UNIVERSITY OF WARSAW

DOCTORAL THESIS

Statistical physics of coevolving networks

Author:
Tomasz RADUCHA

Supervisor:
prof. dr hab. Ryszard KUTNER

Assistant supervisor:
dr Tomasz GUBIEC

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

in the

Faculty of Physics

Warsaw – Palma de Mallorca, May 2020

List of publications

A significant part of results presented in this thesis has been published in the following articles:

- T. Raducha and T. Gubiec, *Coevolving complex networks in the model of social interactions*, **Physica A: Statistical Mechanics and its Applications** vol. 471: 427-435, 2017

- T. Raducha and T. Gubiec, *Predicting language diversity with complex networks*, **PloS one**, vol. 13.4: e0196593, 2018

- T. Raducha, M. Wiliński, T. Gubiec, and H. E. Stanley, *Statistical mechanics of a coevolving spin system*, **Physical Review E**, vol. 98.3: 030301, 2018

- T. Raducha, B. Min, and M. San Miguel, *Coevolving nonlinear voter model with triadic closure*, **Europhysics Letters**, vol. 124.3: 30001, 2018

- T. Raducha and M. San Miguel, *Emergence of complex structures from nonlinear interactions and noise in coevolving networks*, **arXiv:2004.06515**, 2020

"You don't need something more to get something more."

Murray Gell-Mann

Abstract

UNIVERSITY OF WARSAW

Faculty of Physics

Doctor of Philosophy

Statistical physics of coevolving networks

by Tomasz RADUCHA

Statistical physics has introduced key concepts for analyzing systems consisting of a large number of elements. It showed that problems seemingly out of reach for quantitative description can be actually treated in a strict manner. The essential achievement was a shift in the notion of prediction from deterministic to stochastic ground. The approach of statistical physics has been successfully applied in numerous branches of science, of which the most important one for this thesis is the theory of complex systems.

Networks create the core of complexity science. When a studied object contains many interacting parts and the pattern of interaction is not trivial, networks are the most natural tool to describe it. Some recent publications even identify complex systems as coevolving multilayer networks. One can argue with such a rigorous definition, but the coevolution of structure and state on its own has proved to be a crucial feature in complex systems.

Studies on the coevolution's impact on the behavior of particular models are relatively new. Moreover, a general theory of coevolving networks is so far out of reach. In order to get closer to universal laws we need to understand single problems first. Thus, I seek to explore the effects of coevolution in different models, both analytic (or equilibrium) and algorithmic (or non-equilibrium) ones.

The thesis explores outcomes of introducing coevolving mechanisms in three models, namely the voter model, the Axelrod model and the Ising model. The first one is a reference point in the quantitative description of social systems, however it was successfully applied also in physics, biology, and finance. The voter model is extended in this work to integrate coevolution, triadic closure, nonlinear interactions, and noise. The Axelrod model has a purely social interpretation - it's a model of social interactions or dissemination of culture. It is studied in the thesis how different types of rewiring, during the coevolution of the network's structure and state of the nodes, influence the final topology of the system. Results obtained using the Axelrod model are compared with empirical data. The model was improved for better agreement with the empirically observed scaling behavior. Notably, previous extensions of the model didn't resolve the contradiction with

the empirical data, which is solved here. The Ising model was originally constructed to explain ferromagnetism. However, it gained much bigger attention than the first application might have suggested. It became a reference point for network models and has been studied in many variations, not only empirically implied but also theoretically thrilling. In this abstract context, the spin dynamics from the Ising model is combined in the thesis with topological traits of the nodes to analyze the outcome of a coevolving equilibrium model with structural traits included in the Hamiltonian.

In all models studied throughout this thesis new results are obtained. The most important ones are the following. In the nonlinear coevolving voter model with triadic closure a new shattered phase is observed together with high values of the clustering coefficient. When the noise is included two new phases are obtained. One of these phases persists in the thermodynamic limit. The other one contains topological communities driven by state of the nodes, what was not observed before for the coevolving voter model with noise. Additionally, a new analytical description of the model is proposed. As this model is the most general one, it contains previously studied limit cases like the coevolving voter model or the nonlinear coevolving voter model. Similarly, in the Axelrod model high clustering is generated, as well as a power-law degree distribution. But most importantly, due to the implemented changes the model displays a new scaling of the number of domains with the system size. This result, in contrast to results obtained with the original model, is consistent with empirical data. Finally, an equilibrium model of coevolving networks is proposed for the first time. More precisely, the Hamiltonian of the model includes not only states of the nodes and their mutual interactions, but also degree of the nodes, as a local topological trait. A rich phase diagram obtained in this way is described analytically. The observed configurations coincide with those obtained in non-equilibrium models, suggesting a possibility of their equilibrium description.

The results of the thesis provide a new insight into the behavior of coevolving networks from a statistical physics perspective. The theory of complex systems is not yet complete and new building blocks are being discovered. Hopefully, after analyzing and understanding enough separate parts a more universal theory will emerge. The content of this thesis contributes to the development of such theory by providing several new building blocks. Above all, these blocks take into account coevolution of network's structure and state, as one of the crucial properties of complex systems is their adaptive behavior. The obtained results surely represent advancement in particular models that were studied within the scope of the thesis, but can be also seen as another step towards the theory of complex systems.

Acknowledgements

I would like to thank my supervisor prof. Ryszard Kutner for the support and advice throughout my PhD studies. I am also grateful for his patience in reading and correcting this thesis.

I am grateful to my assistant supervisor dr Tomasz Gubiec with whom I conducted my first scientific research. I appreciated his company and sense of humor in any circumstances.

I am in debt to my friends in this amazing journey – Mateusz Wiliński and Jarosław Klamut. I would like to thank Mateusz for the encouragement without which I would probably not finish my PhD and Jarek for invaluable help at the faculty whenever I was abroad.

I would like to thank prof. Maxi San Miguel for giving me an opportunity to work with him. Results of this work create a significant part of this thesis.

Finally, I would like to thank all people who I had the pleasure to interact with during my scientific work. Intellectual stimulus from such interactions built up my scientific awareness and influenced the shape of this thesis. Surely, a big part of the satisfaction from scientific work comes from the possibility to collaborate with wise people.

Contents

List of Abbreviations

BA	**B**arabási-**A**lbert (model or network)
BGS	**B**oltzmann-**G**ibbs-**S**hannon (entropy)
CVM	**C**oevolving **V**oter **M**odel
ER	**E**rdős-**R**ényi (model or network)
FK	**F**okker-**P**lanck (equation)
ERGM	**E**xponential **R**andom **G**raph **M**odel
MaxEnt	**M**aximal **E**entropy (principle)
MC	**M**onte **C**arlo (e.g. steps)
MST	**M**inimal/maximal **S**panning **T**ree
PMFG	**P**lanar **M**aximally **F**iltered **G**raph
WS	**W**atts–**S**trogatz (model or network)

List of Symbols

The following list contains a general denotation used throughout this work. Some of the models, however, can represent different quantities by these symbols. In case of redundancy the definition given in a chapter takes priority over this list.

k_i	degree of a node i
μ, $\langle k \rangle$	average degree in a network
c	connectivity of a network
c_i	local clustering coefficient of a node i
a_i	number of active links of a node i
ρ_i	fraction of active links of a node i
s_i, σ_i	spin/state of a node i
C	global clustering coefficient
d_{ij}	distance between nodes i and j
d_{max}	diameter of a network
l, $\langle d \rangle$	average path length
G	graph object
V	set of vertices
E	set of edges
A	adjacency matrix
B	incidence matrix
a_{ij}	ith row and jth column of adjacency matrix
N	number of nodes in a network
M	number of links in a network
L	linear size of a square lattice
P	path, or probability
N_{ij}^d	number of paths of length d between i and j
S	size of the largest component, or entropy
n_S	number of components
D	largest domain
n_D	number of domains
K_N	complete graph with N nodes
$K_{n,m}$	complete bipartite graph with $n + m$ nodes
$\mathcal{G}$	set of graphs
$H(X)$	Hamiltonian of configuration X
Z	partition function
ρ	density of active links
m	magnetization
τ	convergence or thermalization time

To my grandparents

Chapter 1

Introduction

1.1 Physics of networks

Network science has tremendously developed over the last two decades and a large part of this development has been achieved thanks to work of physicists. A great number of new methods, algorithms, tools, theories, statistics, and empirical results have been published in physics journals. Yet, the connection between physics and the, undoubtedly interdisciplinary, field of complex networks is sometimes questioned[1]. It is true that graphs, as abstract constructs, have been analyzed by mathematicians for a long time before physicists started to contribute [1]. Also sociologists were using networks when describing social systems already for a while. But there was something unique in the techniques of physics that greatly accelerated the advancement. Unlike mathematics, physics is rooted in empirical research, and in contrast to sociology it aims at discovering general rules behind the whole system, rather then describing individuals. Maybe this combination makes physics approach so fruitful, maybe there is something more. Nevertheless, it is clear that physics played a central role in building the foundations of network science.

1.1.1 The concept of network

The central concept of this thesis is network, being at the same time the main tool. Therefore, I feel obliged to specify what it is not only in the mathematically correct definition, but also in an intuitive way. Of course, for a scientist working on the border of statistical physics and complex systems this term is straightforward, there is not much to explain. Nevertheless, the word *network* is being used in many different contexts in different fields, not to mention everyday non-academic conversations. According to the online Oxford Dictionary[2] the noun *network* can mean (i) *an arrangement of intersecting horizontal and vertical lines* or (ii) *a group or system of interconnected people or things*. The first definition is rather narrow and we would call it a square lattice (which is a frequently analyzed type of a network). The second definition

[1] Although, during the panel discussion at the Conference on Complex Systems 2018 in Thessaloniki, Guido Caldarelli, a famous Italian researcher in the field of complex networks, asked *„is this [complex networks] a new science or is it just a new physics?"*.

[2] https://en.oxforddictionaries.com/

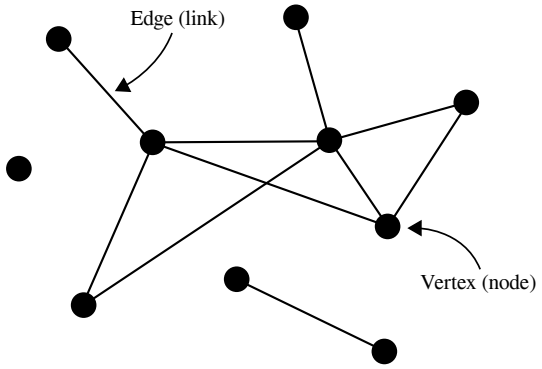


Figure 1.1: A simple graph example.

is much closer to the gist of this work. In fact, the dictionary gives several sub-definitions of the (ii), which are examples of real-world networks studied by many. For instance, *a number of interconnected computers, machines, or operations* is a computer or technological network, *a complex system of railways, roads, or other routes* is a transportation network. These examples also illustrate the every-day intuition. But the basic concept of network is much more elementary – it is a set of objects, called *nodes* or *vertices* (sometimes *agents*), interacting on some level with each other, where every interacting pair can be seen as connected [2]. These *connections* we can also call *links* or *edges*[3]. A network can be further referred to as a *graph* [4]. A graph is usually associated with some kind of a diagram, although in this thesis it is simply a synonym of a network. The jargon here is rich, sometimes even redundant, due to the diversity of fields scientists studying networks come from.

Having an abstract representation of a system we can start exploring it. In the Fig 1.1 you can see a very simple example of a network. Now, put people under the nodes and say the links represent a particular type of interactions between them (daily conversations for instance) and you can analyze social networks within the network science regime. Relate the vertices with a spin of electrons and the edges with pair interactions and you can study ferromagnetic phenomena from the microscopic perspective. For a biologist it can represent proteins interactions, or a part of the food chain, for an engineer it may be a power grid system, a transportation network etc. Obviously, it implies that the description is purely phenomenological. But after all, how one can be sure any theory is not? It is as good as it can teach us about the reality we want to comprehend. And network approach is particularly useful in complex systems. Imagine we have an ultimate theory perfectly describing every part of a system we are studying – given the system is complex enough we have to average or aggregate the description anyways to be able

[3] A strict definition of a network, as a mathematical object, is given in Chapter 2.

[4] Sometimes we can encounter a distinction between sets of names *{network, node, link}* and *{graph, vertex, edge}*, but usually they are used interchangeably. For example, in [3] the difference is noted, but not always obeyed; in [2] these two terminologies are synonymous. Following the second example, in this work there is no formal distinction between the terminologies.

to understand the behavior on the general level and predict the macroscopic outcome. But the ability to predict, as far as advantageous and profitable it can be, is not the only and ultimate measure of a scientific method. The possible advancement in understanding a part of the reality we live in is also important, if not fundamental. In the end, science should serve the truth, not the people.

Some researchers study only properties of individual objects, others study just the interactions between them. Connecting these two approaches can give us a grater insight into the system. Moreover, the *pattern* of interactions can be crucial, and a network representation is a natural way of including this pattern in the study. Going further, one can analyze the dynamics of state of the nodes, or the dynamics of the structure[5] of the network. We have models of double-value spins interacting on a static network and models of the growth of the WWW network with an increasing number of links, ignoring individual properties of each node. Both approaches are valid in their own regime, but we can again make a step towards a more general picture, including both, changes of the structure and states of the nodes, in the description. In fact, in this manner we can grasp much more, as the feedback loop between these two aspects of a network can be and frequently is essential. Such approach is fairly new and still being developed, yet it has its place in the scientific literature. We refer to networks studied in this way as *adaptive networks* or *coevolving networks*. **Coevolving networks lay at the core of this thesis. They are particularly interesting for physicists, since they can be seen as a unification of previously utilized models. Moreover, from statistical physics point of view, adaptive network models can be understood as non-equilibrium processes, however at the end of the thesis also an attempt at equilibrium description of coevolving networks will be made.** The question is whether such representation is useful, can it tell us something about the studied object, can it explain the phenomena? Certainly, the answer is positive and hopefully further chapters of this work will confirm that.

1.1.2 Empirical examples

As we can see, the basic idea of network is very broad. This is probably the reason, why it is used in so many contexts in so many fields with so many applications. It is just an elastic tool, useful in different branches of science [5]. And thanks to the strict mathematical theory behind all the network considerations we can grasp a variety of complex problems with similar or the same set of theoretical ideas.

No wonder networks are gaining recently a lot of popularity, as we can see in Fig 1.2. The increase of usage of the word *network* was enormous at the end of the twenty first century. This may be partially due to the birth of new technological networks of a huge size and a big interest of people in them. To list a few: the Internet, the telephone network, power grids, transportation

[5] *Structure of a network* is an arrangement of the links. Following [3], [4], it is used in this thesis interchangeably with *topology of a network*.

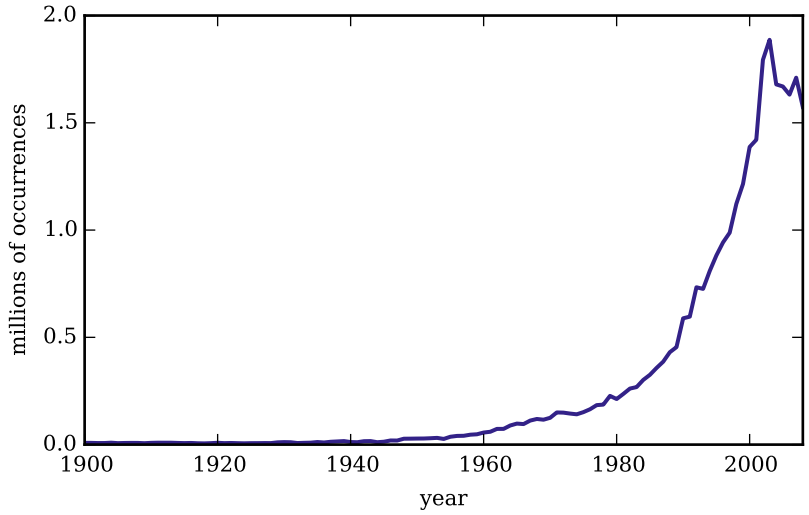


Figure 1.2: The number of occurrences of the word *network* in Google Books [6] since 1900.

networks. The network representation arises here naturally, in the Internet we have computers and related devices creating a set of nodes, and the physical or wireless data connections between them are the links. In power grids we have transmission lines as edges, and generating stations and electricity receivers as nodes. Frequently an important part of a technological network is a software built in order to control the system. Hence, it is possible to obtain virtually full knowledge about the topology, what, combined with the large scale of the structure and the vivid interest in understanding it, provides a fertile area for a scientist.

In social sciences people discovered advantages of the network formalism relatively long time ago. A social network consists of vertices created of people or groups of people and links indicated by a social interaction, such as a friendship, a partnership or a scientific collaboration. People tend to organize in different structures at all possible levels, hence there is a lot of empirical data to study. As in the previous example, recent developments give us the possibility to map huge groups of people in one large network. Facebook, Twitter and other online social platforms grant us a chance of looking at the social systems in a really big scale. The history of the network approach in social sciences, however, is much longer and can be traced back to the nineteenth century. Indeed, „*among researchers who study networks, sociologists have perhaps the longest and best established tradition of quantitative, empirical work*" [2].

Networks are also used in biology, where scientists work on mapping the protein-protein interaction or metabolic networks. Ecologists study interactions between species, e.g. a predator-prey network. The neural network – the diagram of wiring between cells in our brain – is hoped to give us an insight in the amazing computational capability of the brain. Furthermore,

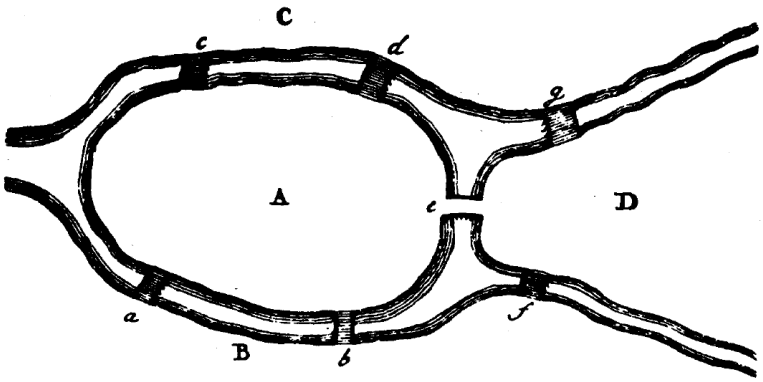


Figure 1.3: Copy of the original drawing of bridges in Königsberg by Leonard Euler in *Solutio problematis ad geometriam situs pertinentis* [7]. Capital letters A-D denote separate parts of land, lower case letters a-g denote 7 bridges on the rivers. Picture available in the public domain, source: Wikimedia Commons.

it inspired a class of artificial intelligence methods known as artificial neural networks.

Different medical data can be better understood when mapped into a network, as it has been done with electroencephalography (EEG) signal or functional magnetic resonance imaging (fMRI). We have information networks, like the World Wide Web or citation networks, financial networks of banks or countries. The list of examples is indeed impressive and the first conclusion one can draw is that the field of complex networks has a truly interdisciplinary character. Different branches of human knowledge has been inspiring and influencing each other and they definitely still do. Nevertheless, physics has brought a wider perspective to the table. It is amazing how many applications networks have in so distinct fields. One could think it is an impossible task to describe it all in a consistent way with common mathematical tools. Yet, many of these systems display similar properties and, once we move to the abstract plane of analyzes, we can collect them under general models and laws. In the twentieth century, this kind of universality had already had a long tradition in physics, but was yet to discover in network science.

1.1.3 Foundations of theory

A point in the history of sciences that is commonly associated with the beginning of the graph theory is Leonard Euler's study of bridges in Königsberg in 1735 [7], [8]. Certainly, it is a very interesting work with a noticeable physics approach – the problem is simplified and presented in an abstract way that one can treat mathematically. But the word *graph* or *network* is not present anywhere in the publication, neither is any diagram like from Figure 1.1 that would be a visual representation of a graph. So, why is this work always pointed at as the origin of graph theory?

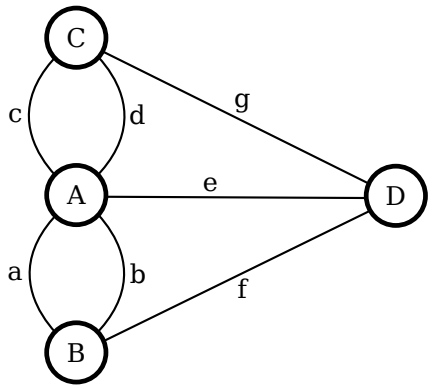


Figure 1.4: The illustration from Figure 1.3 represented in a simpler form – as a graph.

In his publication Leonard Euler was trying to find a general solution to a problem he encountered in Königsberg. The city had at the time seven bridges connecting four separate parts of the land, as it is depicted in Figure 1.3. There was a puzzle popular in the neighborhood: is it possible to walk through all the bridges passing each one exactly once? Euler has heard nobody had done it, however there was no formal proof why it should be impossible. In the first paragraph of his work he said *„This branch [geometry of position] is concerned only with the determination of position and its properties; it does not involve measurements, nor calculations made with them"*. At the time the field was named *geometry of position* by Leibintz, but this sentence contains the essence of reductionism in network science. In this spirit, he created a simplified diagram of rivers and bridges, that today would look like the one in Figure 1.4, if we called points A, B, C, D nodes and a, b, c, d, e, f, g links. Although schemes from Figure 1.3 and Figure 1.4 look different, the idea is the same. This was the first major step made by Euler towards the branch of mathematics called graph theory. The second step was, surprisingly, creating such a representation of the problem that the diagram becomes unnecessary. Once more: he represented the problem in an abstract diagram extracting only the crucial information. This is usually the first step in analysis of many phenomena in network science nowadays. And then he created an even more abstract representation using only letters. Although today one would use rather numbers and the problem would be called *finding a path*, there is a one-to-one projection relation between these approaches. Later, Euler proved that the puzzle of Königsberg's bridges is impossible to solve, there is no path going exactly once through every bridge. He also gave a general necessary condition for this class of problems to be solvable[6]. But the most important was the framework he built to solve it.

A colorful illustration of the reductionism in network science taking into

[6]There can not be more then two nodes with an odd number of connections for the path to exist.

Figure 1.5: Three versions of the London Underground map (top to bottom): from 1908, 1932, and 1933 [9]. Pictures available in the public domain, source: Wikimedia Commons.

account only the most important parts of the system are maps of metro. In Figure 1.5 we can see the evolution of the London Underground map. The first and the oldest version contains all the topography of the city drawn underneath metro lines, making the whole picture vague. In the later update the grid of streets was erased from the map, what made it easier to read, but still the whole diagram looked rather messy due to the growing number of lines. Finally, Henry Charles Beck [9] introduced his famous simplified map of metro, which not only omits unnecessary details of the city, but also represents the metro lines in a simple easier to read manner, using only angles of 90 and 45 degrees. It has little to do with the actual topography of the city, nevertheless it's much more useful and more convenient to extract important information from. For instance, let's have a look at the green line in Figure 1.5. In the middle picture it meanders a lot, but do all of these tiny turns matter when navigating over the metro lines? Not really. And the straight line from the bottom picture, with only 3 turns, is *topologically* equivalent to the previous representation. In other words, it contains the same amount of relevant information. Network approach to real-world systems is similar in this context – it might not always be the perfect reflection of the reality, but it takes into account everything we should potentially focus on and makes it easier to analyze the essence of a problem.

After pioneer work of Euler came many others. A famous example in physics are two laws of electrical circuits proposed by Gustav Kirchhoff [10]. These rules that deal with current and potential differences, that we learn at the beginning of electromagnetism courses, are nothing but pure network analysis. First law – stating that the sum of currents flowing into a node in an electrical circuit is equal to the sum of currents flowing out of that node – is explicitly referring to a node in an electrical circuit, which is a network. The second law – saying that the directed sum of the potential differences around any closed loop is zero – is basically describing a closed path in a (directed) network. The whole methodology of drawing electrical circuits can be seen as a beginning of network visualization.

Later networks have been applied in a growing range of problems. Arthur Cayley identified a correspondence between structural isomers of alkanes and planar graphs [11]. Using networks, he proposed rules limiting the enumeration of alkanes isomers. A similar problem to the one solved by Euler was studied by William Hamilton in 1858 [12]. A Hamiltonian cycle is a path in a network visiting each vertex exactly once, in contrast to the Euler tour, which goes through every edge exactly once. Another famous example is the map coloring problem, or four color map theorem. The problem was: is four colors of ink enough to paint a map in such a way that each of neighboring countries have different color? Empirically three colors were not enough, but a formal proof was given after more than a century. Afterwards it was discovered that the problem can be simplified using networks [13].

A crucial moment, forming the shape of network science in the future to come after it, was the invention the Ising model. The model created by Wilhelm Lenz [14] and solved by his student Ernst Ising [15] is commonly recognized as a milestone in statistical physics. The aim was to explain the

phenomena of ferromagnetism from microscopic point of view and the idea was straightforward – to start from as simple description as possible. This means representing a material as a chain of spins and including only the most important interactions, i.e. nearest-neighbor interactions. This description happened to be too primitive, but the second attempt proved to be more successful. The model was extended into a two-dimensional square lattice. As simple as that, it managed to explain symmetry breaking that had been observed in ferromagnets. Two lessons learned from it by network science were: simplification of the problem focusing on what really matters and talking into account the structure of the analyzed system, which can be essential.

The approach of looking at the microscopic properties of individual elements to later conclude rules and laws of the macroscopic system is a method commonly borrowed from physics in network science. Ising model was later further developed and generalized for different possible configurations [16], continuing to inspire researchers working on other applications of graphs. Ideas from statistical physics have been influencing network science for decades and continue to do so. In particular, this very work is an example of said influence. Ising model is a reference point for one of the models developed in this thesis that will be covered in detail in Chapter 5.

Next big step came with the work of a social scientist Jacob Moreno [17]. First of all, he started a very important branch of network science – analysis of social networks. Moreno was interested in how the social ties and connections influence the psychological functions of an individual. To study such phenomena he developed, and this is the second part of his valuable contribution in network science, a method of representing human contacts, interactions, sympathies and antipathies in a graphical way. Schemes he has created are called *sociograms* and they have basically all traits of the modern network visualization.

Finally, the modern network science started with works of Paul Erdős and Alfréd Rényi [18], [19]. Complex networks waited a long time to be mathematically captured. The breakthrough came with an idea to apply probability theory and treat the system as one of many possible realizations. In this way Erdős and Rényi constructed statistical ensemble, taking into account fluctuations naturally occurring in real-world networks and describing a whole set of networks with similar properties. Again, this approach is unquestionably a reflection of concepts present in statistical physics. Moreover, these two mathematicians discovered a percolation phase transition in random networks, but before describing it more precisely in Chapter 2, strict mathematical definitions of network science shell be introduced.

1.2 Outline of this work

The goal of this thesis is to explore effects of coevolution in network models and to examine its influence on recognized models in general. As coevolution is a step towards more realistic description of networked systems, the work aims at bringing the analysis closer to the empirical observation. In particular, I seek to remedy the problem of the Axelrod model's

disagreement with empirical data, I aim at extending the voter model by essential real-world mechanisms, and I make an attempt at equilibrium description of adaptive networks with an intention of explaining the behavior of the topology.

The thesis has a following structure. Chapter 2 covers the basics of network science. First, the fundamental definitions of graph theory are introduced and different types of graphs are described. In the second part of the chapter fundamental models of network science are presented, together with main numerical[7] and analytical results. The coevolution is discussed. **Chapters 3, 4, and 5 contain original results obtained by myself**. In Chapter 3 the voter model is covered. The chapter starts with an introduction to the general framework of the voter model from a homogeneous case, through complex networks versions, to the coevolving voter model. Finally, **the nonlinear coevolving voter model with triadic closure and with noise are introduced, which are original models created in this work**. Chapter 4 continuous the exploration of algorithmic non-equilibrium models. It is devoted to the Axelrod model. Original definition of the model is described, as well as selected extensions. **The coevolving Axelrod model is then introduced, together with preferential attachment and triadic closure extensions developed in this thesis. At the end of the chapter an original solution dealing with the empirical data discrepancy is proposed**. In Chapter 5 the equilibrium approach to coevolving networks is presented. It starts with a general introduction into statistical mechanics methods and entropy maximization. Then, the Ising model is briefly discussed and similar frameworks describing network's topology are presented. Subsequently, **an original model applying the statistical mechanics approach in a description of coevolving networks is introduced and thoroughly analyzed**. Chapter 6 contains a summary of the work with an outline of main results. **All plots in the thesis were created by myself, based on my own simulations**. Three appendices at the end provide additional information on the issues discussed in the work.

[7] All numerical results presented in the thesis are original, also for fundamental and historical models, where simulations were reproduced to obtain and plot main results.

Chapter 2

Basics of network science

2.1 Graph theory

Graph theory is a branch of mathematics describing networks in a strict manner. It provides many important theoretical discoveries in network science and forms the basic framework for a consistent analyzes. At the foundation of this framework lay definitions, which are especially important in an interdisciplinary field (like complex systems) developed by scientists from many different backgrounds. Many times in the scientific literature we can see different terms referring to the same object, model or phenomena in networks. Here we shall establish the vocabulary, definitions and symbols necessary for the later discussion to be clear and unambiguous.

