Personal Contribution

I produced all analytical calculations of the following paper apart from those about the size-dependence of the relaxation time. The latter calculations were mainly performed together with Dr.D.Vilone. The analogy between social balance and satisfiability problems of computer science came out of discussions and suggestions of Dr.Martin Weigt. All the numerical results reported in the paper were performed personally by myself. In appendix A.2 I report, for completeness, my mean-field analytical prediction of k-cycle dynamics for the case of $k = 5$ (the results of these calculations are reported in the following paper). Moreover, in appendix B.3 I describe in detail the numerical implementation of k-cycle dynamics.

Social balance as a satisfiability problem of computer science

Filippo Radicchi,[1,*] Daniele Vilone,[1,†] Sooeyon Yoon,[2,‡] and Hildegard Meyer-Ortmanns[1,§]
[1]*School of Engineering and Science, International University Bremen, P.O. Box 750561, D-28725 Bremen, Germany*
[2]*Department of Physics and Research Institute of Basic Sciences, Kyung Hee University, Seoul 130-701, Korea*
(Received 1 August 2006; published 23 February 2007)

Reduction of frustration was the driving force in an approach to social balance as it was recently considered by Antal *et al.* [T. Antal, P. L. Krapivsky, and S. Redner, Phys. Rev. E **72**, 036121 (2005)]. We generalize their triad dynamics to k-cycle dynamics for arbitrary integer k. We derive the phase structure, determine the stationary solutions, and calculate the time it takes to reach a frozen state. The main difference in the phase structure as a function of k is related to k being even or odd. As a second generalization we dilute the all-to-all coupling as considered by Antal *et al.* to a random network with connection probability $w<1$. Interestingly, this model can be mapped to a satisfiability problem of computer science. The phase of social balance in our original interpretation then becomes the phase of satisfaction of all logical clauses in the satisfiability problem. In common to the cases we study, the ideal solution without any frustration always exists, but the question actually is as to whether this solution can be found by means of a local stochastic algorithm within a finite time. The answer depends on the choice of parameters. After establishing the mapping between the two classes of models, we generalize the social-balance problem to a diluted network topology for which the satisfiability problem is usually studied. On the other hand, in connection with the satisfiability problem we generalize the random local algorithm to a p-random local algorithm, including a parameter p that corresponds to the propensity parameter in the social balance problem. The qualitative effect of the inclusion of this parameter is a bias towards the optimal solution and a reduction of the needed simulation time.

DOI: 10.1103/PhysRevE.75.026106 PACS number(s): 89.75.Fb, 02.50.Ey, 05.40.−a

I. INTRODUCTION

Recently Antal *et al.* [1] proposed a triad dynamics to model the approach of social balance. An essential ingredient in the algorithm is the reduction of frustration in the following sense. We assign a value of +1 or −1 to an edge (or link) in the all-to-all topology if it connects two individuals who are friends or enemies, respectively. The sign ±1 of a link we call also its spin. If the product of links along the boundary of a triad is negative, the triad is called frustrated (or imbalanced), otherwise it is called balanced (or unfrustrated). The state of the network is called balanced if all triads are balanced. If the balanced state is achieved by all links being positive the state is called "paradise," as in Antal *et al.* [1]. The algorithm depends on a parameter $p \in [0,1]$ called propensity which determines the tendency of the system to reduce frustration via flipping a negative link to a positive one with probability p or via flipping a positive link to a negative with probability $1-p$. For an all-to-all topology Antal *et al.* predict a transition from imbalanced stationary states for $p<1/2$ to balanced stationary states for $p \geqslant 1/2$. Here the dynamics is motivated by social applications so that the notion of frustration from physics goes along with frustration in the psychological sense.

In the first part of this paper (Sec. II) we generalize the triad dynamics to k-cycle dynamics with arbitrary integer k. In the context of *social balance* theory, Cartwright and Harary [2] introduced the notion of balance describing a balanced state with all k cycles being balanced and k not restricted to three. We first study this model on fully connected networks (Sec. III). For given fixed and integer $k \geqslant 3$ in the updating rules, we derive the differential equations of the time evolution due to the local dynamics (Sec. III A) and predict the stationary densities of k cycles, k arbitrary integer, containing $j \leqslant k$ negative links (Sec. III B). As long as k is odd (Sec. III B 1) in the updating dynamics, the results are only quantitatively different from the case of $k=3$ considered in [1]. An odd cycle of length 3, however, is not an allowed loop in a bipartite graph, for which links may only exist between different types of vertices so that the length of a loop of minimal size in a bipartite graph is 4. In addition, it should be noticed that a 4-cycle with four negative links (that is four individuals each of which dislikes two others) is balanced and not frustrated, although it may be called the "hell" (in contrast to the "paradise"), so it does not need to be updated in order to reduce its frustration. (To call the hell with four negative links "balanced" is not specific for the notion of frustration in physics; also in social balance theory it is the product over links in the loop which matters and which decides about balance or frustration [3].) This difference is essential as compared to the triad dynamics, in which a triad of three unfriendly links is always updated. It has important implications on the phase structure as we shall show. For even values of k and larger than 4, again there are only quantitative differences in the phase structure as compared to $k=4$ (Sec. III B 2).

Beyond the social context, the notion of frustration is familiar from the physics of spin glasses. It is the degree of frustration in spin glasses that determines the qualitative features of the energy landscape. A high [low] degree of frustration corresponds to many [few] local minima in the energy landscape. In terms of energy landscape it was speculated by

*Electronic address: f.radicchi@iu-bremen.de
†Electronic address: d.vilone@iu-bremen.de
‡Electronic address: syyun95@gmail.com
§Electronic address: h.ortmanns@iu-bremen.de

Sasai and Wolynes [4] that it is the low degree of frustration in a genetic network which is responsible for the few stable cell states in the high-dimensional space of states.

Calculational tools from spin-glass theory like the replica method [5] turned out to be quite useful in connection with generic optimization problems (as they occur, for example, in computer science) whenever there is a map between the spin-glass Hamiltonian and a cost function. The goal in finding the ground state energy of the Hamiltonian translates to the minimization of the costs. A particular class of optimization problems refers to satisfiability problems. More specifically one has a system of B Boolean variables and Q logical constraints (clauses) between them. In this case, minimizing the costs means minimizing the number of violated constraints. In case of the existence of a nonviolating configuration the problem is said to be satisfiable, it has a zero ground state energy in the Hamiltonian language. Here it is obvious that computer algorithms designed to find the optimal solution have to reduce the frustration down to a minimal value.

The algorithms we have to deal with belong to the so-called incomplete algorithms [6–8] characterized by some kind of Monte Carlo dynamics that tries to find the solution via stochastic local moves in configuration space, starting from a random initial configuration. It either finds the solution "fast" or never (this will be made more precise below). Among the satisfiability problems there is a special class of models [so-called k-SAT (kS) problems [9–11]], for which actually no frustration-free solution exists if the density of clauses imposed on the system exceeds a certain threshold. In this case the unsatisfiability is not a feature of the algorithm but intrinsic to the problem. However, there is a special case of these problems [so-called k-XOR-SAT (kXS) problems] [7,8,11,12] which are always solvable by some global algorithm. The challenge here is to find the solution by some kind of Monte Carlo dynamics, a local stochastic algorithm within a finite time, scaling at most algebraically with the system size. The algorithm is very similar to the one used for solving the kS problem, where actually no solution may exist.

Now it can be easily shown [10–12] that the satisfiability problem kS (and also the subclass kXS) can be mapped onto a k-spin model that is a spin glass. The kS problem is a real spin glass in the sense that it may not be possible to reduce the degree of frustration to zero, while such a frustration-free solution always exists for the kXS problem. Moreover, as we shall show in the second part of this paper, the kXS problem can be mapped to both the k-spin model and the kXS problem. In particular, the three-spin model can be mapped to the triad dynamics of Antal *et al.* [1] if we choose the propensity parameter $p=1/3$. Therefore, in the context of the satisfiability problems, we will generalize the algorithm to include an additional parameter p that has the meaning of the propensity parameter in the social context. As we shall show, a choice of $p \neq 1/3$ accelerates the finding of the solution, since it provides a bias towards the solution. Furthermore, for the kS problem, and similarly for the kXS problem, the phase structure is usually studied as a function of the dilution of the network, parametrized by α, the ratio of the total number of clauses to the total number of links. The reason is that the interesting changes in the static or dynamic features of the problem solution appear at certain threshold parameters in the dilution, while the all-to-all case is not of particular interest there.

Therefore, as a second generalization, we have to generalize the k-cycle dynamics from its all-to-all topology to the topology of a diluted network. This generalization is anyway natural from the social interpretation. Either two individuals may not know each other at all (this is very likely in case of a large population size) or they neither like nor dislike each other, but are indifferent as emphasized in [2] as an argument for the postulated absence of links. We then discuss the phase structure of the 3XS problem as a function of the dilution α, and the propensity parameter p. The main results are additional absorbing states in the k-cycle dynamics if considered on a diluted topology, and the finding of solutions for $p \neq 1/3$ in cases which looked unsolvable before for $p=1/3$. We remark that we do not generalize the k-cycle dynamics to a true spin glass, for which one may no longer expect a frustration-free solution to exist.

II. MODEL FOR SOCIAL BALANCE

We represent individuals as vertices (or nodes) of a graph and a relationship between two individuals as a link (or edge) that connects the corresponding vertices. Moreover, to a link (i,j) between two nodes i and j we assign a binary spin variable $s_{i,j}=\pm 1$, with $s_{i,j}=1$ if the individuals i and j are friends, and $s_{i,j}=-1$ if i and j are enemies. We consider the standard notion of *social balance* extended to cycles of order k [2,13]. In particular a cycle of order k (or a k cycle) is defined as a closed path between k distinct nodes $i_1,i_2,\ldots,i_k$ of the network, where the path is performed along the links of the network $(i_1,i_2),(i_2,i_3),\ldots,(i_{k-1},i_k),(i_k,i_1)$. Given a value of k we have $k+1$ different types $T_0,T_1,\ldots,T_j,\ldots,T_k$ of cycles of order k containing $0,1,\ldots,j,k$ negative links, respectively. A cycle of order k in the network is considered as balanced if the product of the signs of links along the cycle equals 1, otherwise the cycle is considered as imbalanced or frustrated. Accordingly, the network is considered as balanced if each k cycle of the network is balanced.

We consider our social network as a dynamical system. We perform a local unconstrained dynamics obtained by a natural generalization of the local triads dynamics, recently proposed by Antal *et al.* [1]. We first fix a value of k. Next, at each update we choose at random a k cycle T_j. If this k cycle T_j is balanced (j is even) nothing happens. If T_j is imbalanced (j is odd) we change one of its links as follows: if $j<k$, then $T_j \rightarrow T_{j-1}$ occurs with probability p, while $T_j \rightarrow T_{j+1}$ occurs with probability $1-p$; if $j=k$, then $T_j \rightarrow T_{j-1}$ happens with probability 1. During one update, the positive [negative] link which we flip to take a negative [positive] sign is chosen at random between all the possible positive [negative] links belonging to the k cycle T_j. One unit of time is defined to have passed if the number of updates equals L, where L is the total number of links of the network. In Fig. 1 we show a simple scheme that illustrates the dynamical rules in the case of $k=4(A)$ and $k=5(B)$. It is evident from the figure that for even values of k the system remains the same if we simultaneously flip all the spins $s_{i,j} \rightarrow -s_{i,j}$ $\forall$ (i,j) and make the transformation $p \rightarrow 1-p$. The same is not true for

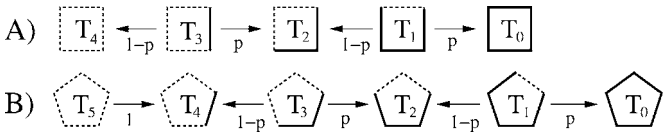


FIG. 1. Dynamical rules in case of $k=4$ (A) and $k=5$ (B). The cycles containing an odd number of "unfriendly" links are considered as imbalanced and evolve into balanced ones. Full and dashed lines represent "friendly" and "unfriendly" links, respectively.

odd values of k. The reason is that a k cycle with only "unfriendly" links is balanced for even values of k, while it is imbalanced for odd values of k. The presence of this symmetry property for even values of k is responsible for very different features in the phase structure as compared to odd values of k. This will be studied in detail in the following sections.

As in [1], for odd values of k, we shall distinguish between stationary states in the infinite volume limit and in the finite volume. In the infinite volume limit, the stationary states can be either balanced (for $p \geqslant 1/2$) or frustrated (for $p<1/2$) since it is not possible to reach the paradise in a finite time. They are predicted as solutions of mean field equations. In the finite volume, such as in numerical simulations, fluctuations about their stationary values do not die out in the phase for $p<1/2$ so that some frustration remains, while for $p \geqslant 1/2$ frozen states are always reached in the form of the paradise. (Other balanced states with a finite amount of negative links are in principle available, but are quite unlikely to be realized during the finite simulation time. They are exponentially suppressed due to their small weight in configuration space.) We calculate the time it takes to reach a frozen state at and above the phase transition (Sec. III C 1). For even values of k we have only two types of stationary frozen states, "paradise" and "hell" with all links being positive and negative, respectively. In this case the time to reach the frozen states at the transition can be calculated in two ways. The first possibility applies for both even and odd values of k and is based on calculating the time it takes until a fluctuation is of the same order in size as the average density of unfriendly links. The second one, applicable to the case of even values of k, can be obtained by mapping the social system to a Markov process known as the Wright-Fisher model for diploid organisms [14], for which the decay time to one of the final configurations (all "positive" or all "negative" genes) increases quadratically in the size N of the system (Sec. III C 2).

III. COMPLETE GRAPHS

We first consider the case of fully connected networks. Later we extend the main results to the case of diluted networks in Sec. IV. In a complete graph every individual has a relationship with everyone else. Let N be the number of nodes of this complete graph. The total number of links of the network is then given by $L=\binom{N}{2}$, while the total number of k cycles is given by $M=\binom{N}{k}$ with $M=\binom{N}{k}$ the binomial coefficient. Moreover we define M_j as the number of k cycles containing j negative links, and $m_j=M_j/M$ the corresponding density of k cycles of type T_j. The total number of positive links L^+ is then related to the number of k cycles according to

$$L^+ = \frac{\sum_{i=0}^{k} (k-i)M_i}{(N-2)!/[(N-k)!(k-2)!]}. \tag{1}$$

A similar relation holds for the total number of negative links L^-,

$$L^- = \frac{\sum_{i=0}^{k} iM_i}{(N-2)!/[(N-k)!(k-2)!]}. \tag{2}$$

In particular, in Eqs. (1) and (2) the numerators give the total number of positive and negative links in all k cycles, respectively, while the same denominator results from the fact that one link belongs to $\binom{N-2}{k-2}=(N-2)!/[(N-k)!(k-2)!]$ different k cycles. Furthermore, the density of positive links is $\rho=L^+/L=1-\sum_{i=0}^{k} im_i$, while the density of negative links is $1-\rho$.

A. Evolution equations

In view of deriving the mean field equations for the unconstrained dynamics, introduced in the former Sec. II, we need to define the quantity M_j^+ as the average number of k cycles of type T_j which are attached to a positive link. This number is given by

$$M_j^+ = \frac{(k-j)M_j}{L^+},$$

while similarly

$$M_j^- = \frac{jM_j}{L^-}$$

counts the average number of k cycles of type T_j attached to a negative link. In term of densities we can easily write

$$m_j^+ = \frac{(k-j)m_j}{\sum_{i=0}^{k} (k-i)m_i} \tag{3}$$

and

$$m_j^- = \frac{jm_j}{\sum_{i=0}^{k} im_i}. \tag{4}$$

Now let π^+ be the probability that a link flips its sign from positive to negative in one update event and π^- the probability that a negative link changes its sign to +1 in one update event. We can express these probabilities in terms of m_j according to

$$\pi^+ = (1-p)\sum_{i=1}^{(k-1)/2} m_{2i-1} \tag{5}$$

and

$$\pi^- = p\sum_{i=1}^{(k-1)/2} m_{2i-1} + m_k, \tag{6}$$

valid for the case of odd k. For even values of k, these probabilities read

$$\pi^+ = (1-p)\sum_{i=1}^{k/2} m_{2i-1} \tag{7}$$

and

$$\pi^- = p\sum_{i=1}^{k/2} m_{2i-1}. \tag{8}$$

Since each update changes $(N-2)!/[(N-k)!(k-2)!]$ cycles, and the number of updates in one time step is equal to L update events, the rate equations in the mean field approximation can be written as

$$\begin{cases} \frac{d}{dt}m_0 = \pi^- m_1^- - \pi^+ m_0^+ \\ \frac{d}{dt}m_1 = \pi^+ m_0^+ + \pi^- m_2^- + - \pi^- m_1^- - \pi^+ m_1^+ \\ \vdots \\ \frac{d}{dt}m_j = \pi^+ m_{j-1}^+ + \pi^- m_{j+1}^- + - \pi^- m_j^- - \pi^+ m_j^+ \\ \vdots \\ \frac{d}{dt}m_{k-1} = \pi^+ m_{k-2}^+ + \pi^- m_k^- + - \pi^- m_{k-1}^- - \pi^+ m_{k-1}^+ \\ \frac{d}{dt}m_k = \pi^- m_{k-1}^- - \pi^- m_k^-. \end{cases} \tag{9}$$

We remark that the only difference between the cases of odd and even values of k comes from Eqs. (5) and (6), and Eqs. (7) and (8), respectively. This difference is the main reason why the two cases of odd and even values of k lead to completely different behavior and why we treat them separately in the following Sec. III B.

B. Stationary states

Next let us derive the stationary states from the rate equations (9) that give a proper description of the unconstrained dynamics of k cycles in a complete graph. Imposing the stationary condition $\frac{d}{dt}m_j=0$, $\forall\ 0\leqslant j\leqslant k$, we easily obtain

$$m_{j-1}^+ = m_j^-, \quad \forall\ 1 \leqslant j \leqslant k. \tag{10}$$

Then, forming products of the former quantities appearing in Eq. (10), we have

$$m_{j-1}^+ m_{j+1}^- = m_j^+ m_j^-, \quad \forall\ 1 \leqslant j \leqslant k$$

and, using the definitions of Eqs. (3) and (4), we finally obtain

$$(k-j+1)(j+1)m_{j-1}m_{j+1} = (k-j)j(m_j)^2, \tag{11}$$

valid $\forall\ 1\leqslant j\leqslant k$. Moreover the normalization condition $\Sigma_i m_i=1$ should be satisfied. Furthermore, in the case of stationarity, the density of friendships should be fixed, so that we should impose $\pi^+=\pi^-$.

1. Case of odd values of k

In the case of odd values of k, the condition for having a fixed density of friendships reads

$$m_k = (1-2p)\sum_{i=1}^{(k-1)/2} m_{2i-1}, \tag{12}$$

where we used Eqs. (5) and (6). In principle the k equations of Eq. (11) plus the normalization condition and the fixed friendship relation (12) determine the stationary solution. For $k=3$ Antal *et al.* [1] found

$$m_j = \binom{3}{j}\rho_\infty^{3-j}(1-\rho_\infty)^j, \quad \forall\ 0\leqslant j\leqslant 3, \tag{13}$$

where

$$\rho_\infty = \begin{cases} 1/[\sqrt{3(1-2p)}+1], & \text{if } p \leqslant 1/2 \\ 1, & \text{if } p \geqslant 1/2 \end{cases} \tag{14}$$

is the stationary density of friendly links. In the same way we can also solve the case $k=5$ exactly with the solution

$$m_j = \binom{5}{j}\rho_\infty^{5-j}(1-\rho_\infty)^j, \quad \forall\ 0\leqslant j\leqslant 5, \tag{15}$$

where

$$\rho_\infty = \left[\sqrt{5(1-2p)\left(1+\sqrt{1+\frac{1}{5(1-2p)}}\right)}+1\right]^{-1} \tag{16}$$

for $p\leqslant 1/2$, while $\rho_\infty=1$ for $p\geqslant 1/2$.

In Fig. 2 we plot the densities m_j given by Eq. (15) and the stationary density of friendly links ρ_∞, given by Eq. (16) as a function of p. Moreover, we verified the validity of the solution performing several numerical simulations on a complete graph with $N=64$ nodes (full dots in the top panel). We compute numerically the average density of positive links after 10^3 time steps, where the average is done over 10^2 different realizations of the system. At the beginning of each realization we select at random the values of the signs of the links, where each of them has the same probability to be positive or negative, so that $\rho_0=0.5$. The numerical results perfectly reproduce our analytical predictions.

As one can easily see, both solutions (13) and (15) are just binomial distributions. This means that the densities of a cycle of order $k=3$ or a cycle of order $k=5$ with j negative links are simply given by the probability of finding these densities on a complete graph in which each link is set equal to 1 with probability ρ_∞ or equal to -1 with probability $1-\rho_\infty$. (As already noticed in [1], this result may come as a surprise, because three-cycle or here the five-cycle dynamics seem to be biased towards the reduction of frustration; on the other hand it is a bias for individual triads without any con-

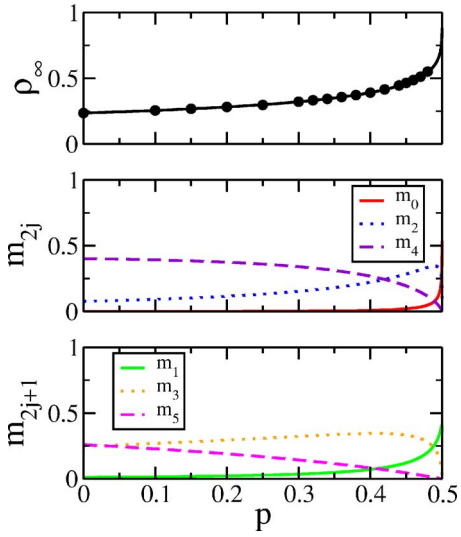


FIG. 2. (Color online) Exact stationary densities m_j for the cycles of order $k=5$ from Eq. (15) (middle and bottom panels, for k cycles with an even and odd number of negative links, respectively) and stationary density of friendly links ρ_∞ from Eq. (16) [top panel], both as a function of the dynamical parameter p. Numerical results are also reported in the top panel for a system with $N=64$ vertices. Each value (full dot) is obtained by averaging the density of friendly links reached after 10^3 time steps over 10^2 different realizations with random initial conditions ($\rho_0=0.5$).

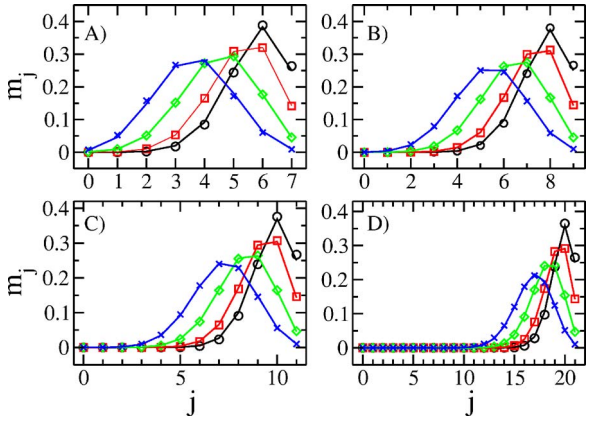


FIG. 3. (Color online) Stationary densities m_j for the k cycles with j negative links and different values of k [$k=7$ (A), $k=9$ (B), $k=11$ (C), and $k=21$ (D)], and for different values of p [$p=0$ (black circles), $p=0.3$ (red squares), $p=0.44$ (green diamonds), and $p=0.49$ (blue crosses)]. The numerical results (symbols) represent the histograms extrapolated from 10^6 samples and over 50 different realizations of the network. In particular the initial values of the spins are equally likely at each realization (so that $\rho_0=0.5$), the distributions are sampled after 5×10^2 time steps and the system size is always $N=64$. The prediction of Eq. (17) is plotted as a full line and the value of ρ_∞ used is taken from the simulations as the average value of the stationary density of positive links.

straint of the type that the frustration of the whole "society" should get reduced.)

For odd values of $k>5$, a stationary solution always exists. This solution becomes harder to find as k increases, because the maximal order of the polynomials involved increases with k (for $k=3$ we have polynomials of first order, for $k=5$ polynomials of second order, for $k=7$ of third order, and so on). So it becomes impossible to find the solution analytically as the maximal order of solvable equations is reached. Nevertheless, we can give an approximate solution using a self-consistent approach as we shall outline in the following. We suppose that the general solution for the stationary densities is of the form

$$m_j = \binom{k}{j} \rho_\infty^{k-j} (1-\rho_\infty)^j, \quad \forall\ 0 \leqslant j \leqslant k, \tag{17}$$

Eq. (17) is an appropriate ansatz as we can directly see from the definition of the density of friendly links $\rho_\infty = 1 - \Sigma_{i=0}^{k} i m_i = 1-(1-\rho_\infty)$, where the last term comes out as mean value of the binomial distribution. (Actually such self-consistency condition is satisfied by any distribution of the m_js with mean value equal to $1-\rho_\infty$.) Moreover, the ansatz for the stationary solution in the form of Eq. (17) has the following features: first it is valid for the special cases $k=3$ and $k=5$, and second, it is numerically supported. In Fig. 3 we show some results obtained by numerical simulations. We plot the densities m_j for different values of k [$k=7$ (A), $k=9$ (B), $k=11$ (C), and $k=21$ (D)] and different values of p [$p=0$ (black circles), $p=0.3$ (red squares), $p=0.44$ (green diamonds), and $p=0.49$ (blue crosses)]. We performed 50 different realizations of a system of $N=64$ vertices, where the densities are extrapolated from 10^6 samples (k cycles) at each realization and after 5×10^2 time steps of the simulations (so that we have reached the stationary state). The initial values of the signs are chosen to be friendly or unfriendly with the same probability ($\rho_0=0.5$). The full lines are given by Eq. (17) for which the right value of ρ_∞ is given by the average stationary density of friendly links and the average is performed over all simulations. Furthermore, we numerically check whether Eq. (17) holds, with the same ρ_∞ if we measure the densities of cycles also of order $k' \neq k$ and, moreover, whether it holds during the time while using the time dependent density of friendly links $\rho(t)$ instead of the stationary one ρ_∞. Since all these checks are positive, we may say that if at some time the distribution of friendly links (and consequently of unfriendly links) is uncorrelated, it will stay so forever.

Let us assume that the ansatz (17) is valid, we then evaluate the unknown value of ρ_∞ self-consistently by imposing the condition that the density of friendly links is fixed at the stationary state

$$\pi^+ = \pi^- \Leftrightarrow (1-2p) \sum_{i=1}^{(k-1)/2} m_{2i-1} = m_k.$$

In particular we can write

$$\sum_{i=1}^{(k-1)/2} m_{2i-1} + m_k = \sum_{i=1}^{(k+1)/2} m_{2i-1} = \xi, \tag{18}$$

and so

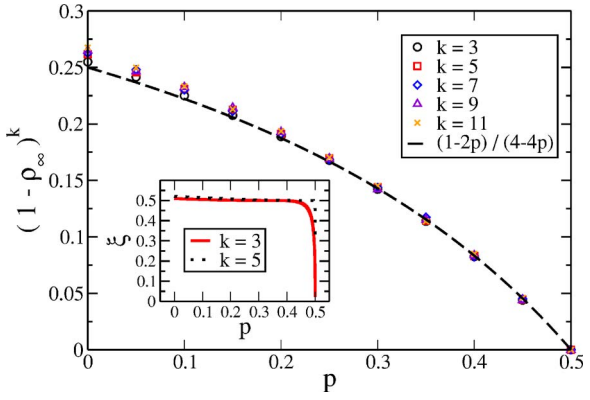


FIG. 4. (Color online) Numerical results (symbols) and approximate solution (dashed line) for the function $(1-\rho_\infty)^k$, depending on the stationary density of positive links ρ_∞ and the parameter k [$k=3$ (black circles), $k=5$ (red squares), $k=7$ (blue diamonds), $k=9$ (violet triangles), $k=11$ (orange crosses)], as a function of the dynamical parameter p. The theoretical result, plotted here as a dashed line, is given by Eq. (19) for $\xi=1/2$. This prediction is in good agreement with the numerical results obtained by averaging the density of friendly links after 10^3 time steps over 10^2 different realizations. The system size is $N=64$. Each simulation starts with random initial conditions ($\rho_0=0.5$). Moreover, as we can see from the inset, the value of ξ calculated for $k=3$ (red full line) and for $k=5$ (black dotted line) is very close to 1/2 for an extended range of p.

$$m_k = (1-2p)(\xi - m_k)$$

from which

$$\rho_\infty = 1 - \left[\frac{\xi(1-2p)}{2(1-p)}\right]^{1/k}, \tag{19}$$

for $p \leqslant 1/2$, while $\rho_\infty=1$ for $p \geqslant 1/2$. In particular we notice that Eq. (19) goes to zero as $k\to\infty$ for $p<1/2$, because $0 \leqslant \xi \leqslant 1$. This means that in the limit of large k the stationary density of friendly links takes the typical shape of a step function centered at $p=1/2$, with $\rho_\infty=0$ for $p<1/2$ and $\rho_\infty=1$ for $p>1/2$. This is exactly the result we find for the case of even values of k (see Sec. III B 2), and it is easily explained since in the limit of large k the distinction between the cases of odd and even values of k should become irrelevant.

Furthermore, it should be noticed that ξ defined in Eq. (18) is nothing else than a sum of all odd terms of a binomial distribution. For large values of k we should expect that the sum of the odd terms is equal to the sum of the even terms of the distribution, so that

$$\xi = \sum_{i=1}^{(k+1)/2} m_{2j-1} \simeq \frac{1}{2} \simeq \sum_{i=0}^{(k-1)/2} m_{2j},$$

because of the normalization. In Fig. 4 we plot the quantity $(1-\rho_\infty)^k$ obtained by numerical simulations for different values of k [$k=3$ (black circles), $k=5$ (red squares), $k=7$ (blue diamonds), $k=9$ (violet triangles), $k=11$ (orange crosses)] as a function of p. Each point represents the average value of the density of positive links (after 10^3 time steps) over 10^2 different realizations. The system size in our simulations is $N=64$, while, at the beginning of each realization, the links have the same probability to have positive or negative spin ($\rho_0=0.5$). From Eq. (19) we expect that the numerical results collapse on the same curve $\xi(1-2p)/(2-2p)$, depending on the parameter ξ. Imposing $\xi=1/2$ (dashed line) we obtain an excellent fit for all values of p. Only for small values of p the fit is less good than for intermediate and large values of p, which is explained by the plot in the inset of Fig. 4. There Eq. (18) is shown as function of p for $k=3$ (black dotted line) and for $k=5$ (red full line). The values of m_j are taken directly from the binomial distribution of Eq. (17) with values of ρ_∞ known exactly from Eqs. (14) and (16) for $k=3$ and $k=5$, respectively. We can see how well the approximation $\xi=1/2$ works already for $k=3$ and how it improves for $k=5$, with the only exception for small values of p where $\xi > 1/2$. Furthermore, we see that $\xi<1/2$ for $p\simeq 1/2$, but in this range the dependence on ξ of Eq. (19) becomes weaker since the factor $\xi(1-2p)$ tends to zero anyway.

2. *Case of even values of k*

The stability of a k cycle with all negative links in the case of even k (see Fig. 1) has far reaching implications on the global behavior of the model. Actually the elementary dynamics is now symmetric. Only the value of p gives a preferential direction (towards a completely friendly or unfriendly cycle) to the basic processes. With odd k, for $p<1/2$ the tendency of the dynamics to reach the state with a minor number of positive links in the elementary processes (without completely unfriendly cycles) is overbalanced by the process $T_k \to T_{k-1}$ which happens with probability 1, so that in the thermodynamical limit the system ends up in an active steady state with a finite average density of negative links due to the competition between the basic processes. Instead, for even k, nothing prevents the system from reaching the "hell," that is a state of only negative links, because here a completely negative cycle is stable. Only for $p=1/2$ we expect to find a nonfrozen fluctuating final state, since in this case the elementary dynamical processes are fully symmetric. Imposing the stationary conditions on the system we do not get detailed information about the final state. As we can see from Eqs. (7) and (8), for $p\neq 1/2$ the only possibility to have $\pi^+=\pi^-$ is the trivial solution for which both probabilities are equal to zero, so that the system must reach a frozen configuration, while for $p=1/2$, π^+ and π^- are always equal, in this case we expect the system to reach immediately an active steady state, cf. Eq. (20) below. (However, in contrast to odd values of k, due to the different time dependence on the size of the system, we shall observe frozen states in a finite volume, either heaven or hell, also in case of $p=1/2$.) In order to describe more precisely the final configuration of this active steady state, it is instructive to consider the mean field equation for the density of positive links. For generic even value of k, it is easy to see that the number of positive links increases in updates of type $T_{2j-1} \to T_{2(j-1)}$ with probability p, whereas it decreases in updates of type $T_{2j-1} \to T_{2j}$ with probability $1-p$, so that the mean field equation that governs the behavior of the density of friendly links is given by

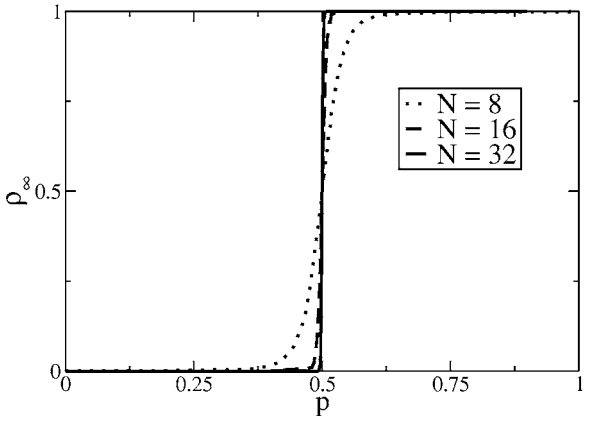


FIG. 5. Behavior of the stationary density of friendly links ρ_∞ as a function of p for three (small) values of N [$N=8$ (dotted line), 16 (dashed line), and 32 (full line)] and for $k=4$. The values of the initial configuration are randomly chosen to be ± 1 with density of friendly links $\rho_0=0.5$. The curves are obtained from averages over 10^3 different realizations.

$$\frac{d\rho}{dt} = (2p-1)\rho(1-\rho)\sum_{i=1}^{k/2}\binom{k}{2i-1}\rho^{k-2i}(1-\rho)^{2(i-1)}. \quad (20)$$

For $p \neq 1/2$ we have only two stationary states, $\rho_\infty=0$ and $\rho_\infty=1$ (the other roots of the steady state equation are complex). It is easily understood that for $p<1/2$ the stable configuration is $\rho_\infty=0$, while for $p>1/2$ it is $\rho_\infty=1$. In contrast, for $p=1/2$ we have $\rho(t)=$const at any time, so that $\rho_\infty=\rho(t=0)=\rho_0$. These results are confirmed by numerical simulations. Moreover, the convergence to the thermodynamical limit is quite fast, as it can be seen in Fig. 5, where we plot the density of friendly links ρ_∞ as a function of p for the system sizes N [$N=8$ (dotted line), $N=16$ (dashed line), and $N=32$ (full line)] and for $k=4$. Each curve is obtained from averages over 10^3 different realizations of the dynamical system. In all simulations the links get initially assigned the values ± 1 with equal probability, so that $\rho_0=0.5$.