A graph (network) $G = (V, E)$ is a pair consisting of a non-empty set V of vertices (nodes) and a set of edges (links) $E \subseteq \{\{u, v\} : u, v \in V\}$, which is a family of two-element subsets of the set V, i.e. every edge starts and ends at one of the vertices of the graph [20], [21]. We denote the sizes of these sets N and M respectively. Consider a simple example of a six-vertex graph $G_{ex} = (V, E)$ with $V(G_{ex}) = \{1, 2, 3, 4, 5, 6\}$ and $E(G_{ex}) = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{4, 5\}\}$, so it contains just four edges. The common way of representing a graph is by drawing a dot for every node and a line connecting two nodes for every link, as is is shown in Figure 2.1 for our simple example. The way in which the links are pictured is irrelevant, they can be straight or bent, they can cross each other. The only and essential information is which nodes do they connect.

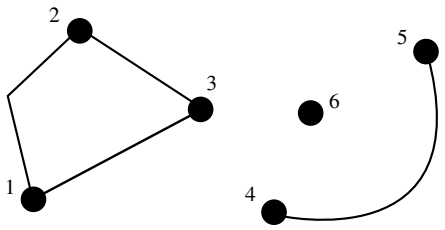


Figure 2.1: A visual representation of the graph G_{ex}.

Another useful manner of defining a network is by an adjacency matrix A. Its elements a_{ij} are equal to 1, if there is a link between the node i and the node j, and 0 otherwise. In simple graphs (see *simple graph* in Section 2.1.2) the adjacency matrix A is symmetrical – if the node i is connected to the node j, then the node j is connected to the node i, but it doesn't always have to be the case (see *directed graph* in Section 2.1.2). The adjacency matrix A for our simple graph G_{ex} can be written as:

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.1}$$

There are advantages and disadvantages of such representation of a network. Firs of all, we can see that in an adjacency matrix we need to keep the information about every virtually possible connection, including auto-connections. Thus, if the graph is sparse, which means it contains few links compared to the maximal possible number of them, we keep a lot of unnecessary information. This might be a difficult dilemma in numerical simulations. For sparse networks it is generally better to keep just the set of edges E, which together with the number of nodes N provides everything required to reconstruct the network. On the other hand, adjacency matrix might be convenient for dense networks, i.e. those containing a number of edges close to the maximal. It might be also advantageous to keep only the list of pairs of vertices that are not connected, it strongly depends on the simulated model, the used algorithm and what we want to achieve. Nevertheless, adjacency matrix is extremely helpful in analytical calculations.

2.1.1 Coefficients and measures

Degree of a node

Degree[1] of a node, denoted by k, is the number of connections with other nodes that it has. If a node is separated from the rest of the network having no connections its degree is equal 0. The maximal possible degree in a simple graph is equal to $N-1$. Having the adjacency matrix A we can calculate the degree of a node i by:

$$k_i = \sum_{j=1}^{N} a_{ij}. \tag{2.2}$$

Often the average degree in the network is an important measure, it can be calculated from $\langle k \rangle = \frac{2M}{N}$. However, many times the essential property of a

[1] Degree of a node is sometimes called *connectivity*, mainly in older publications. Here connectivity refers to a different quantity, namely to $c = M/N$, which is half of the value of the average degree.

network, rendering its structure and thus properties, is the degree distribution $P(k)$ describing the probability of a random node having the degree k. Some of the central network models can be defined by the degree distribution. For instance, we say that a network is scale-free, if the distribution is a power-law function (see Section 2.2.4).

Path and distance

The concept of distance is crucial in formal sciences, but it might be ambiguous how to apply it in networks. Surely, physical or geographical distance is irrelevant in most of the cases – what would it mean when considering the WWW network, or the protein interaction network? What could be the distance between two webpages? The solution is the *path length* [3]. A *path* is a route on a network going through nodes and edges. Or more formally, it is an ordered sequence of edges, for example a path of length 4 between node 1 and 6 in a network can be $P = \{(1,2),(2,5),(5,3),(3,6)\}$ (given all the edges exist). A route going through the same edges, but the opposite direction – from the node 6 to 1 – is a different path as the order matters. Generally, nodes and edges can be repeated in one path, but some alternative definitions do not allow that. A closed path starting and ending at the same node is called a *cycle*.

Having defined the path it is easy to create a convenient distance measure for networks, namely the shortest path length. Therefore, the distance d_{ij} between nodes i and j is equal to the number of edges in the shortest path between them, ranging from 1 (for distinct nodes, for an auto-connected vertex it can be 0) to M, or conventionally even $+\infty$, if a path doesn't exist. Note that there can be more than one shortest path between two nodes.

Another helpful measure is the *diameter* d_{max} of a network, which is the biggest distance between a pair of nodes, i.e. the longest shortest path. However, much more popular and perhaps practical is the *average path length* $\langle d \rangle \equiv l$ (also called the *geodesic*) in the network:

$$\langle d \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}. \tag{2.3}$$

It was the central concept in the famous Milgram experiment [22] and the starting point for creating the idea of the *small world* described in Section 2.2.3.

With the adjacency matrix it is straightforward to calculate the number N_{ij}^d of paths of the length d or the distance d_{ij} between nodes i and j. If $a_{ij} = 1$, there is a direct link between the vertices. If there is a path of length 2, there must be a node k connected to i and j, such that $a_{ik}a_{kj} = 1$. The number of paths of length 2 is the number of such nodes k that satisfy this condition, therefore $N_{ij}^2 = \sum_{k=1}^{N} a_{ik}a_{kj} = a_{ij}^2$ (which stands for the element ij of a matrix A^2, not the element ij to the power of two of a matrix A). Similarly, the number of paths of length d between two nodes is equal to the adequate entry of the dth power of the adjacency matrix $N_{ij}^d = a_{ij}^d$. The distance d_{ij} can be defined as the smallest d for which $a_{ij}^d > 0$. It is also worth noting

that $a_{ii}^2 = k_i$, because every node has exactly as many paths of length 2 from itself to itself back as it has edges, therefore a_{ii}^2 is equal to the degree k_i of the node i. Consequently, the trace of A^2 is equal to the doubled number of edges $\mathrm{Tr}A^2 = \sum_{i=1}^{N} a_{ii}^2 = \sum_{i=1}^{N} k_i = 2M$. Although elegant, this approach is not very useful in computer simulations of big networks.

Worth mentioning are two more quantities associated with path. First one originates from the very first work on graph theory described in Chapter 1 – the Eulerian path is a path that traverses each link in a network exactly once. The second one is the Hamiltonian path, which visits every node in a network exactly once.

Clustering coefficient

In general, the clustering coefficient measures the probability that two neighbors of a given node are connected directly. In other words, it's a density of closed triangles of vertices in a network, with every vertex connected to two others. There are two common parameters widely used and they usually correlate with each other, although the exact values for a given structure can be substantially different. The first one is the local clustering coefficient c_i [2], [23] defined as the number of all connected pairs of neighbors of a node i divided by all possible pairs of neighbors. It can be written as:

$$c_i = \frac{2L_i}{k_i(k_i - 1)}, \tag{2.4}$$

where L_i is the number of links between the k_i neighbors of the node i. To investigate properties of a whole network we want to look at all nodes simultaneously. Therefore, the average local clustering coefficient $\langle c \rangle$ is often very useful in the analysis:

$$\langle c \rangle = \frac{1}{N} \sum_{i=1}^{N} c_i. \tag{2.5}$$

From the equation 2.4 one can see that the value of c_i is not well defined, if the node i has less than 2 neighbors ($k_i < 2$). It is common to exclude such nodes when calculating the average value or to assume their clustering coefficient is equal to 0. Either way, it should be stated how the value was obtained, since depending on our choice it can vary significantly.

The second popular quantity is the global clustering coefficient C, also called transitivity, that directly counts the number of triangles in the network:

$$C = \frac{\text{no. of closed paths of length two}}{\text{no. of paths of length two}} = \frac{3 \times \text{number of triangles}}{\text{no. of connected triples}}. \tag{2.6}$$

The factor 3 in the numerator can be puzzling. It is easy to visualize, if we consider a small network of 3 vertices, where all of them are connected (resulting in 3 edges), like nodes $\{1, 2, 3\}$ in the Figure 2.1. Then, we have 6 paths of length two in such a network (direction matters in paths) all being closed, 3 connected triples (each centered on one of the nodes) and 1 triangle.

The triangle is straightforward, the paths can be represented by all possible permutations of the series 123, and the triples are related to paths of length 2 neglecting the order. The global clustering coefficient is equal 1 in a complete graph or a network consisting of complete subgraphs and 0 in a tree or a square lattice for example. A triangle is the simplest example of a *motif* and the concept of transitivity can be generalized by analysis of different types of motifs [24], [25].

The clustering coefficient is an important measure, because it is often much larger in empirical networks than in corresponding graphs with the same degree distribution, but random connections. It is argued that *triadic closure* is a frequently observed process in social networks [26], [27], where we can measure the highest values of the clustering coefficient [28], [29]. The process consists of literally closing triangles – there is a better chance to become friends with a friend of our friend than with a random person from the whole network. For example, the collaboration network of physicists is said to have $\langle c \rangle = 0.45$, where a network with the same $\langle k \rangle$ and $\langle k^2 \rangle$, but with connections assigned at random, would have $\langle c \rangle = 0.0023$ [2]. For that reason, coevolving models of social networks should be studied extended by the triadic closure, what is done in this work in Chapters 3 and 4.

Component

In empirical networks we can frequently find a pair of nodes i and j, such that it is impossible to reach one from another, i.e. $d_{ij} = \infty$. This is a problem of *connectedness* of a network. Two nodes are *connected*, if there is a path between them and are *disconnected* otherwise. And further, a network is connected, if every pair of nodes in it is connected. If not, the network is disconnected meaning that it contains several separate *clusters* or *components*. A component is a part of a network (a subgraph) where each node is connected to every other node and there is no vertex in the network that could be added to the component retaining this property. More precisely, it can be defined recursively as follows: two vertices i and j belong to the same component, if (i) they are connected by a direct edge or (ii) there is a vertex k such that the vertex i belongs to the same component as the vertex k and the vertex k belongs to the same component as the vertex j.

When the network in the analyzed model becomes fragmented or shattered in many small parts it's intuitive to describe it in terms of components. To estimate the level of fragmentation we can use *the size of the largest component* S, measured in the number of nodes it contains, or the *number of components* n_S. It's also convenient to normalize these quantities by the size N of the network.

In order to understand these definitions we can look at the Figure 2.1. We can see three separate clusters, therefore $n_S = 3$, and the biggest one contains vertices number 1, 2 and 3, so $S = 3$ as well. Since the whole network contains six nodes, both quantities after normalization would be equal to $\frac{1}{2}$. If a network contains more than one component its adjacency matrix can be written in a block diagonal form as the one from Equation 2.1.

Domain

When the studied model allows different states of the nodes, more detailed quantity than the component maybe favorable, such as *domain*. A domain is a group of connected nodes being in the same state. More specifically: two vertices i and j, having states s_i and s_j respectively, belong to the same domain, if (i) they are connected by a direct edge and they have the same state $s_i = s_j$ or (ii) there is a vertex k such that the vertex i belongs to the same domain as the vertex k and the vertex k belongs to the same domain as the vertex j. Obviously, the concept of domain extends the notion of component incorporating state of the nodes. As a result, one component may contain many domains and there can not be fewer domains than components. Additionally, two separate domains having the same state can exist, simply being disconnected. As for the component, two most helpful quantities are *the size of the largest domain* D (in number of nodes) and the *number of domains* n_D. Likewise, it is usually normalized by the size N of the network.

Other measures

There is a number of other measures and coefficients commonly applied in network research, but not of a great importance for this thesis. For example many *centrality measures* exist, of which the simplest one is perhaps the degree. All of them aim at describing the importance of a node in a given context. The famous *PageRank* [30], [31] expressing relevance of a webpage in search engines, *DebtRank* [32], [33] specifying bank's condition from systemic risk point of view, *closeness centrality* [34] measuring the mean distance from a vertex to other vertices etc. Another big concept is the *assortativity* portraying the likelihood of nodes to be connected to similar nodes, in terms of the degree for instance. Additionally, centrality measures may concern edges, like *betweenness centrality* describing the role of a link in information transfer or, more generally, spreading processes. Most of these measures are described in [2]–[4], [35].

2.1.2 Types of graphs

It was shown at the beginning of Section 2.1 how a graph can be defined, but one can construct countless different objects fitting this definition. To be more precise in our description we need narrower categories. This subsection lists all categories necessary to understand the further content of the thesis and a few more to put the work in a broader context of the current research. The following network types are covered in greater detail in [2], [23], [35]–[38].

Simple graph

A *simple graph* is an unweighted, undirected graph with single edges between vertices and no self-loops. An example is presented in Figure 2.2 a) and d). Its adjacency matrix contains only zeros and ones and is symmetrical. Every pair of nodes is either connected by a single link or not, with no distinction

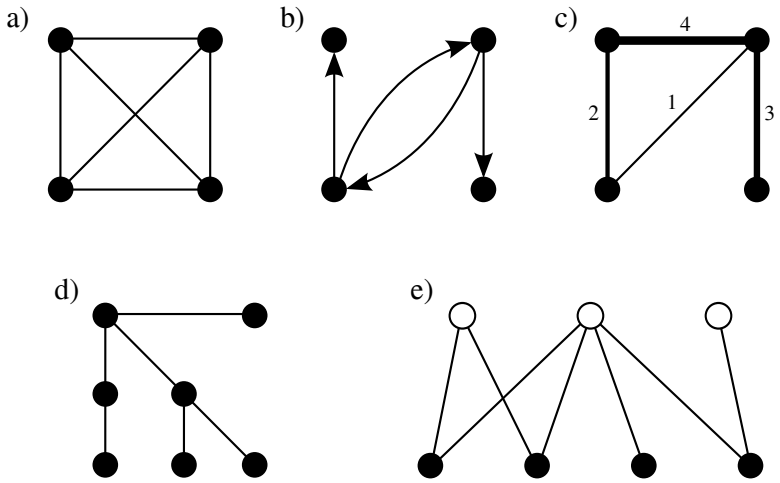


Figure 2.2: Examples of different graph types. a) A simple, complete and regular graph, b) a directed graph, c) a weighted and planar graph (weights are indicated by the line width and a number), d) a tree (consequently also being planar), e) a bipartite graph.

between types of links. A node cannot be connected to itself. A simple graph doesn't have to be connected, i.e. it can contain many components. Usually, the word *graph* or *network* refers to a simple graph, if the type is not specified. Sometimes, a network containing multiple connections between one pair of nodes is called a *multigraph*.

Complete graph

A *complete graph* is a network with every pair of nodes connected by an edge, as in Figure 2.2 a). A complete graph of N nodes is usually denoted by K_N. All nodes have the same degree $k = N - 1$ and the total number of links is equal $M = N(N-1)/2$. The diameter in this graph is $d_{max} = 1$ and so is the average path length $\langle d \rangle = 1$. It is often the first choice to analyze a given model on a complete graph, due to its simplicity when it comes to the analytical description. For numerical simulations it's also convenient, since a full structure is somewhat equivalent to no structure at all – there is no need of keeping any information about it. Unfortunately, it's rather rare among empirical examples. It is argued that very small parts of social networks can take form of a complete graph. But most of the real-world networks are sparse, i.e. they contain much less edges then a corresponding complete graph.

Directed network

A *directed network* is a network with links possessing directions, what can be represented by an arrow, like in Figure 2.2 b). In a directed network a link

from a node i to a node j can exist independently of the opposite link from the node j to i. Therefore, every pair of nodes can be connected by up to two links of opposite directions, and if it is, we say that the connection is *reciprocal*. A consequence of the link direction is a possible asymmetry of the adjacency matrix, in general a_{ij} doesn't have to be equal a_{ji}, what requires a special analytical approach in many applications [39], [40]. A network that is not directed is called undirected network. All models developed in this thesis consider undirected networks.

Weighted network

Sometimes a simple indication of the connection is not enough and the strength or value of this connection is relevant. We call this value a *weight* and a network with weighted connections a *weighted network* [41], [42]. An example can be seen in Figure 2.2 c). Usually, the weight is a real number indicating the strength, the frequency or other feature of the relation between two nodes. The adjacency matrix entries a_{ij} for a weighted network can take any value, in some models even negative, with 0 indicating no link. All models developed in this thesis consider unweighted networks only.

Regular network

A *regular network* is a network where every vertex has exactly the same number of connections $k_i = const$, as in Figure 2.2 a). A regular network with vertices of degree k is called a k-regular graph. An example of such structure is a *square lattice*, where every node has four neighbors (except those on the boundary). The square lattice was very popular in statistical physics and is still frequently used as a null model. Note, that not every regular network displays this level of regularity when plotted, it can even look completely random.

Tree

A *tree* is a graph with no cycles, i.e. no closed loops, like the one in Figure 2.2 d). This structure is well known for applications of *spanning trees* in correlation analysis. A spanning tree of a connected graph $G = (V, E)$ is the minimal possible subset of edges E containing all vertices V and creating a tree. In weighted networks, for example to study a correlation matrix, usually the *maximal* or *minimal spanning tree* (MST) is considered [43]–[45]. MST instead of minimizing the number of edges aims at minimizing/maximizing the total weight of the selected edges.

Bipartite network

A *bipartite network* or a *bigraph*, also called *two-mode network*, is a network whose nodes can be divided into two groups or *modes*, in such a way that there are no connections between nodes belonging to the same mode. In other words, every link goes from one group to the other. These groups are

represented by white and black color in Figure 2.2 e). For instance, a network of banks and financial instruments they are using can be seen as a bipartite network. A bank is connected to a given instrument, say shares of a given company, if it invests in it. Many banks can invest in the same stock and one bank can invest in multiple stocks.

An equivalent of adjacency matrix for bipartite graphs is the *incidence matrix* B. Staying with the financial interpretation, if there is n banks and m instruments, then the incidence matrix is a $n \times m$ matrix, where B_{ij} is 1 if the bank i invests in the instrument j, and 0 otherwise. When every node from one mode is connected to every node from the second mode it's called a complete bipartite graph represented by $K_{n,m}$.

Bipartite networks may give a full representation of particular structures, however it's often helpful to work with nodes of one type only. Such reduction of a bigraph is called *one-mode projection* [46] and is performed by connecting two vertices of one mode, if they have links to the same vertex of the other mode. As there are no cycles of length three in two-mode networks analysis of clustering is not trivial [47]. A structure similar to a bipartite graph is a *hypergraph*, which is actually homeomorphic to a bigraph and contains the same information, but has only one type of nodes labeled adequately.

Planar graph

A *planar graph* is a graph that can be drawn on a plane in such a way that none of its edges intersect. In other words, we can draw a planar graph without edges crossing each other, except for meeting at vertices. For example, graphs in Figure 2.2 c) and d) are planar. A famous Polish mathematician Kazimierz Kuratowski proved a theorem providing a convenient approach for identifying whether a graph is planar [48]. The Kuratowski's theorem states that a graph is planar, if and only if it does not contain any subgraph that is a complete graph K_5 or a complete bipartite graph $K_{3,3}$. It can not contain a *subdivision* of K_5 or $K_{3,3}$ neither. A subdivision of a graph is formed by inserting a vertex in the middle of an edge, or by doing so multiple times. *Planar maximally filtered graphs* (PMFG) are of particular interest due to similar applications as MST [44], [49].

Multilayer network

A *multilayer network* or a *multiplex network* is an object consisting of layers of graphs [50], [51]. In other words, each layer is one network of any kind. Then, every node can exist in any subset of layers, but every representation of the same node in different layers is connected by *inter-layer* connections. Links within one layer are called *intra-layer* connections. An example of a *duplex* – a multiplex network with two layers – is presented in Figure 2.3. Some authors [52], [53] distinguish between multilayer and multiplex networks, by stating that the former allows any node of any layer to connect to any other (in any layer) and the latter needs to contain the same nodes in each layer being connected only to itself via iter-links (with arbitrary intra-links). However, this

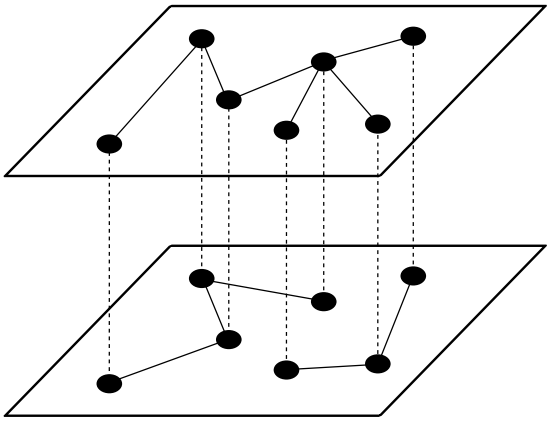


Figure 2.3: An example of a multlilayer network with two layers (a duplex). All nodes exist on both layers, but intra-layer connections are different.

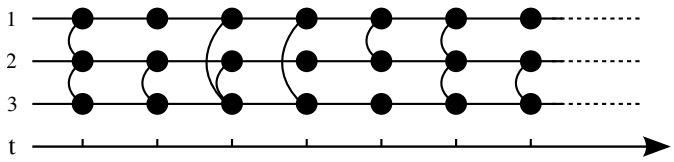


Figure 2.4: One of possible ways of visualizing a temporal network. This simple example considers three nodes 1, 2 and 3 and discrete time of equal intervals. Each tick represents one snapshot of the current state of the network.

branch of network science is relatively young and the consensus about the vocabulary is not yet established. The adjacency matrix for mulilayer networks is generalized to a tensor. There are different methods of aggregating mulilayer networks into a single-layer network, but none of them reproduces all properties [54], [55]. This type of networks is of particular interest in social sciences, since different layers can naturally represent different types of social interactions, or interactions via different media [56], [57].

Temporal network

Another young branch of network science comes from the analysis of *temporal networks* [58]–[60]. A temporal network is a network with a new dimension taken into account, namely time. It contains information about nodes and links, but together with additional information about time of every interaction. Therefore, every connection has its time frames – it can be a single timestamp (giving a discrete time scale) or a period of time. Either way, for

every edge its existence duration has to be specified. An example is given in Figure 2.4. A temporal network can be fully represented by a multilayer network with every layer being the state of the network in between any changes. And like a multilayer network, a temporal network is helpful in the description of real-world social networks as every social interaction has its duration.

2.2 Fundamental models

Graph theory lays within the scope of basic research, but network science additionally to its intellectual value has many applications. Network models are used in physics, chemistry, biology, sociology, linguistics, economics, computer sciences and other disciplines [61]. To cover such a broad range of fields a rich collection of network models is necessary. Indeed, over the years a considerable number has been proposed. Many of them, however, can be seen as an extension of a few fundamental models being the milestones in the network science. In this section the most relevant ones are described.

All network models can be divided into deterministic and probabilistic. For instance, a square lattice is a fully deterministic network. Once it is decided how many nodes it contains its structure is perfectly determined and there is only one way to construct it. Another example can be a triangular lattice, a complete graph, chain-like networks, or fractal networks. In all of them, once we set the parameters values we know exactly how the network will look like, because there is only one possible realization of it. Deterministic models are useful for example in analysis of processes taking place on a crystal structure. The phenomena of ferromagnetism was explained using regular deterministic networks. But in most of the previously mentioned applications there is a certain level of randomness. It can be due to noise or it can be the very nature of the system, i.e. the process of it's creation and evolution is not deterministic. That's why the probabilistic models form much bigger group and are much more frequently applied in description of real-world phenomena.

Going further, the probabilistic models can be divided into algorithmic and analytical ones [62], or into equilibrium and non-equilibrium networks [63], [64]. This division is strongly related to the approach known from statistical physics and thermodynamics. For instance, a static network[2] can be seen as an equilibrium process – their macroscopic properties are constant and independent of time. On the other hand, evolving networks can be understood as a non-equilibrium process in a stationary state – their macroscopic properties keep changing, but particular parameters do not evolve and remain constant (e.g. average degree). The majority of network models are algorithmic (non-equilibrium), like the ones described in Chapters 3 and 4. However, equilibrium approach can also be fruitful, as in the models described in Chapter 5.

[2] A static network is an extreme example given just to emphasize the difference between equilibrium and non-equilibrium models. Chapter 5 contains much more detailed discussion of the topic.

Static networks are often employed to study phenomena taking place *on* a network. For example, a spreading process can be analyzed on a network and its outcomes may, and usually do, significantly vary depending on the topology of the network. That is why the proper description of the structure is so important, but its not always relevant how this structure was created. Contrarily, in evolving network models the dynamics *of* the network is the main subject. Of course, the generated structure can be further studied, for example to examine its influence on mentioned spreading processes. Although, the main question asked in the context of evolving models is what rules give rise to a given structure.

2.2.1 Random networks

By calling a network random we mean that some parameters are fixed, but there is a certain level of randomness in other aspects. For instance, the number of vertices can be fixed and the structure can be accidental. Although there are many models of random networks, the name is normally associated with a particular one, if not specified otherwise. This model was first studied by Salomonoff and Rapoport [65], but largely developed by Erdős and Rényi [18], [19], [66]. Due to the significant contribution of the second authors the model is usually refereed to as *Erdős-Rényi random graph* or just *ER model* (sometimes *Poisson random graph*).

The ER model has two equivalent definitions. In both the number of nodes N is fixed. In the first definition additionally the number of links M is set. Then, to generate the network all links are distributed among the nodes with every pair of nodes having equal probability of being connected. In fact, the model can be alternatively defined as en ensemble of graphs of N vertices and M edges. In this ensemble every possible configuration being a simple graph is equally possible. Therefore, the network generation process consists of picking one of them at random. Although, as in statistical physics, every realization is equally probable there is much higher chance of obtaining homogeneous and disordered one, since there is a greater number of them.

The second definition also has two parameters, this time being the number of nodes N and the probability of link creation p. In this definition the number of links M varies, but on average the model is identical to the first definition, sharing all relevant features. To generate a network one has to create N vertices and for every single pair of them create an edge with a probability p. There is $N(N-1)/2$ pairs, therefore the expected number of links equals $\langle M \rangle = pN(N-1)/2$. Knowing that $\langle k \rangle = 2M/N$, we obtain the average degree for ER network:

$$\langle k \rangle = p(N-1). \tag{2.7}$$

From a single node point of view, the process of obtaining new neighbors can be seen as a Bernoulli trial with probability of success equal p. Every node

has $N-1$ potential neighbors what leads to a binomial degree distribution:

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \tag{2.8}$$

For the thermodynamic limit, i.e. $N \to \infty$ with $\langle k \rangle = const.$, it converges to the Poisson distribution (what is the reason behind the name *Poisson random graph*):

$$P(k) = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}. \tag{2.9}$$

As the probability that any two nodes are connected is p, it is also the probability that two nodes being neighbors of a given node are connected. Therefore, the average local clustering coefficient takes value:

$$\langle c \rangle = p = \frac{\langle k \rangle}{N-1}. \tag{2.10}$$

It can be shown that the diameter of the ER random graph for $\langle k \rangle >> 1$ is equal $d_{max} = \ln N / \ln \langle k \rangle$ and the average path length:

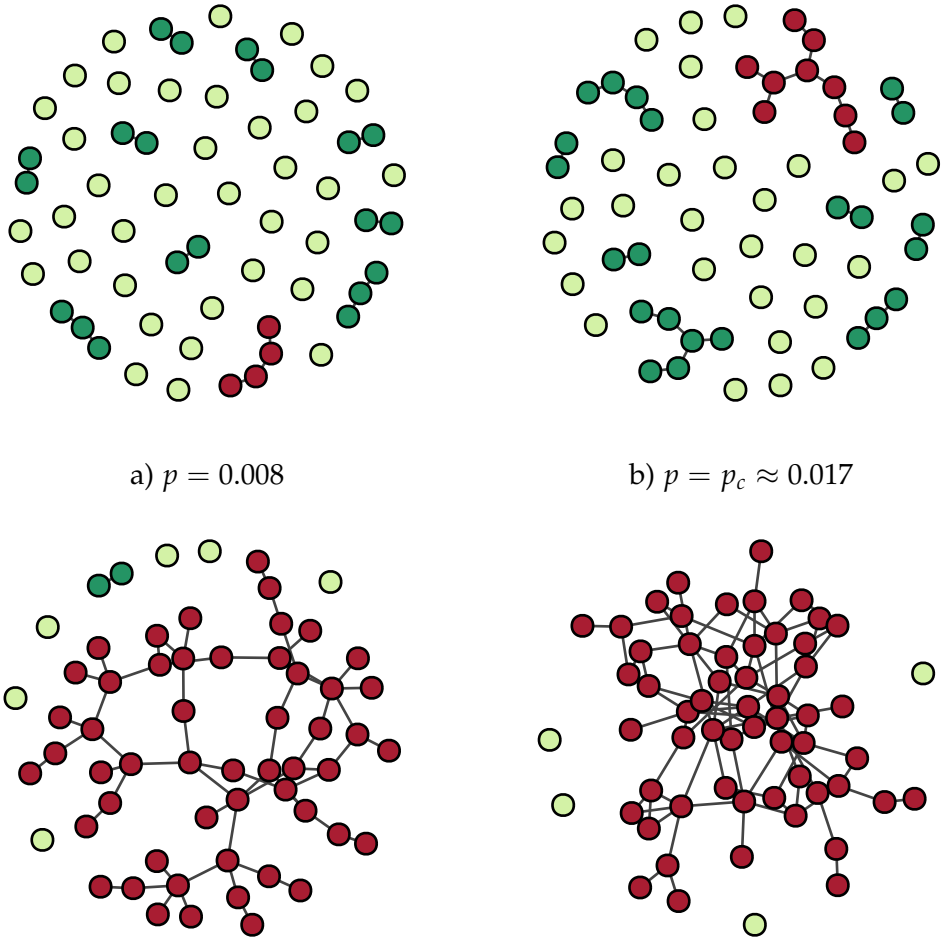
$$\langle d \rangle = \frac{\ln N - \gamma}{\ln \langle k \rangle} + \frac{1}{2}, \tag{2.11}$$

where γ is the Euler constant [67].