C. Frozen configurations

When all k cycles of the network are balanced we say that the network itself is balanced. In particular, in the case of our unconstrained dynamics we can say that if the network is balanced it has reached a frozen configuration. The configuration is frozen in the sense that no dynamics is left since the system cannot escape a balanced configuration. Furthermore, it was proven [2] that if a complete graph is balanced for a fixed odd value of k, then it is balanced for any choice of k and that the only possible balanced configurations are given by bipartitions of the network in two subgroups (or "cliques"), where all the individuals belonging to the same subgroup are friends while every couple of individuals belonging to different subgroups are enemies (this result is also known as *structure theorem* [3]). In the case of even values of k the latter result is still valid if all the individuals of one subgroup are enemies, while two individuals belonging to different subgroups are friends. It should be noticed that one of the two cliques may be empty and therefore the configuration of the paradise (where all individuals are friends) is also included in this result, as well as the hell with all individuals being enemies for the case of even values of k. In the following we will combine our former results about the stationary states (Sec. III B) with the notion of frozen configurations in order to predict the probability of finding a particular balanced configuration and the time needed for freezing our unconstrained dynamical process. For clarity we analyze the cases of odd and even values of k separately, again.

1. Freezing time for odd values of k

$p<1/2$: Let $0 \leqslant N_1 \leqslant N$ be the size of one of the two cliques. Therefore the other clique will be of size $N-N_1$. In such a frozen configuration the total number of positive (L^+) and negative (L^-) links are related to N_1 and N by

$$L^+ = \frac{N_1(N_1-1)}{2} + \frac{(N-N_1)(N-N_1-1)}{2} \quad (21)$$

and

$$L^- = N_1(N-N_1), \quad (22)$$

respectively. As we have seen in the former Sec. III B 1, for odd values of k and $p<1/2$, all k cycles are uncorrelated during the unconstrained dynamical evolution, if we start from an initially uncorrelated configuration. In these cases, we can consider our system as a random process in which the values of the spins are chosen at random with a certain probability. In particular, the probability of a link to be positive is given by ρ, the density of positive links (while $1-\rho$ is the probability for a link to be negative). The probability of reaching a frozen configuration, characterized by two cliques of N_1 nodes and $N-N_1$ nodes, is then given by

$$P(\rho,N_1) = \binom{N}{N_1}\rho^{N(N-1)/2-N_1(N-N_1)}(1-\rho)^{N_1(N-N_1)}. \quad (23)$$

The binomial coefficient $\binom{N}{N_1}$ in Eq. (23) counts the total number of possible bipartitions into cliques with N_1 and $N-N_1$ nodes (i.e., the total number of different ways for choosing N_1 nodes out of N), and each of these bipartitions is considered as equally likely because of the randomness of the process. We should also remark that in Eq. (23) we omit the time dependence of ρ, although the density of positive links ρ follows the following master equation:

$$\frac{d\rho}{dt} = (1-\rho)^k + (2p-1)\sum_{i=1}^{(k-1)/2}\binom{k}{2i-1}\rho^{2i-1}(1-\rho)^{k-2i+1}. \quad (24)$$

Equation (24) follows in analogy to Eq. (20) with the only difference of k being odd. Equation (23) shows that the probability of having a frozen configuration with cliques of N_1 and $N-N_1$ nodes is extremely small, because the number of the other equiprobable configurations with the same number of negative and positive links is equal to $\binom{L}{L^-} \gg \binom{N}{N_1}$, where L^- should satisfy Eq. (22). This allows us to ignore the transient time to reach the stationary state (we expect that the

system goes to the stationary state exponentially fast for any k, as shown in [1] for $k=3$) and consider the probability for obtaining a frozen configurations as

$$P(\rho_\infty) = \sum_{N_1=0}^{N} P(\rho_\infty, N_1). \tag{25}$$

This probability provides a good estimate for the order of magnitude in time τ that is needed to reach a frozen configuration, because $\tau \sim 1/P(\rho_\infty)$. Unfortunately this estimate reveals that the time needed for freezing the system becomes very large already for small sizes N (i.e., τ increases almost exponentially as a function of $L \sim N^2$). This means that it is practically impossible to verify this estimate in numerical simulations.

$p=1/2$: At the transition, for the dynamical parameter $p=1/2$ we can follow the same procedure as used by Antal *et al.* [1]. The procedure is based on calculating the time it takes until a fluctuation in the number of negative links reaches the same order of magnitude as the average number of negative links. In this case the systems happens to reach the frozen configuration of the paradise due to a fluctuation. The number of unfriendly links $L^- \equiv A(t)$ can be written in the canonical form [15]

$$A(t) = La(t) + \sqrt{L}\eta(t), \tag{26}$$

where $a(t)$ is the deterministic part and $\eta(t)$ is a stochastic variable such that $\langle \eta \rangle = 0$. Let us consider the elementary processes

$$A \to \begin{cases} A-1, & \text{rate } M_k \\ A-1, & \text{rate } p \sum_{i=1}^{(k-1)/2} M_{2i-1} \\ A+1, & \text{rate } (1-p) \sum_{i=1}^{(k-1)/2} M_{2i-1} \end{cases} \tag{27}$$

and therefore

$$A^2 \to \begin{cases} A^2-2A+1, & \text{rate } M_k \\ A^2-2A+1, & \text{rate } p \sum_{i=1}^{(k-1)/2} M_{2i-1} \\ A^2+2A+1, & \text{rate } (1-p) \sum_{i=1}^{(k-1)/2} M_{2i-1} \end{cases}. \tag{28}$$

We then obtain the following equations for the mean values of A and A^2:

$$\frac{d\langle A \rangle}{dt} = -\langle M_k \rangle + (1-2p) \sum_{i=1}^{(k-1)/2} \langle M_{2i-1} \rangle$$

and

$$\frac{d\langle A^2 \rangle}{dt} = \langle (1-2A) M_k \rangle + p \left\langle (1-2A) \sum_{i=1}^{(k-1)/2} M_{2i-1} \right\rangle + (1-2p) \left\langle (1+2A) \sum_{i=1}^{(k+1)/2} M_{2i-1} \right\rangle.$$

For $p=1/2$ we find

$$\frac{d\langle A \rangle}{dt} = -\langle M_k \rangle \tag{29}$$

and

$$\frac{d\langle A^2 \rangle}{dt} = \langle M_k \rangle + \sum_{i=1}^{(k-1)/2} \langle M_{2i-1} \rangle - 2\langle A M_k \rangle.$$

Since it is $\langle A \rangle \sim a$ and $\langle M_k \rangle \sim a^k$, we get from Eq. (29)

$$\frac{da}{dt} \sim -a^k, \tag{30}$$

from which

$$a(t) \sim t^{-1/(k-1)}. \tag{31}$$

On the other hand, considering that $d\langle A \rangle^2/dt = 2\langle A \rangle d\langle A \rangle/dt$ and by definition $\sigma = \langle A^2 \rangle - \langle A \rangle^2 = \langle \eta^2 \rangle$, we have

$$\frac{d\sigma}{dt} = \langle M_k \rangle + \sum_{i=1}^{(k-1)/2} \langle M_{2i-1} \rangle - 2(\langle A M_k \rangle - \langle A \rangle \langle M_k \rangle). \tag{32}$$

Moreover, we can write

$$\langle A M_k \rangle - \langle A \rangle \langle M_k \rangle = \langle (La + \sqrt{L}\eta) M_k \rangle - La\langle M_k \rangle = \sqrt{L}\langle \eta M_k \rangle.$$

It is easy to see that $\langle \eta M_k \rangle \sim \langle \eta A^k \rangle = \langle \eta (La + \sqrt{L}\eta)^k \rangle$, so that

$$\langle \eta M_k \rangle \sim \langle \eta (L^k a^k + k L^{k-1/2} a^{k-1} \eta + \cdots + L^{k/2} \eta^k) \rangle \sim k L^{k-1/2} a^{k-1} \langle \eta^2 \rangle + O(\langle \eta^3 \rangle). \tag{33}$$

Dividing Eq. (32) by Eq. (30) and using Eq. (33) we get

$$\frac{d\sigma}{da} = -\left[2k a^{k-1} \sigma - \sum_{i=1}^{(k+1)/2} \binom{k}{2i-1} a^{2i-1} (1-a)^{k-2i+1} \right]. \tag{34}$$

Here we have taken into account that

$$\langle M_j \rangle \sim \binom{k}{j} a^j (1-a)^{k-j}. \tag{35}$$

It is straightforward to find the solution of Eq. (34) as

$$\sigma(a) = C a^{2k} + \frac{\gamma_k}{a} + \cdots + \frac{\gamma_0}{a^{k-2}},$$

with C and γ_j suitable constants. From Eq. (31), for $t \to \infty$ we have

$$\sigma \sim a^{-(k-2)} \sim t^{(k-2)/(k-1)}.$$

For $\eta \sim \sqrt{\sigma}$, we finally obtain

$$\eta \sim t^{(k-2)/[2(k-1)]}.$$

In general, the system will reach the frozen state of the paradise when the fluctuations of the number of negative links

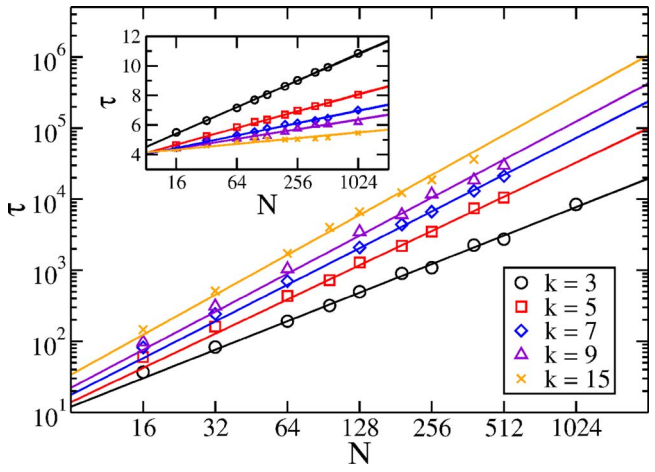


FIG. 6. (Color online) Numerical results (full dots) at $p=1/2$ for the freezing time τ as a function of the system size N and for various k [$k=3$ (black circles), $k=5$ (red squares), $k=7$ (blue diamonds), $k=9$ (violet triangles), and $k=15$ (orange crosses)]. Each point is given by the average value over several realizations (100 realizations for sizes $N\leqslant 64$, 50 realizations for $64<N\leqslant 256$, and 10 realizations for $N>256$). Moreover, as initial configuration of each realization the links are chosen all negative ($\rho_0=0$, antagonistic society) in order to reduce the statistical error (the standard deviation is comparable with the symbol size) caused by the small number of realizations at larger sizes of the system. The full lines have slope $2(k-1)/k$ as expected from Eq. (38). The inset shows the numerical results for the freezing time τ, for different values of k (the same as in the main plot), as a function of the system size N and for $p=3/4$. Each point of the inset is given by the average over 10^3 different realizations with initial antagonistic society.

become of the same order as its mean value. (Note that in this case the mean field approach is no longer valid.) Then, in order of finding the freezing time τ we have just to set equal the two terms on the right hand side of Eq. (26).

$$La(\tau) \sim \sqrt{L}\eta(\tau). \tag{36}$$

Since $L\sim N^2$, we get a power-law behavior

$$\tau \sim N^{\beta} \tag{37}$$

with exponent β as a function of k according to

$$\beta = 2\frac{k-1}{k}. \tag{38}$$

It is worth noticing that in the limit $k\to\infty$ we obtain $\beta=2$, which is the same result as in the case of even values of k as we shall see soon. The analytical results of this subsection are well confirmed by simulations, cf. Fig. 6. There we study numerically the freezing time τ as a function of the system size N for different odd values of k [$k=3$ (black circles), $k=5$ (red squares), $k=7$ (blue diamonds), $k=9$ (violet triangles), and $k=15$ (orange crosses)]. The freezing time is measured until all links have positive sign and paradise is reached. Other frozen configurations are too unlikely to be realized. Each point stands for the average value over a different number of realizations of the dynamical system (100 realizations for sizes $N\leqslant 64$, 50 realizations for $64<N\leqslant 256$, and 10 realizations for $N>256$), where the initial configuration is always chosen as an antagonistic society (all the links being negative so that $\rho_0=0$) to reduce the statistical error. The standard deviations around the averages have sizes comparable to the size of the symbols. The full lines stand for power laws with exponents given by Eq. (38). They perfectly fit with the numerical measurements.

$p>1/2$: For $p>1/2$ the freezing time τ scales as

$$\tau \sim \ln N. \tag{39}$$

We skip the derivation since it would be the same as in the paper of Antal *et al.* [1]. It should be noticed that for $p>1/2$ the paradise is reached the faster the larger k. For simplicity let $p=1$ and imagine that the system is at the closest configuration to the paradise, for which only one link in the system has negative spin. This link belongs to $R=\binom{N-2}{k-2}$ different k cycles. At each update event we select one k cycle at random out of $M=\binom{N}{k}$ total k cycles. This way we have to wait a number of update events $E\sim M/R$ until the paradise is reached, which leads to a freezing time $\tau\sim E/L$, with L the total number of links independent on k, so that

$$\tau \sim \frac{1}{k(k-1)}. \tag{40}$$

For values of $1/2<p<1$ the k dependence of τ should be weaker than the one in Eq. (40), but anyway τ should be a decreasing function of k. The inset of Fig. 6 shows the numerical results obtained for $p=3/4$ as a function of the size of the system N. The freezing time τ is measured for different values of k. We plot the average value over 10^3 different realizations with initial condition $\rho_0=0$.

2. Freezing time for even values of k

$p=1/2$: In the case of even values of k and $p=1/2$ the master equation for the density of positive links [Eq. (20)] reads as $d\rho/dt=0$. Therefore the density of friendly links ρ should be constant during time for an infinite large system. In finite-size systems the dynamics is subjected to non-negligible fluctuations. This allows to understand the scaling features of the freezing time τ with the system size. The order of the fluctuations is $\sqrt{L}$ because the process is completely random as we have seen for the case of odd values of k and $p<1/2$. Differently from the latter case, for even values of k and $p=1/2$ the system has no preferred tendency to go to a fixed point that would be determined by p because $d\rho/dt=0$. In common with Eq. (24) we can view the dynamical system as a Markov chain, here without bias due to p, with discrete steps in time and state space, for which the transition probability for passing from a state with $L^-(t-1)$ negative at time $t-1$ to a state with $L^-(t)$ negative links at time t is given by

$$P[L^-(t)|L^-(t-1)] = \binom{L}{L^-(t)}\left(\frac{L-L^-(t-1)}{L}\right)^{L-L^-(t)}\left(\frac{L^-(t-1)}{L}\right)^{L^-(t)}. \tag{41}$$

Therefore the probability of having $L^-(t)$ negative links at time t is just a binomial distribution where the probability of

having one negative link is given by $\frac{L^-(t-1)}{L}$, the density of negative links at time $t-1$. This includes both the randomness of the displacement of negative links and the absence of a particular fixed point dependent on p. The Markov process, with transition probability given by Eq. (41), is known under the name of the Wright-Fisher model [14] from the context of biology. The Wright-Fisher model is a simple stochastic model for the reproduction of diploid organisms ("diploid" means that each organism has two genes, here named as "−" and "+"), it was proposed independently by Fisher and Wright at the beginning of the 1930s [14]. The population size of genes in an organism is fixed and equal to $L/2$ so that the total number of genes is L. Each organism lives only for one generation and dies after the offsprings are made. Each offspring receives two genes, each one selected with probability 1/2 out of the two genes of a parent of which two are randomly selected from the population of the former generation. Now let us assume that there is a random initial configuration of positive and negative genes with a slight surplus of negative genes. The offspring generation selects its genes randomly from this pool and provides the pool for the next offspring generation. Since the pools get never refreshed by a new random configuration, the initial surplus of negative links gets amplified in each offspring generation until the whole population of genes is "negative." Actually the solution of the Wright-Fisher model is quite simple. The process always converges to a final state with $L^-=0$ $[L^+=L]$ or $L^-=L$ $[L^+=0]$, corresponding to our heaven and [hell] solutions for even values of k. Repeating many times the simulations and starting always from the same density of positive links ρ_0, the fraction of times one ends up in "heaven" in all these trials is just determined by ρ_0 and in "hell" by $1-\rho_0$, so that the number of negative links L^- averaged over different trials equals $(1-\rho_0)L$. Furthermore, on average, the number of negative links decays exponentially fast to one of the two extremal values

$$\langle L^-(t)\rangle \simeq L\begin{cases} e^{-t/L} \\ 1-e^{-t/L} \end{cases}$$

with typical decay time

$$\tau \sim L \sim N^2. \tag{42}$$

This result is perfectly reproduced by the numerical data plotted in Fig. 7. The main plot shows the average time needed to reach a balanced configuration as a function of the size of the system N and for different values of k [$k=4$ (black circles), $k=6$ (red squares), $k=8$ (blue diamonds), and $k=12$ (violet crosses)]. The averages are performed over different numbers of realizations depending on the size N [1000 realizations for sizes $N\leqslant 128$, 500 realizations for $128<N\leqslant 384$ and 50 realizations for $N=384$ and $N=512$ cand 10 realizations for $N=1024$]. The dashed line in Fig. 7 has, in the log-log plane, a slope equal to 2, all numerical data fit very well with this line. Furthermore it should be noticed that there is no k dependence of the freezing time τ, as it is described by Eq. (41). This is reflected by the fact that τ is the same for all the values of k considered in the numerical measurements. Nevertheless, there is a difference between

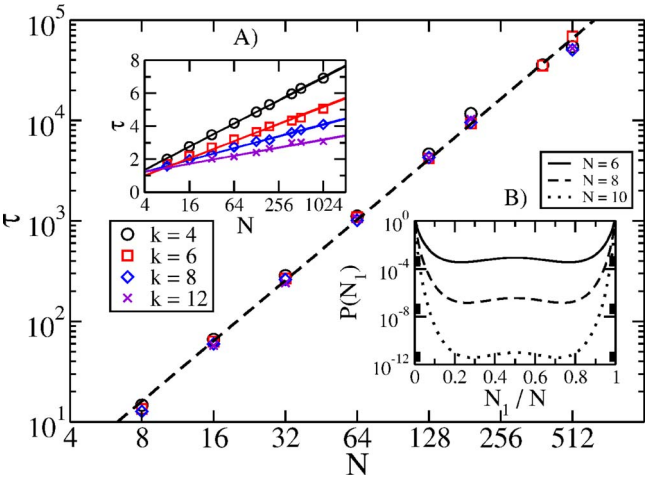


FIG. 7. (Color online) Numerical results for the freezing time τ as a function of the system size N and for various even values of k [$k=4$ (black circles), $k=6$ (red squares), $k=8$ (blue diamonds), and $k=12$ (violet crosses)] and for $p=1/2$. Each point is given by the average value over several realizations [100 realizations for sizes $N\leqslant 64$, 50 realizations for $64<N\leqslant 256$, and 10 realizations for $N>256$]. Moreover, at the beginning of each realization the links are chosen to be positive or negative with the same probability ($\rho_0=0.5$). The dashed line has, in the log-log plane, slope 2 as expected in Eq. (42). The inset (A) shows the numerical results for the freezing time τ, for different values of k (the same as in the main plot), as a function of the system size N and for $p=3/4$. Each point of the inset is given by the average over 10^3 different realizations with random initial conditions. The full lines are all proportional to $\ln N$ as expected. The inset (B) shows the not-normalized probability $P(N_1)$ as a function of the ratio N_1/N and for different values of the system size N [$N=6$ (full line), $N=8$ (dashed line), and $N=10$ (dotted line)]. As one can see, $P(N_1)$ is extremely small for values of $0<N1<N$ already for $N=10$.

our model and the Wright-Fisher model that should be noticed. During the evolution of our model there is the possibility that the system freezes in a configuration different from the paradise ($L^-=0$) or the hell ($L^-=L$). The probability of this event is still given by Eq. (23), with $\rho=L^+(N_1)/L$ as the stationary condition [$L^+(N_1)$ is given by Eq. (21)]. In this way Eq. (23) gives us $P(N_1)$, the non-normalized probability for the system to freeze in a balanced configuration with two cliques of N_1 and $N-N_1$ nodes, respectively. It is straightforward to see that $P(N_1)=1$ for $N_1=0$ or for $N_1=N$, so that the paradise has a nonvanishing probability to be a frozen configuration. Differently for any other value of $0<N_1<N$, $P(N_1)$ decreases to zero faster than $1/N$. This means that for values of N large enough it is appropriate to forget about the intermediate frozen configurations and to consider the features of our model as being very well approximated by those of the Wright-Fisher model. In the inset (B) of Fig. 7 the function $P(N_1)$ is plotted for different values of N [$N=6$ (full line), $N=8$ (dashed line), and $N=10$ (dotted line)] with N_1 a continuous variable for clarity of the figure (we approximate the factorial with Stirling's formula). Obviously $P(N_1)$ goes to zero for $0<N_1<N$ as N increases, already for reasonably small values of N.

The dependence $\tau \sim N^2$ can also be obtained using the same procedure as the one in Sec. III C 1 for the case odd values of k and $p=1/2$. In particular for even values of k we can rewrite Eq. (27) according to

$$A \to \begin{cases} A-1, & \text{rate } p\sum_{i=1}^{k/2} M_{2i-1} \\ A+1, & \text{rate } (1-p)\sum_{i=1}^{k/2} M_{2i-1} \end{cases} \tag{43}$$

and therefore Eq. (28) according to

$$A^2 \to \begin{cases} A^2-2A+1, & \text{rate } p\sum_{i=1}^{k/2} M_{2i-1} \\ A^2+2A+1, & \text{rate } (1-p)\sum_{i=1}^{k/2} M_{2i-1} \end{cases}. \tag{44}$$

For $p=1/2$ we have

$$\frac{d\langle A\rangle}{dt}=0 \tag{45}$$

and

$$\frac{d\langle A^2\rangle}{dt}=\sum_{i=1}^{k/2}\langle M_{2i-1}\rangle.$$

Equation (45) tells us that $a \sim \langle A\rangle = \text{const}$, so that we have

$$\eta \sim \sqrt{t},$$

if we recall Eq. (35). As in the previous case, for determining the freezing time we impose the condition that the average value is of the same order as the fluctuations [Eq. (36)], and, for $L \sim N^2$, we obtain again Eq. (42).

$p \neq 1/2$: For even values of k and for $p \neq 1/2$ the time τ needed for reaching a frozen configuration scales as $\tau \sim \ln N$. In the inset of Fig. 7 numerical estimates of τ for $p=3/4$ and different values of k demonstrate this dependence on the size N of the system. Each point is obtained from averaging over 10^3 different simulations with the same initial conditions $\rho_0=0.5$. Again, as in the case of k odd and $p>1/2$, τ is a decreasing function of k and the same argument used for obtaining Eq. (40) can be applied here, too.

IV. DILUTED NETWORKS

In this section we extend the former results, valid in the case of fully connected networks, to diluted networks. Real networks, apart from very small ones, cannot be represented by complete graphs. The situation in which all individuals know each other is in practice very unlikely. As mentioned in the introduction, links may be also missing, because individuals neither like nor dislike each other but are just indifferent with respect to each other. In the following we analyze the features of dynamical systems, still following the unconstrained k-cycle dynamics, but living on topologies given by diluted networks.

For diluted networks there is an interesting connection to another set of problems usually considered in connection with computer science. It leads to a new interpretation of the social balance problem in terms of a so-called k-SAT (kS) problem (SAT stands for satisfiability) [9–11]. In this type of problem a formula F consists of Q logical clauses $\{C_q\}_{q=1,\ldots,Q}$ which are defined over a set of B Boolean variables $\{x_i=0,1\}_{i=1,\ldots,B}$ which can take two possible values 0 =FALSE or 1=TRUE. Every clause contains k randomly chosen Boolean variables that are connected by logical OR operations ($\vee$). They appear negated with a certain probability. In the formula F, all clauses are connected via logical AND operations ($\wedge$)

$$F=\bigwedge_{q=1}^{Q} C_q,$$

so that all clauses C_q should be simultaneously satisfied in order to satisfy the formula F. A particular formulation of the kS problem is the k-XOR-SAT (kXS) problem [7,8,11,12], in which each clause C_q is a parity check of the kind

$$C_q = x_{i_1}^q + x_{i_2}^q + \cdots + x_{i_k}^q \bmod 2, \tag{46}$$

where $q=1,\ldots,Q$ and $i_1,\ldots,i_k \in \{1,\ldots,B\}$. Parity check is understood in the sense that C_q is TRUE if the total number of true variables which define the clause is odd, while otherwise the clause C_q is FALSE. It is straightforward to map the kXS problem to our former model for the case of odd values of k. Actually, each clause C_q corresponds to a k cycle $[Q \equiv M]$ and each variable x_v to a link (i,j). Furthermore, $[B \equiv L]$ with the correspondence $s_{i,j}=1$ for $x_v=1$, while $s_{i,j}=-1$ for $x_v=0$. For the case of even values of k, one can use the same mapping but consider as clause C_q in Eq. (46) its negation $\overline{C_q}$. In this way, when the number of satisfied variables x_i^q is odd the clause $\overline{C_q}$ is unsatisfied for odd values of k, but satisfied for even values of k.

Moreover, a typical algorithm for finding a solution of the kS problems is the so-called random-walk SAT (RWS). The procedure is the following [7,8]: select one unsatisfied clause C_q randomly, next invert one randomly chosen variable of its k variables $x_{i^*}^q$; repeat this procedure until no unsatisfied clauses are left in the problem. Each update is counted as $1/B$ units of time. As one can easily see, this algorithm is very similar to our unconstrained dynamics apart from two aspects. First, in our unconstrained dynamics we use the dynamical propensity parameter p, which is absent in the RWS. Second, in our unconstrained dynamics we count also the choice of a balanced k cycle as update event, although it does not change the system at all. Because of this reason, the literal application of the original algorithm of unconstrained dynamics is computationally very expensive if it is applied to diluted networks. Apart from the parameter p, we can therefore use the same RWS algorithm for our unconstrained dynamics of k cycles in the diluted case. This algorithm is more reasonable because it selects at each update event only imbalanced k cycles which are actually the only ones that should be updated. In case of an all-to-all topology there are so many triads that a preordering according to the property of being balanced or not is too time consuming so that in this case our former version is more appropriate. In order to count the time as in our original framework of the uncon-

strained dynamics, we should impose that, at the nth update event, the time increases as

$$t_n = t_{n-1} + \frac{1}{L}\frac{\alpha}{\alpha_u^{(n-1)}}. \tag{47}$$

Here $\alpha=M/L$ stands for the ratio of the total number of k cycles of the system (i.e., total number of clauses) to the total number of links (i.e., total number of variables). The parameter α is called the "dilution" parameter, it can take all possible values in the interval $[0,\binom{L}{k}/L]$. $\alpha_u^{(n-1)}=\Sigma_{i=1}^{(k+1)/2}M_{2i-1}/L$ is the ratio of the total number of imbalanced (or "unsatisfied") k cycles over the total number of links, in particular $\alpha_u^{(n-1)}$ is computed before an instant of time at which the nth update event is implemented. Therefore the ratio $\alpha/\alpha_u^{(n-1)}$ gives us the inverse of the probability for finding an imbalanced k cycle, out of all, balanced or imbalanced, k cycles, at the nth update event. This is a good approximation to the time defined in the original unconstrained dynamics. It should be noticed that this algorithm works faster in units of this computational time, but the simulation time should be counted in the same units as defined for the unconstrained dynamics introduced in Sec. II.

The usual performance of the RWS is fully determined by the dilution parameter α. For $\alpha \leqslant \alpha_d$ the RWS always finds a solution of the kS problem within a time that scales linearly with the number of variables L. In particular for the kXS problem $\alpha_d=1/k$. For $\alpha_d<\alpha<\alpha_c$ the RWS is still able to find a solution for the kS problem, but the time needed to find the solution grows exponentially with the number of variables L. For the case of the 3XS problem $\alpha_c \simeq 0.918$. α_d is the value of the dilution parameter for which we have the "dynamical" transition, depending on the dynamics of the algorithm, while α_c represents the transition between the SAT and the UNSAT regions: for values of $\alpha \geqslant \alpha_c$ the RWS is no longer able to find any solution for the kS problem, and in fact no such solution with zero frustration exists for the kS problem (while there still exists a solution for the kXS problem, but it can no longer be found by the RWS algorithm). Furthermore, there is a third critical threshold α_s, with $\alpha_d<\alpha_s<\alpha_c$. For values of $\alpha<\alpha_s$ all solutions of the kS problem found by the RWS are located in a large cluster of solutions and the averaged and normalized Hamming distance inside this cluster is $\langle d\rangle \simeq 1/2$. For $\alpha>\alpha_s$ the solution space splits into a number of small clusters (that grows exponentially with the number of variables L), for which the averaged and normalized Hamming distance inside each cluster is $\langle d\rangle \simeq 0.14$, while the averaged and normalized Hamming distance between two solutions lying in different clusters is still $\langle d\rangle \simeq 1/2$ [12]. For the special case of the 3XS problem α_s was found to be $\alpha_s \simeq 0.818$.

In order to connect the problems of social balance on diluted networks to the kXS problem on a diluted system we shall first translate the parameters into each other. We need to calculate the ratio $\alpha=M/L$ of the total number of k cycles of the network to the total number of links L as a function of w (Sec. IV A). Next we consider the standard RWS applied to the kXS problem taking care on the right way of computing the time as it is given by the rule (47) and the introduction of the dynamical parameter p (Sec. IV B). In particular we focus on the "dynamical" transition at α_d (Sec. IV B 1) and the transition in solution space concerning the clustering properties of the solutions at α_s (Sec. IV B 2). The dynamical parameter p, formerly called the propensity parameter, leads to a critical value p_c above which it is always possible to find a solution within a time that grows at most linearly with the system size. Here we only summarize our results (Sec. IV B 3) without going into detail. We also comment on the validity range of the mean field approximation and translate the phase structure as a function of α back into the phase structure as a function of w.

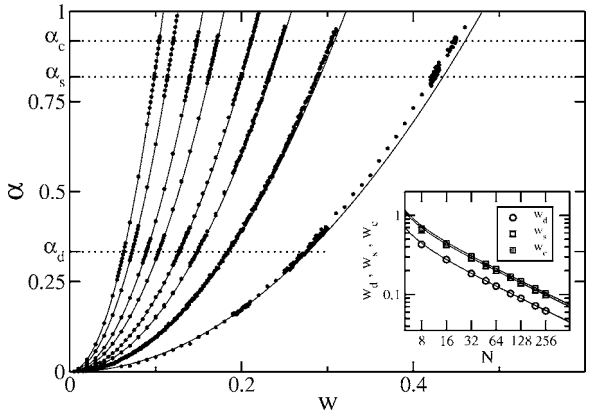


FIG. 8. Numerical results (full dots) for the ratio $\alpha=M/L$ between the total number of cycles M of order $k=3$ and the total number of links L as a function of the probability w for different sizes of Erdös-Rényi networks. In particular the numerical results refer to different network size N: from bottom to top $N=16$, 32, 48, 64, 96, 128, 192, and 256. Each point is given by the average over 10^3 network realizations. The full lines are the predicted values given by Eq. (48), while the dotted lines denote the critical values $\alpha_d=1/3$, $\alpha_s=0.818$, and $\alpha_c=0.918$ as described in detail in the text. In particular the numerical values of the probability w for which these three critical values of α are realized are denoted by w_d (open circles), w_s (open squares), and w_c (gray squares), respectively, they are plotted in the inset, where the full lines are extrapolated by Eq. (48) as $w_i=\sqrt{3\alpha_i/N}$, $i=d,s,c$. The two upper curves for w_s and w_c almost coincide.

A. Ratio α for random networks

Let us first consider Erdös-Rényi networks [16] as a diluted version of the all-to-all topology that we studied before. An Erdös-Rényi network, or a random network, is a network in which each of the $\binom{N}{2}$ different pairs of nodes is connected with probability w. The average number of links is simply $\langle L\rangle=w\binom{N}{2}$. The average number of cycles of order k is given $\langle M\rangle=w^k\binom{N}{k}$, so that the average ratio $\langle\alpha\rangle$ can be estimated as

$$\langle\alpha\rangle \simeq w^{k-1}\frac{2N^{k-2}}{k!}. \tag{48}$$

In Fig. 8 we plot the numerical results obtained for the ratio α as a function of the probability w, in the particular case of cycles of order $k=3$. The reported results, from bottom to

top, have been obtained for values of $N=16$, 32, 48, 64, 96, 128, 192, and 256. Each point is given by the average over 10^3 different network realizations. In particular these numerical results fit very well with the expectations (full lines) of Eq. (48), especially for large values of N and/or small values of w. Furthermore the critical values $\alpha_d=1/3$, $\alpha_s=0.818$, and $\alpha_c=0.918$ (dotted lines) are used for extrapolating the numerical results to w_d (open circles), w_s (open squares), and w_c (gray squares), respectively (see the inset of Fig. 8). w_i, $i=d,s,c$ are the values of the probabilities which correspond to the ratios α_i, $i=d,s,c$, respectively. As expected, they follow the rule $w_i=\sqrt{3}\alpha_i/N$, $i=d,s,c$ predicted by Eq. (48), for $k=3$.

According to the isomorphism traced between the kXS problem and the social balance for k cycles, from now on it is natural to use the following pairs of words synonymously: variable and link, k clause and k cycle, value and sign (or spin), false and negative (or unfriendly), true and positive (or friendly), satisfied and balanced (or unfrustrated), unsatisfied and imbalanced (or frustrated).

B. p-random-walk SAT

So far we have established the connection between the kXS problem and the social balance for k cycles, proposed in this paper. In particular we have determined how the dilution parameter α is related to the parameter w parametrizing (diluted) random networks. In this section we extend the known results for the standard RWS of [7,8] to the p-random-walk SAT (pRWS) algorithm, that is the RWS algorithm extended by the dynamical parameter p that played the role of a propensity parameter in connection with the social balance problem. The steps of the pRWS are as follows:

1. Select randomly a frustrated clause between all frustrated clauses.

2. Instead of randomly inverting the value of anyone of its k variables, as for an update in the case of the RWS, apply the following procedure:

(i) if the clause contains both true and false variables, select a random numer r between 0 and 1. If $r<p$ choose randomly one variable out of the false ones belonging to the clause and flip it. Go then to step 3;

(ii) if the clause contains both true and false variables and $r\geqslant p$ choose randomly one out of the true variables and flip it to the false value, go then to step 3;

(iii) if the clause contains only false values (k should be odd), select with probability 1 one of its false variables, randomly chosen between all the false variables belonging to the clause, and flip it to the true value.

3. Go back to point 1 until no unsatisfied clauses remain in the problem.

The update rules of point 2 are the same used in the case of k cycle dynamics and illustrated in Fig. 1 for the cases $k=4$ (A) and $k=5$ (B). For the special case of the 3XS problem, the standard RWS algorithm and the pRWS algorithm coincide for the dynamical parameter $p=1/3$.

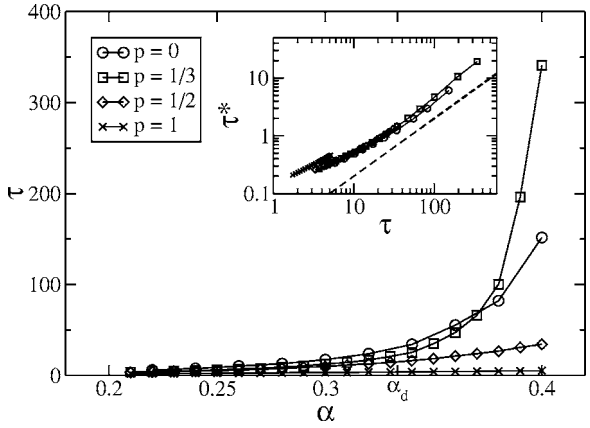


FIG. 9. Time τ for reaching a solution for a system of $L=1000$ variables as a function of the ratio α and for different values of the dynamical parameter p [$p=0$ (circles), $p=1/3$ (squares), $p=1/2$ (diamonds), and $p=1$ (crosses)]. The pRWS performed for $p=1/3$ shows a critical behavior around $\alpha_s=1/3$: for values of $\alpha\leqslant\alpha_s$, τ grows almost linearly with α, while it jumps to an exponential growth with α for $\alpha>\alpha_s$. The same is qualitatively true for $p=0$, but the time τ needed for reaching a solution increases more slowly with respect to the case $p=1/3$ for $\alpha>\alpha_s$. For $p=1/2$ and $p=1$ there seems to be no drastic increment of τ for $\alpha>\alpha_s$. Moreover, the inset shows the dependence of τ^*, the freezing time as calculated in the standard RWS [7,8], on the freezing time τ calculated according to Eq. (47). The almost linear dependence of τ^* on τ (the dashed line has slope 1) explains that there is no qualitative change if we describe the dynamical features of the system in terms of τ or τ^* as time used by the simulations.

1. Dynamical transition at α_d

The freezing time τ, that is the time τ needed for finding a solution of the problem, abruptly changes at the dynamical critical point $\alpha_d=1/k$. Figure 9 reports the numerical estimate of the freezing time τ as a function of the dilution parameter α and for different values of the dynamical parameter p [$p=0$ (circles), $p=1/3$ (squares), $p=1/2$ (diamonds), and $p=1$ (crosses)]. As one can easily see, for $p=1/3$ and $p=0$, τ drastically changes around α_d, increasing abruptly for values of $\alpha>\alpha_d$. For $p=1/2$ and for $p=1$ this drastic change is not observed. This is understandable from the fact that both values of p provide a bias towards paradise, while $p=1/3$ corresponds to a random selection of one of the three links of a triad as in the original RWS and $p=0$ would favor the approach to the hell if it were a balanced state. The simulations are performed over a system with $L=10^3$ variables. Moreover, each point stands for the average over 10^2 different networks and 10^2 different realizations per network of the dynamics on such topologies. At the beginning of each simulation the variables take the values 1 or 0 with the same probability. The inset shows the relation between the time τ^* calculated using the standard RWS and the time τ calculated according to Eq. (47). The almost linear relation (the dashed line has a slope equal to 1) between τ^* and τ means that there is no qualitative change between the two different ways of counting the time.

Following the same argument as in [7], we can specify for the update event at time t the variation of the number of unsatisfied clauses $M_t^{(u)}$ as

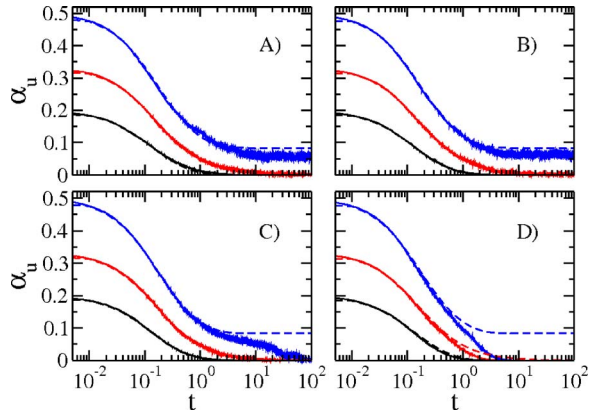


FIG. 10. (Color online) Time behavior of the ratio α_u of unsatisfied clauses for different values of p [(A) $p=0$, (B) $p=1/3$, (C) $p=1/2$, (D) $p=1$] and for different values of the dilution parameter α [$\alpha=0.3$ (black, bottom), $\alpha=0.5$ (red, middle), $\alpha=0.85$ (blue, top)]. Numerical results of simulations (full lines) are compared with the numerical integration of Eq. (49) (dashed lines) leading to a very good fit in all cases, except for $\alpha=0.85$ and for $p=1/2$ and $p=1$. The initial configuration in all the cases is the one of an antagonistic society ($x_i=0$, $\forall\ i=1,\ldots,L$), while the number of variables is $L=10^4$.

$$\Delta M_t^{(u)} = -(k\alpha_u(t)+1) + k\alpha_s(t) = k\alpha - 2k\alpha_u(t) - 1,$$

with α_u the ratio of unsatisfied clauses over L, because, by flipping one variable of an unsatisfied clause, all the other unsatisfied clauses which share the same variable become satisfied, while all the satisfied clauses containing that variable become unsatisfied. In the thermodynamic limit $L\to\infty$, one can use $M_t^{(u)}=L\alpha_u(t)$. Moreover, the amount of time of one update event is given by Eq. (47) so that we can write

$$\dot{\alpha}_u(t) = \frac{\alpha_u(t)}{\alpha}[k\alpha - 2k\alpha_u(t) - 1]. \tag{49}$$

Equation (49) has as stationary state (or a plateau) at

$$\alpha_u = \frac{k\alpha - 1}{2k}. \tag{50}$$

Therefore when the ratio α (that is the ratio of the number of clauses over the number of variables) exceeds the critical "dynamical" value

$$\alpha_d = \frac{1}{k}, \tag{51}$$

the possibility of finding a solution for the problem drastically changes. This result was already found in [7,8]. While for values of $\alpha \leqslant \alpha_d$ we can always find a solution because the plateau of Eq. (50) is always smaller or equal to zero, for $\alpha > \alpha_d$ the solution is reachable only if the system performs a fluctuation large enough to reach zero from the nonzero plateau of Eq. (50). In Fig. 10 we report some numerical simulations for α_u as a function of time for different values of p [(A) $p=0$, (B) $p=1/3$, (C) $p=1/2$, (D) $p=1$], and for different values of the dilution parameter α [$\alpha=0.3$ (black, bottom), $\alpha=0.5$ (red, middle), $\alpha=0.85$ (blue, top)]. The numerical values (full lines) are compared with the numerical integration of Eq. (49) (dashed lines). They fit very well apart from large values of t, for $\alpha=0.85$ and for $p=1/2$ or $p=1$. The initial configuration in all cases is that of an antagonistic society ($x_i=0$, $\forall\ i=1,\ldots,L$), while the number of variables is $L=10^4$.

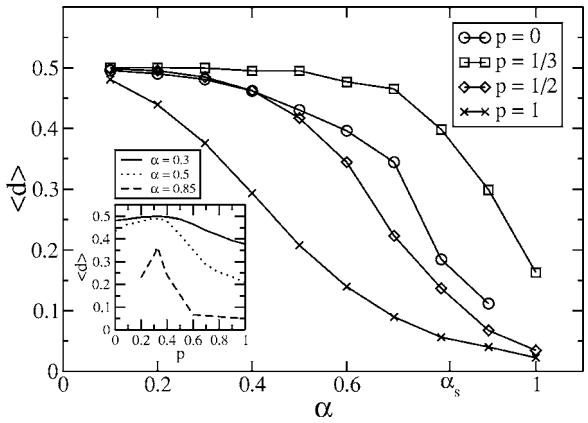


FIG. 11. Normalized Hamming distance $\langle d\rangle$ [Eq. (52)] between solutions as a function of the ratio α and different values of the dynamical parameter p [$p=0$ (circles), $p=1/3$ (squares), $p=1/2$ (diamonds), and $p=1$ (crosses)]. For the standard RWS ($p=1/3$) the distance drops down around the critical point α_s. Different values of p perform not-really-random walks and lead to effective values of α_s smaller than the former one. The inset shows the dependence of $\langle d\rangle$ on the dynamical parameter p. As it is shown for different values of α [$\alpha=0.3$ (full line), $\alpha=0.5$ (dotted line), and $\alpha=0.85$ (dashed line)] the peak of the distance between solutions is for a pRWS which is really random, that is for $p=1/3$. All the points here, in the main plot as well as in the inset, are obtained for a system of $L=20$ variables. Each point is obtained averaging over 10^2 different networks and on each of these networks the average distance is calculated over 10^2 solutions. At the beginning of each simulation the value of one variable is chosen to be 1 or 0 with equal probability.

2. *Clustering of solutions at* α_s

In order to study the transition in the clustering structure of solutions at α_s, we numerically determine the Hamming distance between different solutions of the same problem, starting from different random initial conditions on the same network. More precisely, given a problem of L variables and M clauses, we find T solutions $\{x_i^r\}_{i=1,\ldots,L}^{r=1,\ldots,T}$ of the given problem. This means that we start T times from a random initial configuration and at each time we perform a pRWS until we end up with a solution. We then compute the distance between these T solutions as normalized Hamming distance (normalized over the total number of mutual possible distances)

$$\langle d\rangle = \frac{1}{LT(T-1)}\sum_{r,s=1}^{T}\sum_{i=1}^{L}|x_i^r - x_i^s|. \tag{52}$$

The numerical results for $L=20$ are reported in Fig. 11. We average the distance over $T=10^2$ trials and over 10^2 different problems (i.e., network topologies) for each value of α. As

expected for $p=1/3$ (squares) the distance between solutions drops down around α_s (actually it drops down before α_s because of the small number of variables). For different values of p [$p=0$ (circles), $p=1/2$ (diamonds), and $p=1$ (crosses)], the pRWS is less random and $\langle d\rangle$ drops down before α_s (or at least before the point at which the case of $p=1/3$ drops down). In particular, if we plot (as in the inset) the distance $\langle d\rangle$ as a function of p and for different values of α [$\alpha=0.3$ (full line), $\alpha=0.5$ (dotted line), and $\alpha=0.85$ (dashed line)] we see a clear peak in the distance $\langle d\rangle$ around $p=1/3$. This suggests that a completely random, unbiased RWS always explores a larger region in phase space, it leads to a larger variety of solutions as expected.

3. Further thresholds in α

Differently from the general kS problem, the kXS problem is known to be always solvable [8] and the solution corresponds to one of the balanced configurations as described in Sec. III C for the all-to-all topology. Nevertheless, the challenge here is whether the solutions can be found by a local random algorithm like RWS. In the application of the RWS it can happen that the algorithm is not able to find one of these solutions in a "finite" time, so that the problem is called "unsatisfied." The notion is made more precise in [12]. For practical reasons the way of estimating the critical point α_c that separates the SAT from the UNSAT region is related to the so-called algorithmic complexity of the RWS. A prescription for identifying this threshold was given by [7,8,17]. Here we do not go into any detail, but summarize the results we obtained for studying the analogous question in the vicinity of α_d. In general the ability of the algorithm for finding the solution depends on α, p, L, and the waiting time t_w. As it turned out in our simulations, a choice of $p\neq 1/3$ in the pRWS can strongly improve the performance of the RWS due to the tendency of increasing ($p<1/3$) or decreasing the number ($p>1/3$) of negative links. In particular, as we have seen in the former sections, for $p\geqslant 1/2$ the pRWS approaches the configuration of the paradise for the largest value of $\alpha=\left(\frac{L}{k}\right)/L\gg\alpha_c$ and in a time that scales as $\tau\sim L^{\beta}$, so that there is no UNSAT region at all. If the bias goes in the wrong direction, the performance gets worse, but remains better than in the unbiased case of $p\neq 1/3$.

As a further "threshold" in the dilution parameter we consider the value of α, here called α_m, above which the mean field approximation applies that we used for deriving the results in the all-to-all topology. By construction the "topology" of a kS problem is completely random (for this reason it is sometimes called explicitly the Random k-SAT problem). Each of the L variables can appear in one of the αL clauses with probability $v=\frac{1}{L}+\frac{1}{L-1}+\cdots+\frac{1}{L-k}$. In particular for $L\gg k$ one can simply write $v\simeq\frac{k}{L}$. Then the probability P_r that one variable belongs to r clauses can be described by the Poisson distribution

$$P_r=\frac{(\alpha k)^r}{r!}e^{-\alpha k}, \tag{53}$$

with mean value $\langle r\rangle=\alpha k$ and variance $\sigma_r=\sqrt{\alpha k}$. P_r is plotted in Fig. 12, where the numerical results [symbols, $r=0$ (black circles), $r=1$ (red squares), $r=2$ (blue diamonds), and $r\geqslant 3$ (violet crosses)] are compared to the analytical expectations [lines, $r=0$ (black full line), $r=1$ (red dotted line), $r=2$ (blue dashed line), and $r\geqslant 3$ (violet dotted-dashed line)].

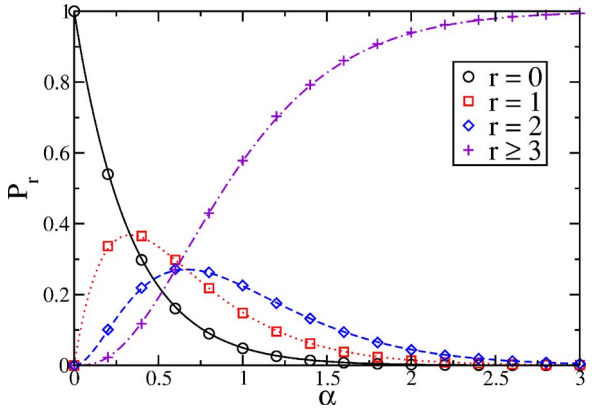


FIG. 12. (Color online) Probability p_r that one variable belongs to r clauses as function of the dilution parameter α. The symbols stand for numerical results obtained over 10^3 different realizations for $L=128$ variables [$r=0$ (black circles), $r=1$ (red squares), $r=2$ (blue diamonds), and $r\geqslant 3$ (violet crosses)]. The lines stand for analytical predictions of Eq. (53) [$r=0$ (black full line), $r=1$ (red dotted line), $r=2$ (blue dashed line), and $r\geqslant 3$ (violet dotted-dashed line)].

If we start from an antagonistic society (all variables false) the minimum value of the dilution α_{min} needed to reach the paradise (if $p\geqslant 1/2$) is that all variables belong to at least one clause. This means that $P_0<1/L$, from which

$$\alpha_{min}=\frac{\ln L}{k}. \tag{54}$$

As we have shown in a comparison of the numerically measured values to the theoretical mean field predictions for the stationary density ${}^{(t)}\rho_{\infty}$ of true variables, α_{min} of Eq. (54) provides already a good estimate for the onset of the mean field regime. Moreover, it should be noticed from Fig. 13 that for $\alpha>\alpha_{min}$ almost all variables belong to at least three clauses. This fact allows the pRWS to explore a larger part of configuration space. Let us assume that one variable belongs to less than three clauses: an eventual update event that flips this variable (so that the one triad becomes balanced) can never increase the number of unsatisfied clauses by frustrating other clauses it belongs to. This reminds us to the situation in an energy landscape in which an algorithm gets stuck in a local minimum when it never accepts a change in the "wrong" direction, i.e., towards higher energy.

4. Effect of the dilution in an Erdös-Rényi network

Let us briefly translate the results we obtained so far for diluted networks as a function of the dilution parameter α, to diluted (social) Erdös-Rényi networks, for which the dilution is parametrized by w, the probability for a randomly chosen pair of nodes to be connected. Here we made the connection explicit only for the case of $k=3$. The mean field description and the results about the phase structure remain valid down

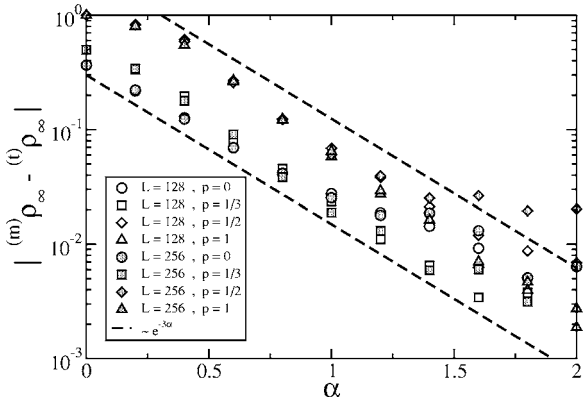


FIG. 13. Difference $|^{(m)}\rho_\infty - {}^{(t)}\rho_\infty|$ between ${}^{(t)}\rho_\infty$ the theoretical prediction for the stationary density of friendly variables [Eq. (14)] and the numerically measured value ${}^{(m)}\rho_\infty$, as a function of the dilution parameter α. ${}^{(m)}\rho_\infty$ is obtained as the average of the density of friendly links (registered after a waiting time $T=200.0$, so that is effectively stationary) over 50 different problems and 50 different pRWS for each problem. The results displayed here are obtained for $L=128$ (open symbols) and $L=256$ (gray filled symbols) and for different values of p [$p=0$ (circles), $p=1/3$ (squares), $p=1/2$ (diamonds), $p=1$ (triangles)]. The initial conditions are those of an antagonistic society. The dashed lines are proportional to $e^{-3\alpha}$.

to a certain degree of dilution, characterized by w_m. This threshold for the validity of the mean field description practically coincides with the criterion whether a single link belongs to at least three triads (for $w>w_m$) or not ($w<w_m$). If it does so, an update event can increase the number of frustrated triads. For $w<w_m$, or more precisely $w<w_d<w_m$ it becomes easier to realize frozen configurations different from the paradise. Isolated links do not get updated at all and isolated triads can freeze to a "+"-"−" configuration. The time to reach such a frozen configuration (in general different from the paradise) grows then only linearly in the system size. Also the solution space, characterized by the average Hamming distance between solutions, has different features below and above another threshold, called w_s with $w_d<w_s<w_m$. Therefore one of the main differences between the all-to-all and the sufficiently diluted topology are the frozen configurations. For the all-to-all case we observed the paradise above p_c for odd and even values of k and the hell for even values of k below p_c, because the probability to find a two-clique-frozen configuration (different from paradise or hell) was calculated to be negligibly small. For larger dilution, also other balanced configurations were numerically found, as mentioned above, and the time passed in the numerical simulations for finding these solutions followed the theoretical predictions.

V. SUMMARY AND CONCLUSIONS

In the first part of this paper we generalized the triad dynamics of Antal *et al.* to a k-cycle dynamics [1]. Here we had to distinguish the cases of even values of k and odd values of k. For all values of integer k there is again a critical threshold at $p_c=1/2$ in the propensity parameter. For odd k and $p<p_c$ the paradise can never be reached in the thermodynamic limit of infinite system size (as predicted by the mean field equations which we solved exactly for $k=5$ and approximately for $k>5$). In the finite volume, in principle one could reach a balanced state made out of two cliques (a special case of this configuration is the "paradise" when one clique is empty). However, the probability for reaching such type of frozen state decreases exponentially with the system size so that in practice the fluctuations never die out in the numerical simulations. For $p>1/2$ the convergence time to reach the paradise grows logarithmically with the system size. At $p=1/2$ paradise is reached within a time that follows a power law in the size N, where we determined the k dependence of the exponent. In particular, the densities of k cycles with j negative links, here evolved according to the rules of the k-cycle dynamics, could be equally well obtained from a random dynamics in which each link is set equal to 1 with probability ρ_∞ or equal to -1 with probability $1-\rho_\infty$. This feature was already observed by Antal *et al.* for $k=3$ [1]. It means that the individual updating rules which seem to be "socially" motivated in locally reducing the social tensions by changing links to friendly ones, end up with random distributions of friendly links. The reason is a missing constraint of the type that the overall number of frustrated k cycles should not increase in an update event. Such a constrained dynamics was studied by Antal *et al.* in [1], but not in this paper.

For even values of k, the only stable solutions are "heaven" (i.e., paradise) and "hell" for $p>1/2$ and $p<1/2$, respectively, and the time to reach these frozen configurations grows logarithmically with N. At $p_c=1/2$ other realizations of the frozen configurations are possible, in principle. However, they have negligible probability as compared to heaven and hell. Here the time to reach these configurations increases quadratically in N, independently of k. This result was obtained in two ways: either from the criterion to reach the stable state when a large enough fluctuation drops the system into this state (so we had to calculate how long one has to wait for such a big fluctuation). Alternatively, the result could be read off from a mapping to a Markov process for diploid organisms, ending up in a genetic pool of either all "+" genes or all "−" genes. The difference in the possible stable states of diploid organisms and ours consists in two-clique stable solutions that are admissible for the even k-cycle dynamics, in principle, however, such clique states have such a low probability of being realized that the difference is irrelevant.

The difference in the exponent at p_c and the stable configurations above and below p_c between the even and odd k-cycle dynamics was due to the fact that "hell," a state with all links negative as in an antagonistic society, is a balanced state for even k, not only by the frustration criterion of physicists, but also according to the criterion of social scientists [2].

As a second natural generalization of the social balance dynamics of Antal *et al.* we considered a diluted network. Here we found a mapping of the k-cycle social balance dynamics of networks to a k-XOR-SAT (kXS) problem, that is a typical satisfiability problem in optimization tasks. We also traced a mapping between the social dynamical rules and the Random-Walk SAT (RWS) algorithm, that is one approach

for solving this problem in a random local way. As we have shown, the diluted version of the three-cycle social dynamics with propensity parameter $p=1/3$ corresponds to a 3XS problem solved by the RWS algorithm in its standard form (as used in [7,8]).

The kXS problem is always solvable like the k-cycle social balance, for which a two-clique solution always exists due to the structure theorem of [2], containing as a special solution the so-called paradise. The common challenge, however, is to find this solution by a local stochastic algorithm. The driving force, shared by both sets of problems, is the reduction of frustration. The meaning of frustration depends on the context: for the k-cycle dynamics it is meant in a social sense as a reduction of social tension, for the kXS problem it corresponds to violated clauses. The mathematical criterion is the same. The local stochastic algorithm works in a certain parameter range, but outside this range it fails. The paradise is never reached for a propensity parameter $p<1/2$, independently of k. Similarly, the solution of the kXS problem is never found if the dilution parameter is larger than α_c, and the RWS algorithm needs an exponentially long time already for $\alpha>\alpha_d$, with $\alpha_d<\alpha_c$.

We generalized the RWS algorithm, usually chosen for solving the kSAT (kS) problem as well as the kXS problem, to include a parameter p that formerly played the role of the propensity parameter in the social dynamics (pRWS). The effect of this parameter is a bias towards the solution so that α_d, the threshold between a linear and an exponential time for solving the problem, becomes a function of p. Problems for which the pRWS algorithm needed exponentially long for $p=1/3$, now become solvable within a time that grows less than logarithmically in the system size for $p>1/2$ and less than powerlike in the system size for $p=1/2$. Along with the bias goes an exploration of solution space that has on average a smaller Hamming distance between different solutions than in the case of the $\frac{1}{3}$RWS algorithm that was formerly considered [7,8].

It would be interesting to generalize the social dynamics to a true spin glass with variables assigned to both nodes and links and to allow for a finite degree of frustration even in the optimal solution.

Our paper has illustrated that the reduction of frustration may be the driving force in common to a number of dynamical systems. So far we were concerned about "artificial" systems like social systems and satisfiability problems. Next one may search for natural networks whose evolution was determined by the goal of reducing the frustration, not necessarily to zero degree, but to a low degree at least.

ACKNOWLEDGMENTS

It is a pleasure to thank Martin Weigt for drawing our attention to random k-SAT problems in computer science and for useful discussions during his visit at the International Center for Transdisciplinary Studies at the International University Bremen.

[1] T. Antal, P. L. Krapivsky, and S. Redner, Phys. Rev. E **72**, 036121 (2005).

[2] D. Cartwright and F. Harary, Psychol. Rev. **63**, 277 (1956); F. Harary, R. Z. Norman, and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs* (John Wiley & Sons, New York, 1965).

[3] F. Harary, Mich. Math. J. **2**, 143 (1953–54); F. S. Roberts, Electronic Notes in Discrete Mathematics (ENDM) **2**, 94 (1999), http://www.elsevier.nl/locate/endm; N. P. Hummon and P. Doreian, Soc. Networks **25**, 17 (2003).

[4] M. Sasai and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **100**, 2374 (2003).

[5] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).

[6] M. R. Garey and D. S. Johnson, *Computer and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979).

[7] W. Barthel, A. K. Hartmann, and M. Weigt, Phys. Rev. E **67**, 066104 (2003).

[8] G. Semerjian and R. Monasson, Phys. Rev. E **67**, 066103 (2003).

[9] S. Cook, in *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 1971), p. 151; J. M. Crawford and L. D. Auton, in *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI-93)* (AAAI Press, Menlo Park, California, 1993), p. 21; B. Selman and S. Kirkpatrick, Science **264**, 1297 (1994).

[10] R. Monasson and R. Zecchina, Phys. Rev. Lett. **76**, 3881 (1996); R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Nature (London) **400**, 133 (1999); G. Biroli, R. Monasson, and M. Weigt, Eur. Phys. J. B **14**, 551 (2000); M. Mézard, G. Parisi, and R. Zecchina, Science **297**, 812 (2002); M. Mezard, T. Mora, and R. Zecchina, Phys. Rev. Lett. **94**, 197205 (2005); A. Braunstein, M. Mezard, and R. Zecchina, Random Struct. Algorithms **27**, 201 (2005).

[11] F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, Phys. Rev. E **63**, 026702 (2001); M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, J. Stat. Phys. **111**, 505 (2003).

[12] S. Cocco, O. Dubois, J. Mandler, and R. Monasson, Phys. Rev. Lett. **90**, 047205 (2003).

[13] F. Heider, Psychol. Rev. **51**, 358 (1944); F. Heider, J. Psychol. **21**, 107 (1946); F. Heider, *The Psychology of Interpersonal Relations* (John Wiley & Sons, New York, 1958); S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge University Press, New York, 1994).

[14] S. Wright, Genetics **16**, 97 (1931); R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, 1930).

[15] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 2005).

[16] P. Erdös and A. Rényi, Acta Math. Acad. Sci. Hung. **5**, 17 (1960); **12**, 261 (1961).

[17] U. Schöning, Algorithmica **32**, 615 (2002).

Paper V

Universality class of triad dynamics on a triangular lattice

Filippo Radicchi, Daniele Vilone and Hildegard Meyer-Ortmanns

Physical Review E **75**, 021118 (2007)

In the following paper, we study the k-cycle dynamics applied to two-dimensional triangular lattices. In particular we focus our attention first on the socially balanced configurations that the lattice reaches when the k-cycle dynamics is applied. It is possible to describe it as a percolation transition, where here the parameter p is not the same as defined in standard percolation theory [see section 2.2] but p is the propensity parameter of k-cycle dynamics [see section 5.1.3]. We make use of all the tools of finite-size scaling analysis [see section 2.2.3] and we properly compute the critical exponents of the transition. Similarly, we study the critical transition as a function of the time [analogously to what was explained in the case of directed percolation in section 2.2.4]. Since, as opposed to a static percolation description our model is dynamical, this allows us to include also the time in our finite-size scaling analysis and to calculate the dynamical critical exponents. This analysis is performed in the second part of the paper.

Personal Contribution

I produced the numerical results reported in the following paper. In particular, in appendix B.3 I describe in detail the fast numerical implementation of k-cycle dynamics used during my simulations on two-dimensional lattices.

Universality class of triad dynamics on a triangular lattice

Filippo Radicchi,* Daniele Vilone,† and Hildegard Meyer-Ortmanns‡
School of Engineering and Science, International University Bremen, P. O. Box 750561, D-28725 Bremen, Germany
(Received 31 October 2006; revised manuscript received 16 January 2007; published 23 February 2007)

We consider triad dynamics as it was recently considered by Antal *et al.* [Phys. Rev. E **72**, 036121 (2005)] as an approach to social balance. Here we generalize the topology from all-to-all to the regular one of a two-dimensional triangular lattice. The driving force in this dynamics is the reduction of frustrated triads in order to reach a balanced state. The dynamics is parametrized by a so-called propensity parameter p that determines the tendency of negative links to become positive. As a function of p we find a phase transition between different kinds of absorbing states. The phases differ by the existence of an infinitely connected (percolated) cluster of negative links that forms whenever $p \leqslant p_c$. Moreover, for $p \leqslant p_c$, the time to reach the absorbing state grows powerlike with the system size L, while it increases logarithmically with L for $p > p_c$. From a finite-size scaling analysis we numerically determine the static critical exponents β and $\nu_\perp$ together with γ, τ, σ, and the dynamical critical exponents $\nu_\parallel$ and δ. The exponents satisfy the hyperscaling relations. We also determine the fractal dimension d_f that satisfies a hyperscaling relation as well. The transition of triad dynamics between different absorbing states belongs to a universality class with different critical exponents. We generalize the triad dynamics to four-cycle dynamics on a square lattice. In this case, again there is a transition between different absorbing states, going along with the formation of an infinite cluster of negative links, but the usual scaling and hyperscaling relations are violated.

DOI: 10.1103/PhysRevE.75.021118 PACS number(s): 05.40.−a, 64.60.Ak, 89.75.Fb

I. INTRODUCTION

Recently Antal *et al.* [1] proposed a so-called triad dynamics to model the approach to social balance [2–4]. An essential ingredient in the algorithm is the reduction of frustration in the following sense. We assign a value of +1 or −1 to a link (or bond) in the all-to-all topology if it connects two individuals who are friends or enemies, respectively. We call the sign ±1 of a link its spin. If the product of links along the boundary of a triad is negative, the triad is called frustrated (or imbalanced), otherwise it is called balanced (or unfrustrated). The state of the network is called balanced if all triads are balanced. The algorithm depends on a parameter $p \in [0,1]$, called the propensity parameter. It determines the tendency of the system to reduce frustration via flipping a negative link to a positive one. For an all-to-all topology Antal *et al.* predict a transition from an imbalanced nonabsorbing stationary state for $p < 1/2$ to a balanced absorbing state for $p \geqslant 1/2$. Here the dynamics is motivated by social applications so that the notion of frustration from physics goes along with frustration in the psychological sense. The mathematical criterion for checking the status of frustration is the same.

In a recent paper [5] we generalized the triad dynamics in two aspects. The first generalization refers to a k-cycle dynamics which contains the triad dynamics for $k=3$. Here it turned out that the main difference comes from the difference of whether k is even or odd, since the phase structure is symmetric about $p=1/2$ for k even. Even in the infinite-volume limit there are only absorbing states, apart from the transition point at $p=1/2$. The second generalization concerned the network topology from all-to-all connections to a diluted network. We studied the phase structure as a function of the propensity p and the dilution. As it turned out, the diluted k-cycle dynamics can be mapped on a certain satisfiability problem in computer science, the so-called k-XOR-SAT problem [6], and socially balanced states in one problem correspond to all logical constraints satisfied in the k-XOR-SAT problem. In both models we have phases of imbalanced nonabsorbing states in the infinite-volume limit, separated by a phase transition from phases of balanced or absorbing states. In a finite volume one only observes balanced states, but as a remnant of the infinite-volume phase structure, the time to reach the absorbing states differs in a characteristic way.