The parameter p in the ER model can take values from 0 to 1, since it is a probability, and it is easy to notice that for these two extreme values the network will display significantly different features. For $p = 0$ there is no edge, the network is just a set of vertices. For $p = 1$ every pair of nodes is connected, therefore we obtain a complete graph. Somewhere in between there must be a change in the network topology, in the sense of the size of the largest component S. We can observe this transformation in Figure 2.5. As connectedness of the network changes, it can be referred to as a percolation transition [68]. Indeed, it displays properties of a percolation process on a square lattice of infinite dimensions [69], [70]. It turns out that the minimal value of the average degree $\langle k \rangle_c$ required to observe macroscopic structures is 1. Therefore, the critical probability at which components of a size comparable to the whole network will start to emerge must be equal $p_c = 1/(N-1)$. The exact size of the normalized giant component in the limit of large network size is given by [18]:

$$S = 1 - \mathrm{e}^{-\langle k \rangle S} = 1 - \mathrm{e}^{-p(N-1)S}. \tag{2.12}$$

The percolation phase transition in the ER network is a continuous phase transition, as it can be seen in Figure 2.6. It was argued that there can exist a discontinuous *explosive* phase transition, but it was proved to be merely a numerical artifact [71], [72]. The percolation phenomena is especially interesting, because it is connected to the question of robustness of networked systems under errors and attacks [73].



a) $p = 0.008$ b) $p = p_c \approx 0.017$

c) $p = 0.035$ d) $p = 0.05$

Figure 2.5: Percolation transition in the ER model visualized for different values of p: a) below the critical point, b) at the critical point p_c, and c), d) above it. Network contains $N = 60$ nodes. The main cluster is colored red, nodes with no connections are colored light green, and the rest dark green.

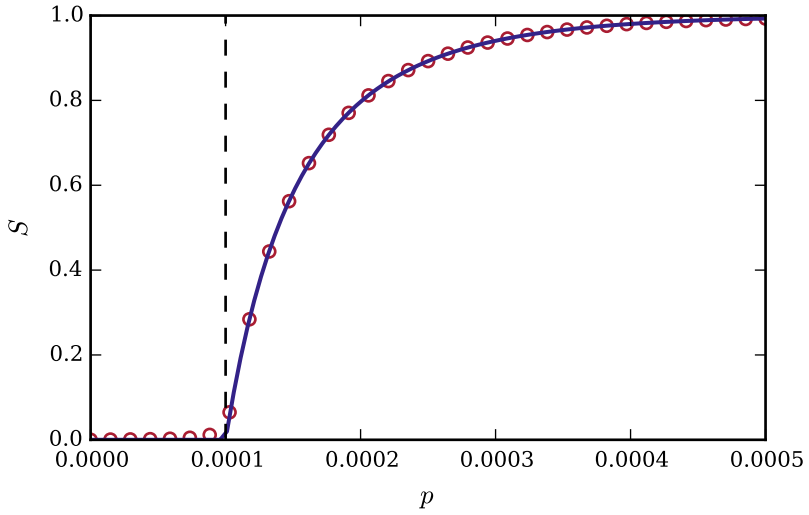


Figure 2.6: Percolation transition in the ER network. Size of the largest component S (normalized) vs. the probability of edge creation p for $N = 10000$. Solid line is a solution of Equation 2.12, circles represent results of a simulation averaged over 100 realizations. Dashed line indicates the critical value p_c.

Unfortunately, as much as the ER random graph is a pioneering and stimulating model, which can provide a convenient framework for network processes analysis, it does not provide an accurate representation of real-world networks. Its structure doesn't display several crucial traits observed in empirical data. The first issue is the shape of the degree distribution. Empirical networks have often power-law, or at least fat-tailed, degree distribution. The exponential decay for ER model is definitely divergent with this observation. The second issue is the clustering coefficient value. As predicted by Equation 2.10, it will tend to zero when the system size grows[3]. Usually, values of clustering coefficient in real networks exceed those in a corresponding ER graph by orders of magnitude. Finally, there is no correlation between the neighboring nodes degrees, since the attachment is purely random. In many networks, assortativity is strong and can not be neglected. Consequently, the only conclusion can be: the real networks are not so random.

2.2.2 Exponential random graphs

As it was mentioned in the previous section, the ER random graph model can be seen as an ensemble of graphs over which we define a simple probability distribution, with every realization having the same probability. This approach proved to be incredibly successful in statistical physics. It would

[3]It will go to zero, if we want to keep the average degree constant, which is reasonable, since real networks are sparse.

be disappointing to utilize it only as a mere comparison. Indeed, such approach was applied in network science in much grater strictness and it is called *exponential random graphs model* (ERGM) [74], [75]. In this model we assume that there is many possible realizations of a network with N vertices, even if only one exits in the empirical world, and we construct a probability distribution over them. This pair – the set of networks $\mathcal{G}$ and the distribution $P(G)$ – creates a statistical ensemble. From normalization we have:

$$\sum_{G \in \mathcal{G}} P(G) = 1. \tag{2.13}$$

We can calculate expected value of any measure on the network $x_i(G)$ over the ensemble as:

$$\langle x_i \rangle = \sum_{G \in \mathcal{G}} x_i(G) P(G). \tag{2.14}$$

Say we can observe certain features x_i of an empirical network and we want to maintain them in the model. Further, having a model reproducing these features we want to ask a question about possibility of observing different features y_i. Or we would like to know how probable is a given structure. ERGM provides natural and very elegant framework to answer these questions. The fixed features are represented by appropriate network measures x_i, or in the simplest form by their average values $\langle x_i \rangle$. Then, in Equation 2.14 the left-hand side is specified and together with Equation 2.13 it forms constraints for the probability distribution $P(G)$. Usually, the number of constraints is small, we want to fix maybe several measures. The total number of achievable graphs of N vertices in the ensemble is equal $2^{N(N-1)/2}$, which is much larger than the number of constraints even for relatively small N. Thus, the constraints don't fully specify the distribution – there is still a lot of flexibility. As in statistical physics, it is argued that the best choice of the distribution with the smallest amount of assumptions is the one that maximizes the Boltzmann-Gibbs-Shannon (BGS) entropy[4] [76], [77]:

$$S = -\sum_{G \in \mathcal{G}} P(G) \ln P(G). \tag{2.15}$$

The maximum entropy principle assures that we are not adding any extra biases or assumptions into our analysis [78] (more thorough description of this principle can be found in Section 5.1). Maximizing Equation 2.15 by the method of Lagrange multipliers leads to:

$$P(G) = \exp\left(\alpha - 1 + \sum_i \beta_i x_i(G)\right), \tag{2.16}$$

[4]Entropy, as usually, is denoted here by S, like the size of the largest component. It was decided to leave this ambiguity to preserve natural symbols of quantities. Everywhere else S stands for the size of the largest component, if not specified otherwise.

where α and β_i are Lagrange multipliers (see Appendix A for the full derivation). The formula can be rewritten as:

$$P(G) = \frac{\mathrm{e}^{-H(G)}}{Z}, \tag{2.17}$$

where $Z = e^{1-\alpha}$ is a *partition function* and $H(G)$:

$$H(G) = -\sum_i \beta_i x_i(G), \tag{2.18}$$

is a *Hamiltonian* of a graph. The partition function Z must fulfill the normalization constraint, therefore $Z = \sum_{G\in\mathcal{G}} e^{-H(G)}$. Parameters β_i are to define from the constraints equations 2.14.

On the other hand, we might be interested in the inverse reasoning. If we treat β_i as free parameters with a certain range of possible values, we can analyze the structures it can produce. We can control the outcome of the model in a similar way as the probability p controls the structure of the ER random network. Such approach is promising, since we can obtain the relation between microscopic measures of a network and its macroscopic properties. Additionally, an analogue of temperature (in the sense of fluctuation size) can be constructed. This reasoning is applied in one of the main models of the thesis in Chapter 5.

In general, it might be difficult to calculate the average value of an arbitrary quantity over the ensemble, but it is straightforward, if the quantity is present in the Hamiltonian:

$$\begin{aligned}\langle x_i\rangle &= \sum_{G\in\mathcal{G}} P(G)x_i(G) = \sum_{G\in\mathcal{G}} \frac{\mathrm{e}^{-H(G)}}{Z} x_i(G) = \frac{1}{Z}\sum_{G\in\mathcal{G}} \mathrm{e}^{\sum_i \beta_i x_i(G)} x_i(G) \\ &= \frac{1}{Z}\frac{\partial}{\partial\beta_i}\sum_{G\in\mathcal{G}} \mathrm{e}^{\sum_i \beta_i x_i(G)} = \frac{1}{Z}\frac{\partial Z}{\partial\beta_i} = \frac{\partial \ln Z}{\partial\beta_i},\end{aligned} \tag{2.19}$$

where the quantity $F = \ln Z$ is called *free energy*. Therefore, the only thing necessary to calculate the value of $\langle x_i\rangle$ is the partition function Z.

En example of ERGM with a simple constraint is a network with a fixed average number of edges $\langle M\rangle$. Then, the Hamiltonian is:

$$H(G) = -\beta M, \tag{2.20}$$

and from Equation 2.17 we obtain:

$$P(G) = \frac{\mathrm{e}^{\beta M}}{Z}, \tag{2.21}$$

where $Z = \sum_{G\in\mathcal{G}} \mathrm{e}^{\beta M}$ from normalization 2.13. It is easy to count the number of links in a network using the adjacency matrix A:

$$M = \sum_{i<j} a_{ij}, \tag{2.22}$$

what allows to compute the exact form of the partition function:

$$\begin{aligned} Z &= \sum_{\{a_{ij}\}} \mathrm{e}^{\beta \sum_{i<j} a_{ij}} = \sum_{\{a_{ij}\}} \prod_{i<j} \mathrm{e}^{\beta a_{ij}} = \prod_{i<j} \sum_{a_{ij}=0,1} \mathrm{e}^{\beta a_{ij}} \\ &= \prod_{i<j} (1+\mathrm{e}^{\beta}) = (1+\mathrm{e}^{\beta})^{N(N-1)/2}, \end{aligned} \tag{2.23}$$

where $\{a_{ij}\}$ indicates all possible adjacency matrix configurations for a simple graph of N vertices. Having Z we can calculate the average value of the chosen measure using Equation 2.19:

$$\langle M \rangle = \frac{\partial \ln Z}{\partial \beta} = \frac{N(N-1)}{2} \frac{1}{1+\mathrm{e}^{-\beta}}. \tag{2.24}$$

Now we can finally find the proper value of the parameter β to satisfy the constraint:

$$\beta = \ln \left(\frac{\langle M \rangle}{\frac{N(N-1)}{2} - \langle M \rangle} \right). \tag{2.25}$$

As we know, the maximal possible number of edges in a simple graph is equal to $N(N-1)/2$, hence from Equation 2.24 we can deduce that the factor $(1+\mathrm{e}^{-\beta})^{-1}$ must be the probability of a single edge to exist:

$$p = \frac{1}{1+\mathrm{e}^{-\beta}}. \tag{2.26}$$

Therefore, the exponential random graphs model with a Hamiltonian given by Equation 2.20 is equivalent to the ER random graph model with the probability p given by Equation 2.26. In other words, the ER model can be seen as a special case of a more general framework of ERGM. Moreover, ERGM can account for several other models, for instance random graphs with arbitrary degree distribution, high clustering or particular motifs [79]–[84]. The model is also referred to as $p*$ (*p-star*), mainly in social sciences [85], [86], where it gained large interest and is frequently applied in the analysis of social networks.

2.2.3 Small-world networks

Models of random networks described in previous sections, although important, suffer a substantial drawback. The real-world networks often times display high values of the clustering coefficient. This effect is particularly visible in social networks [63], [87]–[89]. The ER random graph has insignificant clustering[5]. Moreover, it goes to zero when the network grows, therefore one could expect no clustering at all on the level of whole societies. But clearly this is not the case, hence a better model is necessary.

[5] In the ERGM it is possible to construct a Hamiltonian satisfying high clustering coefficient constraints [81].

A very simple graph that accounts for the higher clustering can be constructed in a following manner. Put N vertices on a line and connect them by edges in such a way that every vertex is connected to k_i nearest neighbors. For simplicity, the degree is kept even and constant, $k_i = \langle k \rangle \equiv k$ for every i. It can be shown that the global clustering coefficient in this network with periodic boundary conditions will be [2]:

$$C_0 = \frac{3(k-2)}{4(k-1)}. \tag{2.27}$$

For instance, $k = 4$ gives clustering $C = \frac{1}{2}$, which is a reasonable value compared to the empirical measurements. The only problem is that such a regular structure generates very long distances between nodes. Indeed, the average path length is given by:

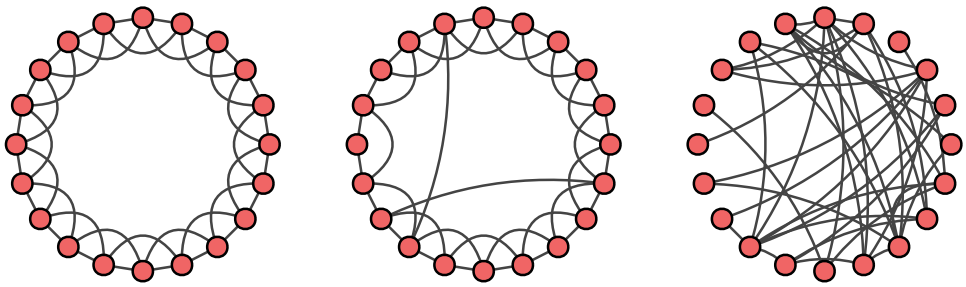
$$l_0 = \frac{N}{2k}, \tag{2.28}$$

which results in $l = 125$ for $k = 4$ and N equal just 1000. As Milgrams experiment showed [22], [90], this is not what we would expect from empirical networks. They are said to be *small*, not in the sense of the number of nodes, but short paths between vertices [91]–[93]. This phenomena is referred to as the *small-world problem* and no regular planar construction can explain it. So the simple model appropriately reproduces the clustering behavior, but fails in terms of the average path length. However, one of the features of the random graph was a very short geodesic (Equation 2.11).

We have two models then. One explains the transitivity very well and fails when it comes to the small-world property, the second the other way around. Why not to merge them? This is exactly what Duncan Watts and Steven Strogatz did creating their simple, yet powerful and famous model [94]. The Watts-Strogatz model (WS)[6] proposes a parameter p, which is a probability of rewiring an edge. More precisely, we start from a regular graph described before and for every link we decide with the probability p to rewire it, choosing new ends at random from the entire network. As it is a simple graph, the rewiring must be performed without generating self-loops or multi-edges. Obviously, for $p = 0$ nothing will change and for $p = 1$ we will obtain a random graph. But the most interesting behavior can be observed somewhere in between these values, when only a fraction of links is rewired, as we can see in Figure 2.7. Originally only one end of every edge was rewired, but this does not lead to the ER random graph for $p = 1$, because the degree distribution will not be Poissonian. Since every node preserves half of its links the minimal degree would be $k/2$, where in the Poissonian random graph there is no such restriction. Elsewhere, the models are identical for both rewiring procedures.

The most striking result of the WS model is the behavior of the network parameters when p is varied. As we can see in Figure 2.8, the average path

[6] The WS model is also called the *small-world model*, but since many models posses the small-world property it's better to stick to the authors names.



a) $p = 0$ b) $p = 0.1$ c) $p = 0.9$

Figure 2.7: Network in the WS model for different values of the rewiring probability p. Between the limiting cases – the regular network and the random graph – a network with the small-world property and high clustering emerges.

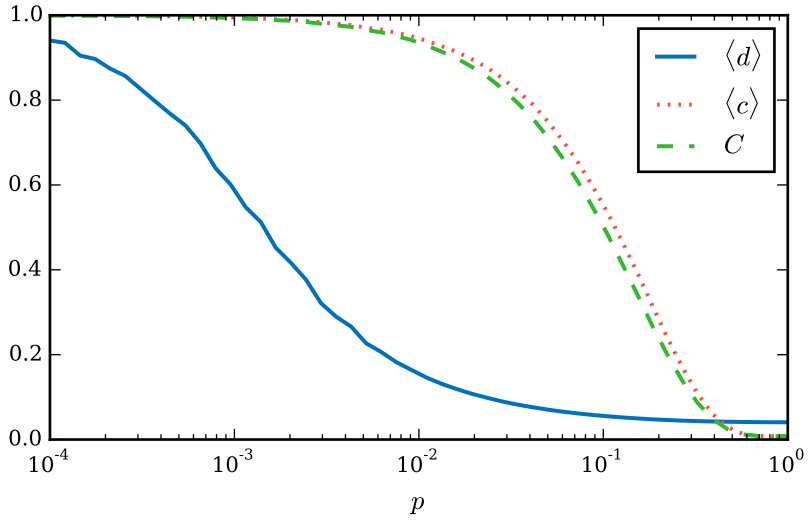


Figure 2.8: Average path length $\langle d \rangle$, global C and average local $\langle c \rangle$ clustering coefficient in the WS model. Results obtained from a simulation for $N = 1000$ and $\langle k \rangle = 4$, averaged over 500 realizations. All values are normalized – divided by the corresponding quantity in $p = 0$ (Equations 2.27 and 2.28).

length rapidly decreases even for small values of p, while the clustering coefficient is still relatively high and drops to zero much later. Therefore, there is a wide area where the generated network has both – the small-world property and significant transitivity. This remarkable outcome explains how it is possible to observe the two properties in empirical networks simultaneously, maintaining them sparse at the same time. It is achievable thanks to the small amount of long-range connections serving as shortcuts between distant vertices.

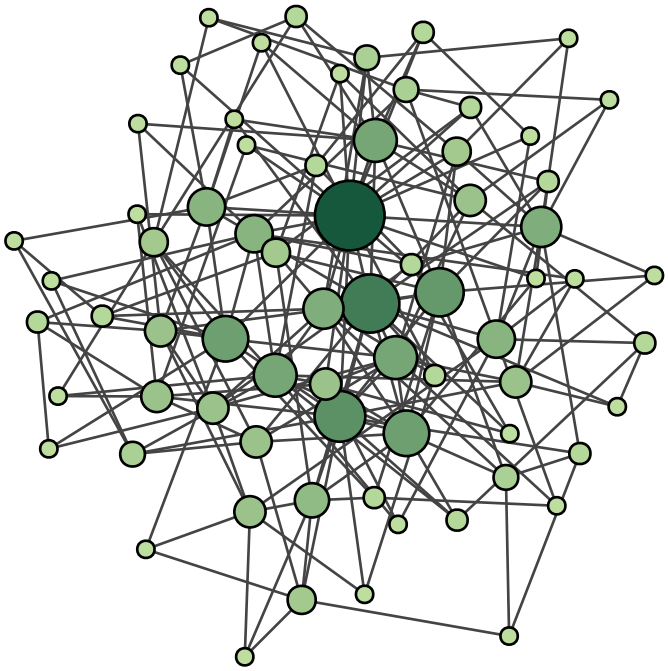
The WS model can be defined also for more complex structures than the one-dimensional chain of nodes. It can be extended to an arbitrary number of dimensions, constructed on a triangular lattice, or a square lattice, although the last one lacks the clustering, but contains many loops of length four. Interesting problems arise when analyzing efficiency [95] and possibility of navigation [96], [97] in small-world networks.

2.2.4 Scale-free networks

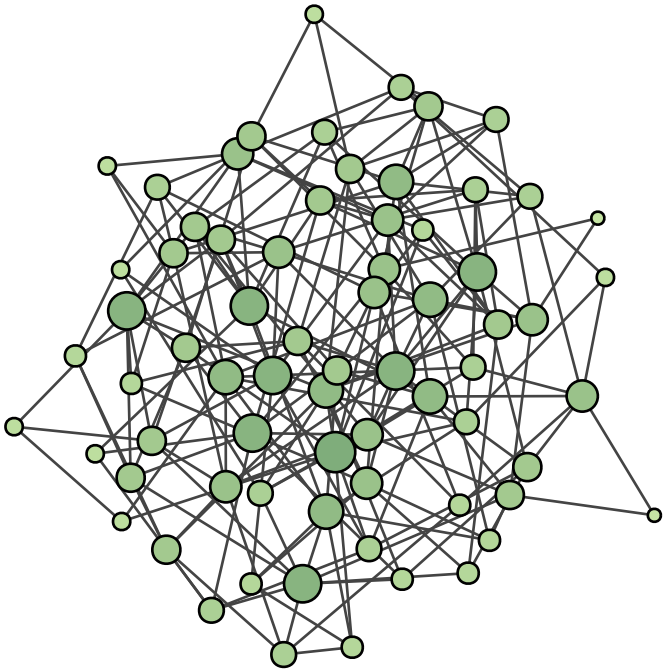
Together with the small-world property and high clustering coefficient a very important topological feature of real-world networks is the *power-law degree distribution* $P(k) \sim k^{-\alpha}$. This distribution and networks that posses it are called *scale-free*, due to lack of a characteristic scale of the phenomena – variance or even the average value may not exist [98]. Such network looks the same before and after zooming in a picture – there is always a lot of weakly connected nodes and a few hubs with a huge number of links. The difference between the scale-free topology and the ordinary random graph can be seen in Figure 2.9. A power-law degree distribution was reported in a number of empirical networks, like the WWW [99], metabolic networks [100], scientific collaboration network [101], the Internet [102], or actors collaboration network [103]. Especially, the World Wide Web case triggered a discussion about how random is the structure of real networks and led to the first scale-free network model.

The scale-free property is a critical aspect in various systems. The first model to explain it was proposed by George Yule to describe the number of species per genus of flowering plants [104]. Later it was applied to demonstrate the city size distribution [105] and scientific citation networks [106]. But the most prominent model was introduced by Alber-László Barabási and Réka Albert [107], [108] in order to explain the structure of the WWW. In all this examples the key to generate a power-law distribution is *growth* and *preferential attachment*. More precisely, over time new nodes are introduced into the network and they connect not completely at random, but with a preference over some types of nodes. Therefore, the Barabási-Albert model (BA) is a model of evolving network defined as follows [3]. We start with m_0 nodes creating a complete graph[7]. At each time step a new node is added into the network and connected to $m \leq m_0$ other nodes in the network. Hence, the minimal degree is equal m. Preferential attachment is achieved by using

[7] The initial configuration doesn't matter, as long as every node has at least one connection.



a) BA network



b) ER network

Figure 2.9: Comparison of a) a scale-free and b) a random structure of a network. Nodes with higher degree are drawn with bigger circles and darker color. It is clear that BA network has much bigger hubs, while ER network is more homogeneous. Both graphs consist of $N = 70$ nodes and $M = 204$ edges. Parameters in the BA model are $m_0 = m = 3$.

a certain probability distribution over existing nodes – the probability $\Pi(k_i)$ that a new node will connect to the node i depends on its degree and is equal:

$$\Pi(k_i) = \frac{k_i}{\sum_{j=1}^{N} k_j}, \tag{2.29}$$

where $N \equiv N(t)$ depends on time, as the network grows. To describe results of many realizations of the BA model we can assume that k_i is a continuous variable, characterizing average over many experiments [108]. Then, the change of the degree k_i in time is:

$$\frac{dk_i}{dt} = m\Pi(k_i) = \frac{mk_i}{\sum_{j=1}^{N} k_j}, \tag{2.30}$$

because at each time step the node i has m possibilities of accumulating a link. The sum $\sum_{j=1}^{N} k_j$ over all degrees is just twice the number of edges in the network $2M$. Every new node introduces m new links at each time step. Therefore, the number of edges added by new nodes in time t is $M(t) = mt$. For a large network, i.e. in the limit $t \to \infty$, we can omit the initial $m_0(m_0 - 1)/2$ edges and write:

$$\frac{dk_i}{dt} = \frac{k_i}{2t}. \tag{2.31}$$

The initial condition for this equation is given by $k_i(t_i) = m$, where t_i is the time when the node i was added to the network. It is straightforward to solve the equation, obtaining a formula for the degree growth in time:

$$k_i(t) = m\sqrt{\frac{t}{t_i}}. \tag{2.32}$$

To obtain the degree distribution in the continuum approximation first we shall find the cumulative probability distribution $\mathcal{P}(k_i \leq k)$. If $k_i \leq k$ at time t, then from Equation 2.32 we have:

$$m\sqrt{\frac{t}{t_i}} \leq k \implies t_i \geq t\frac{m^2}{k^2}, \tag{2.33}$$

which is the condition for the time t_i of the appearance of the node i such that $k_i \leq k$. As the maximal time of appearance is t, this condition requires that $t_i \in [t\frac{m^2}{k^2}, t]$. This interval is well defined, because $k \geq m$ for any node. New nodes are introduced at each time unit, therefore the number of nodes with degree smaller than k is equal to the size if this time interval $\Delta = t - t\frac{m^2}{k^2}$. To obtain probability we have to divide this number by the number of all nodes $N = t + m_0 \approx t$, what gives the final form of the cumulative distribution:

$$\mathcal{P}(k_i \leq k) = 1 - \frac{m^2}{k^2}. \tag{2.34}$$

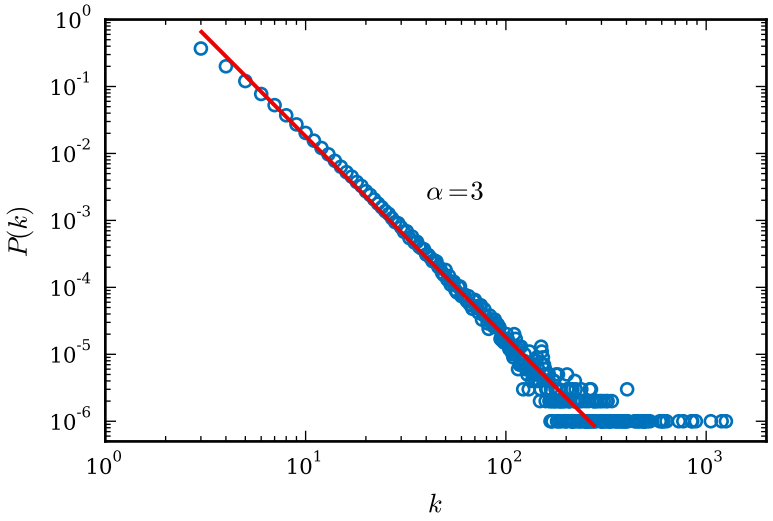


Figure 2.10: Degree distribution of a network in the BA model. Red line represents the analitical result from Equation 2.35 for $m = 3$. Blue circles stand for results of a numerical simulation with $N = 10^6$, $m_0 = m = 3$, averaged over 500 realizations.

After differentiating we arrive at the formula for the degree distribution in the BA model, which is a power law:

$$P(k) = \frac{2m^2}{k^3}. \tag{2.35}$$

The function from Equation 2.35 is presented in Figure 2.10 and compared with results of a numerical simulation. The numerical results slightly deviate from the analytical expression, especially for small k. This is because the continuum approximation is an asymptotic solution. The exact formula can be derived using a master equation [109], [110].

Another intriguing property of a BA network is the behavior of the average path length. As it was reported [111], [112], for a large N limit:

$$\langle d \rangle \sim \frac{\ln N}{\ln \ln N}, \tag{2.36}$$

what is an even slower growth than in a random graph. For that reason scale-free networks are said to display ultra-small-world property. Although the BA model explains fat-tailed degree distribution in networks and preferential attachment is a distinguished process in many real systems, scale-free networks as the random ones fail in explanation of the clustering. The transitivity in the BA model scales as:

$$C \sim \frac{(\ln N)^2}{N}, \tag{2.37}$$

what is better (bigger) than in ER graphs, but still insignificant when compared to the empirical values. There were, however, extensions of the model solving this issue [113], [114]. Recently, there was also a vigorous discussion weather scale-free networks are actually common or rare [115]–[118]. Nevertheless, the BA model is a milestone in network science and even if scale-free networks were rare, they would still be interesting due to their special properties [73].

2.2.5 Constant rewiring

Another example of an evolving network is a graph of fixed number of nodes N and links M with the links constantly being rewired. There is no growth in such model, but preferential attachment can be applied. This kind of evolving network is particularly relevant, because in adaptive networks the structure's evolution is usually achieved exactly by link rewiring. Not all the considerations, however, apply to coevolving networks due to their much higher complexity caused by existence of nodes states and the feedback between the structure and state evolution. Nevertheless, this simple model of link rewiring might be very descriptive. Additionally, it allows to chose arbitrary type of preferential attachment, similarly to the models developed in Chapters 3 and 4.

The model of evolving network with constant rewiring can be defined as follows. We start from constructing a random graph of N vertices and M edges [8]. Then, in every time step [5], [64], [119]:

1. Randomly select a link[9], then randomly select one of its ends. Detach this end from the node.

2. With probability $\Pi(k_i)$ select a node i and attach the rewired link to it.

The procedure is repeated until the graph reaches a stationary state with order parameters fluctuating over the average values. To indentify the degree distribution $P(k)$ in this model we shall use the *rate equation* approach, which is equivalent to the master equation. Let us consider the number of nodes $N(k,t)$ of degree k at time t. There are two possibilities to change this number. (i) If a node of degree k is chosen to be detached from, the number $N(k,t)$ will drop by 1. Probability of this event is equal $k/2M$ (for one particular node), because there is $2M$ ends of edges in the network and k of them are attached to the node. (ii) If a node of degree k is chosen to be attached to, the number $N(k,t)$ will increase by 1. Probability in this case, for one node, is just $\Pi(k)$ by definition. Putting these points together we can write down the

[8] The initial configuration doesn't matter as long as every node has probability of gaining new links grater than zero, i.e. it can be attached back to the network, if the initial configuration happens to be disconnected.

[9] Every time when something is *randomly selected* without specifying the distribution of this random draw, it means the distribution is flat with equal probability of any outcome.

rate equation for $N(k,t)$:

$$\begin{aligned} N(k,t+1) = N(k,t) - \Pi(k)N(k,t) + \Pi(k-1)N(k-1,t) \\ - \frac{k}{2M}N(k,t) + \frac{k+1}{2M}N(k+1,t). \end{aligned} \tag{2.38}$$

Dividing both sides by N and reorganizing factors we arrive at:

$$\begin{aligned} P(k,t+1) - P(k,t) = \left[\Pi(k-1)P(k-1,t) - \frac{k}{2M}P(k,t)\right] \\ - \left[\Pi(k)P(k,t) - \frac{k+1}{2M}P(k+1,t)\right], \end{aligned} \tag{2.39}$$

where $P(k,t) = N(k,t)/N$ is the degree distribution at time t. We are interested in the stationary state at $t \to \infty$ which does not depend on time: $\lim_{t\to\infty} P(k,t) = P(k)$. The left-hand side of Equation 2.39 becomes then zero and we can write:

$$\Pi(k)P(k) - \frac{k+1}{2M}P(k+1) = const. = 0, \tag{2.40}$$

were we assumed that both square brackets from the right-hand side of Equation 2.39 must be equal in the stationary state in order to establish effectively zero flow between $P(k)$ and both neighboring states $P(k-1)$ and $P(k+1)$. Otherwise, the distribution would change its shape and this would mean it's not the stationary state yet. The constant being zero comes from the condition $\lim_{k\to\infty} kP(k) = 0$, which must hold for the distribution to be normalized. Finally, we can obtain the direct relation between the degree distribution and the preferential attachment rule:

$$\Pi(k) = \frac{k+1}{2M}\frac{P(k+1)}{P(k)}. \tag{2.41}$$

The question is, what kind of preferential attachment $\Pi(k)$ in the rewiring process will lead to a power-law degree distribution? Substituting $P(k) \sim k^{-\alpha}$ to Equation 2.41 we obtain:

$$\begin{aligned} \Pi(k) \sim (k+1)\left(\frac{k+1}{k}\right)^{-\alpha} = (k+1)\left(1+\frac{1}{k}\right)^{-\alpha} \\ \approx (k+1)\left(1-\frac{\alpha}{k}\right) = k+1-\alpha+\mathcal{O}(\frac{1}{k}), \end{aligned} \tag{2.42}$$

were we assumed large k and $\lim_{k\to\infty}\mathcal{O}(1/k) = 0$. Therefore, a linear relation can lead to a fat-tailed degree distribution. But not only the asymptotic behavior of $\Pi(k)$ matters. It is crucial how it approaches this asymptote. To

illustrate it let's assume exponential degree distribution this time. If we substitute $P(k) \sim \mathrm{e}^{-k/k_0}$ into Equation 2.41 we obtain:

$$\Pi(k) \sim (k+1)\left(\frac{\mathrm{e}^{-(k+1)/k_0}}{\mathrm{e}^{-k/k_0}}\right) = (k+1)\mathrm{e}^{-1/k_0} \sim k+1. \tag{2.43}$$

This striking result shows how different outcomes we can observe from two linear relations. Second requirement to achieve scale-free property in constantly rewired network is sufficient density of links [64]. Although meaningful, this approach can not be directly applied in coevolving networks, due to the influence of nodes states on the topology and non-trivial feedback loop between them. Also the manner of link rewiring is important. It can be shown that for large networks the results are the same, if the whole edge is detached and rewired with both ends selecting new nodes according to the preferential attachment rule $\Pi(k)$ [64]. However, if we perform rewiring by selecting a node at random and detaching one of its links from the neighbor, the dynamics of the process changes.

2.3 Coevolving networks

Models described so far were either providing rules to construct a static network with given properties or defining dynamics governing evolution of a network. In both cases nodes are indistinguishable, they posses no state other then purely topological, i.e. the number of connections. These kinds of models aim at describing structures that we observe in the empirical data. If there is some dynamics, it's the *dynamics of network*, with no insight into what might be happening on the generated structure.

On the other hand, equally active branch of science focuses on the analysis of the *dynamics on networks*. In this kind of models, an already generated and static network is employed and a given process taking place on this structure is studied. Therefore, even if the graph was initially obtained using one of the evolving network models, there is no evolution of the network anymore. Once it's generated its static version is considered. A complete list of processes studied in this manner on different structures would be massive. To give a few significant examples: opinion formation processes [120]–[122], rumor spreading [123], [124], systemic risk and its connections to topology [123], [125], [126], language evolution and competition [127]–[130].

Obviously, the topology of the network can significantly influence the outcome of the process taking place on it. For instance, one of the most prominent results in network science illustrates it – the epidemic threshold vanishes when changing the topology of the network from random to scale-free [131]–[133]. And vice versa, in real-world networks the state of nodes may affect the topology of the network. It can change the preference in link rewiring or attachment. For example, even a homogeneous initial attractiveness of nodes in the BA model changes the exponent of the power law [109], [110], [134]. For this reason, it is rational to investigate both processes – the

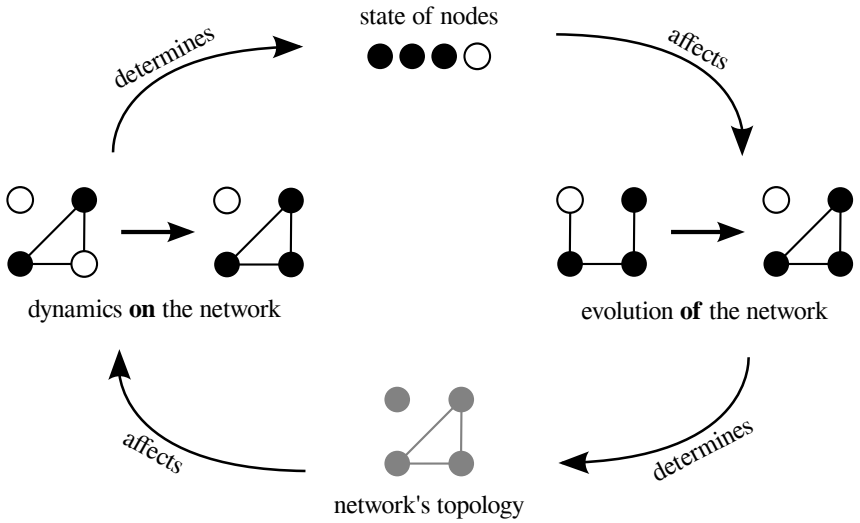


Figure 2.11: A schematic representation of the feedback loop between the topological evolution of a network and the dynamics of nodes states existing in the coevolving network models.

network's evolution and the nodes states dynamics – at once. Moreover, in many empirical systems it is impossible to separate one from another and collective analysis is inevitable. Models incorporating evolution of the network's topology and dynamics of the state of nodes together are called *adaptive* or *coevolving*[10] network models [135].

Although it's still a relatively young field with much to be discovered, a number of coevolving network models was proposed so far, covering phenomena ranging from Boolean network dynamics and synchronization to social opinion formation and ecological food-webs evolution [136]–[157]. Adaptive networks hitherto displayed intriguing properties. Frequently, reported phenomena contains formation of complex topologies, dynamical and sometimes critical self-organization, and diverse phase transitions in the structure and in the state of nodes. A network may become dynamically fragmented, or shattered. High topological heterogeneity among nodes can arise. Often times, seemingly distinct coevolving systems can exhibit similar behavior and can be collectively described within one logical framework. This similarity rationalizes the notion of coevolving network and common analysis of such systems.

Another fascinating aspect of adaptive networks is their potential applicability in autonomous artificial systems. The *internet of things* is becoming real, we posses more and more *smart* devices that have to communicate with each other according to a certain local rules. What emerges from this local interactions defines behavior of the system. It was shown that power grid fluctuations can be reduced thanks to such cooperation [158]. But there is more possible applications. Autonomous robots, self-organizing communication systems, self-adapting production lines etc. In every case all we can

[10]Sometimes a *coevolutionary* network model can be encountered.

define is local rules and all we would wish for is the anticipated global behavior. We want to be sure about the connection between this two opposite levels of a given system.

It is important to understand that the outcome of an adaptive network model is not a simple sum of results of separated underling processes. More precisely, let's say we independently analyze a rewiring mechanism and node states dynamics. We conclude that the rewiring produces a topology A and the dynamics leads to a state B. Then, we can not assume a coevolving model merging the two would display the state B on the topology A. Moreover, often times it's entirely wrong and the system's outcome is something that can't be achieved separating the evolution of the network from the dynamics on the network. It is caused by a non-trivial feedback loop present in adaptive systems (see Figure 2.11). The structure of the network affects the dynamics of state. And the state of the nodes affects the evolution of the network. For instance, in self-organized criticality sandpile model of Bak-Tang-Wiesenfeld [159], [160] the avalanche dynamics and the network's topology may influence each other [161]. When we allow susceptible agents to cut connections with infected ones in epidemic models, we can observe completely new phases [162]. Coevolution gives a whole new dimension to game theory strategies, for example in prisoner's dilemma [163], [164], if we allow agents to rewire their links. All these lessons inspire the extension of existing network models to adaptive versions and provoke a question about the limits of the static approach. They also create the context for this thesis and three models developed within it that are described in the following chapters. It is also an intriguing question, which topological properties of a network can be affected by a given coevolution rule. Answer to this question is one of the central points of the rest of the thesis.

Chapter 3

The voter model

The simplest possible dynamical model on a network needs at least two different states of nodes [165]. Then, rules governing the dynamics must be specified, what can be done in numerous ways. Among equilibrium models, probably the most well known and broadly applied is the Ising model [15], being frequently a starting point for a more complex analysis. When it comes to non-equilibrium or algorithmic models [62], the counterpart of the Ising model is the voter model. The former will be discussed later, the latter is the subject of this chapter.

The voter model has been described, analyzed, simulated and applied under many different names. There is a certain amount of redundancy in the literature reporting results of binary-state models defined by the same time evolution rules. It's no surprise, if one realizes how many potential applications such simple model has. Where by simple I mean a simple definition and basic algorithm, not the outcome of the model which can be and usually is rather complex. To have an idea which models may be associated with the wide definition of the voter model we can distinguish two most important traits of voter-like models. The first one is (i) the binary state of nodes, where the exact numerical values chosen for these two states are not essential. The second feature is (ii) defining the model through possible events for every time step, discrete or continuous, where the basic event is a *state copying*. Further details, like updating one node at a time or the whole network, may vary. We can assume all to all interactions, i.e. a complete graph topology, or a set of agents placed on a one-dimensional square lattice. It can be one-on-one interaction, or a nonlinear interaction, there can be an external field or not etc. But if (i) and (ii) are satisfied, we can refer to the model as a version of the voter model[1].

The first description of a voter-like model, to my best knowledge, was given in 1973 [166]. In this short article authors studied two possible processes taking place on a one- and two-dimensional square lattice of black and white cells. The cells could change their color due to a swapping process or invasion process. The second one is actually consistent with most of the current definitions of the voter model – one node can influence another to change its state. However, the authors proved that both processes are equivalent when it comes to the probability of occupying a given cell by a given

[1] This aggregation under a common name of *voter model* is present in most of the recent literature cited in this chapter. However, in different fields the model can have different names.

color. The probability of event occurrence (swapping or invasion) was defined for continuous time steps. Therefore, it is clear that (i) and (ii) were satisfied in this work.

The first time the name *voter model* was used was two years later in [167]. In this mathematical study exploring ergodicity of infinite systems the possible states are 0 and 1, and the time parameter is desecrate. Although the setup might sound quite distinct from the previous one, it also satisfies (i) and (ii).

To understand the redundancy in many names of the voter model it is best to look at its applications. It has been applied to describe species competing over a territory [166], behavior of ants facing a choice between two food sources, or people selecting a restaurant [168], stylized facts about stock markets [169], opinion dynamics [170], innovation diffusion [171] and voting process [172], to name the main ones. Obviously, the name comes from the last listed application, but the roots of the model can be found in statistical physics of non-equilibrium systems. The model was implemented in a description of a heterogeneous monomer-monomer catalytic process [173] and dimer-dimer surface reactions [174]. In the social context, where the voter model is most frequently applied, two states can correspond to different opinions and the state-copying process to social imitation, i.e. the tendency of people to adapt opinions of others.

3.1 Static network

As for probably all network models, the voter model was thoroughly studied on static networks before the first extension introducing link rewiring came. One of possible divisions of the work done in this context is by the network's topology with one class being a complete graph and the other everything else, i.e. complex networks. When every node is connected to all the others the interaction with neighbors may be considered as an interaction with an aggregated state of the whole network, whereas in complex networks the local structure must be analyzed neighbor by neighbor. Consequently, different mathematical tools are applied to these classes. A brief description of both follows.

3.1.1 Homogeneous case

In the simplest version the voter model is defined and studied on a complete graph, therefore *homogeneous* refers to the structure. Every node is exactly the same from the topological point of view. This approach is sometimes considered a mean-field approximation of the voter model on a complex network with a relatively homogeneous degree distribution.

In some of the articles about the voter model on a complete graph it is not even acknowledged that there is any graph at all, since as mentioned in Section 2.1.2 a complete graph can be seen as no graph at all. There is no underling structure limiting possible interactions – every agent can and

do interact with every other. Therefore, to characterize the playground it's sufficient to specify the number of agents.

The model can be defined in a following manner. Consider a system of N nodes creating a complete graph, i.e. every node can interact with every other. There are only two possible states of each node, as for any version of the voter model, which are $s_i \in \{-1, +1\}$, however the exact values vary between different publications. By n we denote the number of nodes in the state $+1$, and accordingly the number of nodes in the state -1 is $N - n$. Every node is free to change its state. The transition rates per unit time for a single node i are:

$$\begin{aligned} \pi_i^+ &\equiv \pi_i(-1 \to +1) = a + hn, \\ \pi_i^- &\equiv \pi_i(+1 \to -1) = a + h(N - n), \end{aligned} \tag{3.1}$$

where the parameter a represents noise or individual random choices and the parameter h represents the herding mechanism, forcing agents to become more similar to the others. It is therefore clear why the model is also referred to as a *herding model*. The bigger the noise parameter a the more independent are changes in states of the nodes. The higher value of the herding parameter h the more unanimous agents become and we should expect more ordered behavior with most of the nodes being in agreement.

Since it doesn't matter which particular nodes have a given state, but rather how many of them, it is more convenient to analyze the system globally. Accordingly, the parameter $n = 0, 1, 2, ..., N$ fully determines the state of the system. The whole system has in principle many more possible states $(N + 1)$ than a single node (2).

We assume that the probability of any change depends only on the last state of the system, i.e. it is a Markov process, hence we need transition probabilities $P_{1|1}(n', t + \Delta t|n, t)$ for every n and n' to describe it[2]. However, for $\Delta t \to 0$ the probability of two jumps occurring in the same time interval goes to zero and we can neglect the changes of $\Delta n = n' - n > 1$. The dynamics is usually described by transition rates per unit time, which depend on the transition probabilities as $P_{1|1}(n', t + \Delta t|n, t) = \Delta t \pi(n \to n')$. Taking into account possible individual changes from Equation 3.1 we get [169]:

$$\begin{aligned} \pi^+(n) &\equiv \pi(n \to n + 1) = (N - n)(a + hn), \\ \pi^-(n) &\equiv \pi(n \to n - 1) = n[a + h(N - n)]. \end{aligned} \tag{3.2}$$

The transition rate for remaining in the same state $\pi(n \to n)$ may be obtained from the normalization condition $\sum_{n'} P_{1|1}(n', t + \Delta t|n, t) = 1$. Note that transition probabilities should be smaller than 1 per time increment, so discrete simulations of the model are possible up to an upper limit of Δt which is given, for large N, by $\Delta t_{max} = 2/hN^2$. Equations 3.2 provide an *extensive* definition of the noisy voter model on a complete graph. The transition rates

[2] We need also the initial distribution, but it is usually assumed to be equal $P_1(n, t = 0) = \delta_{n,N/2}$, or other arbitrary value.

increase with the system size N. Sometimes an alternative, *intensive* definition is used:

$$\begin{aligned}\pi^{+}_{int}(n) &= \frac{N-n}{N}\left[a + h\frac{n}{N}\right],\\ \pi^{-}_{int}(n) &= \frac{n}{N}\left[a + h\frac{N-n}{N}\right].\end{aligned} \tag{3.3}$$

Here the rates depend only on the relative state of the system, i.e. the fraction of the nodes in the positive state n/N. It can be understood as an equivalent of only local interactions with an assumption that the neighborhood is homogeneous, i.e. identical for every node. In this version the model is also called the *Kirman model* [168], [175]. For any finite size system the difference between two definitions can be eliminated by rescaling the herding parameter h, because rescaling of the whole transition rate by a constant $(1/N)$ changes only the time scale. In the thermodynamic limit, however, the difference is essential. The equilibrium probability distribution for intensive case becomes a Dirac delta. In other words, the fluctuations of the system can be neglected, and its behavior becomes deterministic [176]. Therefore, we will use the extensive definition 3.2 for further considerations.

The transition rates 3.2 imply a Master equation governing the time evolution of probability $P(n,t)$ of having n positive nodes at the time t:

$$\begin{aligned}\frac{\Delta P(n,t)}{\Delta t} =& P(n-1,t)\pi^{+}(n-1) + P(n+1,t)\pi^{-}(n+1)\\ &- P(n,t)[\pi^{-}(n) + \pi^{+}(n)].\end{aligned} \tag{3.4}$$

The master equation gives a very convenient framework for an analytical study of the system, but it is useful to first replace the extensive variable n by an intensive one, namely the magnetization $m = \sum_i s_i/N = 2n/N - 1$. From the definition $m \in [-1,+1]$. Zero magnetization will signify a disordered state of the system with half of the nodes in positive and half in negative states, while $m = \pm 1$ will mean a full agreement with all nodes in the same state. Note that $\Delta m = 2\Delta n/N$, which for a continuous approximation of large N leads to the relation between the probability distributions $P(m,t) = P(n(m),t)N/2$.

Replacing the variable in the Master equation 3.4 for the continuous one we can obtain the Fokker-Planck equation [177] describing the diffusion process visible in the system:

$$\frac{\partial P(m,t)}{\partial t} = -\frac{\partial}{\partial m}\left[\mu(m)P(m,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial m^2}\left[D(m)P(m,t)\right], \tag{3.5}$$

where $\mu(m)$ is the drift and $D(m)$ is the diffusion coefficient. By expanding the transition rates in the Master equation 3.4 and taking into account only

terms up to the second order in $1/N$ we arrive at the final form of the Fokker-Planck equation for the mean-field noisy voter model [169]:

$$\frac{\partial P(m,t)}{\partial t} = \frac{\partial}{\partial m}\left[2amP(m,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial m^2}\left[\left(\frac{4a}{N} + 2h(1-m^2)\right)P(m,t)\right], \quad (3.6)$$

where the drift coefficient is equal $\mu(m) = -2am$ and the diffusion coefficient $D(m) = 4a/N + 2h(1-m^2)$. An alternative way of representing dynamics of a stochastic system is by a stochastic differential equation known as Langevin equation [177]. Having the Fokker-Planck equation, the transformation is straightforward giving:

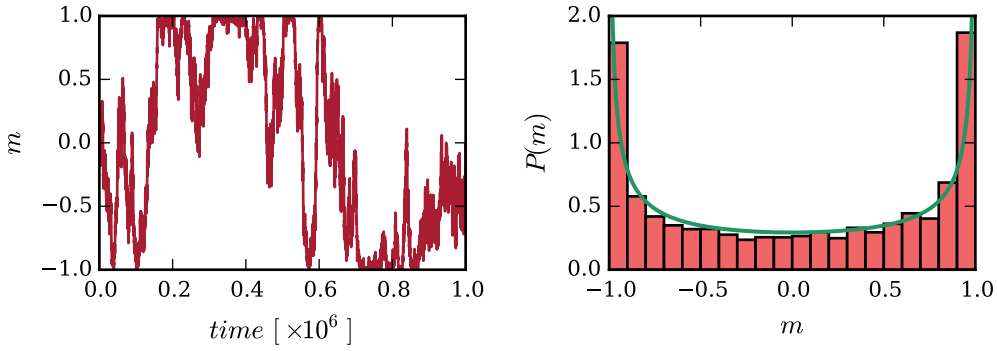
$$\frac{dm}{dt} = \mu(m) + \sqrt{D(m)}\cdot\eta(t) = -2am + \sqrt{\frac{4a}{N} + 2h(1-m^2)}\cdot\eta(t), \quad (3.7)$$

where $\eta(t)$ is a Gaussian white noise, i.e. a random variable with a zero mean value and a Gaussian distribution, $\langle\eta(t)\rangle = 0$, $\langle\eta(t)\eta(t')\rangle = \delta(t-t')$. Analyzing the Langevin equation 3.7 we can see what is the role of a and h parameters. First, the term $4a/N$ vanishes in the continuous approximation $N \to \infty$. It is responsible for finite size fluctuations. The main stochastic component $2h(1-m^2)$ is the biggest for $m = 0$ and goes to zero when $|m| = 1$. Therefore, it shifts the system to one of the boundary consensus states, as it was expected from the herding parameter h. The deterministic term $-2am$, on the other hand, tends to bring the system to the disordered configuration at $m = 0$, where it vanishes. This is also consistent with the intuition for the parameter a, which stood for the random individual choices among agents. From Equation 3.6 one can derive the stationary distribution of the magnetization m [178]:

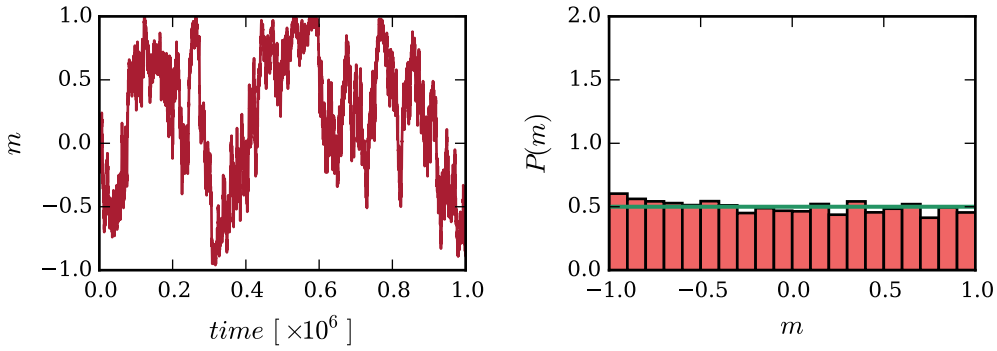
$$P_{st}(m) = \mathcal{Z}^{-1}\left[\frac{a}{2Nh} + \frac{1}{4}(1-m^2)\right]^{\frac{a}{h}-1}, \quad (3.8)$$

where the $\mathcal{Z}$ is a normalization constant. See Appendix B for a full derivation of the stationary solution of the Fokker-Planck equation. The theoretical stationary distribution $P_{st}(m)$ of magnetization is compared with results of a simulation in Figure 3.1. We can observe a phase transition in the system when changing the ratio $\epsilon = a/h$ of the parameters. In the consensus phase (Figure 3.1a)) the stationary distribution is bimodal, meaning that the most probable configuration is either $m = -1$ or $m = 1$, i.e. full consensus with all nodes possessing the same state. The transition happens at the noise rate equal to the herding intensity $\epsilon = 1$ and is marked by a flat distribution with any configuration equally possible. Then, for $\epsilon > 1$ we can observe disordered fully-mixing phase with a unimodal stationary distribution peaked over $m = 0$. Different character of the dynamics is also visible in the trajectory plots in Figure 3.1.

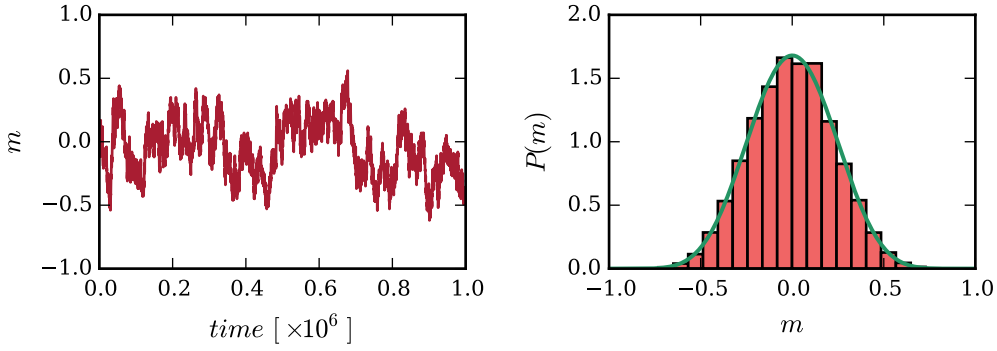
The herding model is likely the simplest version of the voter model, yet it can generate interesting dynamics and was applied in various phenomena, for instance explaining investors behavior at stock markets [169], [176],



a) Results for $\epsilon = 0.4$ – the consensus phase



b) Results for $\epsilon = 1$ – the transition line



c) Results for $\epsilon = 10$ – the disordered phase

Figure 3.1: Single trajectory of the mean-field noisy voter model (left panel) and the stationary probability distribution of magnetization m (right panel) for different vales of $\epsilon = a/h$ and $N = 200$. The histogram is created based on 20 runs of the simulation for 2 million time steps each. The green solid line represents the analytical solution from Equation 3.8. Even for a relatively small N we can observe a good agreement between the theory and simulations.

[178]–[180]. However, it should be noted that financial systems display numerous effects which require more complex analysis [181]–[183]. The model was also extended to encompass asymmetric noise [169], asymmetric herding mechanism and external information effect [178] or existence of zealots in the society [184]. It was also studied for nonlinear interactions [185] and more complex structures [186].

3.1.2 Complex networks

Next step in bringing the model's assumptions closer to real-world networks is changing the topology for more complex one. Complete graphs are present in social groups in a relatively small scale, as a part of the whole system, but definitely they do not describe the global structure. A certain part of a social network can be densely or even fully connected, however large scale real-world networks are sparse.

The voter model on complex networks is normally defined by rules for every desecrate time step, in contrast to the transition rates describing the model on a complete graph. Nevertheless, both approaches have strong similarities in assumptions (binary state, imitation rule) and in results (order-disorder phase transition). Therefore, they are both commonly called *voter model*, although many applications have little to do with a voting process. Still, the most well known application concerns elections and the process of arriving at a certain state of opinions. The model proved to be very effective in this field. One variant of the voter model with effectively just one parameter was able to explain and reconstruct stylized facts about U.S. presidential elections [172].

The model can be defined as follows. We start by creating an initial graph of N nodes and M edges, giving the average degree $\mu = 2M/N$. The initial topology may vary and it will influence the dynamics of the system. Each node is assigned one of the states $s_i \in \{-1, 1\}$ at random[3]. Therefore, on average there is the same number of positive and negative nodes leading to zero magnetization $m = 0$ at the beginning. Then, in every time step:

1. select randomly a node i from the network, it is called *the active node*,
2. select randomly a node j from the neighbors of the active node i,
3. if the node i has a different state than its neighbor j ($s_i \neq s_j$), it changes its state to become the same as the node j, i.e. $s_i \to s_i' = -s_i = s_j$; otherwise nothing happens [4].

The process is continued until a frozen configuration is reached. In the frozen configuration every node has the same state, so the state copying can no

[3] Every time something is done *at random* without specifying the probability distribution it means the distribution is uniform, i.e. probability is constant for every outcome. In this case both states have equal probability $1/2$.

[4] Note, that this point could be rephrased to always copy the state of a random neighbor, what would result in no action, if the neighbor was in the same state. However, for more convenient comparison with other models it is given as above.

longer occur. Note, that the number of nodes N and edges M is constant in time. This algorithm is an example of a *node-update rule*. An alternative is a *link-update rule*, which is equivalent in regular lattices and homogeneous networks, but is significantly different in heterogeneous ones, for instance scale-free networks [187]. Interestingly, the voter model has certain similarities to the zero-temperature Glauber dynamics in the Ising model [188].

The aggregated state of the system is characterized by the magnetization $m = \frac{1}{N}\sum_i s_i$, but for a given value of m the $+1$ and -1 states can be distributed in many different ways over the network. In order to take into account such possibility and measure it an additional order parameter is required. A good choice is the density of *active links* ρ, also called *interface density*. An active link is a link connecting two nodes in opposite states. The number of active links of a node i is denoted by a_i, therefore the local density of active links is given by $\rho_i = a_i/k_i$. Accordingly, the total interface density is $\rho = \frac{1}{2M}\sum_i a_i$. Naturally, when $\rho = 0$ the system has reached an absorbing configuration where nothing can change, what also corresponds to $|m| = 1$, if the network is connected.

Interface density also changes during a simulation in a complete graph, but in that case it is fully determined by the value of magnetization, therefore it would be redundant to use both. For a complete graph with n nodes in the $+1$ state the density of active links is:

$$\rho_{K_N} = \frac{n(N-n)}{N(N-1)/2} = \frac{N}{N-1}\frac{1}{2}(1-m^2) \approx \frac{1}{2}(1-m^2), \tag{3.9}$$

where the dependency $n = (1+m)N/2$ was used. In complex networks this relation is not deterministic. Nevertheless, the stochastic attractor of the dynamics also has a parabolic shape, but with a different maximum. It can be shown , that in random networks with average degree μ the attractor is given by [189]:

$$\rho_a = \frac{\mu-2}{2(\mu-1)}(1-m^2). \tag{3.10}$$

In Figure 3.2 we can see how single trajectories form a shape described by Equation 3.10 to finally arrive at the absorbing state with $m = \pm 1$ and $\rho = 0$.