In this paper we study triad dynamics on a two-dimensional triangular lattice and four-cycle dynamics (called tetrad dynamics) on a square lattice. From the interpretation as an approach to social balance, triad or tetrad dynamics on a regular topology are not more realistic than on an all-to-all topology. Still, triad dynamics shows interesting features in terms of a percolation transition if we compare snapshots of frozen states for different values of p. Also for tetrad dynamics we observe a transition between different absorbing states, but the description in terms of a percolation transition fails. As we shall see, due to the restrictive topology, imbalanced triads and tetrads allow only two elementary processes: either they diffuse or they annihilate each other. As a result, the system always approaches a balanced absorbing state; nevertheless we observe a transition as a function of the propensity parameter p, this time between different absorbing states. The difference is characterized by the presence or absence of an infinite cluster of connected unfriendly (negative) links and for triad dynamics by the time to reach these frozen states. The parameter p should not be confused with the occupation probability of a single bond with a posi-

*Electronic address: f.radicchi@iu-bremen.de
†Electronic address: d.vilone@iu-bremen.de
‡Electronic address: h.ortmanns@iu-bremen.de

FIG. 1. Local dynamics of one update event. The imbalanced triads are represented as filled triangles, while the balanced ones as empty triangles. The shared link is represented as bold; this is the link involved in the update event. (a) When the update event flips the link shared by one imbalanced and one balanced triad we have diffusion. (b) When the update event flips the link shared by two imbalanced triads we have annihilation.

tive sign as it is usually used in connection with bond percolation. Therefore, here for $p \leqslant p_c$ we observe an infinite cluster of negative links, while such a cluster is absent for $p > p_c$. For triad dynamics the time to reach the frozen state grows logarithmically with the system size for $p > p_c$ and as a power of the size for $p \leqslant p_c$. In contrast, for tetrad dynamics it grows as a power of the system size only at the transition point, while it grows logarithmically above and below p_c. For triad dynamics we numerically determine the critical exponents from a finite-size scaling analysis of the frozen patterns. We analyze the probability P_∞ for a link to belong to the infinite cluster that serves as an order parameter as well as the order parameter susceptibility and the distribution of clusters of finite size s. The static critical exponents β, $\nu_\perp$, γ, σ, and τ satisfy the usual hyperscaling relations as well as the fractal dimension d_f and the dynamical critical exponents $\nu_\parallel$ and δ. The critical exponents turn out to be new and completely different from those of standard percolation in two dimensions [7,8].

The outline of the paper is the following. In Sec. II we define the triad dynamics on a triangular lattice. In Sec. III we analyze the finite-volume dependence of the frozen states as a function of the propensity parameter p and present the results for the static critical exponents as well as for the fractal dimension. Moreover, we analyze the finite-time dependence of the geometrical properties of the evolving states as a function of p and report the values of the dynamical critical exponents. Section IV deals with tetrad dynamics on a square lattice for which the hyperscaling relations are violated. In Sec. V we summarize the conclusions.

II. TRIAD DYNAMICS

Our dynamical system is defined on an undirected graph (network) composed of nodes and links. Each link (i,j) between the nodes i and j takes spin values $\sigma_{(i,j)}=-1$ or $\sigma_{(i,j)}=+1$ if the nodes i and j are "enemies" or "friends," respectively. A triad $[\triangle]$ (i,j,k) is characterized by the values assigned to its three links (i,j), (j,k), and (k,i). We have four types of triads, depending on the number of negative links they contain in their boundary: Δ_0, Δ_1, Δ_2, and Δ_3, where the subscript stands for the number of negative (or "unfriendly") links. We use the standard notion of *social balance* as proposed in [2,3] and apply this notion to triads. The sign of a triad is defined as the product of the spins assigned to the links of the triad. A triad is considered as "balanced" or "unfrustrated" if its sign is positive, otherwise it is called

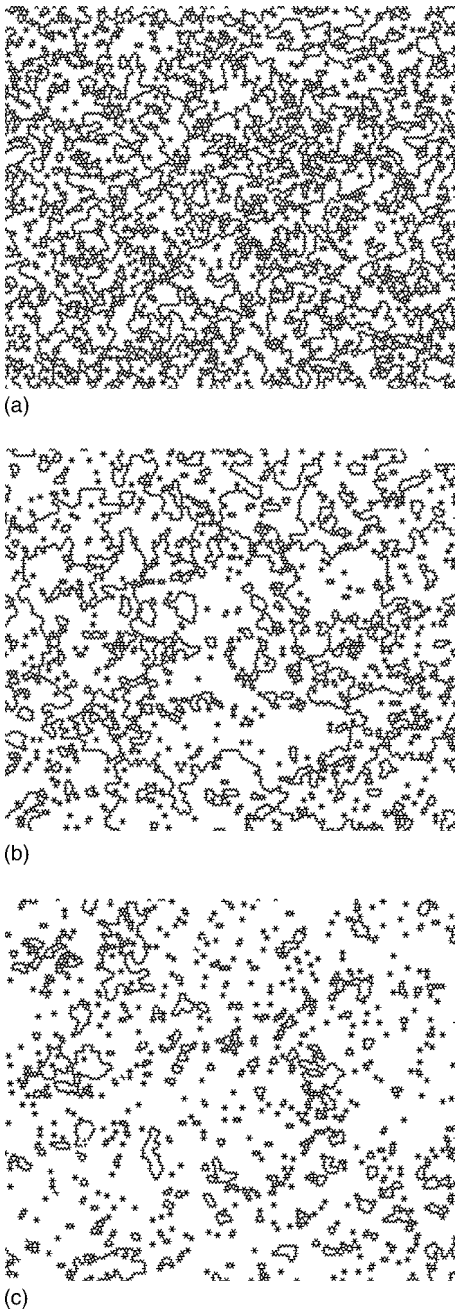


FIG. 2. Typical frozen configurations for a triangular lattice with periodic boundary conditions, with linear size $L=129$ and for different values of p: (a) $p=0.44$, (b) $p=p_c=0.4625$, (c) $p=0.48$. In the plots only negative links are shown.

"imbalanced" or "frustrated." The triads Δ_0 (all friends) and Δ_2 (two friends have the same enemy) are balanced, while the triads Δ_1 and Δ_3 are imbalanced. The network itself is called balanced if and only if all triads belonging to the network are balanced.

As in [1,5] we perform a local unconstrained dynamics in order to reduce the frustration of the network. As it turns out, the local algorithm always drives the network to a fully balanced state without frustrated triads, but the time it needs for reaching the frozen state depends on the choice of parameters. At each update event one triad is selected at random. If

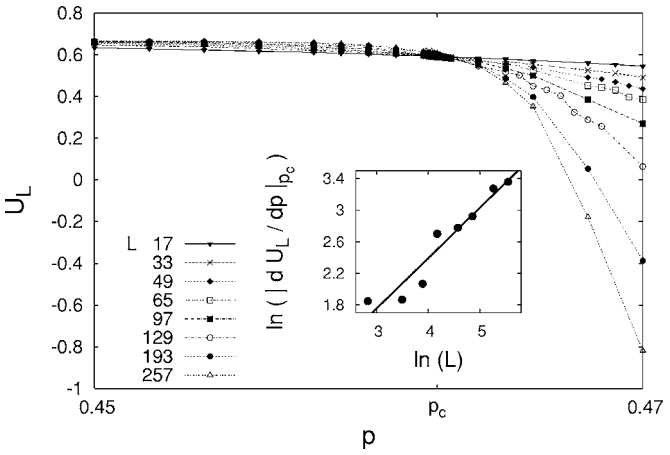


FIG. 3. Binder cumulant U_L as a function of the propensity parameter p. The main plot shows U_L in the vicinity of the critical point. The numerical results for different linear lattice sizes L have a common intersection at $p_c \simeq 0.4625$. The inset shows the ln-ln plot of $|dU_L/dp|_{p_c}$. From the relation $|dU_L/dp|_{p_c} \sim L^{1/\nu_\perp}$ we find $1/\nu_\perp = 0.64(7)$ (solid line).

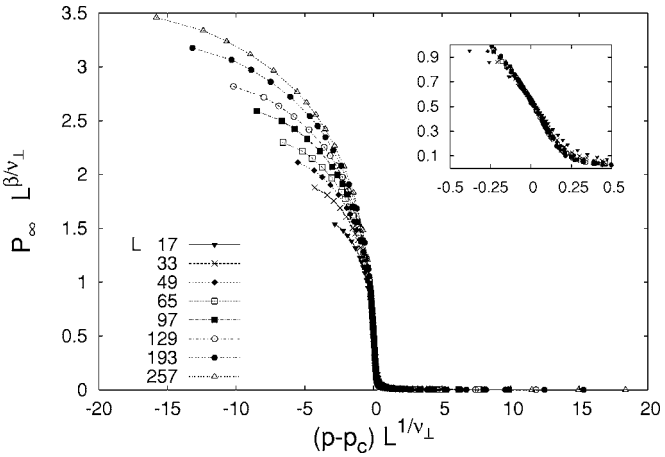


FIG. 5. Same data as in Fig. 4, but now the abscissa is rescaled according to $(p-p_c)L^{1/\nu_\perp}$ with $1/\nu_\perp = 0.64$. The inset shows a zoom around zero. The numerical simulations are the same as in Fig. 3.

the selected triad is balanced (type Δ_0 or Δ_2) nothing happens. If the selected triad is imbalanced (type Δ_1 or Δ_3) it is updated into a balanced one by flipping one of its link. In particular a triad Δ_1 is changed with probability p into a triad Δ_0 (flipping the only negative link of the triad Δ_1), and it is changed with probability $1-p$ into a triad Δ_2 (choosing at random one of the two positive links belonging to the triad Δ_1 and inverting the spin to a negative value). A triad Δ_3 is changed into a triad Δ_2 with probability 1, choosing at random one of its three negative links and reversing its spin to a positive sign. We summarize the updating rules of the local algorithm in the following scheme:

$$\Delta_0 \underset{p}{\leftarrow} \Delta_1 \xrightarrow[1-p]{} \Delta_2 \underset{1}{\leftarrow} \Delta_3. \tag{1}$$

One time unit has passed when the number of single update events equals the total number of links M of the network.

We study the triad dynamics on two-dimensional triangular lattices with periodic boundary conditions. We characterize a triangular lattice using its linear size L. The total number of sites in the lattice is $N=L(L-1)$. The term $L-1$ results from the periodic boundary conditions. The total number of links of the lattice is $M=3N$, while the total number of triads is $N_\Delta=2N$. In particular each link is shared only by two nearest-neighbor triads, so that the triad dynamics cannot increase the total number of imbalanced triads: a single update changes one selected imbalanced triad into a balanced one, while it modifies the other triad, sharing the same updated link, either from balanced to imbalanced [Fig. 1(a)] or from imbalanced to balanced [Fig. 1(b)]. The former we call

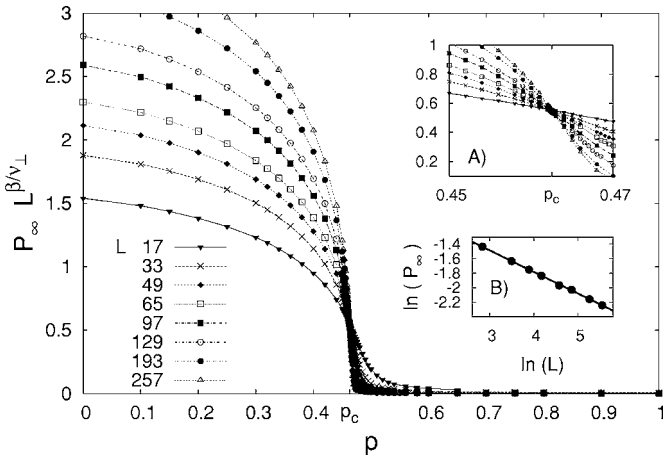


FIG. 4. Probability P_∞ that a bond belongs to the largest cluster as a function of p and for different linear sizes of the lattice. We perform a finite-size scaling for determining the critical point $p_c = 0.4625(5)$ [see the zoom around the critical point in the inset (a)] with the critical exponents $\beta/\nu_\perp = 0.297(3)$ (solid line) as is shown in the inset (b), where P_∞ at the critical point is plotted as a function of the linear size of the lattice L. The numerical simulations are the same as in Fig. 3.

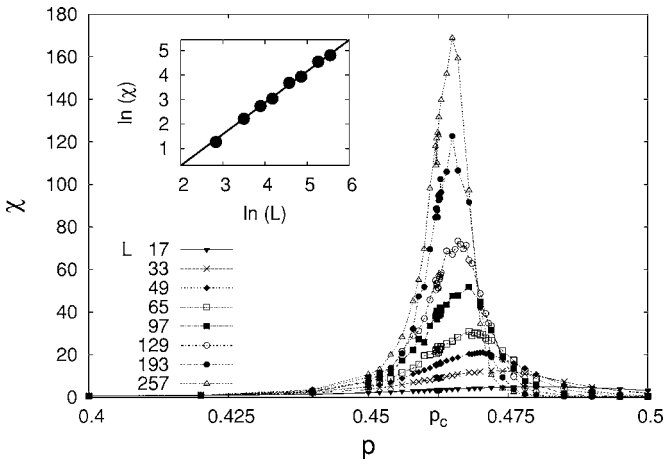


FIG. 6. Susceptibility χ of P_∞ as a function of p. In the main plot we zoom the region around the critical point p_c, while in the inset we plot χ at the critical point as a function of L leading to $\gamma/\nu_\perp = 1.28(3)$ (solid line). The numerical simulations are the same as in Fig. 3.

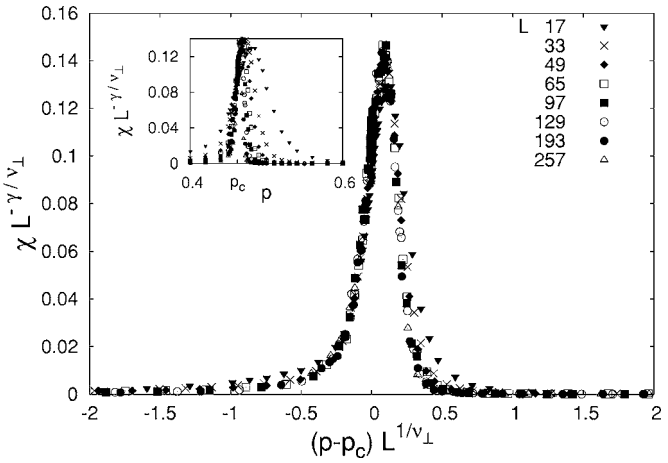


FIG. 7. Finite-size scaling for the susceptibility χ. The data are the same as in Fig. 6. In the inset we zoom the function $\chi L^{-\gamma/\nu_\perp}$ around the critical point p_c, while in the main plot we rescale the abscissa as $(p-p_c)L^{1/\nu_\perp}$. The critical exponents are chosen as $\gamma/\nu_\perp=1.28$ and $1/\nu_\perp=0.64$.

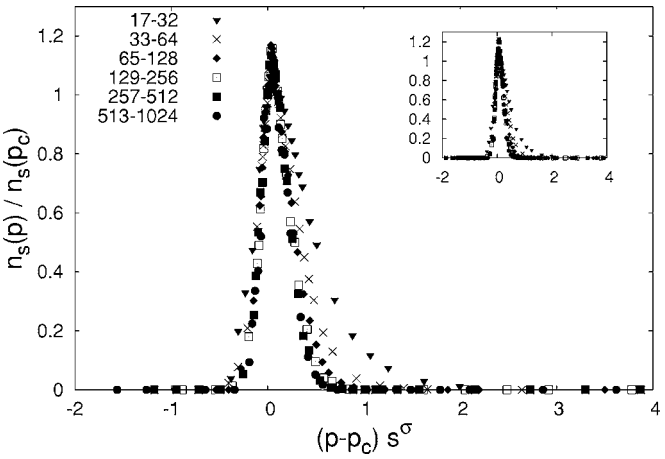


FIG. 9. Numerical test of the validity of Eq. (7). We plot the ratio $n_s(p)/n_s(p_c)$, where n_s is the distribution of the cluster size, as a function of the rescaled quantity $(p-p_c)s^\sigma$ with σ chosen as 0.41 from Eq. (13) knowing β and γ. The main plot refers to $L=257$, the inset to $L=129$.

diffusion, because the imbalanced triad diffuses, the latter annihilation between two imbalanced triads.

Obviously, for a finite-size system we always observe a frozen configuration as the stationary state (see Fig. 2), independently of the initial configuration. From now on we focus on these frozen configurations and study their geometrical properties by using the standard tools of percolation theory [7]. It should be noticed, however, that the dynamical parameter p of the triad dynamics is very different from the occupation probability as it is defined in percolation theory, where it is also called p.

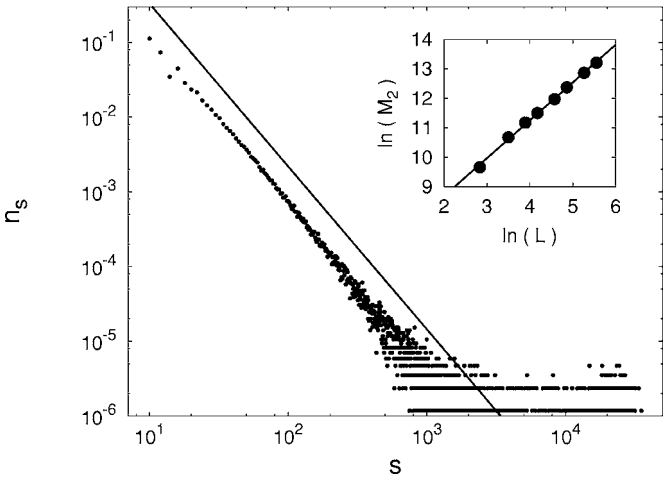


FIG. 8. Distribution of the cluster size $n_s(p_c)$ at the critical point p_c and for $L=257$ (main graph). The distribution is extracted from 10^3 frozen configurations. As expected this distribution follows a power law $n_s \sim s^{-\tau}$; the solid line plotted here corresponds to $\tau=2.19$. The inset shows the second moment of the distribution of the cluster size M_2 at the critical point as a function of the linear size of the lattice. M_2 increases as a power of the linear size L of the lattice with exponent $\gamma/\nu_\perp=1.28(2)$ (solid line).

III. NUMERICAL SIMULATIONS TO DETERMINE THE CRITICAL THRESHOLD AND THE CRITICAL EXPONENTS

In order to study the geometry of the frozen configurations, we consider the distribution of negative links on the lattice. In particular we numerically compute the probability that a negative link belongs to an infinite cluster $P_\infty = M_\infty^- / M$, as the ratio of the number of links belonging to the largest cluster of connected negative links M_∞^- to the total number of links M.

A. Critical propensity parameter p_c

Let us first determine the critical point p_c. Binder's cumulant [9], defined as the fourth-order reduced cumulant of the probability distribution

$$U_L = 1 - \frac{\langle P_\infty^4 \rangle}{3\langle P_\infty^2 \rangle^2}, \tag{2}$$

allows us to determine the critical point p_c without tuning any parameter. In Eq. (2) $\langle\cdot\rangle$ stands for the average over all realizations of the distribution. It is known that the Binder cumulant should satisfy the scaling relation [9]

$$U_L = \tilde{U}[(p-p_c)L^{1/\nu_\perp}], \tag{3}$$

with $\tilde{U}(\cdot)$ a universal function. From Fig. 3 it is obvious that the Binder cumulants for different linear sizes of the lattice L have a common intersection at $p=0.4625(5)$, so that we use $p_c=0.4625$ as the critical propensity parameter in the following analysis. The numerical results of Fig. 3 are extracted from numerical simulations for $L=17$, 33, 49, 65, 97, 129, 193, and 257. The averages are taken over several frozen configurations as they are reached as absorbing states of the triad dynamics, starting from initial conditions where each link is randomly assigned a value of +1 or −1 with the same probability 1/2. The number of realizations is 10^4 for values of L up to 65 and 10^3 for larger values of L.

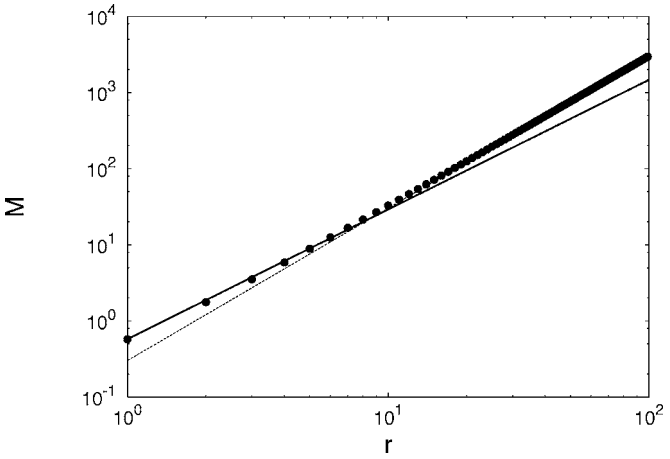


FIG. 10. Mass of negative links M inside a triangular box with r links per side. The mass is measured on a triangular lattice with $L=257$ and periodic boundary conditions. Each point is given by the average over 10^3 frozen configurations, while the number of boxes considered for each r is 10^2. The numerical results fit a power law with $d_f=1.703$ (solid line) and $d=2$ (dotted line).

B. Critical exponent $\nu_\perp$

In order to extract the value of the critical exponent $\nu_\perp$ that characterizes the divergence of the correlation length $\xi_\perp$ in the vicinity of the critical point, we consider the absolute value of the first derivative of the Binder cumulant calculated at p_c, $|dU_L/dp|_{p_c}$. We expect from Eq. (3) to have $|dU_L/dp|_{p_c} \sim L^{1/\nu_\perp}$. This relation is actually satisfied for $1/\nu_\perp=0.64(7)$ as it is seen in the inset of Fig. 3, from which $\nu_\perp=1.6(2)$. For the numerical estimate of the derivative of U_L we use $|dU_L/dp|_{p_c}=|\Sigma_{q=1}^{Q} q a_{L,q} p_c^{q-1}|$ where the coefficients $a_{L,q}$ are obtained by interpolating the cumulant U_L with a polynomial $U_L=\Sigma_{q=0}^{Q} a_{L,q} p^q$. Here we choose $Q=5$ and restrict the interpolation to the interval [0.46,0.465].

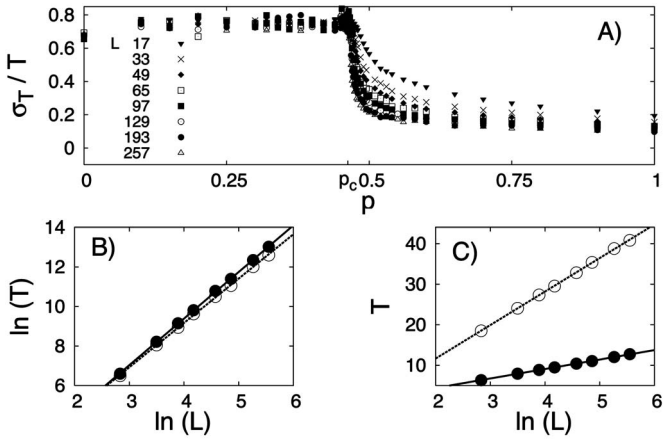


FIG. 11. Average time T needed for reaching a frozen configuration. The data sets are the same as in Fig. 3. In (a) we plot the standard variance σ_T over the average value T as a function of p and for different values of L. The peaks of this ratio are all around the critical point p_c. In (b) we show a ln−ln plot of T as a function of the linear size of the lattice L at the critical point p_c (full circles) and at $p=1/3$ (open circles). It implies that $T\sim L^z$ for $p\leqslant p_c$. We have $z=2.36(1)$ (solid line) at p_c, while $z=2.24(2)$ (dotted line) at $p=1/3$. In (c) the plot of T versus the logarithm of L for $p=1$ (full circles) and 3/4 (open circles) shows that $T\sim \ln(L)$ for $p>p_c$. The numerical simulations are the same as in Fig. 3.

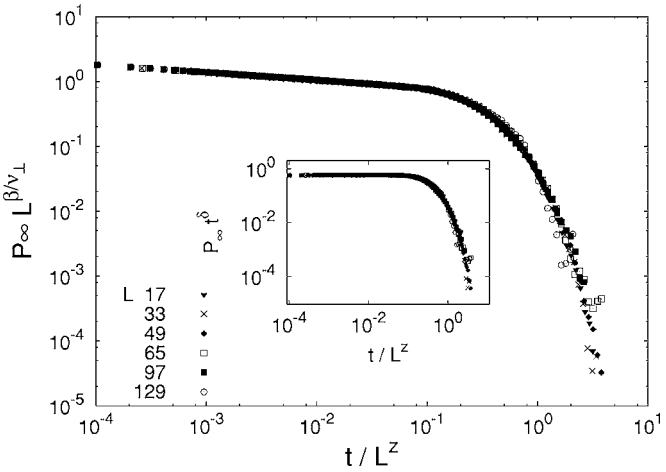


FIG. 12. Finite-size scaling of the time-dependent behavior of the percolation probability $P_\infty(t)$. The values of the critical exponents used here are $\beta/\nu_\perp=0.297$, $z=2.36$, and $\delta=0.126$. The numerical results are obtained by averaging over 10^4 realizations for $L=17$ and 33, 10^3 realizations for $L=49$ and 65, and 10^2 realizations for $L=97$ and 129.

C. Critical exponent β

According to percolation theory, the probability P_∞ satisfies the following finite-size scaling relation:

$$P_\infty = L^{-\beta/\nu_\perp} \tilde{P}[(p-p_c)L^{1/\nu_\perp}], \tag{4}$$

where p_c is the critical value of the propensity parameter p for which the phase transition occurs. (In the infinite-volume limit we have $P_\infty=1$ for $p\leqslant p_c$, while $P_\infty=0$ for $p>p_c$.) $\tilde{P}(\cdot)$ is a universal function. A plot of $P_\infty L^{\beta/\nu_\perp}$ as a function of p and for different values of L is shown in Fig. 4. For a value of $\beta/\nu_\perp=0.297(3)$ [see inset (b)] all curves have an intersection in $p_c=0.4625(5)$ as is more obvious from the inset (a). Using for $\nu_\perp$ the value calculated so far, we obtain $\beta=0.46(5)$.

If we rescale the abscissa as $(p-p_c)L^{1/\nu_\perp}$ with $p_c=0.4625$ and $1/\nu_\perp=0.64$, all curves for different values of L collapse into one (see Fig. 5). This is true especially close to zero as the inset of Fig. 5 clearly shows.

D. Critical exponent γ

Moreover, in Fig. 6 we plot the susceptibility

$$\chi = M[\langle P_\infty^2\rangle - \langle P_\infty\rangle^2], \tag{5}$$

in which we only show a zoom around $p=p_c$, while in the inset we plot the value of χ, at the critical point, as a function of the linear size of the lattice L. From the inset we find $\gamma/\nu_\perp=1.28(3)$, because $\chi\sim L^{\gamma/\nu_\perp}$ at p_c; therefore $\gamma=2.0(3)$. The susceptibility should satisfy the finite-size scaling relation

$$\chi = L^{\gamma/\nu_\perp}\tilde{\chi}[(p-p_c)L^{1/\nu_\perp}], \tag{6}$$

where again $\tilde{\chi}(\cdot)$ is a universal function. This relation is perfectly satisfied (cf. Fig. 7).

E. Critical exponents τ and σ

Furthermore, we consider the probability distribution of having n_s clusters with s negative links. n_s is given by the ratio of the number of clusters of size s to the total number of clusters. As it is known from percolation theory, n_s should satisfy

$$n_s = s^{-\tau}\tilde{n}[(p-p_c)s^{\sigma}], \tag{7}$$

where τ and σ are critical exponents and $\tilde{n}(\cdot)$ is a universal function. We can determine the critical exponent τ (also called the Fisher exponent) by plotting the distribution of the cluster sizes s at the critical point p_c. This distribution is calculated in Fig. 8 for $L=257$. The distribution fits with a power law $s^{-\tau}$, here plotted as solid line with $\tau=2.19$. We checked that the same exponent fits also for smaller values of L.

In Fig. 9 we numerically determine the critical exponent σ. The figure shows the plot of the ratio $n_s(p)/n_s(p_c)$. Using the rescaled variable $(p-p_c)s^{\sigma}$ as abscissa, all the curves corresponding to different values of the cluster size s ($2^4 < s \leq 2^5$, $2^5 < s \leq 2^6$, $2^6 < s \leq 2^7$, etc.) collapse to a single function for $\sigma=0.41$, as expected (see, e.g., [7]).

F. Check of the critical exponent γ

From the cluster distribution it is possible to verify the value of the critical exponent γ that was calculated before. It is known that the second moment of the distribution of the cluster size

$$M_2 = \sum_s s^2 n_s \tag{8}$$

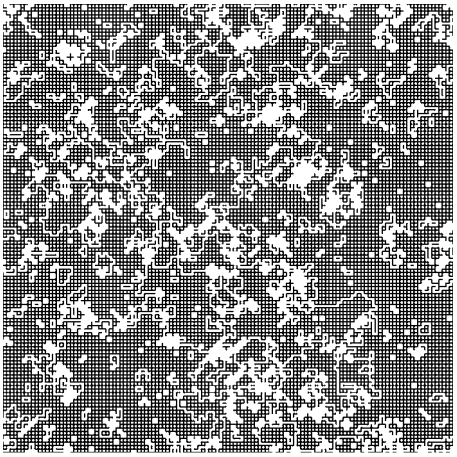
should scale according to $M_2 \sim L^{\gamma/\nu_\perp}$ at the critical point p_c. The numerical value found for the ratio $\gamma/\nu_\perp = 1.28(2)$ is consistent with the former one obtained via the susceptibility (see the inset of Fig. 8).

G. Fractal dimension d_f

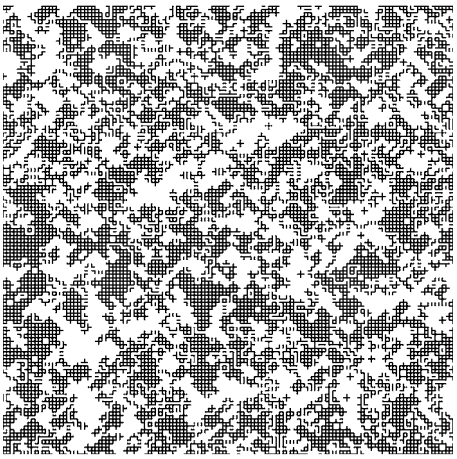
The percolating cluster can be further characterized in terms of a fractal dimension d_f. The fractal dimension d_f is easily computed using a box counting method [8]. We count how many negative links $M(r)$ belong to a triangular box with r links per side. The measurement is performed for $L=257$ at the critical point p_c. We analyze 10^3 frozen configurations, for each of them we consider 10^2 different boxes. The mass $M(r)$ per box is plotted in Fig. 10. As expected we observe the crossover phenomenon

$$M(r) = \begin{cases} r^{d_f} & \text{if } r \ll \xi_\perp, \\ r^{d} & \text{if } r \gg \xi_\perp, \end{cases} \tag{9}$$

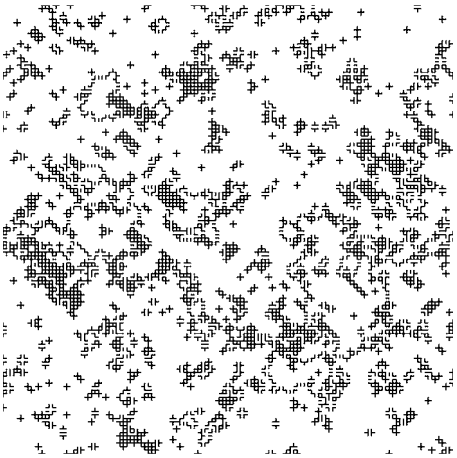
where $d=2$ in this case. The numerical results fit with a power law with $d_f=1.703$ and $d=2$, respectively.



(a)



(b)



(c)

FIG. 13. Typical frozen configurations for a square lattice with periodic boundary conditions, with linear size $L=128$ and for different values of p: (a) $p=0.48$, (b) $p=p_c=0.5$, (c) $p=0.52$. In these plots are shown only the negative links.