In the voter model on a static network we can distinguish two phases. The first one is a disordered fully-mixing phase, where $m = 0$ and $\rho \neq 0$. It is equivalent to the typical initial configuration with every node having the same probability of possessing $+1$ or -1 state. The second phase is an ordered consensus phase, with $m = \pm 1$ and $\rho = 0$. Therefore, it is also an absorbing state. For any finite-size network the system will eventually end up in a frozen configuration. In the thermodynamic limit, however, the consensus phase is not always obtained and the final outcome depends on the topology of the network. In regular lattices dimensionality has a significant influence on the dynamics with meta-stable disordered states prevailing for $d > 2$ [190], [191]. In complex networks the average survival time of meta-stable disordered states decreases with network heterogeneity and with the

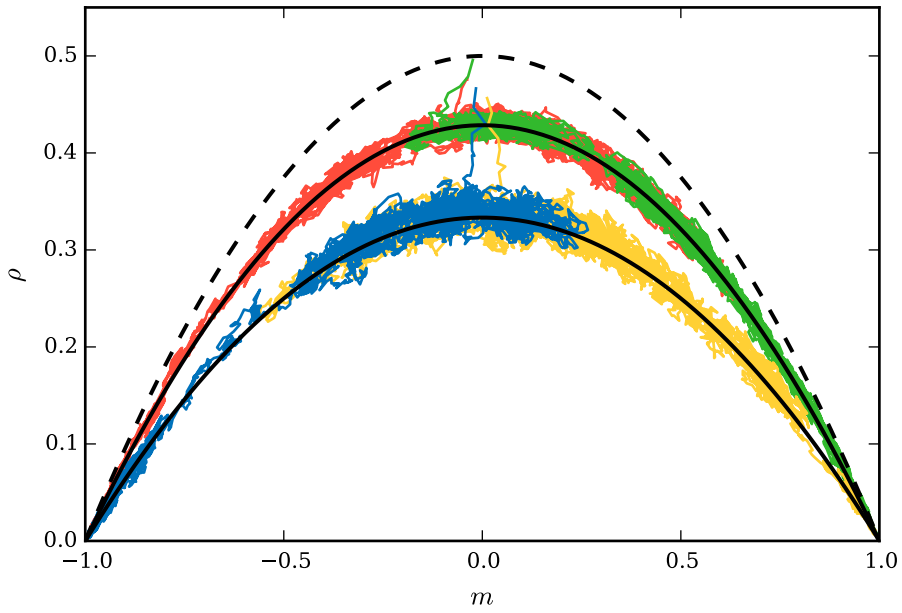


Figure 3.2: Trajectories of the voter model on a static ER graph with $N = 1000$, $\mu = 4$ (blue, yellow) and $\mu = 8$ (red, green). For both values of the average degree there are two trajectories finishing in opposite absorbing states. Solid black lines indicate the attractor ρ_a from Equation 3.10. Dashed black line stand for the complete graph relation from Equation 3.9.

randomness of small-world networks for a finite N [170], [192], [193].

An important extension of above-described dynamics is the noisy voter model [194]. It accounts for inevitable fluctuations observed in real systems. The most important difference is that an absorbing state no longer exists in the presence of noise [195]. The effect of noise shall be described in a greater detail in next sections together with non-linearity.

The voter model on complex networks was also extended by an aging mechanism [196], where nodes having the same state for longer time are less likely to change their opinion. It was also studied on bipartite graphs [197], [198]. Another extension focused on nonlinear interactions [199]. A popular way of introducing non-linearity is done in so-called q-voter model [185], [200]–[204]. In this variant not one but q neighbors of the active node are chosen at random. If they all share the same state, the active node also adapts it, otherwise there is no change. Usually, for the sake of simplicity repetitions are allowed, i.e. the same neighbor may be chosen many times, therefore it is possible to consider q bigger than a given k_i. This is a discrete case of non-linearity. In this thesis the continuous one is developed (Section 3.3).

3.2 Coevolving voter model

The most important extension of the voter model, in the context of this thesis, is the *coevolving voter model* (CVM) [189]. As argued in Section 2.3, it brings the model much closer to the real-world networks and allows to include different topological mechanisms. The CVM is similar to the voter model on static networks, but adds a new possibility of rewiring links. More precisely, we start as previously with a random graph of N nodes and M edges. Here, the initial topology of the network doesn't matter much, because it will change in time due to the rewiring. And the type of rewiring determines the final structure, as in the model described in Section 2.2.5. The initial distribution of states $s_i = \pm 1$ is random. Then, similarly to the previous algorithm, in every time step:

1. select randomly an active node i from the network,
2. select randomly a node j from the neighbors of the active node i,
3. if the node i has a different state than its neighbor j ($s_i \neq s_j$):
   (a) with a probability p the active node i disconnects from the node j and connects to a randomly chosen node l with the same state[5], i.e. $s_l = s_i$,
   (b) with a probability $1 - p$ the node i changes its state to become the same as the node j, i.e. $s_i \to s_i' = -s_i = s_j$,

   otherwise (if $s_i = s_j$) nothing happens.

[5] Rewiring to a node with the same state is performed, if such a node exists and is not a neighbor already. If there is no such node nothing happens.

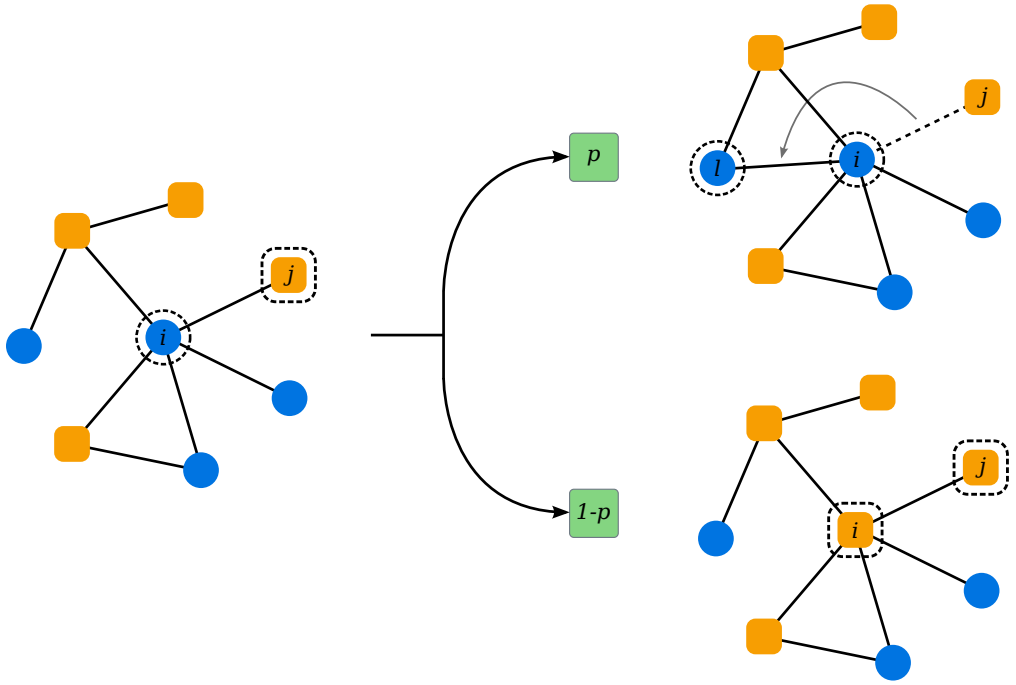


Figure 3.3: A schematic illustration of the CVM algorithm. After choosing the active node i and one of the neighbors j in a different state rewiring is performed with a probability p, or state copying with a complementary probability $1 - p$.

As we can see, now in the case of disparity between two nodes either the state copying or rewiring can occur, with the latter having a probability p. Therefore, for $p = 0$ we have the ordinary voter model on a static network. The probability p is a new parameter of the model called *plasticity*. Note, that the network must be a simple graph at every point with no multi- or auto-connections, thus not every node can be chosen during the rewiring procedure. Number of vertices and edges is constant during the simulation. The algorithm is illustrated on a scheme in Figure 3.3.

For low values of the rewiring probability p trajectories of CVM look qualitatively the same as for the static version, with a small shift of the attractor's line (Figure 3.4). For a large p, however, the trajectory is significantly different – it follows a vertical line straight to the absorbing configuration of $\rho = 0$ (see purple trajectory in Figure 3.4). In this case it arrives at the frozen state of $m = 0$, but the final magnetization depends on its value in the initial configuration. The new frozen state with $\rho = 0$ and $m = 0$ is possible only if the network splits, so no nodes of opposite states are connected. A natural question is how does the final configuration depend on the plasticity p.

Before answering this question it is worth to have a look at the analytical description of the system. In the mean-field approximation it is possible to construct a rate equation for the dynamics of the number of active links [189], [205]. From there it is straightforward to derive the time evolution of the

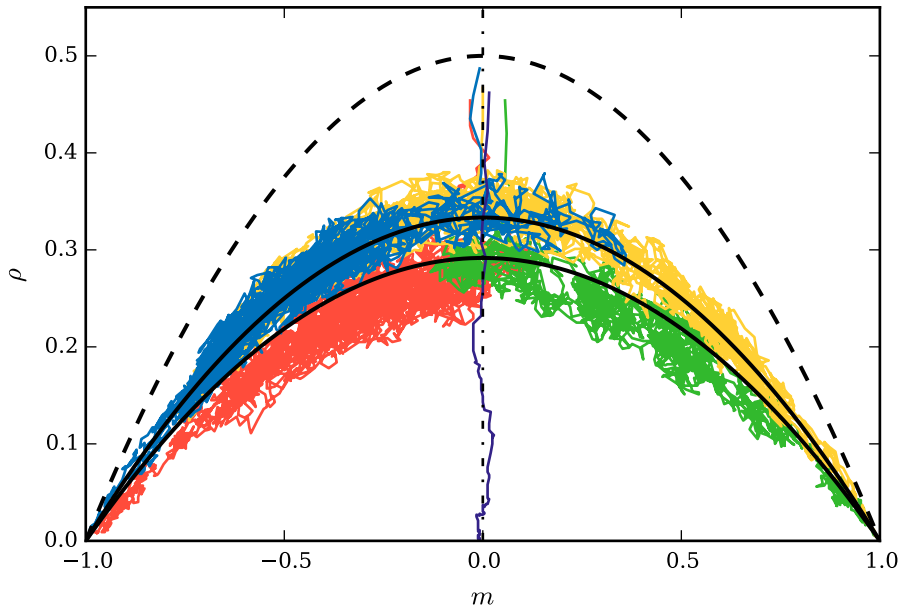


Figure 3.4: Trajectories of the CVM for $N = 500$, $\mu = 4$ and $p = 0$ (blue, yellow), $p = 0.2$ (red, green) and $p = 0.8$ (purple). For $p = 0$ and $p = 0.2$ there are two trajectories finishing in opposite absorbing states. The trajectory for $p = 0.8$ goes straight down. Solid black lines indicate the attractor ρ_a from Equation 3.13. Dashed black line stand for the complete graph relation from Equation 3.9.

interface density in the thermodynamic limit:

$$\frac{d\rho}{dt} = \frac{2\rho}{\mu}\left[(1-p)(\mu-1)(1-2\rho)-1\right], \tag{3.11}$$

where as always $\mu = 2M/N$ is the average degree. This equation has two possible solutions, whose stability depends on p. Let p_c be the boundary value separating regions where different solutions are stable. Then, for $p > p_c$ the stable solution is $\rho = 0$, while for $p < p_c$ the stable solution is given by:

$$\rho^* = \frac{(1-p)(\mu-1)-1}{2(1-p)(\mu-1)}. \tag{3.12}$$

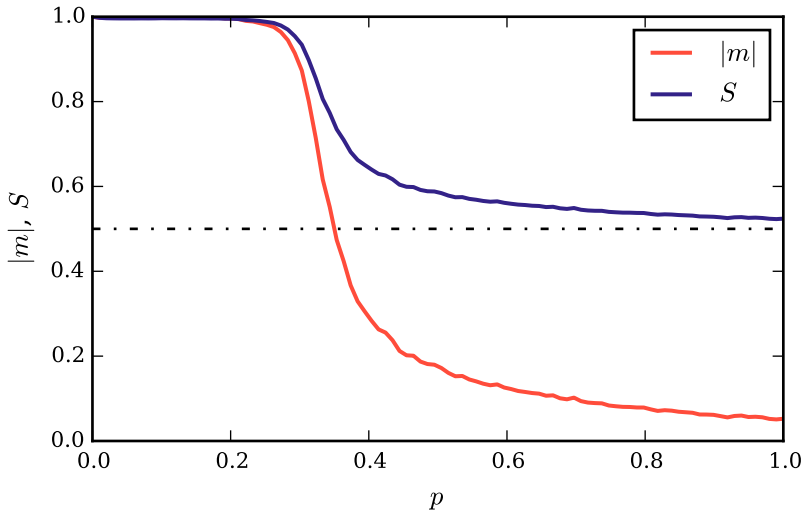
This is yet another indication of two distinct phases in the system. Additionally, the shape of the dynamical attractor can be described using ρ^*:

$$\rho_a = \rho^*(1-m^2), \tag{3.13}$$

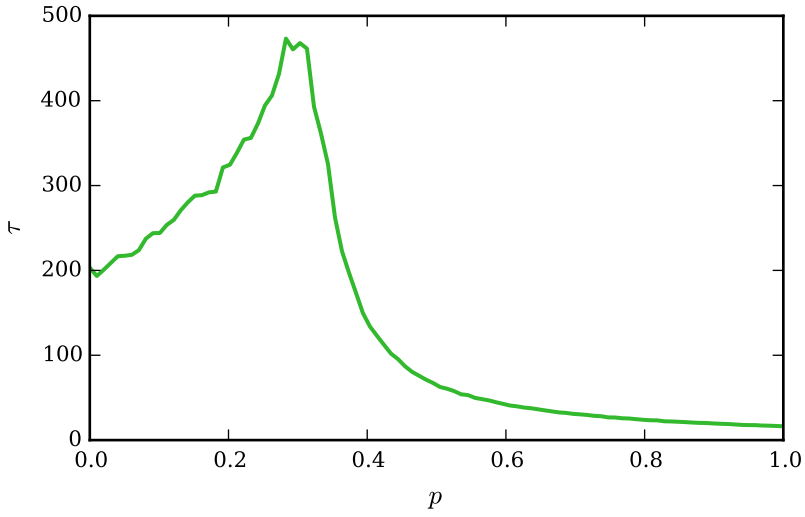
what for $p = 0$ becomes the relation for static networks from Equation 3.10. For any finite size of the network the evolution will finish in the absorbing state of $\rho = 0$, as we can see in Figure 3.4, due to the finite-size fluctuations. Nonetheless, it has been shown [189] that during the dynamics ρ stays for a long time at a plateau for $p < p_c$ before suddenly decreasing. In the thermodynamic limit the system never leaves this plateau, remaining active in a stationary state, where the average value of order parameters does not change in time. A more precise analytical description of the CVM can be obtained using motifs approximation [206].

Since the structure of the network can change in this model it is desirable to use, in addition to the interface density ρ and magnetization m, also topological order parameters such as the size of the largest component S. In Figure 3.5 we can see the behavior of order parameters with a varying plasticity p. It clearly indicates existence of two phases. The first one, for $p < p_c$, is a consensus phase with full magnetization $m = \pm 1$ and one connected component $S = 1$. Increasing p above p_c the network splits into 2 separate components, i.e. a *fragmentation transition* occurs. In addition, magnetization drops to zero, because two components have opposite state. This phase is called a *fragmentation phase*. In the thermodynamic limit the transition becomes discontinuous [189] and only the second phase ends in a frozen configuration, while the first one remains active. For that reason it is also referred to as an *absorbing transition*. The value of p_c can be obtained from the convergence time. It peaks at the transition point, as visible in Figure 3.5b).

It is also worth mentioning that rewiring to a random node despite its state, instead of rewiring to same-state nodes as here, changes properties and position of the phase transition and the final magnetization [207]. Interesting effects occur when analyzing the CVM on a multilayer network as well. Depending on the time scale on each of the layers the model can display the same fragmentation transition as the ordinary CVM with adequate effective average degree, or it can produce a *shattered phase transition*, where many nodes isolate from the network [208]. Finally, non-linearity of interactions



a) absolute magnetization $|m|$ and size of the largest component S



b) average convergence time τ of the simulation in MC steps

Figure 3.5: Dependence of the order parameters (normalized) on the plasticity p in the CVM for $N = 250$ and $\mu = 4$. Results are averaged over 10^3 realizations. A clear fragmentation transition can be observed, indicated also by a drop of magnetization and a peak of the convergence time.

has crucial influence on the dynamics of the system and the characteristics of phases. These effects are described in the subsequent sections.

3.3 Nonlinear coevolving voter model

3.3.1 Global rewiring

The CVM is a very good framework to study influence of different mechanisms on the behavior of a coevolving system, regardless of its interpretation. Nevertheless, the model is mostly associated with social sciences and opinion dynamics. From this perspective an interaction between two nodes, be it rewiring or state copying, is a social interaction between two people, which is not always linear, i.e. the effect of social pressure doesn't have to be proportional to the fraction of people having opposite opinion. This possibility has been acknowledged in social impact theory [209], in language competition dynamics [210], [211], or in language evolution problems [212]. It is therefore reasonable to analyze the effect on non-linearity in microscopic interactions on the macroscopic characteristics of the system.

In order to enable nonlinear interactions the CVM is extended by a parameter q [213]. As always, we begin with a random graph of N nodes and M edges. Every node is randomly assigned one of two states $s_i = \pm 1$. Then, every time step consists of following operations:

1. select randomly an active node i from the network,
2. with probability ρ_i^q an interaction occurs (step 3); otherwise nothing happens, i.e. with probability $1 - \rho_i^q$ go to the next time step,
3. select randomly a node j from the neighbors of the node i in the opposite state, then:
   (a) with probability p the active node i disconnects from the node j and connects to a randomly chosen node l with the same state, i.e. $s_l = s_i$,
   (b) with probability $1 - p$ the node i changes its state to become the same as the node j, i.e. $s_i \to s_i' = -s_i = s_j$.

The local interface density ρ_i gives the fraction of active links attached to the node i, or in other words the fraction of neighbors in a different state. As defined in Section 3.1.2, $\rho_i = a_i/k_i$, where a_i is the number of active links of the node i and k_i is the total number of its links. For $q = 1$ the probability of interaction is proportional to the fraction of neighbors in the opposite state a_i/k_i. This is also the probability of choosing a neighbor in a different state during a blind selection. Therefore, when $q = 1$ the model defined by the above algorithm is fully equivalent to the ordinary linear CVM. For sublinear interactions ($q < 1$) the probability of an action occurring is higher than in the linear case, since $\rho_i \in [0, 1]$. Consequently, the active nodes can be

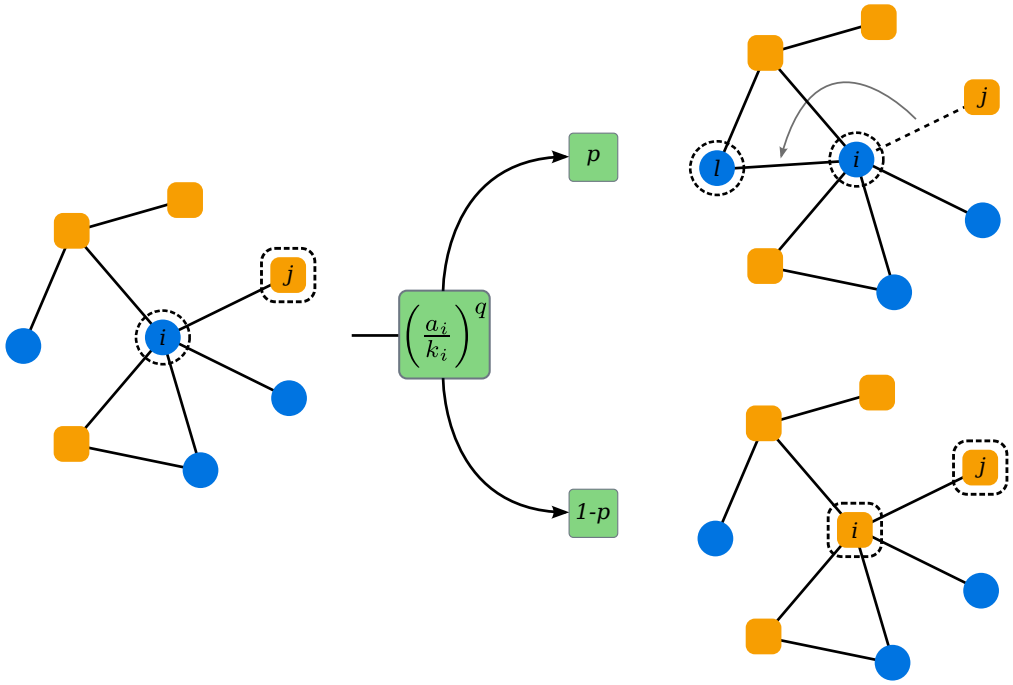


Figure 3.6: A schematic illustration of the nonlinear CVM algorithm. After choosing the active node i an interaction occurs with probability $(a_i/k_i)^q$. Then, one of the neighbors j in a different state is selected and rewiring is performed with a probability p, or state copying with a complementary probability $1-p$.

influenced by a smaller fraction of neighbors in a different state. The opposite happens for super-linear interactions when $q > 1$. The algorithm of the model is illustrated in Figure 3.6.

Conceptually, the difference between the ordinary coevolving voter model and the nonlinear CVM is similar to the difference between simple and complex contagion processes [214], [215]. The influence on the active node comes not from one randomly selected neighbor, but depends on the aggregated state of the neighborhood.

When the parameter q takes an integer value the model becomes equivalent to the q-voter model [199] mentioned in Section 3.1.2. In this variation q neighbors are chosen at random with possible repetitions, and only if all of them share the same state, they can influence the active node. Probability of choosing a neighbor in a different state is equal $a_i/k_i = \rho_i$, and of doing so q times with repetitions ρ_i^q, therefore the q-voter model is a special case of the model defined above for integer q. However, the parameter q can take any real value, in principle from the continuous range $[0, +\infty[$. It was argued to be smaller than 1 in social impact theory [209] and language evolution problems [212]. On the other hand, it was approximated to be 1.3 in language extinction processes [210]. Thus it makes a perfect sense to analyze a continuous spectrum of values.

Using a pair-approximation approach [216], [217] we can describe dynamics of two basic order parameters, namely magnetization m and interface density ρ. Since the network is structurally homogeneous due to the random rewiring, we can assume each node to have the same average degree $\mu = 2M/N$. In the thermodynamic limit evolution of the system is described by [213]:

$$\begin{aligned} \frac{\partial m}{\partial t} &= 2(1-p)(n_- n_q^- - n_+ n_q^+), \\ \frac{\partial \rho}{\partial t} &= \frac{2}{\mu}\left[(1-p)(n_+ n_q^+ \delta_+ + n_- n_q^- \delta_-) - p(n_+ n_q^+ + n_- n_q^-)\right], \end{aligned} \tag{3.14}$$

where $n_+ = (1+m)/2$ and $n_- = (1-m)/2$ are fractions of nodes in the state $+1$ and -1 respectively, and $n_q^\pm \equiv (\rho/2n_\pm)^q$. The symbol $\delta_\pm$ denotes a change in the total number of active links given that a node i such that $s_i = \pm 1$ flipped its state, and $\delta_\pm = \mu - 2q - 2(\mu - q)\frac{\rho}{2n_\pm}$. More detailed discussion of a generalized version of Equations 3.14 and their derivation is given in Section 3.4, where also noise is taken into account.

Equations 3.14 have three kinds of steady state solution with $\frac{\partial m}{\partial t} = \frac{\partial \rho}{\partial t} = 0$, which are $(m, \rho) = (\pm 1, 0)$, $(0, \rho^*)$, and $(m^*, 0)$. The last solution $(m^*, 0)$, where $m^* \neq \pm 1$, corresponds to a fragmentation phase, because the only possibility to obtain no active links and non-zero magnetization is by separating nodes of opposite states in two clusters. However, the exact value of m^* depends on the initial conditions. In our case $m_0 = 0$, leading to clusters of the same size and consequently $m^* = 0$. Therefore, the last stationary solution becomes effectively $(m, \rho) = (0, 0)$.

For $m = 0$ the density of $+1$ nodes and -1 nodes is the same $n_+ = n_- = 1/2$, giving $n_q^\pm = \rho^q$ and $\delta_\pm = \mu - 2q - 2(\mu - q)\rho$. The stationary version of the first equation from 3.14 is then always satisfied, as the right-hand side is also zero. The second equation reduces to:

$$\rho^q\{-p + (1-p)[\mu - 2q - 2(\mu - q)\rho]\} = 0, \tag{3.15}$$

what gives either already discussed $\rho = 0$ or a stationary value ρ^*:

$$\rho^* = \frac{(1-p)(\mu - 2q) - p}{2(1-p)(\mu - q)}. \tag{3.16}$$

Note that for $q = 1$ we recover the stationary solution for the linear CVM (Equation 3.12) as expected. The value of ρ^* decreases with p. Additionally, we know that for $p = 0$, i.e. on a static network, the solution $(m, \rho) = (0, 0)$ can not exist, as the network can't separate into two clusters. On the other hand, for $p = 1$ there is only link rewiring in the system, which can't change the magnetization and must lead to $(m, \rho) = (0, 0)$ state. Therefore, somewhere between these values for $p_c \in]0, 1[$ a phase transition must occur, in which ρ^* drops to zero. The value of p_c can be obtained from Equation 3.16

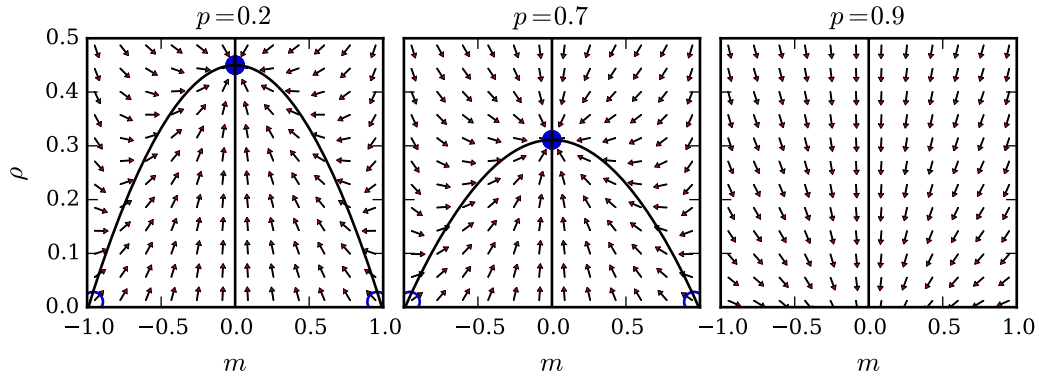


Figure 3.7: Flow diagram of the system dynamics in the (m, ρ) space for $\mu = 8$, $q = 0.5$, and different values of p. Arrows represent the dynamical direction of the system according to the pair approximation solution (Equations 3.14). Fixed points are represented by full circles (stable) or empty circles (unstable).

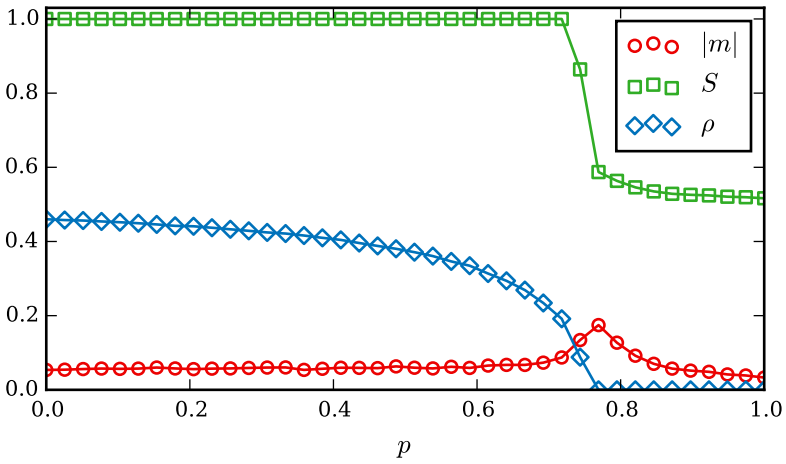
for $\rho^* = 0$. It is equal:

$$p_c = \frac{\mu - 2q}{1 + \mu - 2q}. \tag{3.17}$$

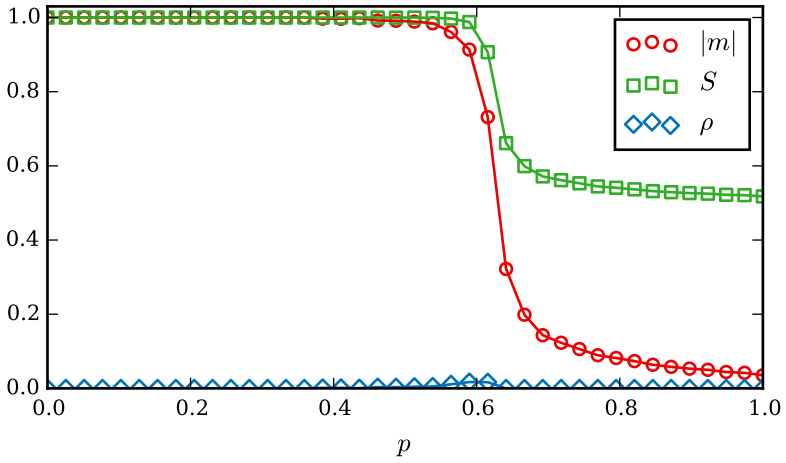
Hence, the position of the phase transition depends on the non-linearity parameter q.

For $p > p_c$ the fragmentation phase $(m, \rho) = (0, 0)$ is a stable solution for any value of q (which has to be fixed first, because $p_c \equiv p_c(q)$). For $p < p_c$ there are two possible solutions: $(m, \rho) = (\pm 1, 0)$, or $(0, \rho^*)$. However, their stability depends on the value of q. For $q < 1$ the solution $(0, \rho^*)$ is stable, corresponding to a dynamically active fully-mixing phase. It is illustrated on a flow diagram in Figure 3.7. But for $q > 1$ the solution $(\pm 1, 0)$ becomes stable, which describes an ordered full-consensus phase. Moreover, in the active disordered phase ($q < 1$) ρ^* decreases to become 0 at the transition point $p = p_c$ and the magnetization is 0 for all values of p. This indicates a continuous absorbing phase transition. However, for $q > 1$ the interface density is 0 for any value of p, but the magnetization is always equal ± 1 for $p < p_c$ and zero for $p > p_c$, indicating a discontinuous fragmentation transition in the thermodynamic limit. As a result, the character of phase transitions strongly depends on the non-linearity of interactions q. Described phenomena occurring when changing p can be observed in numerical simulations as well, what is shown in Figure 3.8. However, due to relatively small size of the network it is difficult to conclude about the continuity of the transitions based solely on these plots. The full scaling analysis was given in [213].

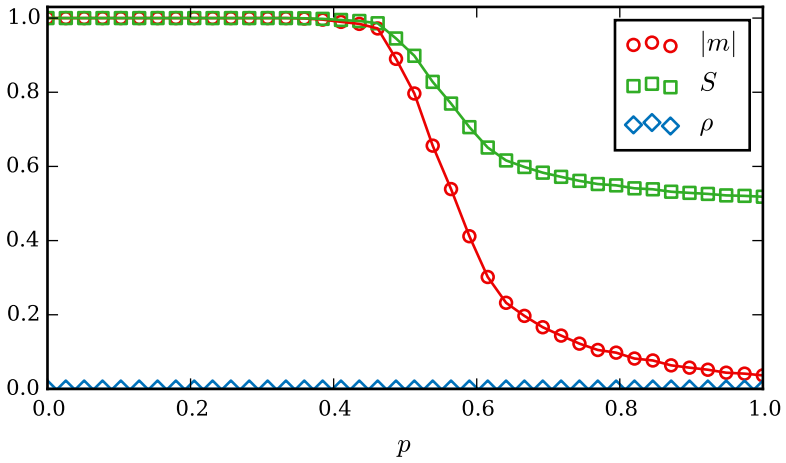
In conclusion, the nonlinear coevolving voter model has three distinct phases (Figure 3.9). For $p < p_c$ and $q \geq 1$ the solution is $(m, \rho) = (\pm 1, 0)$. This corresponds to a consensus phase with all nodes in the same state and the network having one big and connected component. When $q < 1$, still being below p_c, the solution is $(m, \rho) = (0, \rho^*)$, which is a disordered fully-mixing phase. Here both states $+1$ and -1 are equally distributed in the



a) Results for $q = 0.5$



b) Results for $q = 1$



c) Results for $q = 2$

Figure 3.8: Size of the largest component S, magnetization $|m|$, and interface density ρ for $N = 500$, $\mu = 8$, averaged over 500 realizations in the steady state of the nonlinear CVM. All values are normalized.

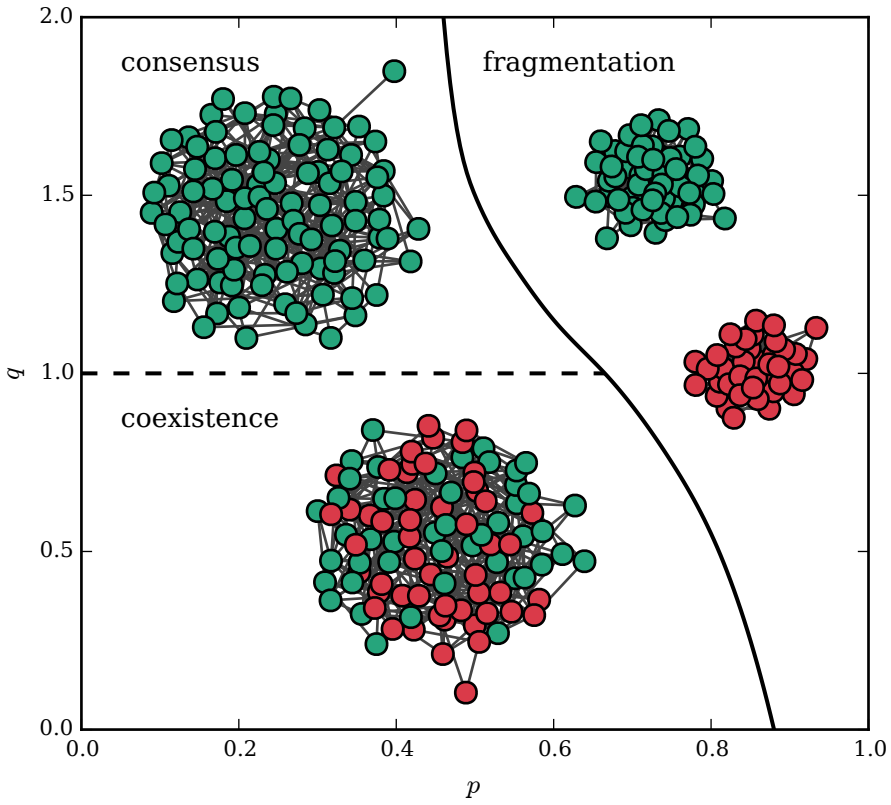


Figure 3.9: Scheme of the phase diagram in the nonlinear coevolving voter model [213] for $N = 10^4$ and $\mu = 8$. Differences between phases are illustrated on smaller networks in a frozen (consensus, fragmentation) or stationary (coexistence) configuration. Green and red colors indicate opposite states of the nodes.

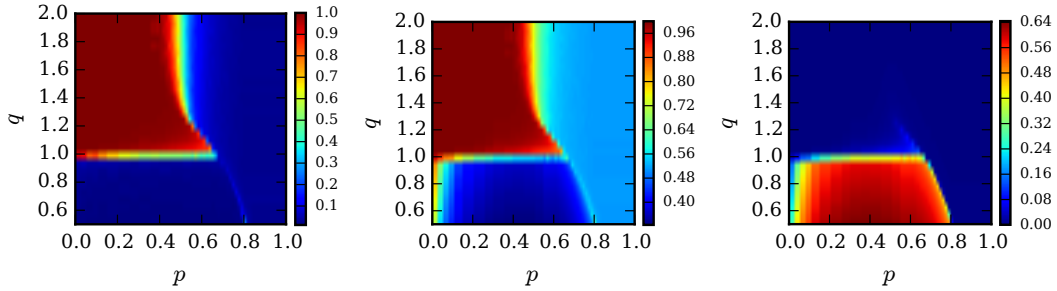
network, which is also connected with one component. Note, that this active phase is predicted in the thermodynamic limit, since for a finite size of the network it will always fall into an absorbing phase. However, as simulations show (Section 3.3.2 and [213]), the convergence time τ in Monte Carlo steps scales for the fully-mixing phase exponentially with the system size $\tau \sim e^N$, making the absorbing state unreachable in practice for relatively small network sizes. For comparison, when $q = 1$ the scaling is linear $\tau \sim N$ or weaker for other phases. The active phase can also exhibit spontaneous symmetry braking [218]. Finally, for $p > p_c$ the solution is $(m, \rho) = (0, 0)$ indicating a fragmentation phase for any value of q. The network splits into two components of similar size and opposite state, as illustrated in Figure 3.9.

3.3.2 Local rewiring

The nonlinear coevolving voter model is a general framework, which can be further extended to include particular effects. It takes into account important features of real-world networked systems and produces results that can be compared with empirical data. In its standard form, however, it lacks a very important mechanism. So far, in described models rewiring was performed randomly with equal probability for any node in the network to be chosen, i.e. it was a global random rewiring. For small networks it is a reasonable assumption, but when increasing the size of the system eventually the possibility of connecting two arbitrary vertices, despite the topological distance between them, ceases to make sense. Then, it becomes natural to include some kind of *local rewiring*, which narrows down the group of possible new connections.

In the context of social sciences, a very important local mechanism of acquiring new connections is so-called *triadic closure*, which reflects the tendency of people to search for new contacts through the current neighbors. It has been proven to play a decisive role in social interactions [87], [219], [220] and has been observed in a variety of systems. It has been also incorporated in a number of models analyzing social dynamics having a significant influence on the outcome [220]–[224].

The mechanism is simple – it assumes that an agent can create a new connection only to a node distant by two edges, i.e. the path length l between the agent and the potential new neighbor must be equal 2. Effectively, it means that agents look for new links via existing ones and connect with neighbors of their neighbors. As a consequence, triadic closure increases value of clustering coefficient, as it literally closes triangles of nodes. This effect is particularly desirable, because social networks display high values of clustering coefficient [63], [87], [88], [225]. Additionally, triadic closure has been recognized as an important psychological and sociological mechanism [28], [29]. It has been also found in empirical studies on social systems [26], [27], [87], [226]. Therefore, it is well motivated to study effects of triadic closure on the behavior of network models, and so is done in this thesis. This section contains original results obtained in the nonlinear CVM extended to account for the triadic closure mechanism.



a) Magnetization $|m|$ b) Largest component S c) No. of components n_S

Figure 3.10: Phase diagram in the (p, q) space for three main order parameters: a) magnetization, b) size of the largest component, and c) number of components. Results are obtained from a simulation of networks with $N = 5000$, $\mu = 8$, averaged over 500 realizations. Based on these order parameters, it is easy to distinguish three phases: a consensus phase, a fragmentation phase, and a shattered phase.

Technically, the nonlinear CVM with triadic closure [227] can be described similarly to the ordinary one with a global rewiring. The main points of the algorithm are the same as in Section 3.3, and Figure 3.6 sums up also this model. The only difference is in the rewiring procedure. Here, if the active node i cuts the connection to a neighbor j and looks for a new neighbor l in the same state, it's not chosen from the entire network. Instead, only nodes distant by two edges are considered, so the node l must be a neighbor of a neighbor of the node i. As previously, if there is no such node l that meets these requirements, nothing happens – rewiring is skipped and algorithm goes to the next time step. Multi-connections and auto-connections are prohibited, as always.

Analytical description of a model with the triadic closure mechanism is hard to obtain, if achievable at all. For this reason such models are analyzed almost exclusively by numerical simulations, and so is done in this thesis. Results can be described in terms of three quantities: the absolute magnetization $|m|$ the relative size of the largest connected component S of the network, and the relative number of separate components n_S. Phase diagrams for these quantities in the (p, q) parameter space are shown in Figure 3.10. Clearly, three distinct phases can be observed. For $q > 1$ and $p < p_c(q)$ an absorbing consensus phase, characterized by the maximal value of the absolute magnetization and a single large component, is visible. Except much higher clustering, this phase is equivalent to the one found in the previous section. When $q < 1$ and $p < p_c(q)$ we obtain a shattered phase, where the absolute magnetization drops to zero and the network is composed of an active component and a number of nodes separated from the network (see Figure 3.11). In this shattered phase many nodes that initially belong to the main component are detached from the network during the evolution, resulting in a big number of components. Therefore, the shattered phase can be identified

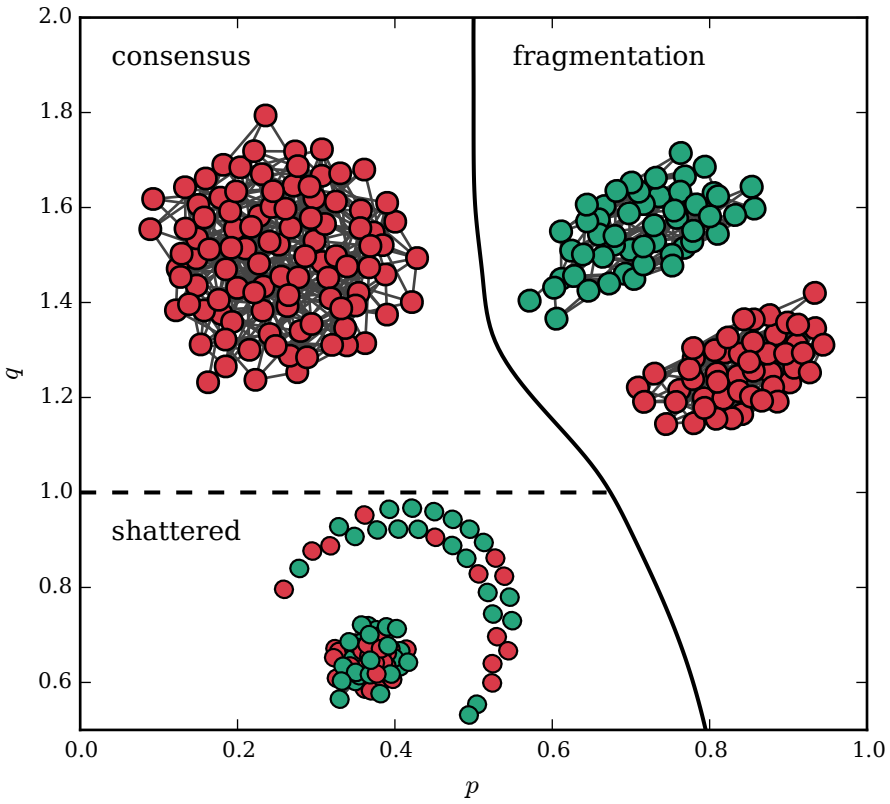


Figure 3.11: Scheme of the phase diagram in the nonlinear CVM with triadic closure, summarizing results form Figure 3.10. Differences between phases are illustrated on smaller networks in a frozen (consensus, fragmentation) or stationary (shattered) configuration. Green and red colors indicate opposite states of nodes. Note the new shattered phase, which was not observed before.

using the new order parameter n_S. Together with higher clustering, this is the main difference between the model with local rewiring and the one with global rewiring from the previous section. Increasing the value of plasticity above p_c for any q, we obtain a fragmentation phase. Here, the network splits into two components of approximately the same size being in opposite consensus states as previously. The absolute magnetization is therefore close to zero and the size of the largest component around $1/2$. Schematic illustration of each phase can be found in Figure 3.11.

The new shattered phase is a result of local rewiring. When a node is connecting to a new neighbor it can only choose from among the nodes distant by two edges. It means that a new candidate to be attached must have at least one connection and be a part of the same component. Therefore, nodes once detached from the main component cannot be attached to it again. Nevertheless, the shattered phase has finite boundaries. That is because for $q > 1$ interaction is less frequent and does not occur so easily when ρ_i is small. This

facilitates an increase of local homogeneity via state copying. On the other hand, when $p \approx 1$ rewiring is the leading mechanism, and it quickly causes fragmentation. In consequence, the shattered phase encounters boundaries when increasing q or p.

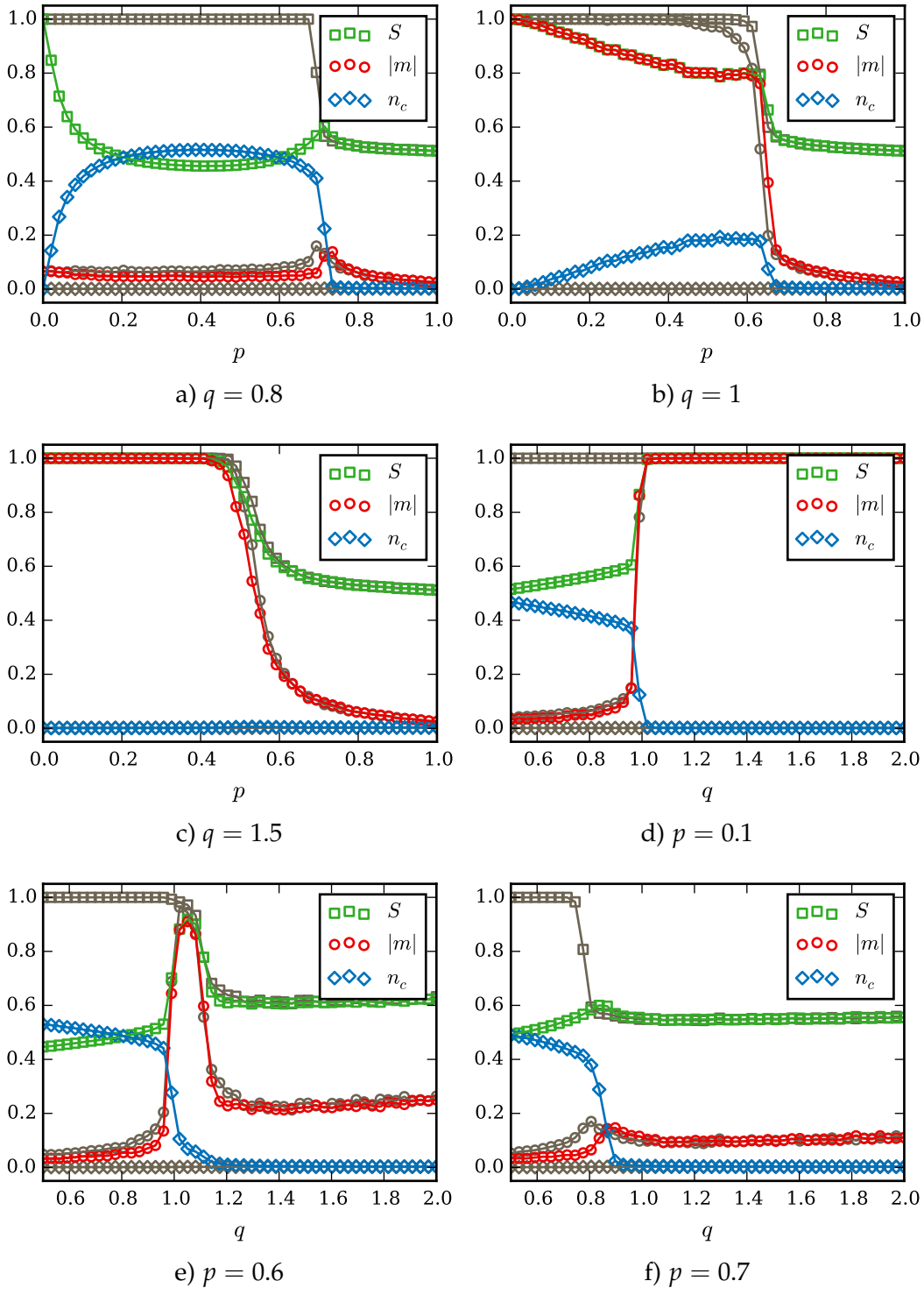
The biggest difference between the nonlinear CVM with and without triadic closure is visible in the clustering and the number of components. When looking at the transitions, however, one can see other more subtle, though visible, differences. Comparison of both models is presented in Figure 3.12. Disproportion in the number of components is clear, but we can see differences also in the magnetization and the size of the largest component. Some of them are caused by a shift of the transition to other spectrum of q and p parameters. Nonetheless, behavior at $q = 1$, for instance, is different. In the presence of triadic closure the level of consensus in the consensus phase depends on the plasticity and decreases with growing p.

For $p = 0.6$ (Figure 3.12e)) we can observe all three phases and two transitions when changing q. For small q we observe the active shattered phase. Increasing q above 1 we first obtain a transition to the consensus phase with a high magnetization and a big main component. The second transition occurs for larger values of q to the fragmentation phase with low values of $|m|$ and $S \approx 1/2$. When p increases further, i.e $p = 0.7$ (Figure 3.12f)), the consensus phase disappears and the shattered phase changes directly into the fragmentation phase with increasing q, what can be observed in the decreasing value of n_S at the transition point.

One may expect a strong dependence of the transition behavior on the average degree in the network. As a matter of fact, for $\mu \geq 4$ the transition is independent of μ in the super-linear region $q > 1$, while for $q \leq 1$ results for $\mu = 8$ are representative for a broad range of average degrees. For too sparse network $\mu \approx 4$ there is a large shift of the absorbing transition towards smaller values of p [227].

A finite system is always bound to reach an absorbing state for any combination of parameters in the limit of $t \to \infty$. But for the shattered phase the convergence time τ (in MC steps) to a frozen configuration grows exponentially with the system size $\tau \sim e^N$. That is why this phase is called *active*. In practice, it is expected to remain active even for a relatively small network due to the exponential divergence of τ. It is shown in Figure 3.13, together with the behavior for other values of parameters. For $q = 1$, corresponding to a linear interaction, and $p < p_c$ we reproduce a linear scaling with the system size $\tau \sim N$, as reported in the ordinary CVM [189]. For the consensus and fragmentation phases the growth of τ is sub-linear – either logarithmic or power-law with a small exponent.

Triadic closure should lead to a significantly higher clustering coefficient than random rewiring. This assumption is investigated in Figure 3.14 for different values of p and q in the stationary state. For nodes having less than two links local clustering coefficient is not well defined, therefore they are excluded from the analysis. While the value of the clustering coefficient for the global rewiring remains at similar level as for a random graph for all tested parameter configurations, it clearly reaches values far from zero for

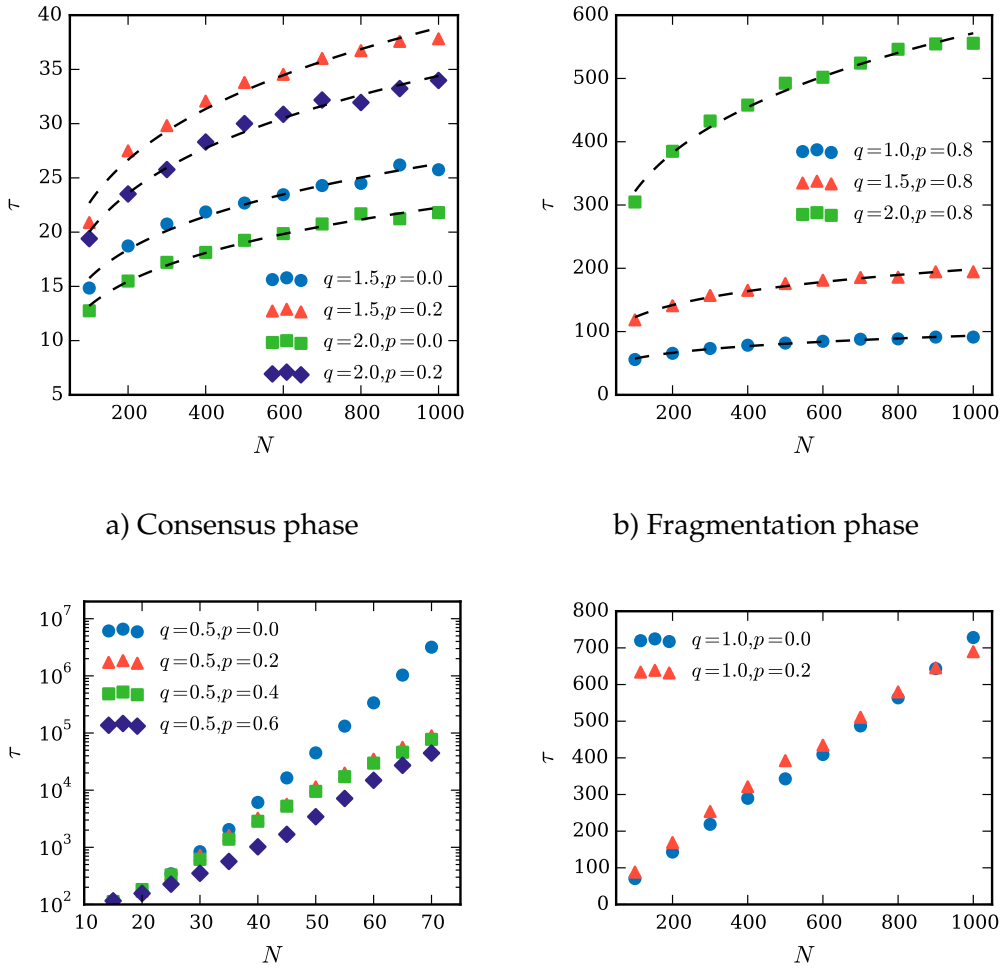


a) $q = 0.8$ b) $q = 1$

c) $q = 1.5$ d) $p = 0.1$

e) $p = 0.6$ f) $p = 0.7$

Figure 3.12: Phase transition with respect to p and fixed values of q, and with respect to q and fixed values of p. The size of the largest component S (blue squares), the magnetization $|m|$ (red circles), and the number of components n_S (green diamonds) are calculated for $N = 1000$, $\mu = 8$, averaged over 500 simulation runs. Every quantity for global rewiring (i.e. without triadic closure) is also plotted with the same symbols in gray. These plots can be understood as horizontal and vertical cross-sections of the phase diagram from Figure 3.10.



a) Consensus phase

b) Fragmentation phase

c) Shattered phase

d) Linear case

Figure 3.13: Convergence time τ in the number of Monte Carlo steps to a frozen configuration as a function of the network size N in different phases. Results are obtained for $\mu = 8$ and averaged over 500 realizations. Values of parameters are given in the plots. Dashed lines represent the best power-law fit with $\alpha \in [0.21, 0.25]$, which is slightly better than a logarithmic fit.

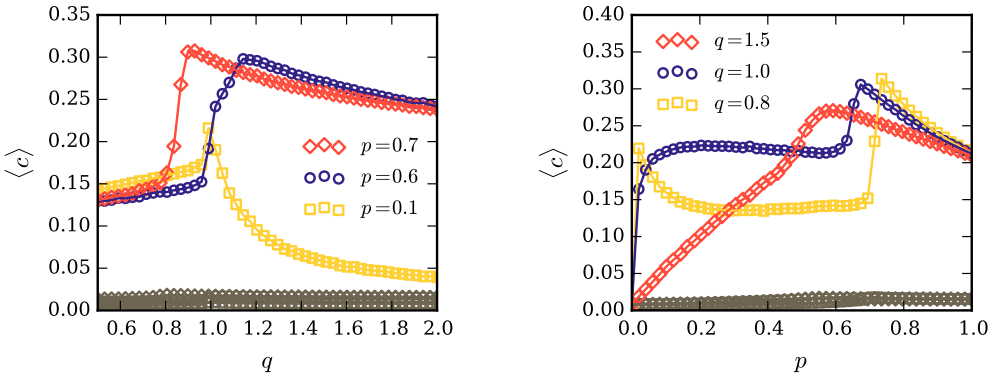


Figure 3.14: Average local clustering coefficient $\langle c \rangle$ as a function of p and q parameters, for $N = 1000$, $\mu = 8$, averaged over 500 simulation runs. Every quantity for global rewiring (i.e. without triadic closure) is also plotted with the same symbols in gray. Note significantly larger values of the clustering coefficient in a presence of the triadic closure in comparison with previous models.

the local rewiring, exceeding 0.3 in extreme cases. Additionally, the position of peaks in $\langle c \rangle$ is coincident with transition points, meaning that a raise in the clustering coefficient can be an indicator of a phase transition.

The nonlinear coevolving voter model with triadic closure displays new interesting effects that were not reported before in voter-like models. We can identify three different phases, namely consensus, fragmentation and dynamically active shattered phase. They are characterized by different topological properties and magnetization. Shattered phase, which doesn't exist for a global random rewiring, suggests that triadic closure is the origin of many isolated parts in complex adaptive systems, such as social systems. Keeping in mind the fact that real-world social systems are known to have high clustering, the nonlinear CVM with triadic closure can provide a plausible way to reconstruct such structures and describe social phenomena.

3.4 Nonlinear coevolving voter model with noise

So far we discussed the voter model starting from the simplest case on a static complete graph up to a complex modification taking into account coevolution of structure and state, non-linearity of interactions and tridic closure local rewiring. However, one very important feature of physical systems was not covered yet. This feature is noise. Any large empirical system, especially a complex one, exhibits randomness of some kind. Even if noise comes from the deterministic chaos, inevitably it is there. Therefore, regardless the answer to a question about the true nature of the subject, whether it is deterministic or stochastic, one should consider random fluctuations in the model.

Certainly, fully deterministic networks don't exist in social systems, where noise is unquestionable [165], [228]. It can manifest itself on various levels. First, people chose other people to interact with at random. The exact form of this randomness can take different forms, nevertheless the structure's evolution is never hard-coded. But the most fundamental part of randomness lays probably within individual choices. Having exactly the same influence on two people's opinions we can not be sure of the outcome. This mechanism is sometimes referred to as non-conformism [229].

The nonlinear CVM with noise is a natural extension accommodating above considerations [230]. The model is an original contribution into the landscape of voter models developed in this thesis. It's the most general version of the CVM. It can be seen as a unification of the nonlinear CVM [213] and the CVM with noise [231]. The nonlinear CVM with noise extends the first one by the noise rate ε and the second one by the non-linearity parameter q. All results presented in this section, numerical and analytical, were obtained for the first time by myself. The model is defined as follows. Firstly, a random graph of N nodes and M edges is generated and every node is assigned a state $s_i = \pm 1$ at random. Then, following operations are performed in every time step:

1. select randomly an active node i from the network,
2. with a probability ρ_i^q an interaction occurs (step 3); otherwise nothing happens, i.e. with a probability $1 - \rho_i^q$ go to the step no. 4,
3. select randomly a node j from the neighbors of the node i in the opposite state, then:
   (a) with probability p the active node i disconnects from the node j and connects to a randomly chosen node l with the same state, i.e. $s_l = s_i$,
   (b) with probability $1 - p$ the node i changes its state to become the same as the node j, i.e. $s_i \to s_i' = -s_i = s_j$,
4. with a probability ε the active node i draws a random state.