H. Time to reach the frozen states

So far we have characterized the geometrical properties of the final absorbing configuration. As a next step we focus on the dynamical features and determine the time the system

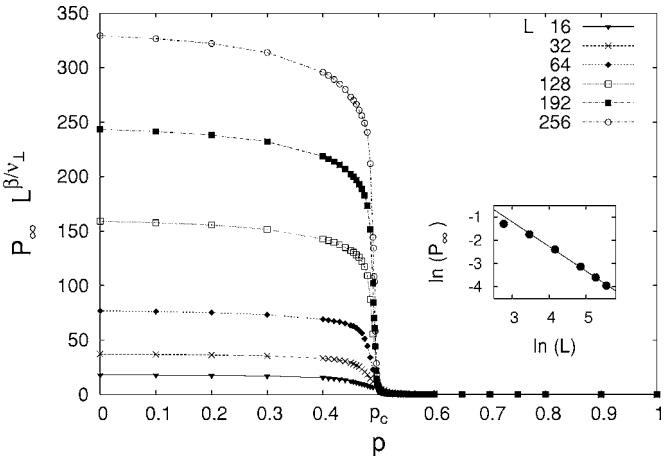


FIG. 14. Probability P_∞ that a link belongs to the largest cluster as a function of p and for different linear sizes of the lattice. The ratio of critical exponents $\beta/\nu_\perp=1.05(2)$ at the critical point $p_c=0.5$ is shown in the inset.

needs to reach the final frozen configurations. In Fig. 11(a) we plot the ratio of the variance σ_T and its average value T as a function of p. It is interesting to notice that σ_T/T has a maximum around p_c. Below the critical point p_c, the relaxation time T is governed by a power-law relation in the linear size of the lattice L: $T\sim L^z$. Our numerical analysis reveals $z=2.36(1)$ at p_c and $z=2.24(1)$ at $p=1/3$ [see Fig. 11(b)]. Furthermore, above the critical point p_c, T grows logarithmically with L: $T\sim\ln(L)$ [see Fig. 11(c), where we show the L dependence of T for $p=3/4$ and for $p=1$]. The different size dependence of the time the system needs to reach a frozen configuration in both phases can be understood in qualitative terms. For $p>p_c$ there is a lower probability for having negative links. The stable configurations forming out of these links are mainly local objects like star-like configurations which do not percolate through the lattice [see Fig. 1(c)]. Therefore the time to reach such a state characterized by local objects depends weakly on the system size, that is logarithmically. For $p\leqslant p_c$ there is a high probability of having negative links. However, here large loops that finally percolate through the lattice make up the stable configurations [see Figs. 1(a) and 1(b)]. So the time to reach this configuration is more sensitive to the system size: it grows as a power of the size.

I. Dynamical critical exponents $\nu_\parallel$ and δ

The existence of the dynamical critical exponent z suggests the possibility to analyze the percolation transition of our model from a dynamical point of view. Therefore we performed other numerical simulations with the aim of determining some dynamical critical exponents. We followed a procedure similar to the one usually applied for other time-dependent critical phenomena [10]. Starting from a fully occupied lattice [i.e., all the links of the triangular lattice are set to be negative so that $P_\infty(t=0)=1$], we observe how P_∞ evolves in time. According to finite-size scaling analysis we should expect, at the critical point p_c, a behavior of the type

TABLE I. Critical threshold and critical exponents, numerically determined for triad dynamics on two-dimensional triangular lattices. The critical exponents of triad dynamics on triangular lattices are compared with those of standard percolation in two dimensions (which does not have any dynamical exponents).

	Triad dynamics	Percolation in two dimensions[7]
p_c	0.4625(5)	
β	0.46(5)	$5/36\simeq0.138$
γ	2.0(3)	$43/18\simeq2.388$
$\nu_\perp$	1.6(2)	$4/3\simeq1.333$
σ	0.41(6)	$36/91\simeq0.396$
τ	2.19(1)	$187/91\simeq2.055$
d_f	1.703(3)	$91/48\simeq1.895$
$\nu_\parallel$	3.8(5)	None
δ	0.126(1)	None

$$P_\infty(t)\sim L^{-\beta/\nu_\perp}f(t/L^z), \tag{10}$$

where $f(\cdot)$ is a suitable universal function. We numerically test the validity of this relation. In Fig. 12 we plot $P_\infty(t)L^{\beta/\nu_\perp}$ versus t/L^z, for $p=p_c$ and various lattice sizes L. The values of $\beta/\nu_\perp$ and z calculated above are consistent with Eq. (10), since all curves for different values of L collapse into a single curve. In analogy with other time-dependent critical phenomena [10,11], we may look at z as the ratio between the critical exponent $\nu_\parallel$ of the typical time-correlation length $\xi_\parallel$ of the percolation cluster and $\nu_\perp$ as the typical spatial correlation length $\xi_\perp$. This means $z=\nu_\parallel/\nu_\perp$, therefore knowing $\nu_\perp$ and z, we estimate $\nu_\parallel$ to be 3.8(5).

Moreover we determine the critical exponent δ of the time decay of P_∞ at the critical p_c: $P_\infty(t)\sim t^{-\delta}$. [$P_\infty(t)$ is expected to decay to zero at p_c, because there the ratio of links belonging to the infinite cluster over all links is still vanishing in the infinite volume. Choosing a different initial value of P_∞ at time $t=0$ leads to the same powerlike long-time behavior in time, so $P_\infty(0)=1$ is chosen for convenience.] Therefore, by means of the same simulations as used before it is possible to test the validity of the scaling relation

$$P_\infty(t)\sim t^{-\delta}g(t/L^z), \tag{11}$$

where $g(\cdot)$ is a proper universal function. We show the result of this numerical test in the inset of Fig. 12, where we impose $\delta=0.126$. The actual value of δ used in this plot is obtained from the known values of β and $\nu_\parallel$ and the hyperscaling relation of Eq. (15) that connect the three critical exponents β, $\nu_\parallel$ and δ. The numerical results reported in both plots of Fig. 12 are obtained by averaging over a large number of realizations (10^4 for $L=17$ and 33, 10^3 for $L=49$ and 65, 10^2 for $L=97$ and 129).

J. Universality class of triad dynamics in two dimensions

We list the critical exponents for the phase transition between different absorbing states in Table I. The critical exponents satisfy the known hyperscaling relations. For ex-

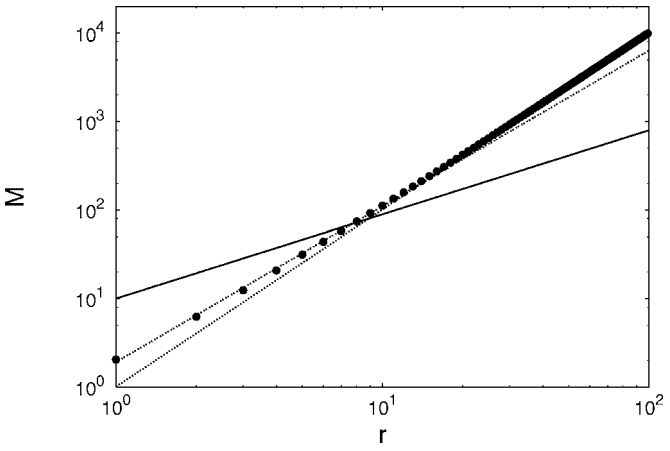


FIG. 15. Mass of negative links M inside a square box with r links per side. This measure is used on a square lattice with $L=256$ and periodic boundary conditions. Each point is given by the average over 10^3 frozen configurations, while the number of boxes considered at each r is 10^3. The solid line has the slope 0.95, a value that is expected by the former measure of $\beta/\nu_\perp$. Actually it is not consistent with the direct fit of the data from which we find $d_f=1.76(2)$ (dashed line). The dotted line has slope $d=2$.

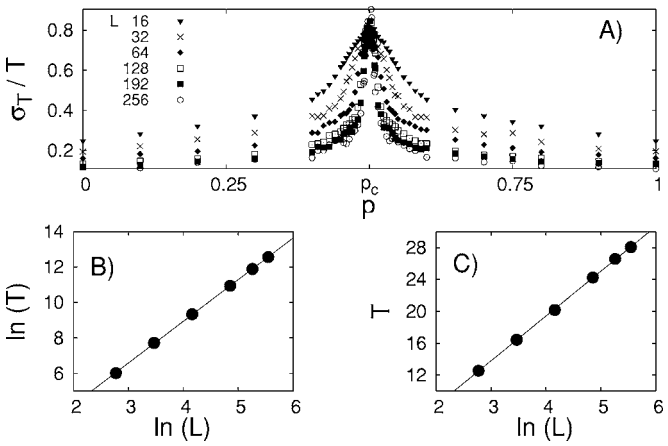


FIG. 16. Average time T needed for reaching a frozen configuration. The simulations are the same as in Fig. 14. In (a) we plot the standard variance σ_T over the average value T as a function of p and for different values of L. The peaks of this ratio are all in the vicinity of the critical point p_c. In (b) we show a ln-ln plot of T as a function of the linear size of the lattice L at the critical point leading to $T\sim L^z$ with $z=2.35(1)$. In (c) the plot of T versus the logarithm of L for $p=1$ shows that $T\sim\ln(L)$.

ample, it is known from percolation theory that the critical exponents β, γ, τ, and σ are related by

$$\tau=\frac{1}{2}\left(5-\frac{\gamma-\beta}{\gamma+\beta}\right) \tag{12}$$

and

$$\sigma=\frac{1}{\gamma+\beta}. \tag{13}$$

Calculating τ and σ from the former equations, using the numerically obtained values of β and γ, we find $\tau=2.19(1)$ and $\sigma=0.41(6)$. Both values are consistent with Figs. 8 and 9.

The fractal dimension is related to the ratio $\beta/\nu_\perp$ by the hyperscaling relation

$$d_f=d-\frac{\beta}{\nu_\perp}. \tag{14}$$

From this relation we find $d_f=1.703(3)$ in agreement with Fig. 10.

The exponent of the time-decay of the percolation probability $P_\infty(t)\sim t^{-\delta}$ at the critical point is related to the critical exponents β and $\nu_\parallel$ by the hyperscaling relation [10]

$$\delta=\frac{\beta}{\nu_\parallel}. \tag{15}$$

From this relation we calculate $\delta=0.126(1)$ in agreement with Fig. 12.

The errors of p_c and of the critical exponents as indicated in Table I arise as follows. The critical point p_c is directly estimated from Fig. 3, verified in Fig. 4, and supported by Fig. 10. The only source of error here is given by the step size used for varying the propensity parameter p. Of course this step size is related to the parameters of the simulations. It can be decreased by increasing the system size and the total number of simulations.

The critical exponents $\beta/\nu_\perp$ of Fig. 4(b), $1/\nu_\perp$ of Fig. 3, $\gamma/\nu_\perp$ of Figs. 7 and 8, and z of Fig. 11 have errors due to the linear fit in a double-logarithmic plane. In principle one should also account for the propagation of the error entering the value of the critical point $p_c=0.4625(5)$, which we have neglected here.

In order to check the validity of the hyperscaling relations we evaluate the corrections to all derived quantities (τ, σ, d_f, $\nu_\parallel$, and δ) by using the standard formula for error propagation. The hyperscaling relations are then said to be satisfied if they hold within these error bars.

To our knowledge the critical exponents calculated so far are new [7,8,11]. Therefore, the percolation transition between different absorbing states of triad dynamics on two-dimensional lattices can be described by standard percolation theory, but the transition seems to belong to a new universality class.

IV. TETRAD DYNAMICS

The notion of social balance can be extended to any geometric figure with k links (in graph theory denoted as k-cycles) [3,5]. As it turned out in [5] for an all-to-all topology and a diluted topology, generalizing the triad dynamics to a k-cycle dynamics leads to qualitative differences in the phase structure of which the main differences are due to k being even rather than being larger than three. In this section we therefore focus on the case of $k=4$ and call the four-cycles tetrads defined on square lattices. According to the number of negative links belonging to a particular tetrad, we distinguish five types of tetrads: $\square_0$, $\square_1$, $\square_2$, $\square_3$, $\square_4$. The updating rules of the unconstrained dynamics for triads of (1)

are naturally extended to tetrads in the following way:

$$\square_0 \underset{p}{\leftarrow} \square_1 \xrightarrow[1-p]{} \square_2 \underset{p}{\leftarrow} \square_3 \xrightarrow[1-p]{} \square_4. \tag{16}$$

The local tetrad dynamics is then applied to square lattices with periodic boundary conditions. L denotes the linear size of the lattice, i.e. the number of sites per row or per column. The total number of sites of the lattice is $N=L^2$, the total number of links $M=2N$, and the total number of tetrads $N_{\square}=N$.

Similarly to the case of triad dynamics on two-dimensional triangular lattices, at a first glance there seems to be a percolation transition also in case of tetrad dynamics. The critical point p_c should be equal to $p_c=0.5$ due to the symmetry of the system (16) under the simultaneous transformation $\sigma_i \rightarrow -\sigma_i \forall i$ and $p \rightarrow 1-p$ (see Fig. 13).

However, differently from the triad dynamics, no scaling relations hold in the vicinity of the transition at which again an infinite cluster of negative links forms. The hyperscaling relations are violated. For example, the relation (14) does not hold because one can measure $\beta/\nu_{\perp}=1.05(2)$ (Fig. 14) and expect to have $d_f=0.95(2)$, while it is directly seen from Fig. 15 that $d_f=1.76(2)$.

Tetrad dynamics resembles triad dynamics. Either each unstable tetrad diffuses, or two unstable tetrads annihilate if they meet. Finite-size systems always reach a frozen configuration within a finite time, so we can measure this time T. As we can see from Fig. 16(a), we find a symmetric behavior around the critical point p_c. At the critical point T scales according to $T \sim L^z$ with $z=2.35(1)$ [see Fig. 16(b)], this value is actually consistent with the one found for the triad dynamics. Away from the critical point $T \sim \ln(L)$ [see Fig. 16(c)].

V. SUMMARY AND CONCLUSIONS

The driving force in triad dynamics is the reduction of the number of frustrated triads. A state of zero frustration is called a state of social balance. Imposed on a triangular lattice, frustrated triads can diffuse or annihilate each other. Depending on the value of the propensity parameter p the final absorbing state can be characterized by the absence ($p > p_c$) or presence ($p \leqslant p_c$) of an infinite cluster of negative links. The time to reach the frozen configurations scales with the system size in a way that further characterizes the phases: it scales logarithmically for $p > p_c$ and in a powerlike way for $p \leqslant p_c$. The static critical exponents $\nu_{\perp}$, β, γ, τ, and σ as well as the fractal dimension d_f and the dynamical critical exponents $\nu_{\parallel}$ and δ satisfy hyperscaling relations within the error bars. The values of these exponents seem to characterize a new universality class. The essential difference that we observe for tetrad dynamics on a square lattice is the symmetry between the absorbing states in the different phases. Again, an infinite cluster of negative links forms for $p \leqslant p_c$, but the time to reach the frozen configurations shows the same dependence on the system size in both phases, with the only exception at the transition point. The percolation picture breaks down in the sense that no scaling and therefore no hyperscaling relations are satisfied.

[1] T. Antal, P. L. Krapivsky, and S. Redner, Phys. Rev. E **72**, 036121 (2005).

[2] F. Heider, Psychol. Today **51**, 358 (1944); Sankhya **21**, 107 (1946); *The Psychology of Interpersonal Relations* (John Wiley & Sons, New York, 1958); S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge University Press, New York, 1994).

[3] D. Cartwright and F. Harary, Psychol. Rev. **63**, 277 (1956); F. Harary, R. Z. Norman, and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs* (John Wiley & Sons, New York, 1965).

[4] F. Harary, Mich. Math. J. **2**, 143 (1953-54); F. S. Roberts, Electronic Notes in Discrete Mathematics (ENDM) **2**, 94 (1999); N. P. Hummon and P. Doreian, J. Ship Prod. **25**, 17 (2003).

[5] F. Radicchi, D. Vilone, S. Yoon, and H. Meyer-Ortmanns, Phys. Rev. E **75**, 026106 (2007).

[6] M. R. Garey and D. S. Johnson, *Computer and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979).

[7] D. Stauffer, *Introduction to Percolation Theory* (Taylor and Francis, London, 1985).

[8] A. Bunde and S. Havlin, *Fractals and Disordered Systems* (Springer, New York, 1996); D. ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, Cambridge, U.K., 2000).

[9] K. Binder, Z. Phys. B: Condens. Matter **43**, 119 (1981); K. Binder and D. W. Heerman, *Monte Carlo Simulation in Statistical Physics* (Springer-Verlag, Berlin, 1997).

[10] H. Hinrichsen, Adv. Phys. **49**, 7 (2000).

[11] G. Ódor, Rev. Mod. Phys. **76**, 663 (2004).

Chapter 6

Updating Dynamics on Networks

Introduction

Numerical simulations are nowadays powerful tools in statistical physics. The high computational speed of the actual computers allows scientists to simulate complex systems in a reasonable time, and the future progress towards even more powerful machines will extend the range of applicability of computer simulations. Numerical analysis is useful for two reasons. First of all, it allows the confirmation of theoretical predictions and sometimes leads the scientists towards the right theoretical approach for the solution of the problem. Second, when the complexity of the system is too high and no analytic approach is available, the numerical analysis gives a solution. The former reason is more serious than the latter, but, for practical applications, purely numerical solutions are clearly important.
In the context of numerical simulations, a highly topical discussion arises in the context of the updating rules for the evolution of the systems. In this chapter, I will propose my contribution to this discussion. Using a simple and standard model of statistical physics I will show how different updating schemes for the dynamics can lead to completely different behavior and features of the model itself (see **Paper VI** [122]). In order to make the topic a bit more understandable, I will introduce generic updating schemes, considered in **Paper VI** [122] , in sections 6.1.1 and 6.1.2. Moreover in section 6.2 I will give some basic notions of the Ising model, since this model is used in **Paper VI** [122] .

6.1 Updating Schemes

A dynamical model is characterized by a set of variables evolving in time. In principle, both the time and the state variables can be defined as discrete or continuous. A dynamical model

is generally only classified via the time, whether time is continuous or discrete. This leads to the definition of the two main categories already introduced in section 1.6. Continuous dynamical models are described by ordinary differential equations, and discrete dynamical models by finite-difference equations. In this chapter, I specialize only on discrete dynamical models, since they are most often used in applications of statistical physics. In particular, I focus on the simplest case in which both the time and the state variables take discrete values. I will distinguish between two possible ways of implementing the evolution of the system: synchronous or asynchronous dynamics.

6.1.1 Synchronous Updating Dynamics

Consider, for simplicity, a system of N units and assign to every unit i of the system a discrete state variable σ_i. One does not need to specify any underlying topology for his dynamical system (this can be either a lattice or a network). The updating rules for synchronous dynamics are such that all the units evolve simultaneously. Therefore the state variables σ_i at time $t+1$ are functions $f_i(\cdot)$ of all the variables of the system at time t:

$$\sigma_i(t+1) = f_i\left[\sigma_1(t), \sigma_2(t), \ldots, \sigma_i(t), \ldots, \sigma_N(t), \vec{E}\right] \quad , \; \forall\, i = 1, \ldots, N \quad . \tag{6.1}$$

In the system of equations (6.1), one supposed that the dynamical system is affected by some external influence. The external factors are represented by a set of parameters $\vec{E}$ and enter as arguments of the functions $f_i(\cdot)$. These parameters can be of various origins and stand for different external forces, temperature of the system or noise. Synchronous updating algorithms are used for the numerical simulations in many systems. A typical example is represented by Boolean networks [74]. Also, the numerical integration of ordinary differential equations corresponds effectively to a synchronous algorithm [see, for example, appendix B.2].

6.1.2 Asynchronous Updating Dynamics

An "asynchronous" dynamics is by definition a dynamics which is "non-synchronous". This literally means that not all the state variables of the system evolve at the same time. In this thesis, I focus only on a particular kind of asynchronous dynamics that is quite often used in statistical physics. At each instant of time, only one unit i of the system is selected at random out of all the units of the system and the variable σ_i assigned to this unit is updated according to the rule

$$\sigma_i(t+dt) = f_i\left[\sigma_1(t), \sigma_2(t), \ldots, \sigma_i(t), \ldots, \sigma_N(t), \vec{E}\right] \quad , \tag{6.2}$$

with time increment dt in general different from 1 as chosen in Eq.(6.1) [see below]. None of the other state variables evolve, in the sense that $\sigma_j(t+dt) = \sigma_j(t)$, $\forall\ j \neq i$. At each step of the algorithm, each unit has generally the same probability of being selected for being updated according to the evolution rule of Eq.(6.2). This means that after N selections, on average, all units have been updated once. The increment of time dt is set equal to $1/N$, since this allows comparison of the asynchronous version of the algorithm [i.e., Eq.(6.2)] with its synchronous version [i.e., Eq.(6.1)] on the same time scale. I have already considered an asynchronous algorithm of this type in **Paper IV** [124] and **Paper V** [123] . In statistical physics one refers to this kind of asynchronous algorithms as local stochastic algorithms [107]. Local stochastic algorithms have a broad range of applications from spin models to social models and are the most often used algorithms in statistical physics.

6.1.3 Remarks

There is an open debate about which of the two updating schemes is the more "realistic." For some models, an asynchronous updating scheme is more appropriate, while for other models an application of synchronous dynamics is preferred. There are few cases in which the particular choice of the updating scheme does not change the features of the model [23, 111], while in general, in out-of-equilibrium processes, one can say that the dynamical properties of the model are strongly affected by the particular choice of the dynamics [24, 51].

6.2 The Ising Model

In this section, I briefly review the definition and some basic properties of the Ising model, since this model is chosen in **Paper VI** [122] . I note that this section is very far from being an exhaustive introduction to the Ising model since the applications of this model are the main topic of several books of statistical physics like [85, 107].
The Ising model was proposed as a model that can explain the magnetic properties of matter and in particular to predict the right critical exponents of the ferromagnetic-paramagnetic transition at the Curie temperature of iron [see chapter 2 for detailed description of phase transitions and critical exponents]. The Ising model is the simplest formulation of a general class of models called spin models. It is named after the physicist E.Ising who solved the one-dimensional case in 1925.
In the Ising model, the variable σ_i associated with the unit i of the system can take two possible values or spins, generally denoted by $+1/2$ and $-1/2$, or $+1$ and -1, or simply "up" and "down". One can assume, without loss of generality, that $\sigma_i = \pm 1$. The temperature

T of the system is included in the problem. In principle, one can also take into account the effect of an external magnetic field h, but here I do not consider this effect. Given a particular configuration of the system $\vec{s} = \{\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_N\}$, the energy of the configuration $\vec{s}$ is given by the Ising Hamiltonian

$$H(\vec{s}) = -\sum_{i \neq j} J_{i,j}\sigma_i\sigma_j \quad . \tag{6.3}$$

The factor $J_{i,j}$ represents the coupling between the units i and j of the system. When the coupling between the units i and j is present, either $J_{i,j} = 1$ or $J_{i,j} = -1$ are generally chosen. In case of the absence of the coupling between the units i and j, $J_{i,j} = 0$ is chosen. For $J_{i,j} = 1$ there is ferromagnetic interaction between the units i and j, while for $J_{i,j} = -1$ the interaction is anti-ferromagnetic. The matrix J can therefore be viewed as the adjacency matrix of the weighted graph associated with the system (i.e., the weights are just $+1$ or -1). I invite the interested reader to refer to the definition of weighted graphs given in section 1.2.1. All configurations with the same energy have the same probability of being realized. At temperature T, the probability $P(H, T)$ of finding the system in a particular configuration with energy H obeys the Boltzmann distribution

$$P(H, T) = \exp\left(\frac{-H}{kT}\right) \quad , \tag{6.4}$$

where k is the Boltzmann constant. Theoretically, the model can be solved by finding the explicit expression for the **partition function** associated with the Hamiltonian. This is in general a very hard task. At the moment, exact solutions of the Ising model are known only for the case in which the underlying topology is a one-dimensional or a two-dimensional lattice or in case of all-to-all topologies.

Nevertheless, it is possible to solve the Ising model either analytically in certain approximation or numerically. The algorithm that is often used for numerical solutions is known as the Metropolis algorithm [91]. I do not want to give a detailed description of this algorithm here since it is very well described in every textbook on statistical physics and also mentioned in detail in my **Paper VI** [122] . I would only like to note that the Metropolis algorithm is an asynchronous algorithm and it always leads to the Boltzmann distribution of the equilibrium configurations, which are representative for the thermodynamic equilibrium. In my **Paper VI** [122] I am, however, not interested in the thermodynamic equilibrium.

Paper VI

Phase transition between synchronous and asynchronous updating algorithms

Filippo Radicchi, Daniele Vilone and Hildegard Meyer-Ortmanns

submitted to Journal of Statistical Physics (2007)

In the following paper, we propose a particular analysis of the ferromagnetic Ising model on a one-dimensional lattice. The purpose of this study is not the analysis of the equilibrium dynamics of the Ising model itself. We use this model only to point out the big difference that can arise from a different choice in the updating rules for the algorithm. In a simple way, we tune the algorithm from its standard asynchronous realization to a synchronous one. We find a phase transition between an absorbing and an active state as a function of the parameter that tunes the updating scheme. For a specific choice of this parameter, the algorithm amounts to the Metropolis algorithm. The phase transition is further studied by numerically measuring its critical exponents which characterize the universality class of parity conservation.

Personal Contribution

I produced the numerical results reported in the following paper. In particular, in appendix B.4 I describe in detail the numerical implementation of our model.

Phase transition between synchronous and asynchronous updating algorithms

Filippo Radicchi,[1,*] Daniele Vilone,[1] and Hildegard Meyer-Ortmanns[1,†]

[1] *School of Engineering and Science , Jacobs University Bremen , P.O.Box 750561 , D-28725 Bremen , Germany.*

We update a one-dimensional chain of Ising spins of length L with algorithms which are parameterized by the probability p for a certain site to get updated in one time step. The result of the update event itself is determined by the energy change due to the local change in the configuration. In this way we interpolate between the Metropolis algorithm at zero temperature when p is of the order of $1/L$ and L is large, and a synchronous deterministic updating procedure for $p = 1$. As a function of p we observe a phase transition between the stationary states to which the algorithm drives the system. These are non-absorbing stationary states with antiferromagnetic domains for $p > p_c$, and absorbing states with ferromagnetic domains for $p \leq p_c$. This means that above this transition the stationary states have lost any remnants of the ferromagnetic Ising interaction. A measurement of the critical exponents shows that this transition belongs to the universality class of parity conservation.

PACS numbers: 05.70.Ln, Nonequilibrium and irreversible thermodynamics ; 05.50.+q, Lattice theory and statistics ; 64.90.+b, Other topics in equations of state, phase equilibria and phase transitions

I. INTRODUCTION

The issue of synchronous versus asynchronous updating algorithms has attracted much attention in connection with Boolean networks [1, 2], neural networks [3, 4], biological networks [5, 6] and game theory [7, 8]. In a synchronous updating scheme all the units of the system are updated at the same time. Asynchronous updating means that not all units are updated at the same time. The algorithm can be asynchronous in the sense that each unit is updated according to its own clock, as in distributed systems for parallel processing [9], or it is asynchronous in the sense that only one randomly chosen unit is updated at each step, as in Monte Carlo algorithms [10]. In the context of thermal equilibrium dynamics, updating algorithms like the Metropolis algorithm [11] are designed in a way that they drive the configurations to a set that is representative for the Boltzmann equilibrium distribution. The dynamics of the algorithm then enters only in an intermediate step, it is not representative for the intrinsic equilibrium dynamics of the system that is determined by the Hamiltonian. In out-of-equilibrium systems the updating scheme plays a more prominent role. The number of attractors in Boolean networks, for example, increases exponentially with the system size [12] for synchronous update, and with a power for critical Boolean networks [13] for asynchronous update. The phase diagrams of the Hopfield neural network model [3, 14] and the Blume-Emery-Griffiths model [15] depend on the updating mode as well, while those of the Q-state Ising model [16, 17], and the Sherrington-Kirkpatrick spin glass [18] are independent on the used scheme.

Probably neither a completely synchronous nor a random asynchronous update is realistic for natural systems. Here we interpolate between these two extreme cases not in a more realistic way, but in a way that allows to identify a phase transition between the stationary states. Synchronous updates may appear as non-physical updates for an Ising model as we consider later, since they do not lead to an equilibrium at minimal energy. Our main goal here is not to propose the synchronous update as another physically relevant mode, but to illustrate the sensitive dependence of the stationary states on the updating scheme even in case of a simple model like a chain of Ising spins. In the context of biological networks such as the yeast cell cycle [19], it depends in general on the concrete dynamics and the function of the network whether the biological "updating" has to proceed in a well ordered deterministic way, or whether it can proceed more or less randomly. Therefore it seems to us not to be meaningful to call synchronous updates the non-physical and asynchronous the physical modes, it depends on the actual case considered. As we shall show, our interpolation by means of a single parameter p is equivalent to the cellular automaton version of the non-equilibrium kinetic Ising model (NEKIMCA) introduced in [20] and further considered in [21]. The equivalence refers to the elementary processes in terms of diffusion, annihilation and even-branching processes that preserve parity. Therefore the phase transition between the stationary states should belong to the parity conservation universality class. This is the universality class shared also by the models of branching and annihilation random walk with an even number of offspring [22], the non-equilibrium kinetic Ising model (NEKIM) [20, 23] and the cellular automaton version NEKIMCA [20]. Our analysis of critical exponents confirms this expectation.

*f.radicchi@gmail.com

†h.ortmanns@jacobs-university.de

II. THE DYNAMICS

We consider a one-dimensional lattice of length L. To each site i, $i = 1, \ldots, L$, of this chain we assign a spin variable σ_i, where σ_i takes the values $+1$ or -1. The Hamiltonian of the system is given by $H = -J\sum_{i=1}^{L}\sigma_i\sigma_{i+1}$, where J is the coupling constant between neighboring sites. Here we consider ferromagnetic couplings so that $J = 1$. Moreover, for definiteness we choose periodic boundary conditions so that $\sigma_{L+1} = \sigma_1$. The results will not depend on this choice.

A canonical local stochastic updating algorithm for studying thermodynamic equilibrium features of the Ising model is the Metropolis algorithm [11]. Given a configuration of the system $\Sigma(t) = \{\sigma_i(t)\}$ at time t, we pick up one site j at random and flip its spin with probability $P_j(t) = \min\{1, \exp[-\beta\Delta E_j(t)]\}$. Here $\beta = 1/kT$, with k the Boltzmann constant and T the temperature, while $\Delta E_j(t) = 2\sigma_j(t)[\sigma_{j-1}(t) + \sigma_{j+1}(t)]$ is the difference in local energy that a flip of σ_j would induce. In particular, for zero temperature an increase in energy of the resulting configuration is always rejected. After the single update of the j-th site the time increases by $t \to t + 1/L = t'$ (and L single updates are considered as one time unit). The new configuration is given by $\Sigma(t') = \{\sigma_i(t); \sigma_j(t')\}$, where all sites $i \neq j$ have the same spin value as they have at time t and only the spin of the j-th site may be flipped. In the context here, we emphasize that the Metropolis algorithm is fully asynchronous in the sense that we have at most one spin flip per single update event.

Here we are no longer interested in the equilibrium properties of the Ising model, but in the algorithmic dynamics itself when it is applied to Ising spins. Therefore we give up the asynchrony of the algorithm. Given the configuration $\Sigma(t) = \{\sigma_i(t)\}$ at time t, we visit all sites and select each of them with probability p as a candidate to get flipped. The selected sites are $j_1, \ldots, j_m$ and m, the total number of selected sites, is a random integer obeying the binomial distribution $B(m, L, p) = \binom{L}{m}p^m(1-p)^{L-m}$. Each of the m selected sites is then updated according to the Metropolis rule at temperature zero, so that the spin of the j_v-th site is flipped with probability $P_{j_v}(t)$, $\forall\ v = 1, \ldots, m$. After one step of the algorithm, the time increases as $t \to t + p = t'$, and the new configuration is $\Sigma(t') = \{\sigma_i(t); \sigma_{j_1}(t'); \ldots; \sigma_{j_m}(t')\}$, where all sites $i \neq j_1, \ldots, j_m$ have the same spin value as they have at time t, while the spins of the m selected sites are attempted to be synchronously flipped. One time unit has passed when the average number of update events equals to the total number of sites L. By varying p we can "tune" the algorithm from asynchrony, for p of the order of $1/L$, to synchrony, for $p = 1$. In particular, for p of the order of $1/L$ and for sufficiently large values of L we recover the usual Metropolis algorithm, that is, when two or more simultaneous selections, which occur at most with probability of the order of $1/L^2$, become negligible. Differently from the standard Metropolis algorithm at zero temperature, for general values of p, the total energy of the new configuration $\Sigma(t')$ can be increased with respect to the total energy of the old configuration $\Sigma(t)$. Moreover, the large-time configurations $\Sigma(t \to \infty)$ do no longer obey the Boltzmann distribution. In particular, in the fully synchronous update, all spins (due to $p = 1$) having a neighboring spin of opposite sign before the synchronous update, will be updated and flipped at the same time so that the global energy of the resulting configuration may increase.

In this paper, for simplicity, we focus on the case of zero temperature. The equilibrium ground-state of the one-dimensional Ising ferromagnet at zero temperature is one of the two ferromagnetic states with all spins positive or negative. In contrast, the completely synchronized dynamics does not drive the system to the ground state, but acts as parallel algorithm and amounts to a deterministic map T: $\Sigma(t+1) = T\Sigma(t)$, for all t. After a transient time $t_0 \leq L/2$, the algorithm drives the system into a cycle of length two [17], where the system "flips" between two configurations, $\Sigma_1 = \{{}^1\sigma_i\}$ and $\Sigma_2 = \{{}^2\sigma_i\}$, such that $\Sigma_2 = T\Sigma_1$ and $\Sigma_1 = T\Sigma_2$, for all $t \geq t_0$. In particular, these configurations result from each other by an overall flip of signs in the sense that $\Sigma_2 = \{-{}^1\sigma_i\}$ and $\Sigma_1 = \{-{}^2\sigma_i\}$. Therefore, it is natural to study intermediate values of p, in particular to focus on the transition between the ferromagnetic and the anti-ferromagnetic configurations of the final state.