At the and of the time step, regardless of what happened before, the active node with a probability ε draws a random state, so each state has in that case the same probability equal $1/2$. Note, that this is equivalent to changing the current state with a probability $\varepsilon/2$. The algorithm of the model is illustrated in Figure 3.15. The simulation is ran until a stationary configuration or a frozen state is reached. Obviously, a frozen configuration can exist only for $\varepsilon = 0$.

Numerical simulations of the model reveal a complex behavior patterns. Figure 3.16 contains the (p, ε) phase diagram for three different values of the q parameter – the sub-linear case $q = 0.5$, the ordinary linear case $q = 1$, and the super-linear case $q = 2$. We can distinguish three general phases in the model. The first one, indicated by the red area in the figure, is a consensus phase. In this range of parameters magnetization m is close to ± 1 and the

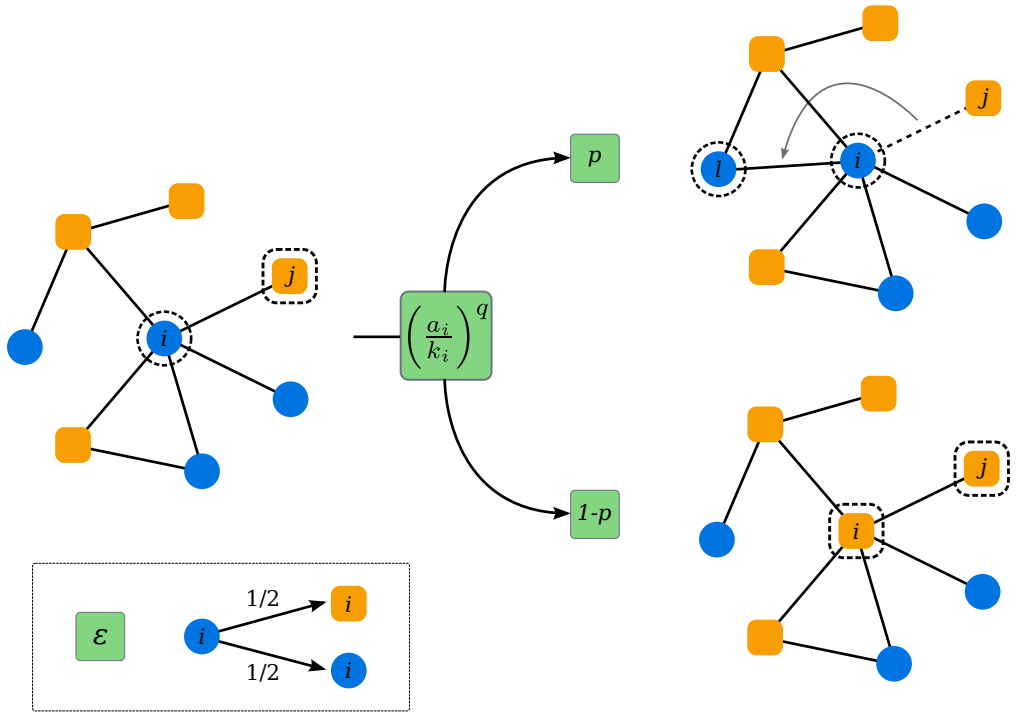
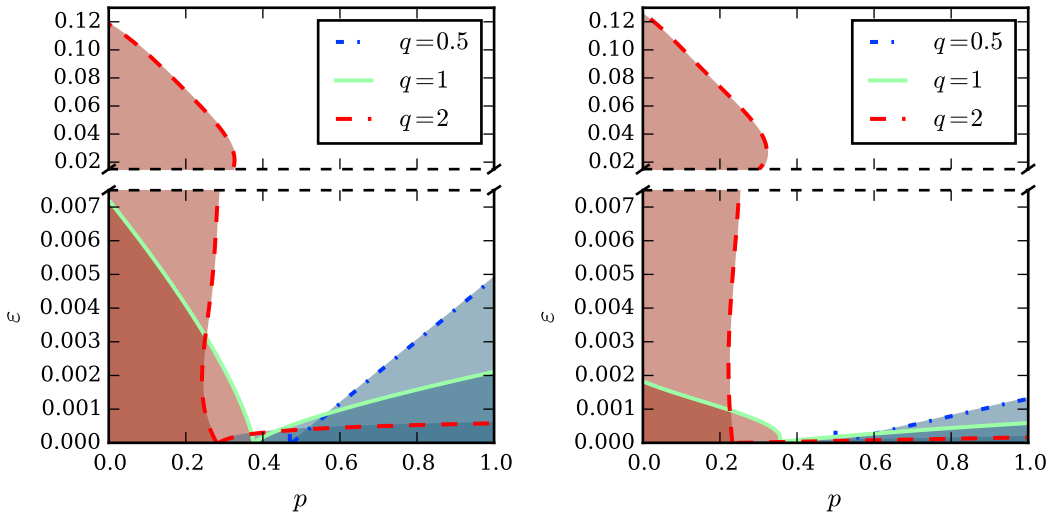


Figure 3.15: A schematic illustration of the nonlinear CVM with noise algorithm. After choosing the active node i an interaction occurs with probability $(a_i/k_i)^q$. Then, one of the neighbors j in a different state is selected and rewiring is performed with a probability p, or state copying with a complementary probability $1-p$. At the end of each time step with a probability ε the active node draws a random state.



a) $N = 250$ b) $N = 1000$

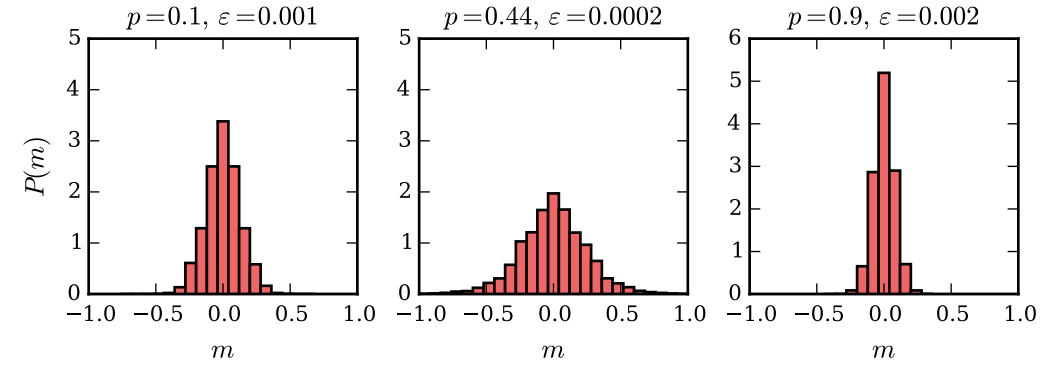
Figure 3.16: Phase diagram in (p, ε) space for $\mu = 4$ and different values of N and q. Picture is made based on simulations averaged over 500 realizations. Red area represents the consensus phase, white the coexistence phase, and blue the dynamical fragmentation phase. The border between two first phases is defined as a line of the medium absolute magnetization $|m| = 0.5$. Accordingly, the border between two last phases is a line of the medium size of the largest component $S = 0.75$.

network is connected having one large component. If we increase the noise rate ε or the plasticity p sufficiently, we obtain a coexistence phase, indicated by the white area. Here, the magnetization drops to zero $m = 0$, hence there is no consensus in the system anymore. But the network is still connected. Finally, for high values of the rewiring probability above p_c and not too big noise rates fragmentation arises. It is marked by the blue area in the figure. In this model, however, it is a dynamical fragmentation – the network consists of two separate components being in the opposite states, but it is possible that they get connected for a moment due to the noise and random rewiring, creating again one big network. We can observe constant switching between these two arrangements. If the network stays more than half of the time in the disconnected configuration, we assume it is a (dynamical) fragmentation phase.

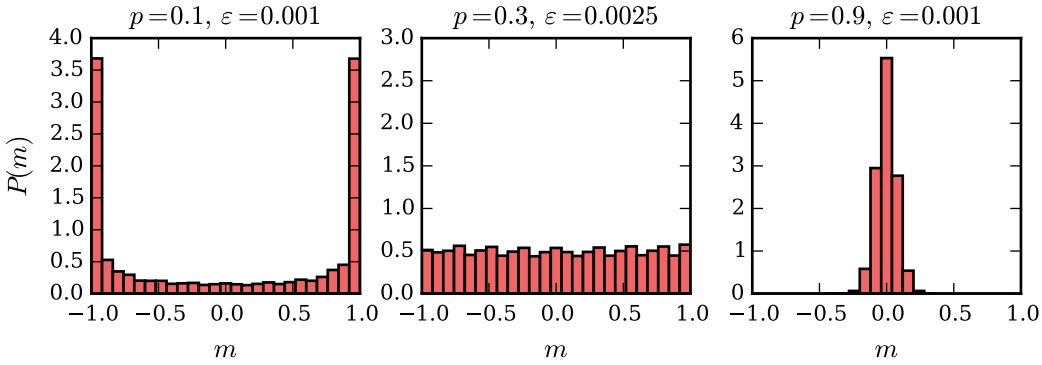
As already reported [231], for the linear case $q = 1$ consensus and fragmentation phases exist only for a finite size of the network and their size in the parameter space decreases with a growing number of nodes. As we can see in Figure 3.16, the same holds for the sub-linear scenario for the fragmentation phase. The consensus phase does not exist for $q < 1$ at all. The only point where absolute magnetization raises slightly above zero is at p_c and for $\varepsilon \approx 0$, due to stronger fluctuations. On the other hand, the fragmentation phase prevails for even twice larger noise rate in comparison to the linear scenario.

In the super-linear case $q = 2$ the fragmentation phase is much smaller and disappears faster. But the consensus phase endures for much bigger noise. We can observe the consensus state even for ε greater by almost two orders of magnitude than in the linear model. Most importantly, the scaling behavior is different for super-linear case. Indeed, non-linearity has a significant influence on the nature of the consensus phase. For $q < 1$ it does not exist, for $q = 1$ it exists only in finite networks, and for $q > 1$ the phase persists in the thermodynamic limit, what shall be also proven analytically in next paragraphs. We can further observe in Figure 3.16 that p_c decreases with q. All together it is a strong evidence that the character of interactions in coevolving networks can be crucial for the shape and properties of the stationary state.

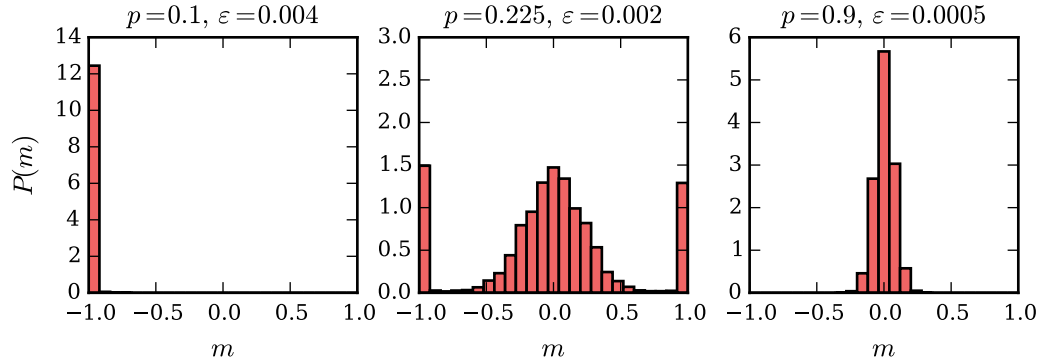
As said before, we can easily distinguish three phases by looking at the magnetization and the size of the largest component. However, the consensus phase is significantly different for $q = 1$ and $q > 1$. It can be described by a consensus in both cases, as the system displays $m = \pm 1$ most of the time. But the stability of this state changes with varying q. It is clearly visible from the probability distribution of magnetization in Figure 3.17 and individual trajectories in Figure 3.18. For the linear case the magnetization is bimodal with two equal peaks at values $+1$ and -1. Whilst the super-linear case displays a single peak of the magnetization at either of the boundary values, depending on the run. Once the consensus is reached the system stays in it, where for $q = 1$ the system goes back and forth between opposite consensus states. Additionally, the consensus phase for $q > 1$ is independent from the the system size, but it vanishes in the thermodynamic limit for the linear case



a) $q = 0.5$



b) $q = 1$



c) $q = 2$

Figure 3.17: Probability distribution of the magnetization m for $N = 250$ and $\mu = 4$ averaged over 10^7 MC steps after thermalization. a) For $q = 0.5$ magnetization always displays unimodal distribution with a varying standard deviation. b) For $q = 1$ we can distinguish the consensus phase, where the distribution is bimodal with maximum at ± 1. c) For $q = 2$ the consensus phase is also clearly visible, however it has a unimodal distribution with the peak at either of the boundary values (when sampling over time). The difference between the consensus phase for $q = 1$ and $q = 2$ is clearly visible.

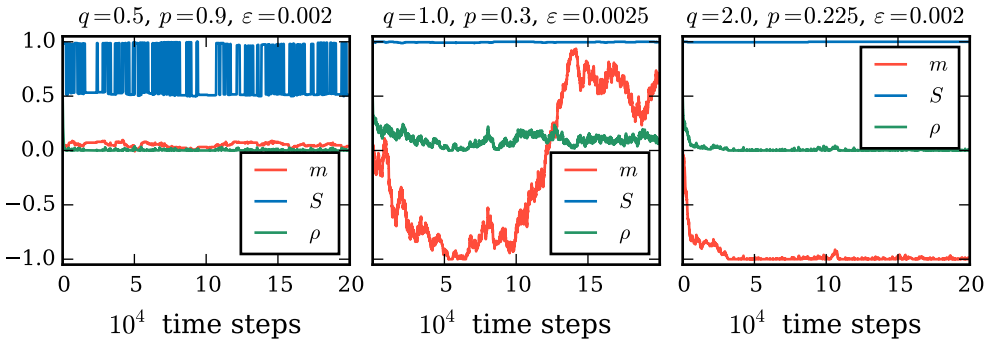


Figure 3.18: Exemplary trajectories of m, S, and ρ for $N = 250$ and $\mu = 4$. For $q = 0.5$ we can see the dynamical fragmentation, for $q = 1$ the plot shows the transition line with every value of m equally possible, for $q = 2$ we can see the stability of $m = -1$ state.

[231] (see Figure 3.16).

The consensus phase does not exist in the sub-linear case ($q < 1$), thus the magnetization always displays a single peak at 0. The variance, however, may vary depending on the parameters combination and it takes the maximal value for noise going to zero and $p = p_c$, i.e. close to the transition point between coexistence and fragmentation phases in the nonlinear CVM [213].

The final indication of the strong influence of non-linearity on the system's behavior is the magnetization distribution for $q > 1$ at the transition line. As we can see in Figure 3.17c), it becomes trimodal. A trimodal magnetization distribution was reported before in the noisy voter model on a static network [232], but only for the non-linearity parameter equal 5 or larger. Here it is obtained already for $q = 2$.

The fragmentation phase can be defined based on the size of the largest component. In other phases it is equal to the size of the whole network ($S = 1$), while the fragmentation is characterized by a dynamical separation into two equal components of opposite states. Due to the noise manifested in random changes of nodes' states and the rewiring the two components are constantly being reconnected and disconnected (see Figure 3.18). Following [231], the phase boundary can be defined as the line at which the network is half of the time fragmented and half of the time contains only one big component. The probability distribution of the largest component size is presented in Figure 3.19.

Similarly to the consensus phase, the coexistence phase can take different forms as well. In general, it is defined by zero magnetization (also zero absolute magnetization) and by existence of one big component. Nonetheless, this description leaves room for different possible configurations. Previously, in the linear model only fully-mixing phase was discovered, with nodes of states $+1$ and -1 well mixed inside a random graph. On the other hand, we can satisfy conditions for the coexistence phase having two evident communities highly connected internally and of opposite states, with only a few

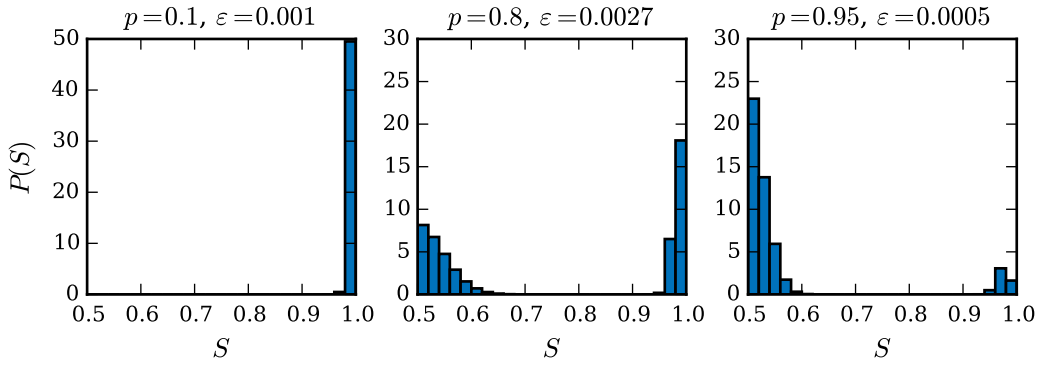


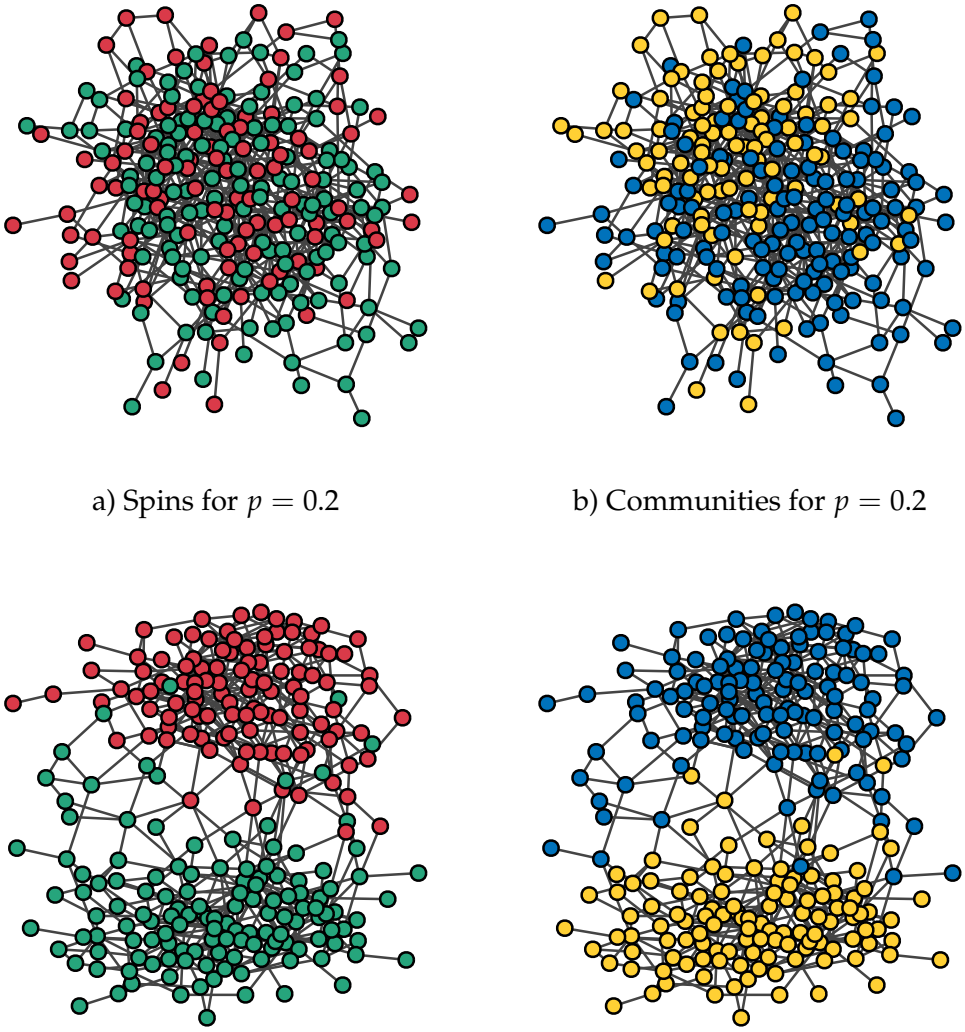
Figure 3.19: Probability distribution of the size of the largest component S for $N = 250$, $\mu = 4$, and $q = 0.5$ averaged over 10^7 MC steps after thermalization. Results show how the network goes from being connected (left panel) to being half of the time fragmented at the transition line (middle panel), to finally stay mostly disconnected (right panel).

inter-links between them. There is still zero magnetization and one cluster in such configuration. One can clearly see the difference between these two possibilities in Figure 3.20.

The difference between them can be also seen in the density of active links, although it is not obvious what value of the density to chose for the boundary. Another quantitative measure is an overlap between *real* communities – defined by the state of the nodes – and topological communities found by a community detection algorithm. Here a classical algorithm from [233] is used, but the result does not differ much when using different ones. Each node is assigned to the real community by its state and to a structural community by the algorithm's result. The relative overlap between these two communities can uncover the difference between the structured coexistence and the fully mixing configuration[6]. For a random assignment or no community structure the overlap will be close to 0.5. This corresponds to the fully-mixing coexistence phase. For the structured coexistence phase the overlap should be close to 1 (Figure 3.20).

The dynamics of the system can be described using an approach similar to the one for the nonlinear coevolving voter model [213] (Section 3.3.1). This approach takes into account magnetization m and density of active links ρ, therefore it can not detect the dynamical fragmentation phase, as it is defined by the largest component size S. Since the network is structurally homogeneous due to the random rewiring, we can assume each node to have the same average degree $\mu = 2M/N$. Let us denote by $n_+ = (1+m)/2$ and $n_- = (1-m)/2$ the fraction of nodes in the state $+1$ and -1 respectively.

[6]Note that one has to take the maximum overlap from two possible community assignments. If we have structural communities a and b, we can associate community a with the state $+1$ and community b with -1, or the other way around. Therefore, in a perfect overlap with a wrong assignment one can get zero overlap. Trying both possibilities and taking maximum solves this issue.



a) Spins for $p = 0.2$

b) Communities for $p = 0.2$

c) Spins for $p = 0.55$

d) Communities for $p = 0.55$

Figure 3.20: Examples of the network in a stationary state of the fully-mixing (a, b) and structured (c, d) coexistence phase. Spin $+1$ is indicated by green color, -1 by red (a, c). Communities found by the algorithm [233] are colored blue and yellow (b, d). In the structured coexistence phase the real communities defined by the state are well overlapping with the structural communities. Networks obtained for $q = 0.5$, $\varepsilon = 0.004$, $N = 250$ and $\mu = 4$.

When we pick a node i in the state $s_i = \pm 1$ as the active node the probability of choosing a neighbor in the opposite state is given by $\rho/2n_\pm$. In other words, the coefficient $\frac{\rho}{2n_\pm}$ gives the density of active links ρ_i for the node i being in the state $s_i = \pm 1$. Therefore, the probability of an interaction is given by $\rho_i^q = (\rho/2n_\pm)^q \equiv n_q^\pm$, which is at the same time the probability of choosing a neighbor in the opposite state q times. Hence, knowing that this probability was fulfilled when an interaction occurs, we can approximate that there is at least q neighbors in the opposite state. For the rest of them the probability of being in a different state than the focal node is $\rho/2n_\pm$, all together rising[7] $a_i \approx q + (\mu - q)\frac{\rho}{2n_\pm}$, which is an approximate number of active links of the node i under the condition that an interaction occurred.

To describe the evolution of the interface density ρ we must estimate contributions of different events in the network, which can result in a change of ρ. Analyzing the model's algorithm given at the beginning of the section we can distinguish four such events: (i) rewiring, followed by a change of the state through noise, (ii) rewiring, without a change of the state through noise, (iii) changing the state of the node through the state copying with no further action from noise, and (iv) changing the state of the node only as a result of noise, with no state copying or rewiring before. Let $\delta_\pm$ be the change in the total number of active links given that a node i, such that $s_i = \pm 1$, flipped its state. Then, the total change in the number of active links in the four possible scenarios is: (i) $1 + \delta_\pm$, (ii) -1, and for events (iii) and (iv) just $\delta_\pm$. When changing the state all a_i active links of the node i become inactive and all other $\mu - a_i$ inactive links become active, therefore the total change in the number of active links is $\delta_\pm = \mu - 2a_i$. Using previous approximation we can write $\delta_\pm \approx \mu - 2q - 2(\mu - q)\frac{\rho}{2n_\pm}$.

Magnetization changes only when the node's state is flipped via copying or noise in three possible scenarios: (i) state copying with no noise action, (ii) link rewiring followed by the noise action, and (iii) no interaction – neither state copying nor link rewiring – but the noise action alone. When the focal node having a state $s_i = \pm 1$ flips, the total change in the magnetization is equal $\Delta m = \mp 2/N$. Putting it all together, we can write down the formulas for expected changes in m and ρ:

$$\begin{aligned}\Delta m = \frac{2}{N}\Big\{ & (1-p)(1-\frac{\varepsilon}{2})(n_- n_q^- - n_+ n_q^+) \\ & + p\frac{\varepsilon}{2}(n_- n_q^- - n_+ n_q^+) \\ & + \frac{\varepsilon}{2}\left[n_-(1-n_q^-) - n_+(1-n_q^+)\right]\Big\},\end{aligned} \tag{3.18}$$

[7] This is a rough estimate, more precise one could be obtained using Bayes' theorem, however it doesn't display a significant difference in the results.

where $1/N$ accounts for the fact that m is a normalized quantity. Similarly, for the interface density we have:

$$\begin{aligned}\Delta\rho = \frac{1}{M}\Big\{ & p\frac{\varepsilon}{2}\left[n_+ n_q^+(1+\delta_+) + n_- n_q^-(1+\delta_-)\right] \\ & - p(1-\frac{\varepsilon}{2})(n_+ n_q^+ + n_- n_q^-) \\ & + (1-p)(1-\frac{\varepsilon}{2})(n_+ n_q^+\delta_+ + n_- n_q^-\delta_-) \\ & + \frac{\varepsilon}{2}\left[n_+(1-n_q^+)\delta_+ + n_-(1-n_q^-)\delta_-\right]\Big\},\end{aligned} \tag{3.19}$$

where $1/M$ is a change in ρ when one link becomes active, since it is also a normalized quantity. Using the fact that $\Delta t = 1/N$ and dividing both sides by it we obtain in the thermodynamic limit the following rate equations:

$$\begin{aligned}\frac{\partial m}{\partial t} = & \ 2(1-p)(1-\frac{\varepsilon}{2})(n_- n_q^- - n_+ n_q^+) \\ & + 2p\frac{\varepsilon}{2}(n_- n_q^- - n_+ n_q^+) \\ & + 2\frac{\varepsilon}{2}\left[n_-(1-n_q^-) - n_+(1-n_q^+)\right],\end{aligned} \tag{3.20}$$

and for the interface density:

$$\begin{aligned}\frac{\partial \rho}{\partial t} = \frac{2}{\mu}\Big\{ & p\frac{\varepsilon}{2}\left[n_+ n_q^+(1+\delta_+) + n_- n_q^-(1+\delta_-)\right] \\ & - p(1-\frac{\varepsilon}{2})(n_+ n_q^+ + n_- n_q^-) \\ & + (1-p)(1-\frac{\varepsilon}{2})(n_+ n_q^+\delta_+ + n_- n_q^-\delta_-) \\ & + \frac{\varepsilon}{2}\left[n_+(1-n_q^+)\delta_+ + n_-(1-n_q^-)\delta_-\right]\Big\},\end{aligned} \tag{3.21}$$

where we also used the relation $M = \mu N/2$. After few simple algebraic transformations it can be rewritten as:

$$\begin{aligned}\frac{\partial m}{\partial t} &= 2(1-p)(1-\varepsilon)(n_- n_q^- - n_+ n_q^+) + \varepsilon(n_- - n_+), \\ \frac{\partial \rho}{\partial t} &= \frac{2}{\mu}\Big[(1-p)(1-\varepsilon)(n_+ n_q^+\delta_+ + n_- n_q^-\delta_-) - p(n_+ n_q^+ + n_- n_q^-) \\ &\qquad + \frac{\varepsilon}{2}(n_+\delta_+ + n_-\delta_-)\Big].\end{aligned} \tag{3.22}$$

If we substitute magnetization $m = 0$, the first equation is automatically fulfilled and the second one, after putting $\frac{\partial \rho}{\partial t} = 0$, leads to a stationary equation

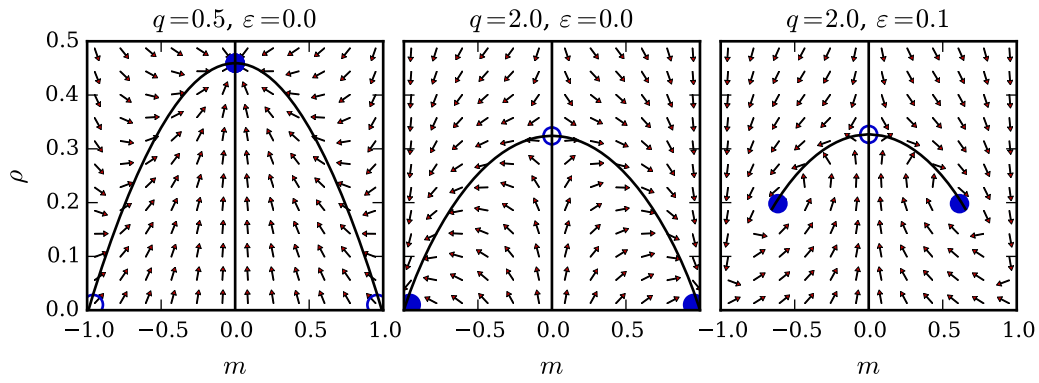


Figure 3.21: Flow diagram of the system dynamics in the (m, ρ) space for $\mu = 8$, $p = 0.1$, and different values of q and ε. Arrows represent the dynamical direction of the system according to the pair approximation solution (Equations 3.22). Fixed points are represented by full circles (stable) or empty circles (unstable). Note how non-linearity and noise can change the stability and position of fixed points.

for the interface density:

$$\begin{aligned} -\rho^{q+1} 2(\mu - q)(1-p)(1-\varepsilon) + \rho^q[(1-p)(1-\varepsilon)(\mu - 2q) - p] \\ -\rho\varepsilon(\mu - q) + \frac{\varepsilon}{2}(\mu - 2q) = 0. \end{aligned} \tag{3.23}$$

Solution of 3.22 for $m = \pm 1$ does not exist with a finite noise rate ε, whilst for $\varepsilon = 0$ it raises $\rho = 0$ (see Section 3.3.1). Setting the noise rate to zero together with the magnetization we obtain the stationary solution of the nonlinear CVM:

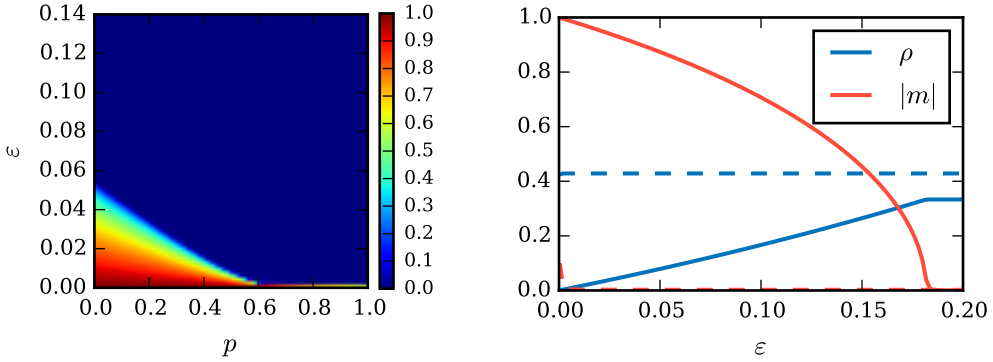
$$\rho^*_{\varepsilon=0} = \frac{(1-p)(\mu - 2q) - p}{2(1-p)(\mu - q)}, \tag{3.24}$$

which is consistent with Equation 3.16. For $\varepsilon = 0$ and $q = 1$ the solution of the standard CVM is recovered (as in Equation 3.12):

$$\rho^*_{\varepsilon=0,q=1} = \frac{(1-p)(\mu - 1) - 1}{2(1-p)(\mu - 1)}. \tag{3.25}$$

It is therefore justified to say that the analytical description developed in this thesis is the most general one among those considering the coevolving voter model. Previous models [189], [213], [231] are special cases of the nonlinear CVM with noise and so Equations 3.22 include a description of these cases.