Let us consider the active bonds of the system, where we define a bond as active if it connects two sites with opposite spins. As remnant of the zero-temperature Ising model only sites belonging to at least one active bond can flip and do flip if they are selected as candidates for the updating. Only a few elementary processes, that involve active bonds, can take place: diffusion, annihilation and creation in the sense of even-branching processes. To have a clear notation, let us indicate as $\uparrow$ a site with positive spin and as $\downarrow$ a site with negative spin. Consider, for example, a local configuration such as $\cdots\uparrow\uparrow\downarrow\downarrow\cdots$ at time t: at time $t+p$ it can evolve to $\cdots\uparrow\uparrow\uparrow\downarrow\cdots$ or to $\cdots\uparrow\downarrow\downarrow\downarrow\cdots$, depending on whether they are selected for an update that happens with probability $2p(1-p)$ [diffusion or random walk], or it evolves to $\cdots\uparrow\downarrow\uparrow\downarrow\cdots$ with probability p^2 [branching], or it remains unchanged with probability $(1-p)^2$. Using the same rules, a local configuration such as $\cdots\uparrow\uparrow\downarrow\uparrow\uparrow\cdots$ at time t, later, at time $t+p$, can become $\cdots\uparrow\downarrow\uparrow\downarrow\uparrow\cdots$ with probability p^3 [branching], or $\cdots\uparrow\uparrow\uparrow\uparrow\uparrow\cdots$ with probability $p(1-p)^2$ [annihilation], etc.. These three processes, diffusion (random walk), annihilation and branching, correspond to the three elementary processes in the NEKIM [20, 23], that is random walk and annihilation due to a Glauber-spin flip transition rate, and spin exchange with branching due to a Kawasaki-exchange rate, and in common with them, they preserve parity of active bonds (called kinks in the NEKIM). Parity is the number of active bonds modulo 2. The Glauber-spin flip transition rate, as it is used in NEKIM, is a generalization of the spin-flip

rate as it was used in the seminal paper by Glauber [24]. Glauber predicted the time evolution of the expectation value of a single spin and spin products at temperature T if the spin-flip rate is parameterized by two parameters, γ and α; the parameter γ determines the tendency of a spin towards alignment with the neighbors (it is one for zero temperature, so that our algorithm corresponds to $\gamma = 1$), and α sets the time scale on which the transitions take place; in our parametrization this scale is determined by p. In addition to the processes in Glauber's dynamics we have branching processes. Their frequency of occurrence depends on the choice of the parameter p. For p of the order of $1/L$ and $T = 0$ we obtain a ferromagnetic stationary state in agreement with Glauber's predicted time behavior of the total magnetization for $\gamma = 1$, the case in which the magnetization does not decay with time.

III. RESULTS

The former considerations suggest that the transition between ferromagnetic and anti-ferromagnetic behavior (without active bonds and with dominance of active bonds, respectively) is given by the competition of annihilation and branching of active bonds. In particular, the branching is favored by a synchronous updating scheme, because a new couple of active bonds can be created only if two neighboring sites simultaneously flip their spins, so that the flips amount to an exchange of spins. Therefore the transition between stationary states (ferromagnetic and anti-ferromagnetic ones) can be considered as a transition between the asynchronous/synchronous updating schemes. Similarly, in case of the NEKIMCA, that is a stochastic cellular automaton version of the NEKIM, parity-conserving branching of active bonds is generated due to the synchronous update of neighboring sites (without an explicit spin exchange rate as in the mere NEKIM). Favoring active-bond branching due to synchronous updates in the NEKIMCA, also there supports the convergence to an active state as in our case. Figure 1 with active bonds represented as black dots shows a qualitative picture of the phase transition.

We plot the time evolution of an isolated active bond (that is not the only one in the system) for three different values of p. We only display the intermediate part of the lattice around the initial active bond. The initial configuration is chosen as $\cdots\uparrow\uparrow\downarrow\downarrow\cdots$. In the supercritical regime $p > p_c$, the average number of branching is larger than the average number of annihilations, so that the active bonds spread over the entire system. In the critical and subcritical regime $p \leq p_c$, annihilation and branching processes are balanced and the active bonds do not spread over the system, but remain confined to a finite region.

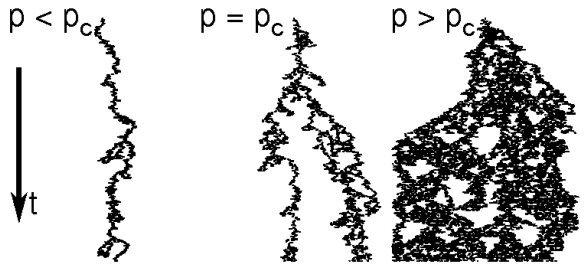


Figure 1: Evolution of an isolated active bond in the subcritical, critical and supercritical regimes. We show only the part around the initial active bond. From left to right: $p = 0.39$, $p = 0.41 = p_c$, $p = 0.43$.

A. The critical point

In order to obtain a quantitative description of the transition, we use as order parameter the density of active bonds

$$\rho = \frac{1}{2L} \sum_{i=1}^{L} 1 - \sigma_i \sigma_{i+1} \quad , \tag{1}$$

given by the ratio of the number of active bonds and the total number of bonds in the lattice. The initial condition is chosen as $\rho(0) = 1$, so that the lattice is fully occupied with active bonds. All data points are obtained from averaging over at least 10^3 realizations and up to 10^5 realizations for small sizes of the lattice. Here, these values of L are considered as the large-volume limit in the following. Let us first determine the critical probability p_c.

In Figure 2A) we plot the time behavior of ρ for three different values of p and $L = 10^4$. As we can see, for $p = 0.41$ $\rho(t) \sim t^{-\delta}$ with $\delta = 0.286(1)$, while for $p = 0.40$ $\rho(t)$ decreases with negative curvature, for $p = 0.42$ $\rho(t)$ increases with positive curvature. The positive curvature characterizes the different phase. Therefore, within the given accuracy, we locate the critical threshold p_c in the large-volume limit as the largest value for which the curvature is non-positive. In this way we obtain $p_c = 0.41(1)$ as the critical point. In order to determine p_c with higher precision we should increase the linear size of the lattice L and the time for observing the positive curvature of ρ. Here we do not perform this kind of computationally expensive simulations, but calculate those critical exponents which do not require a higher precision in p_c. The fact that we observe a phase transition at $p_c > 0$, corresponding to a finite branching rate $\sigma_c > 0$, can be understood from the analytic theory of branching and annihilating random walks by Cardy and Täuber [25]. By means of a field-theoretic dynamical renormalization group, Cardy and Täuber identified the fluctuations which are responsible for shifting the critical value of the branching rate σ_c to values larger than zero for dimensions $d < d'_c = 4/3$, in particular for $d = 1$

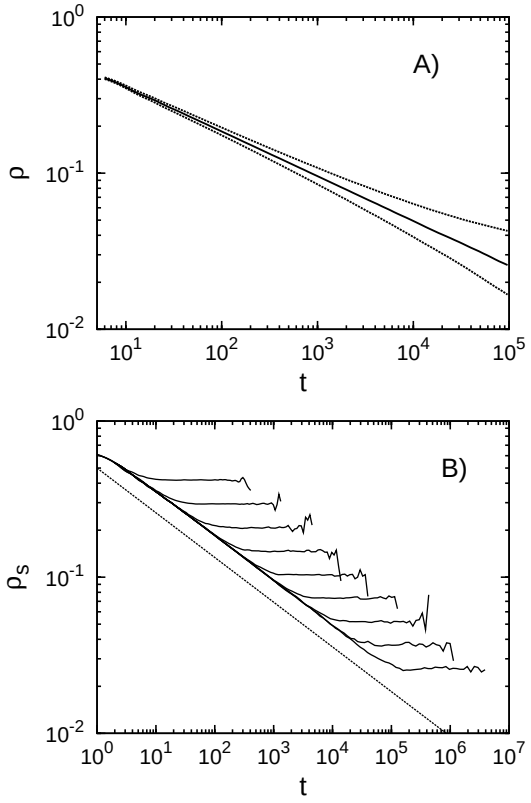


Figure 2: **A)** Time decay of the density of active bonds ρ in the subcritical ($p = 0.40$), critical ($p = 0.41 = p_c$) and supercritical ($p = 0.42$) regime, from bottom to top, respectively. **B)** Average density of active bonds ρ_s surviving up to time t over samples, plotted as a function of time at the critical point for $L = 16, 32, 64, 128, 256, 512, 1024, 2048, 4096$ and 6144, from top to bottom respectively. The dotted line has a slope of $-\delta = -0.286(1)$.

and m even as in our case, where m enters the branching according to $A \to (m+1)A$, and the number of "particles" (active bonds) is locally conserved modulo two. (The measured values of the critical exponents that we consider in the next section cannot be reproduced by this analytic approach.)

B. Critical exponents

a. Critical exponent δ Along the determination of p_c we have already read off the exponent δ from the time evolution of ρ: for $p \leq p_c$ and in the thermodynamic limit $L \to \infty$, we observed $\rho(t) \sim t^{-\alpha(p)}$, with $\alpha(p)$ a continuous and monotonically decreasing function of p and enclosed by curves with $\alpha(p) = 1/2$ for $p \to 0^+$, as in the case of the standard Metropolis algorithm [10, 26], and $\alpha(p_c) = \delta$, as observed in our numerical simulations. Note that the power-law decay does not only hold at p_c, but also below p_c in the large-volume limit. Therefore the negative curvature in the subcritical regime is actually a finite-size effect and not an effect of $p < p_c$. Also in the supercritical regime the curve bends down to zero as a finite-size effect after a sufficiently large time, but the positive curvature signals the onset of the new (supercritical) phase.

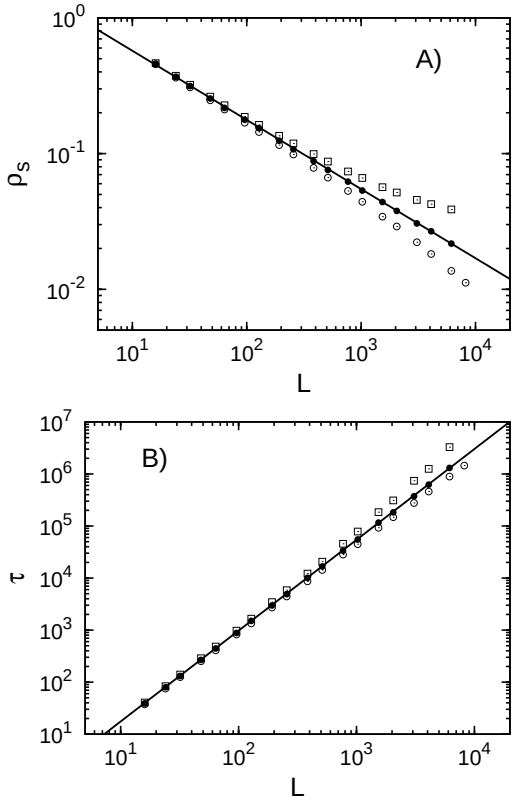


Figure 3: **A)** Average density of active sites over the surviving samples ρ_s as a function of size L of the lattice for $p = 0.40$ (empty dots) 0.41 (full dots) and 0.42 (empty squares). The full line has slope $-\beta/\nu_\perp = -0.51(1)$. **B)** Average relaxation time τ as a function of the size L of the lattice for $p = 0.40$, 0.41 and 0.42 [the symbols are the same as in A)]. The full line has slope $z = \nu_\parallel/\nu_\perp = 1.746(2)$.

1. *Finite-size scaling analysis*

Finite-size effects in the density of active bonds are manifest in two ways: if we follow the time evolution of a certain configuration of a chain of length L, we observe a power-law decay of active bonds up to a certain time τ_d for $p \leq p_c$. After that, either the density drops to zero faster than a power-law, this happens for most configurations, for which a fluctuation drives the system into the absorbing state, or, in the minority of evolutions, the number of active bonds fluctuates around a plateau, before the plateau drops to zero in the end at time τ_s.

Configurations of this minority are called *surviving* configurations up to time τ_s. Now it is easier to locate the onset of the plateau than the onset of a faster decay, and therefore to study the finite-size scaling of the density of active bonds ρ_s of surviving configurations as a function of L, averaged only over the surviving realizations [22].

a. Critical exponents β and $\nu_\perp$ In Figure 2B) we plot the time behavior of ρ_s at the critical point for several values of the size L. As we can see, after an initial transient in which ρ_s decreases as $t^{-\delta}$, it reaches a stationary value depending on L. This value vanishes in the thermodynamical limit, since the plateau is a finite-size effect. We average along the values of the plateau (since averaging over time and over different realizations are assumed to be equivalent), for different system sizes. In Figure 3A) this average value is plotted as a function of L for three values of p ($p = 0.40$, $p = 0.41 = p_c$ and $p = 0.42$). Again, at the critical point we find a power law decay $\rho_s \sim L^{-\beta/\nu_\perp}$, with $\beta/\nu_\perp = 0.51(1)$, while the decay deviates from the power-law behavior in both the subcritical and the supercritical regimes. (Here the exponent β characterizes the behavior of the order parameter $\rho(p) \sim |p_c - p|^\beta$, $\nu_\perp$ the spatial correlation length $\xi(p) \sim |p_c - p|^{-\nu_\perp}$ both in the vicinity of p_c.)

b. Critical exponents $\nu_\parallel$, η and γ Moreover, from the finite-size scaling analysis we calculate the dynamical exponent $z = \nu_\parallel/\nu_\perp$, where $\nu_\parallel$ characterizes the time-like correlation length $\tau(p) \sim |p_c - p|^{-\nu_\parallel}$. The exponent z is derived from the relaxation time τ needed by a finite system to reach the absorbing configuration $[\rho\,(t \geq \tau) = 0]$, starting from a chain fully occupied with active bonds. The unit of the relaxation time is again given by the time that has passed until the average number of update events equals to L. In Figure 3B) τ is plotted as a function of L for $p = 0.40$, $p = 0.41 = p_c$ and $p = 0.42$. At $p = p_c$ we find again a power-law dependence $\tau \sim L^z$, with $z = \nu_\parallel/\nu_\perp = 1.746(2)$, while for different values of p τ behaves differently from a power-law behavior, as it is seen in Figure 3B). Furthermore, in Figure 4 we verify the finite-size scaling relation at p_c

$$\rho\,(L,t) \sim L^{-\beta/\nu_\perp} f\,(t/L^z) \quad , \tag{2}$$

where $f(\cdot)$ is a suitable universal function.

Finally, we determine the static exponent γ from the growth of fluctuations in the order-parameter susceptibility, defined as $\chi_s = L(\langle\rho_s^2\rangle - \langle\rho_s\rangle^2)$, via $\chi_s \sim L^{\gamma/\nu_\perp}$ at p_c (where the index s again refers to an average over the surviving configurations). We find $\gamma \simeq 0$, as it is shown in Figure 5.

In the large-volume limit, we further determine the dynamical exponent η and check the exponents δ and z in the following way. Starting from a configuration like $\cdots\uparrow\uparrow\downarrow\uparrow\uparrow\cdots$, that is a ferromagnetic configuration with only two active bonds, we numerically compute the survival probability $P(t)$ (that is the probability that the system had not entered the absorbing state up to time t), the average number of active bonds $\bar{n}(t)$ and the average mean square distance of spreading $\bar{R}^2(t)$ from an

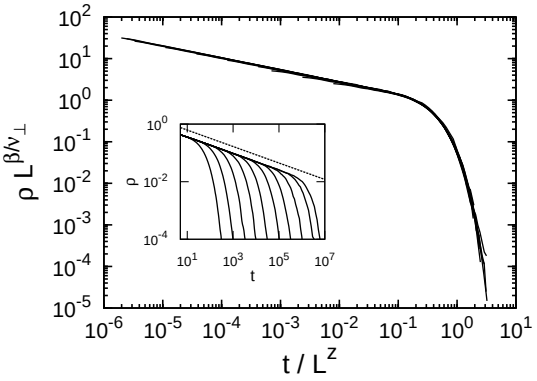


Figure 4: The main plot shows the finite-size scaling of the density of active bonds ρ [Eq.(2)]. The inset shows the unscaled data, where the dotted line has a slope equal to $-\delta$. The datasets are the same as in Figures 2 and 3.

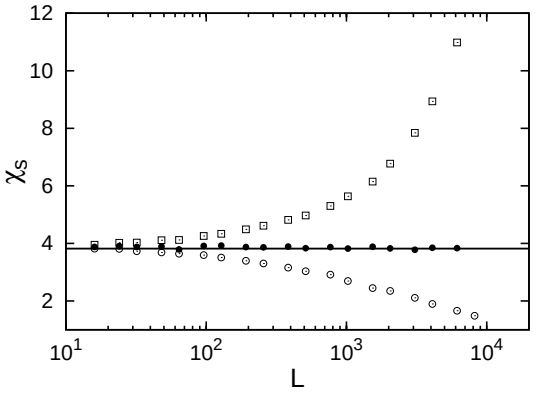


Figure 5: Susceptibility χ_s as a function of the linear size of the lattice L. χ_s is plotted for $p = 0.40, 0.41$ and 0.42, where the symbols adopted are the same as in Figure 3. As it is shown at the critical point $p_c = 0.41$, χ_s is a constant, that means $\gamma \simeq 0$. The datasets are the same as in Figures 2 and 3.

arbitrary selected site. These quantities are expected to behave at the critical point p_c like $P(t) \sim t^{-\delta}$, $\bar{n}(t) \sim t^\eta$ and $\bar{R}^2(t) \sim t^{2/z}$ [27]. The critical exponent δ can be easily calculated by looking at the time behavior of $P(t)$. In Figure 6A) we report the *effective exponent* [27], calculated according to the formula

$$-\delta(t) = \frac{\log_{10}\left[P(t)/P(t/b)\right]}{\log_{10}(b)} \quad , \tag{3}$$

where b is a fit parameter. The right value of the critical exponent δ is then given by

$$\delta = \lim_{t\to\infty} \delta(t) \quad . \tag{4}$$

In our numerical simulations we consider a lattice of size $L = 6144$ and we average over 10^4 realizations. The same

simulations and similar definitions for the effective exponents $\eta(t)$ and $2/z(t)$ lead to the numerical evaluation of η [Figure 6B)] and $2/z$ [Figure 6C)]. From these simulations we find $\eta \simeq 0$ and values of δ and z consistent with the former ones within the error bars. In particular in Figure 6 we choose $b = 5$ for the evaluation of the effective exponents, but we verified the consistency of our numerical measurements also for other values of b.

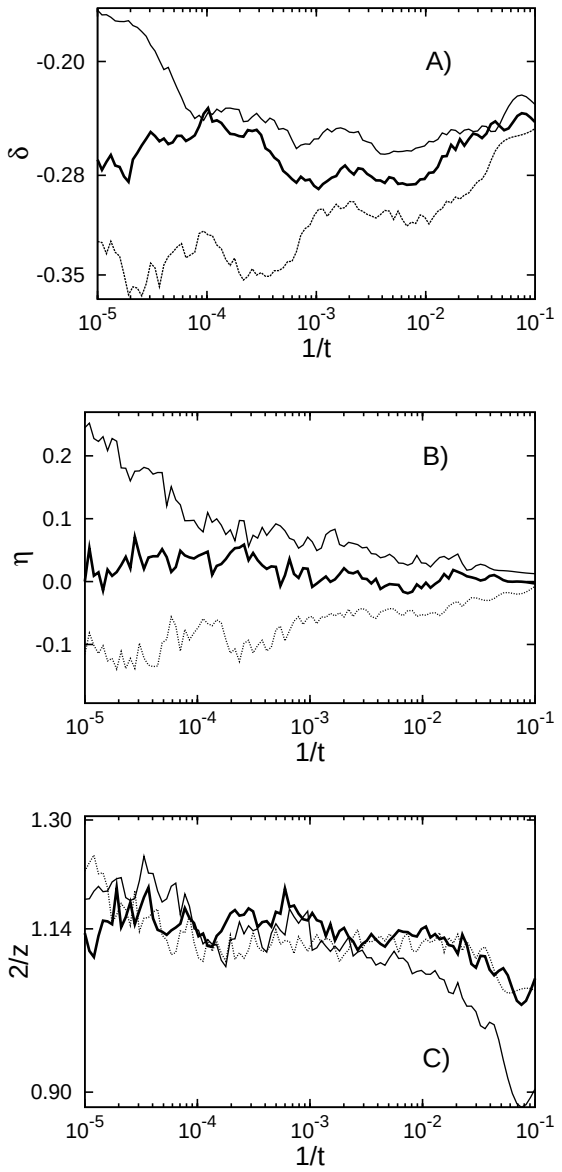


Figure 6: Time dependent behavior of the *effective exponents* $\delta(t)$ (**A**), $\eta(t)$ (**B**) and $2/z(t)$ (**C**). We plot these quantities for different values of p: $p_c > p = 0.4$ (dashed line), $p_c = p = 0.41$ (bold line) and $p_c < p = 0.42$ (solid line).

	$\beta/\nu_\perp$	γ	δ	$\nu_\parallel/\nu_\perp$	η
PC	1/2	0	2/7	7/4	0
SAIM	0.51(1)	0.0	0.286(1)	1.746(2)	0.0

Table I: Critical exponents of the parity conservation (PC) universality class in comparison to the exponents of the synchronous/asynchronous Ising model (SAIM) measured in this paper.

IV. CONCLUSIONS

In conclusion, the critical exponents we have measured and reported in Table I are consistent with the ones conjectured for the parity conservation (PC) universality class [23, 27, 28, 29, 30, 31, 32].

This is expected because periodic boundary conditions for Ising spins as well as the updating rules we used preserve the parity of the number of active bonds. Moreover, for free boundary conditions only the boundaries may violate the conservation of parity that leads to a negligible effect. In terms of elementary processes (diffusion, annihilation and branching) the model is equivalent to the branching and annihilation random walk with an even number of offsprings. Moreover we use the same elementary processes as those used in NEKIM or NEKIMCA, for which the phase transition between active and absorbing states is known to belong to the PC universality class. In contrast to the NEKIM(CA), our algorithm is parameterized in a different, but equivalent way, for which the frequency of branching processes is tuned via the parameter p, that was introduced to directly interpolate between synchronous and asynchronous updates, and not between Glauber and Kawasaki dynamics as in case of the NEKIM. Also the background for both parameterizations is different. The combination of Glauber and Kawasaki dynamics was originally introduced for studying temperature driven non-equilibrium phase transitions. Our focus was the effect of synchronous/asynchronous updating modes on the stationary states in dynamical processes on networks, for which the Ising chain is the simplest case to start a systematic study.

An increase in the dimensionality of the lattice would allow a mean-field description of the transition, since the parity conservation universality class has the critical dimension $d_c = 2$ [27, 28]. Our preliminary numerical simulations for this case show qualitatively similar phases as in $d = 1$, but with the transition point shifted towards 1. Our work represents a systematic interpolation between a synchronous and asynchronous updating scheme. The existence of a phase transition sheds some light on the interpretation of stationary states whenever they depend on the updating mode. Above the transition these states may have lost any remnants of the dynamics below the transition (in our case we end up with an antiferromagnetic state above and a ferromagnetic below the transition). Instead, the stationary states are representative for the dynamics of the updating algorithm itself.

[1] K. Klemm and S. Bornholdt, pre-print arXiv:q-bio/0309013 (2003)
[2] F. Greil and B. Drossel, Phys. Rev. Lett. **95**, 048701 (2005)
[3] J.J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982)
[4] R.O. Grondin, W. Porod, C.M. Loeffler and D.K. Ferry, Biol. Cybern. **49**, 1 (1983)
[5] K. Klemm and S. Bornholdt, Proc. Natl. Acad. Sci. USA **102**, 18414 (2005)
[6] K. Klemm and S. Bornholdt, Phys. Rev. E **72**, 055101 (2005)
[7] B.A. Huberman and N.S. Glance, Proc. Natl. Acad. Sci. USA **90**, 7716 (1993)
[8] H.J. Block and B. Bergersen, Phys. Rev. E **59**, 3876 (1999)
[9] E. Arjomandi, M.J. Fischer and N.A. Lynch, J. ACM **30**, 449 (1983)
[10] M.E.J. Newman and G.T. Barkema, *Monte Carlo Methods in Statistical Physics*, Oxford University Press, Oxford (1999)
[11] N.A. Metropolis, M.N. Rosenbluth, A.H. Rosenbluth, E. Teller and J. Teller , J. Chem. Phys. **21**, 1087 (1953)
[12] B. Samuelsson and C. Troein , Phys. Rev. Lett. **90** , 098701 (2003)
[13] B. Drossel, T. Mihaljev and F. Greil , Phys. Rev. Lett. **94** , 088701 (2005)
[14] J.F. Fontanari and R. Köberle, J. Physique **49**, 13 (1988)
[15] D. Bollé and J. Busquets Blanco, Eur. Phys. J. B **47**, 281 (2005)
[16] R. B. Potts, Proc. Camb. Phil. Soc. **48**, 106 (1952)
[17] D. Bollé and J. Busquets Blanco, Eur. Phys. J. B **42**, 397 (2004)
[18] H. Nishimori, TITECH Report (1997)
[19] F.Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang, Proc. Natl. Acad. Sci. USA **101**, 4781 (2004)
[20] N. Menyhárd and G.Ódor, J.Phys.A: Math.Gen.**28**, 4505 (1995)
[21] G.Ódor and N.Menyhárd, Phys.Rev. E **73**, 036130 (2006)
[22] I. Jensen, Phys. Rev. E **50**, 3623 (1994)
[23] N. Menyhárd and G. Ódor, J. Phys. A: Math. Gen. **29**, 7739 (1996)
[24] R.J. Glauber, J. Math. Phys. **4**, 294 (1963)
[25] J. Cardy and U.C. Täuber, Phys. Rev. Lett. **77**, 4780 (1996)
[26] V. Privman, *Nonequilibrium Statistical Mechanics in One Dimension*, Cambridge University Press, Cambridge (1997)
[27] H. Hinrichsen, Adv. Phys **49**, 815 (2000)
[28] G. Ódor, Rev. Mod. Phys. **76**, 663 (2004)
[29] I. Jensen, J. Phys. A: Math. Gen. **26**, 3921 (1993)
[30] D. ben-Avraham, F. Leyvraz and S. Redner, Phys. Rev. E **50**, 1843 (1990)
[31] D. Zhong and D. ben-Avraham, Phys. Lett. A **209**, 333 (1995)
[32] H.K. Janseen, Phys. Rev. Lett. **78**, 2890 (1997)

Summary, Conclusions and Perspectives

In this thesis, I analyzed phase transitions and critical phenomena on networks. I considered different contexts ranging from purely topological studies to the analysis of dynamical models defined on networks. My work therefore connects two different fields of research in statistical physics: the more recent one of complex networks and the one with a long tradition of phase transitions. These topics were introduced in chapters 1 and 2, respectively, with the aim of giving the reader the basic background needed for a good understanding of the rest of the thesis.
This thesis represents an example of how the ability of statistical physicists to generate and study models of many interacting particles can be useful for understanding problems apparently not directly related to physics, such as biological and social problems. I give here a summary about the topics studied during this thesis. Moreover, I report some considerations and indicate perspectives for each topic.

In chapter 3, I introduced fractal networks and I defined a method for calculating the fractal dimension of real networks. In **Paper I** [165] , I demonstrated that real networks with degree-disassortive mixing are generally fractal and self-similar, while real networks with degree-assortative mixing are not. The fact that a large number of real networks are fractal was a surprising result since fractality and self-similarity were apparently at odds with the small-world feature, but the contradiction was solved by a suitable method of measuring the fractal dimension.
The results are quite amazing and call for further explanations in terms of biological mechanisms why scale-free degree distributions often seem to go along with self-similar features.
In this context, there are several perspectives. Analogously to critical phenomena at the critical point, there is the possibility of defining other critical exponents apart from the fractal dimension. From a theoretical point of view, the conjecture of hyper-scaling relations valid for "critical" networks and, moreover, the identification of different universality classes for different real networks could be extremely interesting. Extensions are possible in case of weighted and directed networks, after an appropriate redefinition of distances and therefore

of boxes for the tiling procedure of the network. I already investigated the case of weighted networks without finding new interesting results, since the fractality of the networks seems to be unchanged by considering the weighted or the unweighted version of the network itself. To my knowledge, the case of directed networks is open.

In chapter 4, I started with the analysis of dynamical systems defined on networks. In particular, during this chapter I considered the case in which every node of the network is a Kuramoto oscillator and the underlying network is representative of the interactions. In **Paper II** [120] , I specialized in the study of d-dimensional lattices of Kuramoto oscillators in the presence of a pacemaker. I showed analytically and numerically that the system of interacting oscillators can reach a phase-locked state, in which all the oscillators rotate with the same frequency, only for a certain range of parameters which depends on the linear size and the dimensionality of the lattice. By sending the dimension of the lattice to infinity I was able to interpret the right scaling (as a function of the system size) of the condition for having phase-locking on networks, a condition that has been numerically studied before in other papers [78]. Moreover in **Paper III** [121] , I studied the same system of Kuramoto oscillators in the presence of a pacemaker, but with interactions of variable range. I showed that in this case the range of parameters for which the system can reach the phase-locked state behaves non-monotonically as a function of the parameter that tunes the range of the interactions and gives rise to reentrant synchronization. In addition in **Paper III** [121] , I proposed a simple explanation of how a pacemaker can be dynamically created by just a simple linear local gradient in the distribution of the natural frequencies of the oscillators of the system. This result is practically important because it justifies the description of an *ad-hoc* defective oscillator as pacemaker, since a dynamically induced one behaves similarly.
In view of practical applications, my results may serve to decide whether a given ensemble of oscillators can be entrained by one pacemaker, if the size of the system is fixed and the pacemaker is placed at a certain position. It is then the depth of the network which one should calculate in order to determine whether one pacemaker is sufficient. Moreover, my results have shown that the Kuramoto dynamics and the possibility of phase entrainment are rather sensitive to the topology of the network so that the topology should be optimized in practical applications.
In this context, there are several possible extensions. First of all, the possibility of studying the same system but in the presence of more than one pacemaker, since the presence of just one defect is quite unrealistic in real systems. I have already performed some numerical simulations of this kind of model. When the pacemakers are not too different, the "strongest" pacemaker of the system wins (strongest in the sense of having the largest natural frequency): the whole system behaves as in the case in which there is only one pacemaker. For pacemak-

ers differing considerably in their frequencies, I observed several interesting phenomena like spirals and vortex formations which destroy the phase-locking of the whole system. There a more accurate analysis is still needed. Moreover, it could be of extreme interest to study the role of pacemakers in different kind of oscillating systems such as pulse-coupled oscillators [95], excitable media like active rotators [135] or FitzHugh-Nagumo oscillators [46, 100] or combinations of them.

In chapter 5, I introduced the concept of social balance and an evolutionary algorithm (k-cycle dynamics) driving a network to a socially balanced configuration. I discussed in detail the broad interest in this model of social scientists, mathematicians and physicists. Moreover, I remark on the strong analogies between the notion of social balance and the one of "satisfiability" in computer science. All these arguments were made precise in my **Paper IV** [124] , in which I completely solved the case of k-cycle dynamics in fully connected networks and randomly diluted networks. Moreover in **Paper V** [123] , I extended the model to two-dimensional lattices, where I numerically found the evidence of the existence of a percolation phase transition with critical exponents that suggest the discovery of a new universality class.
In the social context, these results allow an understanding of real historical alliances, like the bipartite state in the Triple Alliance formed by Germany, Austria-Hungary and Italy and the Triple Entente formed by Britain, Russia and France in the World War I, and provide an explanation of the stability of bipartitions, which actually are frequently observed in real political systems (i.e., democrats and republicans in USA, or left and right parties in most of all the European democracies like in Italy, Germany, etc.). In the context of satisfiability problems in computer sciences, my results show that a biased search for the solution may induce a quantitative change in the time to find the solution: it may turn from unsolvable to solvable in the sense that the time increases algebraically instead of exponentially with the system size.
Also in this context, several future perspectives are possible. First of all, it could be interesting to study the same model on the more realistic scale-free networks. Also, more realistic would be the introduction of short-cuts in two-dimensional lattices in order to study the effect of the small-world feature on the notion of social balance. Possible modifications in the model could be useful. The possibility of considering directed and multi-weighted networks instead of simple undirected and binary-weighted networks will be more realistic since real social relationships are generally not symmetric and do not just represent friends or enemies. Moreover, the introduction of signs also on the nodes of the network and not only on the links would lead to the realization of truly frustrated systems, and analogies with spin glasses would become evident.

In chapter 6, I considered a highly topical debate in the scientific community interested in networks. Since numerical simulations are frequently used for dynamical models, it is extremely important to take care of the way in which dynamical systems are implemented on computers. In this context, the role of the updating procedure for the numerical simulation of a dynamical model is crucial and there is still an open debate. In my **Paper VI** [122] , I showed how the features of a simple model can be affected by a different choice in the updating scheme for the dynamics. In a simple way I tuned the dynamics of the model from an asynchronous scheme to a synchronous one and I numerically demonstrated the existence of a transition between two different phases in the long-time configurations reached by the model. I numerically evaluated the critical exponents of this phase transition: they turned out to be consistent with the ones of a well known universality class.
My results provide a warning not to jump to conclusions in interesting stationary states. The stationary states may sensitively depend on the order of updating events.
Since this topic is very interesting, there are several possibilities for extending my analysis. Essentially, these extensions can be made following two different directions: the modification of the dynamical model and the modification of the underlying network of interactions. A similar numerical analysis of Random Boolean networks [74, 73] would be an example, since it was already shown that the number of attractors follows a different scaling with the size of the network whether one considers asynchronous [51] or synchronous [75] updating schemes, so I would interpolate between these cases.

Appendix A

Analytic Calculations

I describe here the analytic calculations which lead to results already cited in the thesis but for which I did not report the full derivations. I report in section A.1 the full derivation of the depth for Cayley trees and d-dimensional lattices: these calculations are used in **Paper II** [120] . Moreover, in section A.2 I derive the exact solution for the mean-field case of 5-cycle dynamics, analyzed in **Paper IV** [124] .

A.1 Depth for Cayley Trees and d-dimensional Lattices

In **Paper II** [120] , I have calculated the absolute value of the critical threshold $|\Delta\omega/K|_C$ as a function of the radius R of a Cayley tree and as a function of the linear size N of a d-dimensional lattice. I have shown that for d-dimensional lattices $|\Delta\omega/K|_C \sim N^{-d}$, while for a Cayley tree $|\Delta\omega/K|_C \sim \exp(-R)$. The depth D of a lattice is defined as the mean distance over all the oscillators from the pacemaker. Here I explicitly show that for Cayley trees one has $D \simeq R$ (subsection A.1.1), while for d-dimensional lattices one has $D \sim N$, independently of the dimensionality d of the lattice (subsection A.1.2).

A.1.1 Depth for Cayley Trees

As stated in **Paper II** [120] , the pacemaker is placed at the center of the Cayley tree. The maximal distance from the pacemaker is R, while the coordination number is z. Order the nodes per shell and indicate with $n(r)$ the number of nodes per shell. By construction of the Cayley tree one can write

$$\begin{cases} n(0) = 1 \\ n(1) = z\,n(0) = z \\ n(2) = (z-1)n(1) = (z-1)z \\ n(3) = (z-1)n(2) = (z-1)(z-1)z \\ \vdots \end{cases} ,$$

so that

$$n(r) = z(z-1)^{r-1} \quad . \tag{A.1}$$

By definition, the depth D of the network is given by the average distance of the network from the pacemaker, so that in this case one can write

$$D = \frac{\sum_{r=1}^{R} r\,n(r)}{\sum_{r=1}^{R} n(r)} = \frac{\sum_{r=1}^{R} r(z-1)^{r-1}}{\sum_{r=1}^{R}(z-1)^{r-1}} \quad . \tag{A.2}$$

Eq.(A.2) can be easily calculated using the truncated geometric series

$$\sum_{m=0}^{M} x^m = \frac{x^{M+1}-1}{1-x} \quad .$$

One can write

$$\sum_{r=1}^{R}(z-1)^{r-1} = \sum_{r=0}^{R-1}(z-1)^{r} = \frac{(z-1)^R - 1}{z-2} \quad ,$$

$$\sum_{r=1}^{R} r(z-1)^{r-1} = \begin{array}{l} \frac{\partial}{\partial z}\sum_{r=0}^{R}(z-1)^r \\ \frac{\partial}{\partial z}\frac{(z-1)^{R+1}-1}{z-2} \\ \frac{(R+1)(z-1)^R(z-2)-(z-1)^{R+1}+1}{(z-2)^2} \end{array} .$$

Eq.(A.2) can be now written as

$$D = \frac{(R+1)(z-1)^R(z-2)-(z-1)^{R+1}+1}{(z-2)\left[(z-1)^R-1\right]} \quad .$$

In particular for $z \gg 2$ one can find the expected result

$$D \simeq R \quad . \tag{A.3}$$

A.1.2 Depth for d-dimensional Lattices

Consider d-dimensional lattices with $N+1$ sites for each of the d sides. The total number of nodes is $(N+1)^d$, so that the number of nodes per shell should obey the relation

$$\sum_{r=1}^{\tilde{N}} n(r) = (N+1)^d - 1 \quad , \tag{A.4}$$

where $\tilde{N}$ gives the maximal distance from the pacemaker.
The depth can be written as

$$D = \frac{\sum_{r=1}^{\tilde{N}} r\,n(r)}{\sum_{r=1}^{\tilde{N}} n(r)} \quad . \tag{A.5}$$

In particular, I consider only lattices with periodic boundary conditions and even value of N. Since one is interested in the limit of large N, the former restrictions are not significant. Moreover, I calculate analytically the depth D only for the simplest cases of $d = 1$ and $d = 2$. For the case of general values of d, I propose a simple explanation and show numerically that the same relation $D \sim N$ applies to $d > 2$.

One-dimensional Lattice

Consider the following picture, in which a one-dimensional lattice with linear size $N = 6$ and periodic boundary conditions is presented. Each box corresponds to a site and in each box the distance from the pacemaker is reported, when the pacemaker is placed in the middle.

3	2	1	0	1	2	3

r	0	1	2	3	4	5	6
$n(r)$	1	4	8	12	12	8	4

Table A.1: Number of nodes per shell $n(r)$ in case of a lattice with periodic boundary conditions, in $d = 2$ dimension and with linear size $N = 6$.

In this case, one simply writes $n(0) = 1$ and $n(r) = 2$, $\forall\, 1 \leq r \leq N/2$. Therefore,

$$\sum_{r=1}^{\tilde{N}} n(r) = 2 \sum_{r=1}^{N/2} = N \quad ,$$

while the depth in Eq.(A.5) becomes

$$D = \frac{2\sum_{r=1}^{N/2} r}{N} = \frac{(N/2+1)(N/2)}{N} = \frac{N+2}{4} \quad . \tag{A.6}$$

Two-dimensional Lattice

Consider the following figure, in which a two-dimensional lattice of linear size $N = 6$ and periodic boundary conditions is plotted. Moreover I report in each box, corresponding to a site of the lattice, its distance from the pacemaker, placed in the middle.

6	5	4	3	4	5	6
5	4	3	2	3	4	5
4	3	2	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	2	3	4
5	4	3	2	3	4	5
6	5	4	3	4	5	6

As one can see from Table A.1 the number of nodes per shell $n(r)$ can be written as

$$n(r) = 4a(r) \quad ,$$

where

$$a(r) = \min\left[r \ , \ (N+1) - r\right] \quad . \tag{A.7}$$

Note that Eq.(A.7) is symmetric at $r = N/2$. First of all, one can confirm the bound given in Eq.(A.4) [$\tilde{N} = N$ in this case]:

$$\sum_{r=1}^{N} n(r) = 8 \sum_{r=1}^{N/2} r = 8 \frac{N/2\,(N/2+1)}{2} = N^2 + 2N = (N+1)^2 - 1 \quad .$$

Next one calculates

$$\begin{aligned}\sum_{r=1}^{N} r n(r) &= 4\sum_{r=1}^{N/2} r \quad +4\sum_{r=N/2+1}^{N} r(N+1-r) \\ &= 4\sum_{r=1}^{N/2} r \quad +4(N+1)\sum_{r=N/2+1}^{N} r - 4\sum_{r=N/2+1}^{N} r^2 \\ &= 4\sum_{r=1}^{N/2} r \quad +4(N+1)\left(\sum_{r=1}^{N} r - \sum_{r=1}^{N/2} r\right) - 4\left(\sum_{r=1}^{N} r^2 - \sum_{r=1}^{N/2} r^2\right) \\ &= 8\sum_{r=1}^{N/2} r \quad +4(N+1)\left(\sum_{r=1}^{N} r - \sum_{r=1}^{N/2} r\right) - 4\sum_{r=1}^{N} r^2\end{aligned} \quad .$$

In particular one obtains

$$\begin{aligned}8\sum_{r=1}^{N/2} r^2 &= 8\frac{N/2(N/2+1)(N+1)}{6} \\ &= \frac{N+1}{3} 4\left(N^2/4 + N/2\right) \\ &= \frac{N+1}{3}\left(N^2 + 2N\right)\end{aligned} \quad ,$$

$$\begin{aligned}4(N+1)\left(\sum_{r=1}^{N} r - \sum_{r=1}^{N/2} r\right) &= 4(N+1)\left[\frac{N(N+1)}{2} - \frac{N/2(N/2+1)}{2}\right] \\ &= 2(N+1)\left(N^2 + N - N^2/4 - N/2\right) \\ &= 2(N+1)\left(3N^2/4 - N/2\right)\end{aligned}$$

and

$$\begin{aligned}8\sum_{r=1}^{N} r^2 &= 8\frac{N(N+1)(2N+1)}{6} \\ &= \frac{2(N+1)}{3}\left(2N^2 + N\right)\end{aligned} \quad .$$

Putting everything together one finds

$$\begin{aligned}\sum_{r=1}^{N} r n(r) &= (N+1)\left(N^2/3 + 2N/3 + 3/2N^2 + N - 4/3N^2 - 2/3N\right) \\ &= (N+1)\left[N^2\left(1/3 + 3/2 - 4/3\right) + N\left(2/3 + 1 - 2/3\right)\right] \\ &= (N+1)\left(N^2/2 + N\right) \\ &= (N+1)/2\left(N^2 + 2N\right)\end{aligned} \quad ,$$

therefore, inserting the former results into Eq.(A.5), one has

$$D = \frac{N+1}{2} \quad . \tag{A.8}$$

d-dimensional Lattice

I do not calculate explicitly here the case of general dimension d. One can easily deduce the scaling of the depth D. One can say that the number of nodes per shell $n(r)$ grows as $n(r) \sim r^{d-1}$, as the reader has already seen for $d = 1$ and $d = 2$. This scaling is clear since $n(r)$ has the same dimension $(d-1)$ as that of a perimeter in d-dimensions. Using

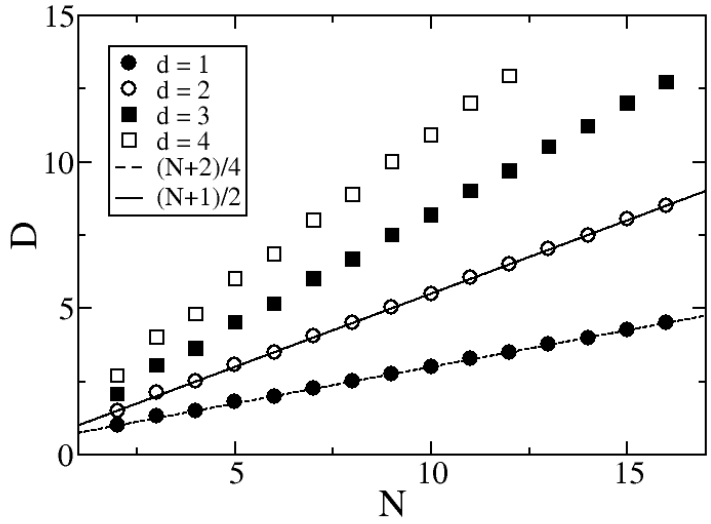


Figure A.1: Numerical determination of the depth D for d-dimensional lattices as a function of the linear size N of the lattice . Here I report the results for $d = 1$ (full circles) , $d = 2$ (empty circles) , $d = 3$ (full squares) and $d = 4$ (empty squares). The dashed line is given by Eq.(A.6), while the full line by Eq.(A.8). From this plot the linear dependence $D \sim N$, independently of the dimension d of the lattice, is evident.

this scaling assumption, passing to the continuum limit and forgetting about the boundary conditions, one can then write

$$D \sim \frac{\int_1^N r\, r^{d-1} dr}{\int_1^N r^{d-1} dr} \sim N \quad .$$

In Figure A.1 I plot the depth D as a function of the linear size N on d-dimensional lattices. D is calculated numerically (symbols); for $d = 1$ and $d = 2$ I also plot the expected exact solutions of Eq.s (A.6) and (A.8). It is obvious that the depth D grows linearly with the linear size N of the lattice, independently of its dimension d.

A.2 Mean Field Solution for 5-cycle Dynamics

In this section, I solve exactly the case of 5-cycle dynamics applied to a fully connected network. This result is part of **Paper IV** [124] , but in this paper no explicit calculation about this result is reported. For clarity, it should be noted that the case of $k = 3$ was solved by Antal *et al.* [9], while the solution for the general value of k is not exact, but just an approximation and is fully described in my **Paper IV** [124] .
I do not repeat here all the calculations used to arrive at the evolution equations of the system, since they are clearly stated in **Paper IV** [124] . Imposing the stationary condition ($\dot{n}_j = 0, \forall\, j$) they can be rewritten as

$$\begin{cases} 5n_0n_2 = 2\left(n_1\right)^2 \\ 2n_1n_3 = \left(n_2\right)^2 \\ 2n_2n_4 = \left(n_3\right)^2 \\ 5n_3n_5 = 2\left(n_4\right)^2 \end{cases} , \tag{A.9}$$

while the condition for a fixed number of friendships at the stationary state is

$$n_5 = \epsilon\left(n_1 + n_3\right) \quad , \tag{A.10}$$

where I defined $\epsilon = 1-2p$ for shortness of notation. I remark the fact that p is the propensity parameter of the model, while I indicate with n_j the density of 5-cycles containing j negative links. Substituting Eq.(A.10) into the last equation of the former system (A.9), one has

$$5\epsilon\left(n_3\right)^2 + 5\epsilon n_1n_3 = 2\left(n_4\right)^2$$

and using the second and the third Eq.s of the the system (A.9) one has

$$10\epsilon n_2n_4 + 5\epsilon/2\left(n_2\right)^2 = 2\left(n_4\right)^2 \quad ,$$

from which

$$4\left(n_4\right)^2 - 20\epsilon n_2n_4 + 5\epsilon\left(n_2\right)^2 = 0$$

and

$$n_4 = n_2\frac{10\epsilon + \sqrt{100\epsilon^2 + 20\epsilon}}{4} = n_2\beta/2 \quad , \tag{A.11}$$

where I defined

$$\beta = 5\epsilon\left(1 + \sqrt{1 + \frac{1}{5\epsilon}}\right) \quad . \tag{A.12}$$

Using Eq.(A.11) in the third equation of system (A.9), one obtains

$$n_3 = n_2\sqrt{\beta}$$

and so on

$$n_1 = \frac{n_2}{2\sqrt{\beta}}$$

$$n_0 = \frac{n_2}{10\beta}$$

$$n_5 = \frac{n_2\beta^2}{10\sqrt{\beta}}$$

Putting all these results into the normalization condition, one has

$$\begin{aligned} 1 = \sum_{j=0}^{5} n_j &= n_2\left(\tfrac{1}{10\beta} + \tfrac{1}{2\sqrt{\beta}} + 1 + \sqrt{\beta} + \tfrac{\beta}{2} + \tfrac{\beta^2}{\sqrt{\beta}}\right) \\ &= n_2\tfrac{\sqrt{\beta}+5\beta+10\beta\sqrt{\beta}+10\beta^2+5\beta^2\sqrt{\beta}+\beta^3}{10\beta\sqrt{\beta}} \\ &= n_2\tfrac{1+5\beta^{1/2}+10\beta+10\beta^{3/2}+5\beta^2+\beta^{5/2}}{10\beta} \\ &= n_2\tfrac{\left(\sqrt{\beta}+1\right)^5}{10\sqrt{\beta}} \end{aligned} \quad .$$

Therefore

$$n_2 = \frac{10\sqrt{\beta}}{\left(\sqrt{\beta}+1\right)^5} \tag{A.13}$$

and furthermore,

$$n_0 = \frac{1}{\left(\sqrt{\beta}+1\right)^5} \tag{A.14}$$

$$n_1 = \frac{5\sqrt{\beta}}{\left(\sqrt{\beta}+1\right)^5} \tag{A.15}$$

$$n_3 = \frac{10\beta\sqrt{\beta}}{\left(\sqrt{\beta}+1\right)^5} \tag{A.16}$$

$$n_4 = \frac{5\beta^2}{\left(\sqrt{\beta}+1\right)^5} \tag{A.17}$$

$$n_5 = \frac{\beta^2\sqrt{\beta}}{\left(\sqrt{\beta}+1\right)^5} \tag{A.18}$$

Looking at Eq.s (A.14), (A.15), (A.13), (A.16), (A.17) and (A.18), it is natural to write them as

$$n_j = \binom{5}{j}\rho_\infty^{5-j}\left(1-\rho_\infty\right)^j \quad , \tag{A.19}$$

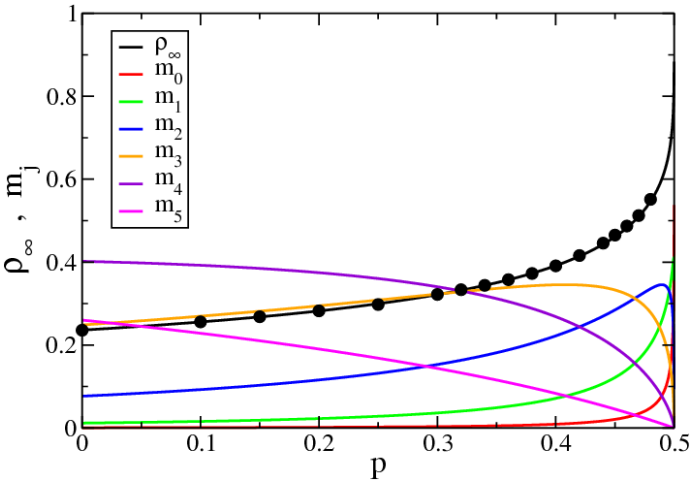


Figure A.2: Stationary densities of 5-cycles with j negative links n_j [from Eq.(A.19)] and stationary density of positive links ρ_∞ [from Eq.(A.20)] as functions of the propensity parameter p. The full dots are the results of numerical simulations performed on a network with $N = 64$ nodes. (from **Paper IV** [124])

with

$$\rho_\infty = \begin{cases} 1/\left[\sqrt{5(1-2p)\left(1+\sqrt{1+\frac{1}{5(1-2p)}}\right)}+1\right] & \text{, if } p \leq 1/2 \\ 1 & \text{, otherwise} \end{cases} . \qquad \text{(A.20)}$$

In Figure A.2, I plot Eq.s(A.19) and (A.20). The theoretical results are in perfect agreement with the results obtained from numerical simulations (see **Paper IV** [124] for more details).

Appendix B

Pseudo-codes of the Numerical Simulations

In this appendix, I present the pseudo-codes of all the numerical simulations performed in papers accumulated in the present thesis. In particular, section B.1 refers to the simulations performed in **Paper I** [165] ; section B.2 refers to the numerical results reported in **Paper II** [120] and **Paper III** [121] ; section B.3 refers to the numerical implementation of the model described in **Paper IV** [124] and **Paper V** [123] ; and finally, section B.4 describes the numerical simulation performed in **Paper VI** [122] . The purpose of this appendix is to allow a reproduction of my numerical measurements and eventually an extension of my results. I chose to report the descriptions of my numerical simulations as pseudo-codes because these descriptions are independent of the specific programming language used for their implementation. The original programs were written in *C* language [76]. The codes are licensed under GPL[1] and can be directly obtained from me by asking to `filippo.radicchi@gmail.com`.

[1] http://www.gnu.org/licenses/gpl.html

B.1 Box Counting Method for Calculating the Fractal Dimension of a Network

In **Paper I** [165] , I implement the box counting method introduced by C.Song *et al.* [140]. This method allows to distinguish whether a network is a fractal or not and, if it is, to determine its fractal dimension d_B [see section 3.2 for the definitions of fractal networks and their fractal dimension d_B]. This short appendix is devoted to the description of the algorithm that I implement in my numerical measurements. A detailed description of possible algorithms able to tile a network according to the formulation of C.Song *et al.* [140] can be found in [139]. The following text describes just one of the possible ways of implementing the box covering method.

Consider a network of size N on which one wants to apply the box covering method. I assume the network is undirected and unweighted [definitions given in section 1.2.1], but in principle the following algorithm can be simply modified for directed and/or weighted networks. First of all, one should store the topological information about the network. This can be made by writing its adjacency matrix [see section 1.2]. Since the majority of real networks are sparse (i.e., almost all the entries of the adjacency matrix are zero), one writes the adjacency matrix in a compact way. One assigns to the i-th node of the network a vector $\vec{M}_i$ in which one collects the indices $j_1, j_2, \ldots, j_{k_i}$ of the k_i nodes directly connected to the node i [$A_{i,j_v} = 1$, $\forall$ $v = 1, \ldots, k_i$]: therefore one has $\vec{M}_i = (j_1, j_2, \ldots, j_{k_i})$. It is appropriate to consider a vector $\vec{S} = (s_1, \ldots, s_i, \ldots, s_N)$, where the generic entry s_i can be a positive integer value or zero meaning the node i has already been assigned to a box or not, respectively. At the beginning of the tiling procedure, one sets $s_i = 0$, $\forall i$. For speeding up the algorithm it is also appropriate to consider a vector $\vec{T}_t = (r_1, \ldots, r_{N_t})$, where the indices of the N_t nodes, not yet at the step of the algorithm t included in any box, are stored. At the beginning, all the nodes of the network are in this vector: therefore, $N_1 \equiv N$. Then one can start with the following procedure, assuming that the linear size of boxes is fixed *a priori* equal to ℓ_B.

1. If the vector $\vec{T}_t$ is not empty, select a node i from the vector $\vec{T}_t$: set $s_i = t$ and delete the node i from $\vec{T}_t$. The node i is the "seed" of the t-th box. Initialize the vector $\vec{Q}$ in which all the nodes of the network at distances less than ℓ_B from the node i and not yet included in any box are stored.

2. As long as the vector $\vec{Q}$ is not empty, select a node j from $\vec{Q}$. Set $s_j = t$ and delete the node j from $\vec{T}_t$.

3. Consider the vector $\vec{Q}_j$ in which all the nodes of the network at a distance from the node j less than ℓ_B are stored. Then rewrite the vector $\vec{Q}$ as the intersection between

the set of nodes reported in $\vec{Q}$ itself and the set of nodes listed in $\vec{Q}_j$. Go back to point 2.

4. Increase $t \to t+1$. Go back to point 1.

At the end of the algorithm, the total number of boxes needed to cover the network is given by $N_B = t - 1$. One should consider all the values of ℓ_B up to $\ell_B = d + 1$, with d equal to the diameter [see section 1.3.2], because one automatically has $N_B(\ell_B) = N$, $\forall\, \ell_B > d$. By plotting $N_B(\ell_B)$ versus ℓ_B, one can see whether the network satisfies the power-law behavior of Eq.(3.1) or whether it does not. In the former case it makes sense to determine the exponent d_B by a simple linear fit in a double-logarithmic plot.
A couple of remarks about the algorithm are in order. The selections, used at point 1 and at point 2 of a node in the vector $\vec{T}_t$ and of a node in the vector $\vec{Q}$, respectively, are in general made according to the degrees of the nodes. I always select the node with the largest degree since this choice should reduce the total number of boxes needed for tiling the network. In reality the power-law behavior of Eq.(3.1) is not affected by this particular criterion of selection, since a different choice, like a purely random selection, changes only the prefactor of the power-law.

B.1.1 Computation of the Distance between Nodes in a Network

The most important remark is about the algorithm able to calculate the distance between two nodes and needed for the construction of the vectors $\vec{Q}$ and $\vec{Q}_j$ of the point 1 and 2, respectively. This algorithm uses the topological information stored in the vectors $\vec{M}_i$, defined above. Imagine that one wants to construct the vector $\vec{Q}_j$ of all the vertices at distance less than ℓ_B from the node j. One makes use of the vector $\vec{C}$ of length N. The i-th entry of this vector, namely C_i, is used for storing the distance between the node j and the node i. At the beginning, one sets $C_i = -1$, $\forall i \neq j$ and $C_j = 0$, where $C_i = -1$ is used as default value indicating that the distance between the nodes j and i has not yet been evaluated. Moreover, one uses the variable c for indicating the actual value of the distance considered during the algorithm. At the beginning one sets $c = 0$. In addition one defines the vector $\vec{V}_c$ as the vector composed of all nodes at distance c from the node j. At the beginning, when $c = 0$, one sets $\vec{V}_c = (j)$. Finally, one sets $\vec{Q}_j = (j)$ at the beginning, since the node j is at distance $0 < \ell_B$ from itself. Then one starts with the following algorithm.

1. Until $c < \ell_B$, consider all the nodes i inside the vector $\vec{V}_c$. Look all the neighbors of i: this can be done by analyzing all the nodes stored in $\vec{M}_i$. Consider the v-th component of the vector $\vec{M}_i$, namely the node $r = M_{i,v}$:

(a) If the distance of r has not yet been computed (if $C_r = -1$), then set $C_r = c + 1$ and add r to the vectors $\vec{V}_{c+1}$ and $\vec{Q}_j$. Go to point 2;

(b) If the distance of r has already been computed (if $C_r \neq -1$), then do nothing.

2. Increase $c \to c + 1$ and go back to point 1.

The former algorithm clearly ends when $c = \ell_B - 1$, since one is interested only in nodes which have a maximal distance $\ell_B - 1$ from the j. At the end the vector $\vec{Q}_j$ contains all the nodes at distance less than ℓ_B from the node j.

B.2 Numerical Integration of the Kuramoto Model

In both **Paper II** [120] and **Paper III** [121] , I performed extensive numerical simulations for the study of the Kuramoto model in the presence of a pacemaker. Here I report the pseudo-code of the simulations and the way in which I determine the phase transition between synchronized and non-synchronized regimes from the numerical simulations.
The system which I am interested is a system of coupled Kuramoto oscillators assigned to a general topology. The topology, either regular or complex, can be easily modelled as a network. One assigns to the generic node i of the network a phase variable $\varphi_i \in \mathbb{R}$. The topology of the underlying network is fully specified by its adjacency matrix [see section 1.2]. Since all the networks which one considers are sparse networks (i.e., almost all the entries of the adjacency matrix are zero), one uses a compact way for storing the adjacency matrix in order to reduce the computation-time and the memory-space needed by the computer program. One assigns to the generic vertex i of the network a vector $\vec{M}_i = (j_1, j_2, \ldots, j_{k_i})$ in which the k_i nodes $j_1, j_2, \ldots, j_{k_i}$ connected to the node i are listed [k_i is the degree of the node i as defined in section 1.3.1]. Once the topology is specified and stored in the $(N+1)$ vectors $\vec{M}_i$, with $i = 0, \ldots, N$, one can start with the simulation.
The Kuramoto model is a dynamical model continuous in time [remember section 1.6.2]. This means that one should numerically integrate the differential equations. One fixes the time increment dt, used for the numerical integration, *a priori*. The initial values for the phases of the oscillators $\varphi_i\,(t=0)$ are fixed in an arbitrary way, generally uniformly distributed between $[0, 2\pi]$. One then applies the following procedure:

1. Run a cycle over all the oscillators of the system. Specifically the phase φ_i of the i-th oscillator evolves as

$$\varphi_i(t+dt) = \varphi_i(t) + \frac{dt}{6}\left[k_1^{(i)}(t) + 2k_2^{(i)}(t) + 2k_3^{(i)}(t) + k_4^{(i)}(t)\right] \quad ,$$

   the coefficients $k_j^{(i)}(t)$, $j = 1, \ldots, 4$, are explained below.

2. The time increases by the fixed amount dt. This means that the actual value of φ_i is replaced by $\varphi_i(t+dt)$, the value calculated at point 1.

This algorithm can be run for an arbitrary time (i.e., number of iterations). The four coefficients $k_1^{(i)}(t)$, $k_2^{(i)}(t)$, $k_3^{(i)}(t)$ and $k_4^{(i)}(t)$ enter the evolution equation, since I use the so-called **Runge-Kutta** algorithm of the fourth-order [54] for the numerical integration. Without specifying the theory beyond this algorithm, one can easily describe them in the specific case of my model by the formulas

$$k_1^{(i)}(t) = \Gamma_i\left[\varphi_i(t), \vec{M}_i\right] \quad ,$$

$$k_2^{(i)}(t) = \Gamma_i \left[\varphi_i(t) + k_1^{(i)}(t)/2, \vec{M}_i\right] \quad ,$$

$$k_3^{(i)}(t) = \Gamma_i \left[\varphi_i + k_1^{(i)}(t)/2 + k_2^{(i)}(t)/2, \vec{M}_i\right] \quad ,$$

and

$$k_4^{(i)}(t) = \Gamma_i \left[\varphi_i + k_1^{(i)}(t)/2 + k_2^{(i)}(t)/2 + k_3^{(i)}(t), \vec{M}_i\right] \quad ,$$

where

$$\Gamma_i \left[\varphi_i, \vec{M}_i\right] = \omega_i + \frac{K}{k_i} \sum_{j=1}^{k_i} \sin\left(\varphi_i - \varphi_{M_{i,j}}\right) \quad , \tag{B.1}$$

and $M_{i,j}$ stands for j^{th}-component of the vector $\vec{M}_i$. In Eq.(B.1) I report the explicit function for the case of interactions between nearest neighbors, while in the case of interactions which depend on the distance between the oscillators, as in the case studied in **Paper III** [121] , this function must be modified.

In order to understand whether the system of oscillators is synchronized or not I used a very simple method. In my numerical simulations the phase of the oscillator i can assume any real value ($\varphi_i \in \mathbb{R}$). Moreover, I indicate by $\Delta\varphi_{i,j}(t) = \varphi_i(t) - \varphi_j(t)$ the phase lag between the oscillators i and j at time t. Two oscillators i and j, which move with the same frequency, should have by definition a constant phase lag : $\Delta\varphi_{i,j}(t') - \Delta\varphi_{i,j}(t) = 0 \quad , \quad \forall\ t', t$. Therefore, when the system is in the synchronized regime the phase lag should be a time-constant for all the pairs of oscillators. When the system is in the non-synchronized regime, however, then at least one phase lag should vary in time. In particular, in the case of my model the absolute value of the phase difference between the pacemaker s and another oscillator i increases in time, which means $|\Delta\varphi_{s,i}(t')| - |\Delta\varphi_{s,i}(t)| > 0 \quad , \quad \forall\ t' > t$. Therefore the procedure is the following. One observes, over a sufficiently large interval of time, the phase lags between each oscillator and the pacemaker: if the absolute values of all these quantities never explode but stay constant, one says that the system is synchronized; if at least one absolute value of these phase lags constantly increases, then the system is in the non-fully synchronized regime.

B.3 Numerical Implementation of the k-cycle Dynamics

In both **Paper IV** [124] and **Paper V** [123] , I reported on numerical results of computer simulations. Here I would like to give a general pseudo-code description of how the k-cycle dynamics is implemented on a computer. Both parameters of the simulation, which are the values of the propensity parameter p and of the length of cycles k used for the dynamics, are fixed *a priori*. The same is true for the topology of the particular social network on which k-cycle dynamics is applied, but not for the signs assigned to the links of the social network since they are permanently changing during k-cycle dynamics.

A convenient way to store the information about the topology of the network is to assign to each node i a vector $\vec{M}_i$ which contains the indices of the m_i nodes $j_1, \ldots, j_{m_i}$ connected to the i-th node. Therefore one can write $\vec{M}_i = (j_1, j_2, \ldots, j_{m_i})$. The sign of the respective links between the i-th node and the connected nodes $j_1, \ldots, j_{m_i}$ can be stored in the same way in a vector $\vec{S}_i = \left(A_{i,j_1}, A_{i,j_2} \ldots, A_{i,j_{m_i}}\right)$, where A is the adjacency matrix of the social network as defined in section 5.1.2. For practical reasons, it is convenient to consider also the k-cycle network associated with the social network. In this network, each k-cycle of the social network is a node and two nodes are connected if the respective k-cycles share at least one link in the social network. The s-th node of the k-cycle network is stored in a vector $\vec{k}_s = ({}^s i_1, {}^s i_2, \ldots, {}^s i_k)$, where ${}^s i_1, \ldots, {}^s i_k$ are the nodes of the social network which compose the s-th k-cycle. In the vector $\vec{k}_s$, the nodes of the social network should be ordered. This means that they should not be sorted in an increasing or decreasing way but ordered in the sense that one should follow the closed path that forms the actual s-th k-cycle [see Figure 5.1(b)]. The adjacency of the k-cycle network is without signed links. Therefore, one needs only one vector $\vec{K}_s$, where the adjacent nodes are stored exactly as previously stated for the nodes of the social network, for the s-th k-cycle of the social networks. For example, if one fixes $k = 3$ and one considers the social network of Figure 5.4, then one has: for the social network $\vec{M}_1 = (2, 4)$ and $\vec{S}_1 = (1, 1)$, $\vec{M}_2 = (1, 3, 4)$ and $\vec{S}_2 = (1, -1. - 1)$, $\vec{M}_3 = (2, 4)$ and $\vec{S}_3 = (-1, 1)$, $\vec{M}_4 = (1, 2, 3)$ and $\vec{S}_4 = (1, -1, 1)$; for the k-cycle network $\vec{k}_1 = (1, 2, 4)$ and $\vec{K}_1 = (2)$, $\vec{k}_2 = (2, 3, 4)$ and $\vec{K}_2 = (1)$, since there are only two triads in the social network.

Moreover, it is useful to introduce a vector $\vec{I}$ in which the imbalanced k-cycles are stored and marked by the same indices as the ones used in the k-cycle network. Again, taking as an example the social network of Figure 5.4, $\vec{I} = (1)$ because the second triad is balanced. Then one can start the following algorithm that simulates the k-cycle dynamics.

1. If $\vec{I}$ is empty, there is no dynamics, since one has already reached a fully balanced configuration. If the vector $\vec{I}$ is not empty, select one index s^* at random out of the

vector $\vec{I}$. The s^*-th k-cycle then performs the following dynamics:

2. If the s^*-th k-cycle contains only negative links, choose two nodes of the social network inside the vector $\vec{k}_{s^*}$, say ${}^{s^*}i_v$ and ${}^{s^*}i_w$, at random and go to point 4

3. If the s^*-th k-cycle contains both negative and positive links, choose at random a number $0 \leq r \leq 1$ out of a uniform distribution, then

   - If $r < p$, then select two nodes of the social network inside the vector $\vec{k}_{s^*}$, say ${}^{s^*}i_v$ and ${}^{s^*}i_w$, at random and such that $A_{{}^{s^*}i_v,{}^{s^*}i_w} = -1$, and go to point 4
   - If $r \geq p$, then select two nodes of the social network inside the vector $\vec{k}_{s^*}$, say ${}^{s^*}i_v$ and ${}^{s^*}i_w$, at random and such that $A_{{}^{s^*}i_v,{}^{s^*}i_w} = 1$, and go to point 4

4. Flip the link of the social network between the nodes ${}^{s^*}i_v$ and ${}^{s^*}i_w$, so that both vectors $\vec{S}_{{}^{s^*}i_v}$ and $\vec{S}_{{}^{s^*}i_w}$ are changed, since the variables $A_{{}^{s^*}i_v,{}^{s^*}i_w} \to -A_{{}^{s^*}i_v,{}^{s^*}i_w}$ and $A_{{}^{s^*}i_w,{}^{s^*}i_v} \to -A_{{}^{s^*}i_w,{}^{s^*}i_v}$ change, respectively.

5. Delete from the vector $\vec{I}$ the selected k-cycle s^*, since after the flip of the link the s^*-th k-cycle is balanced. Then check all its neighbor k-cycles stored in K_{s^*}: add to $\vec{I}$ all k-cycles which became imbalanced after the single flip of the link, while deleting from $\vec{I}$ the ones that became balanced after the single spin flip. Go back to point 1.

At each selection at point 1, the time is increased by $1/L \cdot (I + C)/I$. I stands for the total number of imbalanced k-cycles before the selection [that is the length of the vector $\vec{I}$], while C is the number of balanced ones. Therefore, $(I + C)$ is the total number of k-cycles in the social network, while $(I + C)/I$ is the inverse of the probability of selecting an imbalanced k-cycle out of all possible k-cycles, which approximates very well the time increment as defined in the original application of k-cycle dynamics [see section 5.1.2].
It should be noted that the former algorithm is definitely faster than a mere application of the k-cycle dynamics of section 5.1.2 only for sufficiently diluted social networks, such as random networks with low connectivity or two-dimensional triangular lattices. In fact, if in the vectors K_{s^*} too many neighbor k-cycles are stored, it does not make much sense to apply this same procedure. For example, in the case of fully connected social networks there are exactly $k\binom{N-2}{k-2}$ elements in K_{s^*}, so that the original application of the k-cycle dynamics of section 5.1.2 is much more convenient.

B.4 Numerical Implementation of the Asynchronous and Synchronous Updating Dynamics

In **Paper VI** [122] , I studied the phase transition between the ferromagnetic phase and the anti-ferromagnetic phase of a one-dimensional Ising model, with periodic boundary conditions, at temperature zero. This phase transition is studied as a function of the probability p, which enters in the definition of the updating algorithm. Here I would like to give to the reader a clear knowledge of the implementation of this model in a computer program.
The lattice is composed of L sites and to every site a spin variable is assigned. Therefore it is convenient to store all the spin variables into a vector $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_L)$. Since the temperature of the system is zero, during the evolution only sites with at least one neighbor with different spin are updated. Therefore, one can call these sites "active" and the others "inactive." It is natural to make use of a vector $\vec{a}$ in which all the active sites of the lattice are stored during the evolution. The updating scheme depends on the choice of the probability p that one assumes as fixed *a priori*. At time $t = 0$, one assigns some prescribed values to the spins. Then one starts with the following algorithm.

1. Set $\vec{a} = \emptyset$, then visit all the sites i of the lattice: if $\sigma_i \neq \sigma_{i+1}$, then add i to $\vec{a}$. The periodicity of the boundaries is imposed by the assumption $\sigma_{L+1} = \sigma_1$.

2. Visit all the sites stored in $\vec{a}$. For each of them, say the site i, extract a random number $0 \leq r \leq 1$ out of a uniform distribution: if $r < p$, the site i is selected and its spin flipped $\sigma_i \to -\sigma_i$.

3. Time increases as $t \to t + p$.

This algorithm can run until the vector $\vec{a}$ is empty after the procedure of point 1, which means no active sites are left in the lattice and therefore all spins are aligned, or up to an arbitrary maximal time of observation.

Bibliography

[1] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, *The Kuramoto model: a simple paradigm for synchronization phenomena*, Review of Modern Physics **77** (2005), 137.

[2] A. A. Adronov and A. A. Vitt, *Theory of oscillators*, Gostekhizdat, Moskow, 1937.

[3] R. Albert and A. Barabási, *Statistical mechanics of complex networks*, Reviews of Modern Physics **74** (2002), 47.

[4] R. Albert and A. L. Barabási, *Emergence of scaling in random networks*, Science **286** (1999), 509.

[5] R. Albert and A. L. Barabási, *Mean-field theory for scale-free random networks*, Physica A **272** (1999), 173.

[6] R. Albert, H. Joeng, and A. L. Barabási, *Diameter of the world wide web*, Nature **401** (1999), 130.

[7] R. Albert, H. Joeng, and A. L. Barabási, *Scale-free characteristics of random networks: The topology of the world wide web*, Physica A **281** (2000), 69.

[8] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley, *Classes of small-world networks*, Proceedings of the National Academy of Sciences USA **97** (2000), 11149.

[9] T. Antal, P. L. Krapivsky, and S. Redner, *Dynamics of social balance on networks*, Physical Review E **72** (2005), 036121.

[10] A. Arenas, A. Diaz-Guilera, and C. J. Pérez-Vicente, *Synchronization processes in complex networks*, 2006.

[11] A. Arenas, A. Díaz-Guilera, and C. J. Pèrez-Vicente, *Synchronization reveals topological scales in complex networks*, Physical Review Letters **96** (2006), 114102.

[12] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, *The architecture of complex weighted networks*, Proceedings of the National Academy of Sciences USA **101** (2004), 3747.

[13] A. Barrat and M. Weigt, *On the properties of small-world network models*, The European Physical Journal B **13** (1999), 547.

[14] W. Barthel, A. K. Hartmann, and M. Weigt, *Solving satisfiability problems by fluctuations: The dynamics of stochastic local search algorithms*, Physical Review E **67** (2003), 066104.

[15] C. C. Barton and P. R. La Pointe, *Fractals in petroleum geology and earth processes*, Plenum Press, New York, 1995.

[16] C. C. Barton and P. R. La Pointe, *Fractals in the earth sciences*, Plenum Press, New York, 1995.

[17] G. Bianconi and A. Capocci, *Number of loops of size h in growing scale-free networks*, Physical Review Letters **90** (2003), 078701.

[18] G. Bianconi and M. Marsili, *Loops of any size and hamilton cycles in random scale-free networks*, 2005.

[19] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The theory of critical phenomena: An introduction to the renormalization group*, Oxford University Press, Oxford, 1992.

[20] G. Biroli, R. Monasson, and M. Weigt, *A variational description of the ground state structure in random satisfiability problems*, The European Physical Journal B **14** (2000), 551.

[21] B. Blasius and R. Tönjes, *Quasiregular concentric waves in heterogeneous lattices of coupled oscillators*, Physical Review Letters **95** (2005), 084101.

[22] S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares, and C. S. Zhouc, *The synchronization of chaotic systems*, Physics Reports **366** (2002), 1.

[23] D. Bollé and J. Busquets Blanco, *Two-cycles in spin-systems: sequential versus synchronous updating in multi-state ising-type ferromagnets*, The European Physical Journal B **42** (2004), 397.

[24] D. Bollé and J. Busquets Blanco, *The Blume-Emery-Griffiths neural network with synchronous updating and variable dilution*, The European Physical Journal B **47** (2005), 281.

[25] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, *Graph structure in the web*, Computer Networks **33** (2000), 309.

[26] J. Buck and E. Buck, *Biology of synchronous flashing of fireflies*, Nature **211** (1966), 562.

[27] A. Bunde, W. Dieterich, and H. E. Roman, *Dispersed ionic conductors and percolation theory*, Physical Review Letters **55** (1985), 5.

[28] A. Bunde and S. Havlin, *Fractals and disordered systems*, Springer-Verlag, Berlin, 1991.

[29] D. Cartwright and F. Harary, *Structural balance: a generalization of Heider's theory*, Psychological Review **63** (1956), 277.

[30] C. Castellano, F. Cecconi, V. Loreto, D. Parisi, and F. Radicchi, *Self-contained algorithms to detech communities in networks*, The European Physical Journal B **38** (2004), 311.

[31] C. Castellano and R. Pastor-Satorras, *Non-mean-field behavior of the contact process on scale-free networks*, Physical Review Letters **96** (2006), 038701.

[32] C. Castellano, D. Vilone, and A. Vespignani, *Incomplete ordering of the voter model on small-world networks*, Europhysics Letters **63** (2003), 153.

[33] S. Cocco, O. Dubois, J. Mandler, and R. Monasson, *Rigorous decimation-based construction of ground pure states for spin-glass models on random lattices*, Physical Review Letters **90** (2003), 047205.

[34] K. Wiesenfeld P. Colet and S. H. Strogatz, *Synchronization transitions in a disordered Josephson series array*, Physical Review Letters **76** (1996), 404.

[35] H. Daido, *Lower critical dimension for populations of oscillators with randomly distributed frequencies: A renormalizzation-group approach*, Physical Review Letters **61** (1988), 231.

[36] H. Daido, *Order function and macroscopic mutual entrainment in uniformly coupled limit-cycle oscillators*, Progress in Theoretical Physics **88** (1992), 1213.

[37] R. de Castro and J. W. Grossman, *Famous trails to Paul Erdös*, Mathematical Intellingencer **21** (1999), 51.

[38] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Potts model on complex networks*, The European Physical Journal B **38** (2004), 177.

[39] S. N. Dorogovtsev and J. F. F. Mendes, *Language as an evolving word web*, Proceedings of the Royal Society B (London) **268** (2001), 2603.

[40] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of networks: From biological nets to the internet and www*, Oxford University Press, Oxford, 2003.

[41] R. Engelman and Z. Jaeger, *Fragmentation, form and flow in fractured media*, Adam Hilger, Bristol, 1986.

[42] P. Erdös and A. Rènyi, *On random graphs*, Publicationes Mathematicae **6** (1959), 290.

[43] P. Erdös and A. Rènyi, *On the evolution of random graphs*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences **5** (1960), 17.

[44] M. Faloutsos, P. Faloutsos, and C. Faloutsos, *On power-law realtionships of the internet topology*, Computer Communications Review **29** (1999), 251.

[45] I. J. Farkas, E. Ravasz, A. Schurbert, A. L. Barabási, and T. Vicsek, *Networks in life: Scaling properties and eigenvalue spectra*, Physica A **381** (2003), 601.

[46] R. FitzHugh, *Impulses and physiological states in theoretical models of nerve membrane*, Biophysical Journal **1** (1961), 445.

[47] M. R. Garey and D. S. Johnson, *Computer and intractability: A guide to the theory of NP-completeness*, Freeman, San Francisco, 1979.

[48] M. Girvan and M. E. J. Newman, *Community structure in social and biological networks*, Proceedings of the National Academy of Sciences USA **99** (2002), 7821.

[49] P. Grassberger, *On the spreading of two-dimensional percolation*, Journal of Physics A: Mathematical and General **18** (1985), L215.

[50] P. Grassberger, *Spreading of percolation in three and four dimensions*, Journal of Physics A: Mathematical and General **19** (1986), 1681.

[51] F. Greil and B. Drossel, *The dynamics of critical kauffman networks under asynchronous stochastic update*, Physical Review Letters **95** (2005), 048701.

[52] J. W. Grossman and P. D. F. Ion, *On a portion of the well.known collaboration graph*, Congressum Numerantium **108** (1995), 129.

[53] N. Guelzim, S. Bottani, P. Bourgine, and F. Kepes, *Topological and causal structure of the yeast transcriptional regulatory network*, Nature Genetics **31** (2002), 60.

[54] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations i: Nonstiff problems, second edition*, Springer-Verlag, Berlin, 1993.

[55] F. Harary, *On the notion of balance of a signed graph*, The Michigan Mathematical Journal **2** (1953-54), 143.

[56] F. Harary, R. Z. Norman, and D. Cartwright, *Structural models: An introduction to the theory of directed graphs*, John Wiley & Sons, New York, 1965.

[57] F. Heider, *Social perception and phenomenal causality*, Psychological Review **51** (1944), 358.

[58] F. Heider, *The psychology of interpersonal relations*, John Wiley & Sons, New York, 1958.

[59] C. P. Herrero, *Ising model in small-world networks*, Physical Review E **65** (2002), 066110.

[60] C. P. Herrero, *Ising model in scale-free networks: A Monte Carlo simulation*, Physical Review E **69** (2004), 067109.

[61] H. Hinrichsen, *Nonequilibrium critical phenomena and phase transitions into absorbing states*, Advances in Physics **49** (2000), 815.

[62] H. Hong, M. Y. Choi, and B. J. Kim, *Synchronization on small-world networks*, Physical Review E **65** (2002), 026139.

[63] H. Hong, M. Y. Choi, and B. J. Kim, *Collective phase synchronization in locally-coupled limit-cycle oscillators*, Physical Review E **70** (2004), 045204(R).

[64] H. Hong, B. J. Kim, and M. Y. Choi, *Comment on "ising model on a small world network"*, Physical Review E **66** (2002), 018101.

[65] M. Huxman, S. Beaney, and D. Raffaelli, *Do parasites reduce the chances fo triangulation in a real food web?*, Oikos **76** (1996), 284.

[66] C. Huygens, *Horologium oscillatorium*, Apud F. Muguet, Parisiis, France, 1673.

[67] R. Ferrer i Cancho and R. V. Solè, *The small world of human language*, Proceedings of the Royal Society B (London) **268** (2001), 2261.

[68] H. Jeong, H. Mason, A. L. Barabási, and Z. N. Oltvai, *Lethality and centrality in protein networks*, Nature **411** (2003), 41.

[69] H. Jeong, B. Tombor, R. Albert, and A. L. Barabási, *The large-scale organization of metabolic networks*, Nature **407** (2003), 651.

[70] L. P. Kadanoff, *Scaling laws for ising models near T_c*, Physica **2** (1966), 263.

[71] Y. Y. Kagan and L.Knopoff, *Spatial distribution of earthquakes: the two-point correaltion function*, Geophysical Journal of the Royal Astronomical Society **62** (1980), 303.

[72] V. K. Kalapala, V. Sanwalani, and C. Moore, *The structure of the United States road network*, Preprint, University of New Mexico, 2003.

[73] S. A. Kauffman, *Homeostasis and differentiation in random genetic control networks*, Nature (London) **224** (1969), 177.

[74] S. A. Kauffman, *Metabolic stability and epigenesis in randomly constructed genetic nets*, Journal of Theoretical Biology **22** (1969), 437.

[75] S. A. Kauffman, *The origins of order: Self-organization and selection in evolution*, Oxford University Press, Oxford, 1993.

[76] B. W. Kernighan and D. M. Ritchie, *The C programming language*, Prentice-Hall, Englewood Cliffs, New Jersey, 1978.

[77] D. Knuth, *The art of computer programming: Sorting and searching*, vol. 3, Addison-Wesley, Reading, Massachusetts, 1998.

[78] H. Kori and A. S. Mikhailov, *Entrainment of randomly coupled oscillator networks by a pacemaker*, Physical Review Letters **93** (2004), 254101.

[79] G. Kozyreff, A. G. Vladimirov, and P. Mandel, *Global coupling with time delay in an array of semiconductor lasers*, Physical Review Letters **85** (2000), 3809.

[80] Y. Kuramoto, *Chemical oscillations, waves, and turbolence*, Springer-Verlag, New York, 1984.

[81] S. H. Lee, H. Jeong, and J. D. Noh, *Random field ising model on networks with inhomogeneous connections*, Physical Review E **74** (2006), 031118.

[82] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang, *The yeast cell-cycle network is robustly designed*, Proceedings of the National Academy of Sciences USA **101** (2004), 4781.

[83] T. M. Ligget, *Interacting particle systems*, Springer, New York, 1985.

[84] T. Mora M. Mézard and R. Zecchina, *Clustering of solutions in the random satisfiability problem*, Physical Review Letters **94** (2005), 197205.

[85] S. Ma, *Modern theory of critical phenomena*, WA Benjamin, London, 1976.

[86] B. B. Mandelbrot, *How long is the coast of britain? statistical self-similarity and fractional dimension*, Science **155** (1967), 636.

[87] B. B. Mandelbrot, *Les objets fractals: forme, hasard et dimension*, Flammarion, Paris, 1975.

[88] B. B. Mandelbrot, *The fractal geometry of nature*, W.H. Freeman and Company, New York, 1982.

[89] J. Marro and R. Dickman, *Nonequilibrium phase transitions in lattice models*, Cambridge University Press, Cambridge, 1999.

[90] N. Masuda and N. Konno, *Multi-state epidemic processes on complex networks*, Journal of Theoretical Biology **243** (2006), 64.

[91] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equation of state calculations by fast computing machines*, The Journal of Chemical Physics **21** (1967), 1087.

[92] M. Mézard, G. Parisi, and V. A. Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore, 1987.

[93] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, *Two solutions to diluted p-spin models and xorsat problems*, Journal of Statistical Physics **111** (2004), 505.

[94] S. Milgram, *The small world phenomenon*, Phychology Today **2** (1967), 60.

[95] R. E. Mirollo and S. H. Strogatz, *Synchronization of pulse-coupled biological oscillators*, SIAM Journal on Applied Mathematics **50** (1990), 1645.

[96] R. Monasson and R. Zecchina, *Entropy of the k-satisfiability problem*, Physical Review Letters **76** (1996), 3881.

[97] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, *Determining computational complexity from characteristic 'phase transitions'*, Nature **400** (1999), 133.

[98] Y. Moreno and A. F. Pacheco, *Synchronization of Kuramoto oscillators in scale-free networks*, Europhysical Letters **68** (2004), 603.

[99] Y. Moreno, M. Vazquez-Prada, and A. F. Pacheco, *Fitness for synchronization of network motifs*, Physica A **343** (2004), 279.

[100] J. Nagumo, S. Arimoto, and S. Yoshizawa, *An active pulse transmission line simulating nerve axon*, Proceedings of the Institute of Radio Engineers **50** (1962), 2061.

[101] Z. Néda, E. Ravasz, T. Vicsek, Y. Brechet, and A. L. Barabási, *Physics of the rhythmic applause*, Physical Review E **61** (2000), 6987.

[102] M. E. J. Newman, *Scientific collaboration networks: I. network construction and fundamental results*, Physical Review E **64** (2001), 016131.

[103] M. E. J. Newman, *The structure of scientific collaboration networks*, Proceedings of the National Academy of Sciences USA **98** (2001), 404.

[104] M. E. J. Newman, *Assortative mixing in networks*, Physical Review Letters **89** (2002), 208701.

[105] M. E. J. Newman, *The structure and function of complex networks*, SIAM Review **45** (2003), 167.

[106] M. E. J. Newman and G. T. Barkema, *Monte Carlo methods in statistical physics*, Oxford University Press, Oxford, 1999.

[107] M. E. J. Newman and G. T. Barkema, *Monte Carlo methods in statistical physics*, Oxford University Press, Oxford, 1999.

[108] M. E. J. Newman and J. Park, *Why social networks are different from other types of networks*, Physical Review E **68** (2003), 036122.

[109] M. E. J. Newman and D. J. Watts, *Scaling and percolation in the small-world network model*, Physical Review E **60** (1999), 7332.

[110] E. Niebur, H. G. Schuster, and D. M. Kammen, *Collective frequencies and metastability in networks of limit-cycle oscillators with time delay*, Physical Review Letters **67** (1991), 2753.

[111] H. Nishimori, TITECH Report, 1997.

[112] G. Ódor, *Universality classes in nonequilibrium lattice systems*, Reviews of Modern Physics **76** (2004), 663.

[113] E. Oh, K. Rho, H. Hong, and B. Kahng, *Modular synchronization in complex networks*, Physical Review E **72** (2005), 047101.

[114] E. Oh, K. Rho, H. Hong, and B. Kahng, *Modular synchronization in complex networks*, Physical Review E **72** (2005), 047101.

[115] H. Park, S. Kim, H. B. Pyo, and S. Lee, *Effects of time-delayed interactions on dynamic patterns in a coupled oscillator system*, Physical Review E **60** (1999), 4962.

[116] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A universal concept in nonlinear science*, Cambridge University Press, Cambridge, 2001.

[117] A. S. Pikovsky, M. Zaks, M. Rosenblum, G. Osipov, and J. Kuths, *Phase synchronization of chaotic oscillators*, Chaos **7** (1997), 680.

[118] F. Radicchi, *Metodi di idenficazione di strutture comunitarie in reti complesse*, Master's thesis, Universitá degli studi di Roma "Tor Vergata", Rome, Italy, 2003.

[119] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, *Defining and identifying communities in networks*, Proceedings of the National Academy of Sciences USA **101** (2004), 2658.

[120] F. Radicchi and H. Meyer-Ortmanns, *Entrainment of coupled oscillators on regular networks by pacemakers*, Physical Review E **73** (2006), 036217.

[121] F. Radicchi and H. Meyer-Ortmanns, *Reentrant synchronization and pattern formation in pacemaker-entrained kuramoto oscillators*, Physical Review E **74** (2006), 026203.

[122] F. Radicchi, D. Vilone, and H. Meyer-Ortmanns, *Phase transition between synchronous and asynchronous updating algorithms*, Journal of Statistical Physics (2007), in press.

[123] F. Radicchi, D. Vilone, and H. Meyer-Ortmanns, *Universality class of triad dynamics on a triangular lattice*, Physical Review E **75** (2007), 021118.

[124] F. Radicchi, D. Vilone, S. Yoon, and H. Meyer-Ortmanns, *Social balance as a satisfiability problem of computer science*, Physical Review E **75** (2007), 026106.

[125] A. Rapoport, *Contribution to theory of random and biased nets*, Bulletin of Mathematical Biophysics **19** (1957), 257.

[126] E. Ravatz and A. L. Barabási, *Hierarchical organization in complex networks*, Physical Review E **67** (2003), 026112.

[127] J. Rayleigh, *The theory of sound*, Dover Publishers, New York, 1945.

[128] S. Redner, *How popular is your paper? an empirical study of the citation distribution*, The European Physical Journal B **4** (1998), 131.

[129] J. Reichardt and S. Bornholdt, *Detecting fuzzy community structures in complex networks with a potts model*, Physical Review Letters **93** (2004), 218701.

[130] P. J. Reynolds, H. E. Stanley, and W. Klein, *Large-cell Monte Carlo renormalization group for percolation*, Physical Review B **21** (1980), 1223.

[131] F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, *Simplest random k-satisfiability problem*, Physical Review E **63** (2001), 026702.

[132] F. Ritort, *Solvable dynamics in a system of interacting random tops*, Physical Review Letters **80** (1998), 6.

[133] F. S. Roberts, *On balanced signed graphs and consistent marked graphs*, Electronic Notes in Discrete Mathematics (ENDM) **2** (1999), 94.

[134] I. Rodriguez-Iturbe and A. Rinaldo, *Fractal river basins: Chance and self-organization*, Cambridge University Press, New York, 1997.

[135] H. Sakaguchi, S. Shinomoto, and Y. Kuramoto, *Phase transition and their bifurcation analysis in a large population of active rotators with mean-field coupling*, Progress of Theoretical Physics **79** (1988), 600.

[136] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna, *Small-world properties of the indian railway network*, Preprint, cond-mat/0208535, 2002.

[137] R. Solomonov and A. Rapoport, *Connectivity of random nets*, Bulletin of Mathematical Biophysics **13** (1951), 107.

[138] S. Son, H. Jeong, and J. D. Noh, *Random field ising model and community structure in complex networks*, The European Physical Journal B **50** (2006), 431.

[139] C. Song, L. K. Gallos, S. Havlin, and H. A. Makse, *How to calculate the fractal dimension of a complex network: the box covering algorithm*, 2007.

[140] C. Song, S. Havlin, and H. A. Makse, *Self-similarity of complex networks*, Nature (London) **433** (2005), 392.

[141] C. Song, S. Havlin, and H. A. Makse, *Self-similarity of complex networks*, Nature **433** (2005), 392.

[142] C. Song, S. Havlin, and H. A. Makse, *Origins of fractality in the growth of complex networks*, Nature Physics **2** (2006), 275.

[143] V. Sood and S. Redner, *Voter model on heterogeneous graphs*, Physical Review Letters **94** (2005), 178701.

[144] D. Stauffer and A. Aharony, *Introduction to percolation theory*, Taylor & Francis, London, 1985.

[145] S. H. Strogatz, *From Kuramoto to Crawford: exploring the onset of synchronization in population sof coupled nonlinear oscillators*, Physica D **143** (2000), 1.

[146] S. H. Strogatz, *Sync: The emerging science of spontaneous order*, Hyperion, New York, 2003.

[147] S. H. Strogatz and R. E. Mirollo, *Collective synchronisation in lattices of nonlinear oscillators with randomness*, Journal of Physics A: Mathematical and General **21** (1988), L699.

[148] S. H. Strogatz and R. E. Mirollo, *Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies*, Physica D **31** (1988), 143.

[149] K. Suchecki, V. M. Eguiluz, and M. San Miguel, *Conservation laws for the voter model in complex networks*, Europhysics Letters **69** (2005), 228.

[150] M. Timme, F. Wolf, and T. Geisel, *Coexistence of regular and irregular dynamics in complex networks of pulse-coupled oscillators*, Physical Review Letters **89** (2002), 258701.

[151] M. Timme, F. Wolf, and T. Geisel, *Prevalence of unstable attractors in networks of pulse-coupled oscillators*, Physical Review Letters **89** (2002), 154105.

[152] J. Travers and S. Milgram, *An experimental study of the small world problem*, Sociometry **32** (1969), 425.

[153] B. R. Trees, V. Saranathan, and D. Stroud, *Synchronization in disordered Josephson junction arrays: Small-world connections and the Kuramoto model*, 2004.

[154] B. van der Pol, *Forced oscillations in a circuit with non-linear resistence (reception with reactive triods)*, Philosophical Magazine **3** (1927), 64.

[155] A. G. Vladimirov, G. Kozyreff, and P. Mandel, *Synchronization of weakly stable oscillators and semiconductor laser arrays*, Europhysics Letters **61** (2003), 613619.

[156] T. J. Walker, *Acoustic synchrony: Two mechanisms in the snowy tree cricket*, Science **166** (1969), 891.

[157] D. J. Watts, *Networks, dynamics and the small world phenomenon*, American Journal of Sociology **105** (1999), 493.

[158] D. J. Watts, *Small world*, Princeton University Press, Princeton, 1999.

[159] D. J. Watts and S. H. Strogatz, *Collective dynamics of small world networks*, Nature **393** (1998), 440.

[160] J. G. White, E. Southgate, J. N. Thompson, and S. Brenner, *The structure of the nervous system of the nematode C. Elegans*, Philosophical Transactions of the Royal Society (London) **314** (1996), 1.

[161] K. G. Wilson, *The renormalization group: critical phenomena and the Kondo problem*, Reviews of Modern Physics **47** (1975), 773.

[162] A. T. Winfree, *The geometry of biological time*, Springer-Verlag, New York, 1980.

[163] H. Yamada, *Phase-locked and phase-drift solutions of phase oscillators with asymmetric coupling strengths*, Progress in Theoretical Physics **108** (2002), 13.

[164] M. K. S. Yeung and S. H. Strogatz, *Time delay in the Kuramoto model of coupled oscillators*, Physical Review Letters **82** (1999), 648.

[165] S.H. Yook, F. Radicchi, and H. Meyer-Ortmanns, *Self-similar scale-free networks and disassortativity*, Physical Review E **72** (2005), 045105(R).

[166] R. A. York and R. C. Compton, *Quasi-optical power combining using mutually sychronized oscillator arrays*, IEEE Transactions on Microwave Theory and Techiniques **39** (1991), 1000.

[167] A. N. Zaikin and A. M. Zhabotinsky, *Concentration wave propagation in two-dimensional liquid-phase self-oscillating system*, Nature **225** (1970), 535.

[168] M. A. Zaks, A. B. Neiman, S. Feistel, and L. Schimansky-Geier1, *Noise-controlled oscillations and their bifurcations in coupled phase oscillators*, Physical Review E **68** (2003), 066206.

[169] R. Zallen, *The physics of amorphous solids*, John Wiley, New York, 1983.

Acknowledgements

I would like to thank **Prof. Hildegard Meyer-Ortmanns** for being my PhD supervisor. I am very grateful for all the things I learnt from her, for her good suggestions and advises and for being constantly motivating and directing me to an interesting work. I thank **Prof. Jürgen Fritz**, **Prof. Herbert Jaeger** and **Prof. Albert Díaz-Guilera** for serving on my comprehensive examination committee. Moreover, I thank **Margaret Parks** for her useful corrections in order to have a well written thesis.

I would like to express my gratitude to **Dr. Soon-Hyung Yook** for the time spent together and for his teaching in computer programming. Many thanks also to **Dr. Daniele Vilone** for the good and productive cooperation we had during last year. A particular thank to **Prof. Michael Zaks** for the time he dedicate me and for his lectures in non-linear dynamics. I would like to thank also **Dr. Satya N. Majumdar**, **Dr. Martin Weigt**, **Dr. Martin R. Evans**, **Dr. Sooyeon Yoon**, **Stefan Braunewell** and **Yong-Yeol Ahn** for discussions during their visits at Jacobs University Bremen as well as **Dr. Claudio Castellano** for some discussions in Rome.

I am particularly grateful to the **Jacobs University Bremen** (formerly International University Bremen) for having given me the possibility to perform my PhD studies. Moreover I thank all the members of the university for maintaining a nice atmosphere in the campus, since a nice place is fundamental for the production of a good work.

My biggest thanks are for my family. In particular to my wife **Katia**: I could never complete my PhD studies without her constant love, care and support. A loving thank to my daughter **Marlene**: her birth was the most important moment in my life and since then I am extremely motivated to do the best I can do. A final thank to the rest of my family, my **parents** and my sister **Marzia**. They always support my personal choices and they were always close to me in spite of the geographical distance.

Financial support during the three years of my thesis from the **DFG** (Deutsche Forschungsge-

meinschaft) under the contract number ME 1332/10-1 and 10-2 is gratefully acknowledged.

Curriculum Vitae

FILIPPO RADICCHI

Italy (home):	
via Pupinia , 15	filippo.radicchi@gmail.com
00133 , Roma	phone +39-06-2053369
	mobile +39-349-2123326
Germany (home):	
Süderstrasse, 26	
28199 , Bremen	+39-0421-7942851
Germany (office):	
Jacobs University Bremen	
Campus Ring, 1	f.radicchi@iu-bremen.de
28759 , Bremen	+39-0421-200-3527

Research Interests

Complex networks, complex systems, non-linear dynamics, neural networks, biological physics, information technology.

Academic History

1998-2003 Laurea (Master of Sciences) in Physics
University of Rome II *Tor Vergata*, Rome, Italy

2004-present PhD Student
Jacobs University Bremen
(formerly International University Bremen)
Bremen, Germany

Teaching Experience

2000 Assistant
Didactic Laboratory of Physics
University of Rome II *Tor Vergata*, Rome, Italy

2004 Lecturer
Dynamics of Non-linear Systems: Applications for Social and Biological Problems
Institute of Cognitive Sciences and Technology
Consiglio Nazionale delle Ricerche, Rome, Italy

2004 Teaching Assistant
Statistical Mechanics of Complex Networks
Jacobs University Bremen, Bremen, Germany

2005 Teaching Assistant
Computer Science and Statistical Physics
Jacobs University Bremen, Bremen, Germany

Publications

Journals

1. F.Radicchi, C.Castellano, F.Cecconi, V.Loreto and D.Parisi
Defining and identifying communities in networks
Proc. Natl. Acad. Sci. USA, vol. 101, no. 9, 2658-2663
2004

2. C.Castellano, F.Cecconi, V.Loreto, D.Parisi and F.Radicchi
   Self-contained algorithms to detech communities in networks
   Eur. Phys. J. B, vol. 38 , 311
   2004

3. S.H.Yook, F.Radicchi and H.Meyer-Ortmanns
   Self-similar scale-free networks and disassortativity
   Phys. Rev. E , vol. 72 , 045105(R)
   2005

4. F.Radicchi and H.Meyer-Ortmanns
   Entrainment of coupled oscillators on regular networks by pacemakers
   Phys. Rev. E , vol. 73 , 036217
   2006

5. F.Radicchi and H.Meyer-Ortmanns
   Reentrant synchronization and pattern formation in pacemaker-entrained Kuramoto Oscillators
   Phys. Rev. E , vol. 74 , 026203
   2006

6. F.Radicchi, D.Vilone, S.Yoon and H.Meyer-Ortmanns
   Social Balance as a Satisfiability Problem of Computer Science
   Phys. Rev. E , vol. 75 , 026106
   2007

7. F.Radicchi, D.Vilone and H.Meyer-Ortmanns
   Universality class of triad dynamics on a triangular lattice
   Phys. Rev. E , vol. 75 , 021118
   2007

8. F.Radicchi, D.Vilone and H.Meyer-Ortmanns
   Phase transition between synchronous and asynchronous updating algorithms
   J. Stat. Phys., in press
   2007

Thesis

F.Radicchi
Metodi di identificazione di strutture comunitarie in reti complesse
Master of Sciences Thesis (in Italian)
Department of Physics, University of Rome II *Tor Vergata* , Rome , Italy
2003

Seminars and Summer Schools

Seminars

1. *Defining and Identifying Communities in Networks*
   Complex Systems Research Group
   Medical University of Vienna
   Vienna, Austria
   $1^{st} - 2^{nd}$ of April 2004

2. *Reducing Frustration on Regular and Random Topologies*
   Institute for Scientific Interchange Foundation
   Torino, Italy
   21^{st} of December 2006

3. *Reducing Frustration in Social Systems: Regular and Random Networks*
   Cross-Disciplinary Physics (IMEDEA)
   Mediterranean Institute for Advanced Studies
   Mallorca, Spain
   $22^{nd} - 23^{rd}$ of January 2007

Summer Schools

1. *Design and Control of Self-Organization in Physical, Chemical and Biological Systems*
   The Abdus Salam : International Center for Theoretical Physics
   Trieste, Italy
   25^{th} of July till 5^{th} of August 2005

Personal Data

Date of birth	12^{th} July 1978
Citizenship	Italian
Languages	Italian (native) , English (good) , Spanish (scholastic), German (scholastic)

Computer Knowledge

Operating Systems

Linux	Advanced
Microsoft Windows	Advanced

Programming Languages

C/C++	Advanced
Perl, Php, Java, SQL, HTML	Good

Other **(Advanced/good knowledge)**

Matlab, LaTeX, Macromedia Dreamweaver and Flash, Drupal.