The stability of possible solutions (m^*, ρ^*) of Equations 3.22 depends on the configuration of p, q, and ε. The best way to visualize the stability is by drawing a flow diagram, as in Figure 3.21. The line of $\rho = 0$ is not necessarily stable as in the nonlinear CVM, because it doesn't create a frozen configuration due to noise. Since the analytical description is derived for the thermodynamic limit, we don't observe stable fixed points at non-zero



a) Magnetization $|m|$ b) Static network $p = 0$

Figure 3.22: Numerical solutions of the pair approximation description of the system (Equation 3.22). a) Magnetization $|m|$ for $\mu = 5$ and $q = 2$ in the (p, ε) phase diagram, showing existence of the consensus phase in the thermodynamic limit. b) Magnetization $|m|$ and interface density ρ in the static case $p = 0$ and $\mu = 8$. Solid lines represent solution for $q = 2$, dashed lines for $q = 1$. A transition is visible for only for $q = 2$.

magnetization for $q \leq 1$. This finding is consistent with the scaling behavior of numerical results (Figure 3.16), indicating survival only of the coexistence phase in the large network limit. For the super-linear case, although the fixed points are placed at the same values of the magnetization ($m = -1, 0, 1$), their stability is inverted – now only the solutions of $|m| = 1$ are stable. It is well visible in the analytical prediction of the phase diagram in the Figure 3.22a). This implies that the consensus phase should be observed for any N when $q > 1$, which is in agreement with numerical results. Separate mean-field prediction of the consensus phase disappearance in the thermodynamic limit was given by Diakonova et. al. [231] for the special case of $q = 1$.

The non-linear noisy voter model (without the coevolution, on a static network) has been thoroughly analytically studied [232] and $q = 1$ was reportedly the bordering value between unimodal and bimodal distribution of the magnetization m. It other words, it is a border between existence and nonexistence of the consensus phase. As the non-linear noisy voter model is a special case of the model studied here, one should expect the same properties for $p = 0$. The agreement with this work is evident after analyzing the border of phase diagrams for zero plasticity (indicating a static network). This behavior is also separately illustrated in Figure 3.22b).

As mentioned before, the dynamical fragmentation phase becomes smaller when increasing the system size to finally disappear in the limit of $N \to \infty$. It also becomes smaller for bigger values of the q parameter. This behavior can be understood in a following way. Having two separate, but homogeneous inside, clusters of opposite state the only way of connecting them is by a random change of a node's state and one of its links rewiring to the second component. Probability of the first event is independent of q and is simply given by $\varepsilon/2$. When the contrasting node is selected as the active node the

probability of an interaction is ρ_i^q. Since $\rho_i \in [0,1]$, for smaller q the probability of an interaction is higher, except for boundary cases of $\rho_i = 0, 1$. To reconnect the clusters rewiring must occur, but this happens always with the probability p, despite the value of q. Therefore, for a single node of the opposite state than the whole cluster the probability of connecting to the other cluster is constant (since $\rho_i = 1$). However, once two clusters are connected, the higher probability of an interaction for lower q means a higher probability of rewiring causing fragmentation again. Consequently, we should expect the fragmentation phase to persist for larger noise when q is lower.

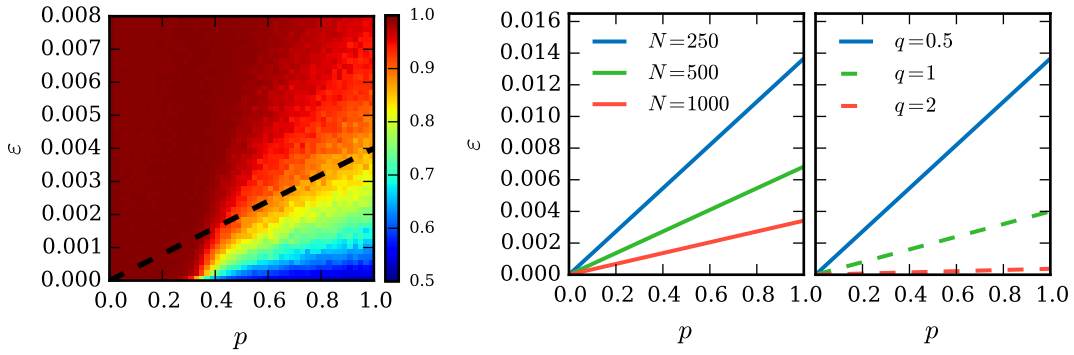
In order to describe the behavior of dynamical fragmentation region more precisely we shall first approximate probabilities of reconnecting two separate clusters and disconnecting clusters sharing at most two links. In this approach events of probability proportional to $(1/N)^3$ or ε^2 (and of higher orders) are omitted.

Imagine two separate and internally homogeneous components of opposite states, as it happens for the fragmentation phase phase. As said before, the simplest way of connecting them under the rules of the nonlinear CVM with noise involves two steps. Firstly, one of the nodes, call it i, must change it's state, what is possible only due to noise. Probability of this is equal $\varepsilon/2$. Secondly, the node i that previously changed its state must rewire one of its links to the opposite cluster, with which it currently shares the spin. This can happen with a probability equal $p\rho_i^q/N$, because we need to select this particular node as the active node $(1/N)$, an interaction must occur (ρ_i^q), and the rewiring must be performed (p). Since the node i is the only node of a different state in its component we have $\rho_i^q = 1$. Finally, it rises the probability of reconnecting two components equal:

$$P_r = \frac{\varepsilon}{2}\frac{p}{N}. \tag{3.26}$$

Note, that there are other possible scenarios of reconnecting two components, but they contain more steps, therefore are less probable.

Crossing the transition line between coexistence and fragmentation phases from the other side we will see one large cluster disconnecting into two equal components. As previously, imagine a situation two time steps before a possible fragmentation – the network consists of two components of opposite states, but homogeneous within each other. This time, however, they are connected by two links. One of the nodes i is a bridge, i.e. it is connected to two nodes in the opposite cluster. Now, for the fragmentation to occur we need both of the edges between the components to be rewired. The probability of rewiring the first one is $P_{d_1} = \frac{1}{N}(2/\mu)^q p(1-\frac{\varepsilon}{2}) + \frac{2}{N}(1/\mu)^q p$. We have to select the node i $(1/N)$ or one of its two neighbors in the second cluster $(2/N)$. An interaction must occur, what happens with probability $(a_j/\mu)^q$, where number of active links is 2 for the node i and 1 for its neighbors. At the end, the rewiring must be performed with a probability p, and additionally if the node i was selected, it can not change its state due to noise $(1-\frac{\varepsilon}{2})$. Otherwise fragmentation could not be achieved in two steps. The transition



a) Largest component S

b) Scaling of Equation 3.28

Figure 3.23: Analytical prediction from Equation 3.28 of the phase transition between coexistence and dynamical fragmentation phases. a) Comparison of the numerical results for $N = 250$, $\mu = 4$, and $q = 1$ with the analytical prediction indicated by a black line. b) Scaling with the network size for $q = 0.5$ (left panel) and with the q parameter for $N = 250$ (right panel).

occurs, however, for very small values of noise and therefore we can approximate $1 - \frac{\varepsilon}{2} \approx 1$. To rewire the second link we have to select one of two nodes $(2/N)$ at its ends, an interaction must occur $(1/\mu^q)$, which must be a rewiring (p). Therefore, the probability of losing the last connection between two components is $P_{d_2} = \frac{2}{N}(1/\mu)^q p$. Finally, we obtain the probability of disconnecting two clusters sharing only two links:

$$P_d = P_{d_1} P_{d_2} = \left[\frac{1}{N}\left(\frac{2}{\mu}\right)^q p + \frac{2}{N}\left(\frac{1}{\mu}\right)^q p\right] \frac{2}{N}\left(\frac{1}{\mu}\right)^q p. \tag{3.27}$$

Along the transition line a constant disconnecting and reconnecting of the network can be observed. By definition, it happens at such a rate that half of the time the system consists of two separate components, and half of the the time the network is connected. Therefore, at the transition line we expect $P_r = P_d$, what leads to the equation for the critical density of noise:

$$\varepsilon_S(p) = \frac{4}{N}\left(\frac{1}{\mu}\right)^{2q} (2^q + 2)p. \tag{3.28}$$

Based on this approximation we expect the fragmentation phase fading away with growing non-linearity parameter q and with growing system size N, as it is depicted in Figure 3.23. Both predictions are in a good agreement with the phase diagram, where we can see the same nature of scaling (Figure 3.16). A direct comparison of Equation 3.28 with numerical results is presented in Figure 3.23a). This simple approach manages to asses the transition line. It doesn't predict the value of p_c, but this is achieved by the pair approximation (Equation 3.22), when putting $\varepsilon = 0$ (see Section 3.3.1).

Depending on the values of the three parameters – the rewiring probability p, the noise intensity ε, and the non-linearity parameter q – we can observe three main phases previously reported separately: the consensus phase, the coexistence phase, and the fragmentation phase. There are, however, significant differences within two first ones. The consensus phase has a different behavior depending on the non-linearity parameter. For $q = 1$ it is a consensus with absolute magnetization equal 1 on average, but real magnetization switching between -1 and $+1$ states, leading to a bimodal distribution within one realization. For $q > 1$ we have a constant consensus, i.e. states $+1$ and -1 are stable. Therefore, during one realization the system stays at either $m = +1$ or $m = -1$, rising unimodal distribution with a peak at the boundary magnetization.

The coexistence phase can be also divided into two [230]. The first one is the already reported fully-mixing phase with a random network structure and random states of the nodes, giving zero magnetization. But for a higher plasticity and low noise intensity we can observe the structured coexistence phase, which has the same magnetization value, but different topology. In this phase one can easily distinguish two communities of opposite states connected by just a few inter-links. The structural difference can be confirmed by community detection algorithms.

These findings fill a gap in the studies of the CVM. They provide a binding bridge between the CVM with noise [231] and the nonlinear CVM [213]. Additionally, the nonlinear noisy voter model [232] and the ordinary CVM [189] can be seen as special cases of the model described in this section [230]. There is a full consistency with those limit cases. This work brings the analysis of the voter model to a greater complexity by taking into account many possible effects. It may provide a tool in evaluation of the relevance of different factors in description of opinion dynamics, but can be also a reference point in the study of coevolving network models.

Chapter 4

The Axelrod model

We have seen already that models created to describe physical phenomena, or more generally the way of reasoning developed in physics, can be applied in social sciences. The very idea of linking micro and macro scale by studying how local interaction rules shape the global characteristics of a system was established in statistical physics. However, one can ask similar questions in sociology, biology, ecology, economics. One of them could be – why do we observe so much diversity in societies even though people tend to become more similar when they interact? And that is exactly the question Robert Axelrod asked.

Some scientists try to understand the way in which people pick up behavior patterns from others, or how the social influence can change habits of individuals. On the other hand, researchers also seek to understand how cultural diffusion shapes whole societies. How geographical, political, demographic, or linguistic barriers are formed on a global scale. But it is particularly interesting to connect these two scales and find out how one can determine the other. The global pattern, after all, must be a result of the local rules, even if the connection is complex and hard to recognize.

Robert Axelrod came up with an idea to analyze it in a *model of social interactions* [234] (which is also called a *model of dissemination of culture*). We can see many examples of cultural influence and adaptation in the World. The Moorish architecture in southern Spain, the style of Japanese woodblock prints visible in paintings of van Gogh and Matisse, the rhythms of Latin America on European dance floors etc. But there are also some minorities resisting to merge with the surrounding culture, like Basque and Welsh languages [235]. Social scientists could give a long explanation of every single example incorporating a historical background, geographical constraints, local environment and many subtle variables one should take into account. The Axelrod model, however, provides the answer to general questions of what is more likely, what can we expect in a given situation, and what mechanisms are the backbone of social change. It might not provide a precise answer for a particular problem, but it helps understanding the overall phenomena.

One can ask how do we know that a model is trustworthy, if it depends on assumptions we have made. And the assumptions can not be so easily validated by experiment, as in say classical mechanics. Nevertheless, they are based on empirical evidence and scientific reasoning. But this is also a question how to distinguish a result of a particular model from a result that every good model should produce. It sounds exactly like something

that physics should know about. Perhaps for this reason Axelrod model was greatly developed by physicists.

4.1 Static network

4.1.1 Original definition

Originally, the Axelrod model was defined for a static square lattice [234]. Every node in the network can be interpreted either as an individual or a group of people being homogeneous enough to represent it as a single entity. A connection between two nodes corresponds to the possibility of interaction. Neighbors in the network can exchange their cultural traits becoming more similar. The higher similarity they already exhibit, the more likely it is to become even more alike. This mechanism is called homophily and was argued to be a crucial element of social interactions [236], [237]. Connected vertices can be interpreted as co-workers, friends, family or two neighboring villages. Each of them is described by F features embedded in a state vector $\sigma_i = (\sigma_{i,1}, \sigma_{i,2}, ..., \sigma_{i,F})$. Every features can take any integer value from 1 to q, i.e. $\sigma_{i,j} \in \{1, 2, ..., q\}$. Therefore, there are q^F distinct possible states of a node. The resemblance to the Potts model [238], [239], especially in the case of $F = 1$, can't be unnoticed. However, the Potts model is an equilibrium model and the Axelrod model is an algorithmic non-equilibrium model. Yet, they can share common behavior in some circumstances, e.g. at the transition point.

The features or traits of the nodes can be interpreted in different ways. They can describe preferences, interests, attitude, opinion, language and so on. Collectively, all of this is aggregated under a general term *culture*, and so the model is said do describe the *dissemination of culture*. In words of the author „*culture is taken to be what social influence influences*". However, more narrow interpretations are also possible and one of them treats the Axelrod model as a model of language change. The author addressed this interpretation in the original publication, as it caused serious problems. It shall be described separately later on with a solution to the problem given in Section 4.3.

The model is defined by its algorithm. First, we must randomly select initial values of traits for every of N nodes. Each value from 1 to q have the same probability. The linear size of the lattice is given by L, therefore $N = L^2$. Then, in every time step:

1. select randomly an active node i from the network,
2. select randomly a node j from the neighbors of the active node i,
3. compare the state σ_i with σ_j and determine the overlap m, i.e. the number of traits such that $\sigma_{i,k} = \sigma_{j,k}$,
   (a) if all traits are equal ($m = F$), or all traits are different ($m = 0$), nothing happens,

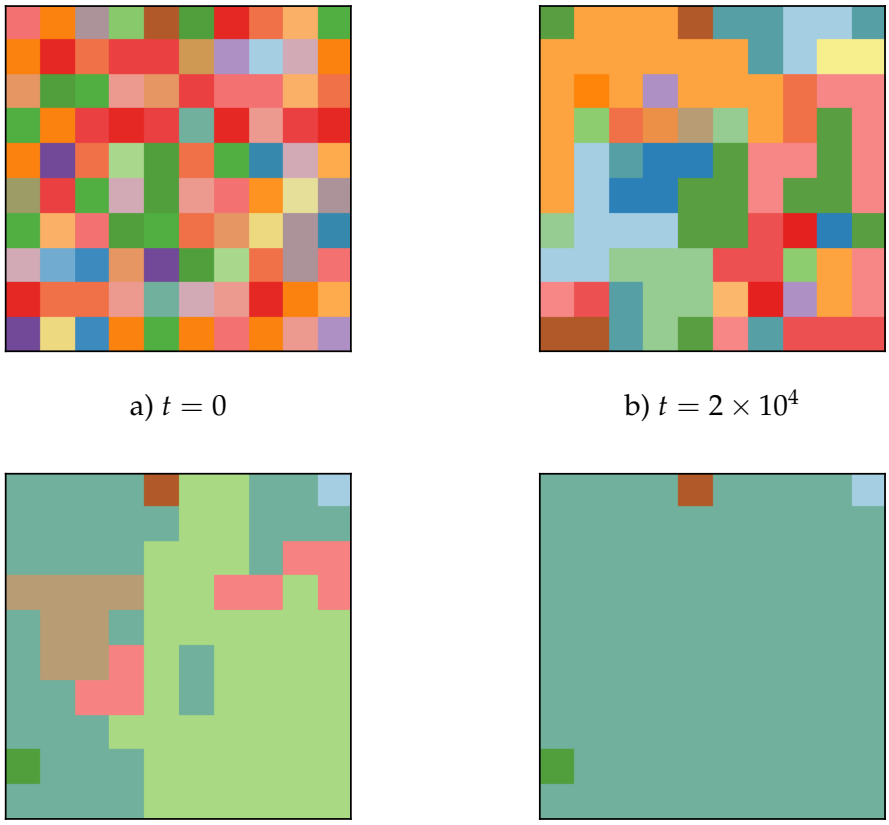
(b) otherwise, with probability equal m/F a positive interaction occurs, in which we randomly select one of not shared features f (from among $F - m$) and the active node i adopts its value from the node j, i. e. $\sigma_{i,f} \to \sigma'_{i,f} = \sigma_{j,f}$.

Simulation is ran until an absorbing state is obtained, what always happens in finite-size networks. In this state every pair of neighbors shares either none or all the traits, therefore nothing can change anymore. Sociological evidence suggests that the more in common two groups have the more likely is the transmission of next cultural ideas. It is reflected in the probability of interaction m/F increasing with a bigger overlap. Axelrod doesn't go into details of how an interaction actually happens, what it exactly consists of. Instead, he assumes that on average people tend to become more similar, and the tendency increases with the amount of the already present common ground.

The simplest way to describe heterogeneity of the system is by counting nodes in the same state creating uniform areas. One of the coefficients defined in Section 2.1.1 perfectly fits this task, namely the domain. Therefore, a natural order parameter for the Axelrod model is the relative size of the largest domain D and the number of domains n_D. A domain might be interpreted as a culturally homogeneous region. Note, that there could be more than one domain with the same state of nodes in the network, simply being separated by regions with a different state.

What we could expect from a model defined as above is a gradual erosion of cultural barriers and eventual convergence to a monoculture. This kind of behavior can be observed in Figure 4.1. At the beginning of the simulation almost every node has a different state, but there are many neighbors sharing at least one common feature. This is enough to start the process of assimilation, which ends with one big domain and three nodes possessing different states. This is also the behavior many expect to eventually observe in the real world – the convergence to one global culture. However, it doesn't seem to be happening and cultural diversity persists. This is exactly what the model was supposed to prove, that despite the local convergence we can observe global diversity. It is sufficient to increase q from 10 to 20 to observe what the author was aiming to illustrate. In Figure 4.2 we can see another result of a simulation leading to a quite different configuration. The dominant culture now takes less than a half of the population and there is ten times more distinct domains.

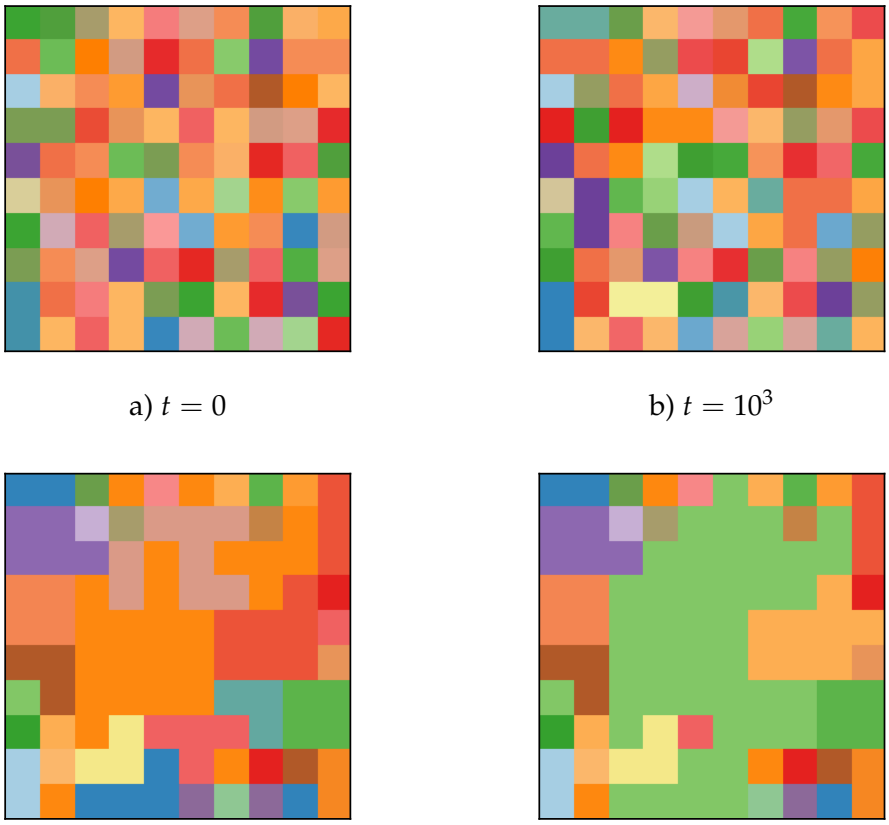
It's fair to say that one can predict the general increase of homogeneity merely from the rules that govern the system's evolution. But first of all, it is a surprising outcome (yet awaited based on the empirical observations) to preserve high level of diversity in the final configuration. And secondly, we can use the model to obtain the dependence of the final diversity on the initial configuration and parameters values. And so it was reported [234] that the final homogeneity increases with F, decreases with q, increases with the



a) $t = 0$ b) $t = 2 \times 10^4$

c) $t = 5 \times 10^4$ d) $t = 10^5$

Figure 4.1: Snapshots of the network during evolution of the Axelrod model from a) the initial configuration to d) the absorbing state. Simulation was ran for $N = 100$, $F = 5$, and $q = 10$. Time is given in a number of small time steps. Every state is indicated by a different color.



a) $t = 0$ b) $t = 10^3$

c) $t = 2 \times 10^4$ d) $t = 5 \times 10^4$

Figure 4.2: Snapshots of the network during evolution of the Axelrod model from a) the initial configuration to d) the absorbing state. Simulation was ran for $N = 100$, $F = 5$, and $q = 20$. Time is given in a number of small time steps. Every state is indicated by a different color.

number of neighbors[1], and increases with the system size.

The first two dependencies can be deduced analysing the formula for the probability of two neighbors sharing at least one common trait in the initial configuration. Let A_{ij} be an event where two vertices i and j share at least one feature, $A_{ij} = \{(\sigma_i, \sigma_j) \mid \exists_f \ \sigma_{i,f} = \sigma_{j,f}\}$. Then $P(A_{ij}) = 1 - P(A'_{ij})$, where A'_{ij} is the complementary event in which all features are different, $A'_{ij} = \{(\sigma_i, \sigma_j) \mid \forall_f \ \sigma_{i,f} \neq \sigma_{j,f}\}$. We have q^F possible states for each node, therefore there are q^{2F} possibilities for a pair of them. If we want two nodes not to share any trait, after choosing one of q^F states for the first one, the second one has only $q-1$ allowed values for each trait excluding the one chosen for the first node, hence there are $(q-1)^F$ possibilities. All together, it gives the complementary probability equal $P(A'_{ij}) = [(q-1)/q]^F$ and the probability of two nodes sharing at least one common feature:

$$P(A_{ij}) = 1 - \left(\frac{q-1}{q}\right)^F. \tag{4.1}$$

The higher this probability is the more homogeneous outcome we should expect, because having a trait in common means that neighbors can start interacting. And only interaction can lead to convergence. For instance, going with q to infinity the initial configuration would be already frozen, since there couldn't be any overlap.

Assuming that there are no correlations, we can get a rough number of partially matching pairs of neighbors by multiplying the probability 4.1 by the number of connections. Looking at the relation 4.1 it is clear that this number decreases with q and increases with F. And so does the final homogeneity of the system, as mentioned before. In the exemplary realisations the value of $P(A_{ij})$ was approximately 0.41 for Figure 4.1 and 0.23 for Figure 4.2.

The most unexpected and problematic behavior of the model is the decrease of heterogeneity with the system size. It means that the larger the network, the fewer domains we will observe in the final configuration, as we can see in Figure 4.3 and 4.4 with exception of the initial finite size effect[2]. Additionally, the relative size of the largest domain will increase with growing N. Of course, for q large enough there would be a linear growth of n_D with the network size, simply because every node would create a separate domain. However, this is a trivial dependence and we are interested in a complex behavior where the interaction actually shapes the system. This result is surprising, because one would expect higher cultural diversity for bigger regions, when there is more niches to develop its own local culture. After all, we do observe more distinct cultures on whole continents than on

[1] A square lattice was modified to increase the number of neighbors from 4 to 8 and 12, maintaining its regular character.

[2] Existence of the maximum number of domains for small networks can be easily explained. For a 2×2 lattice there are 4 nodes and therefore maximally 4 domains, for 3×3 lattice there can be 9 domains at most etc. Additionally, the maximal possible number of domains is only obtained, if no interaction occurred. So there must be an initial growth of the number of domains, simply because of the growing capacity to contain them.

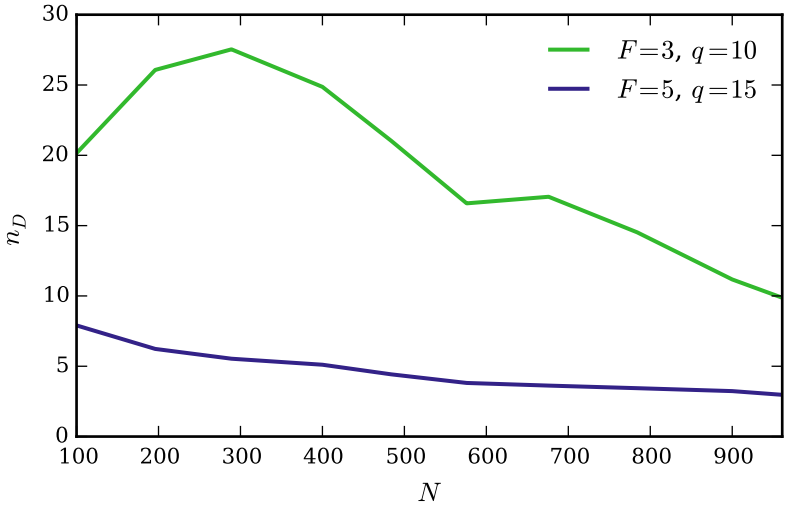


Figure 4.3: Absolute number of domains n_D for different system sizes $N = L^2$ in the Axelrod model. Results obtained for two parameter configurations and averaged over 250 realizations with periodic boundary conditions. For $F = 3$ we can see a finite size effect causing the initial growth of n_D for $N < 300$, which quickly vanishes for bigger networks.

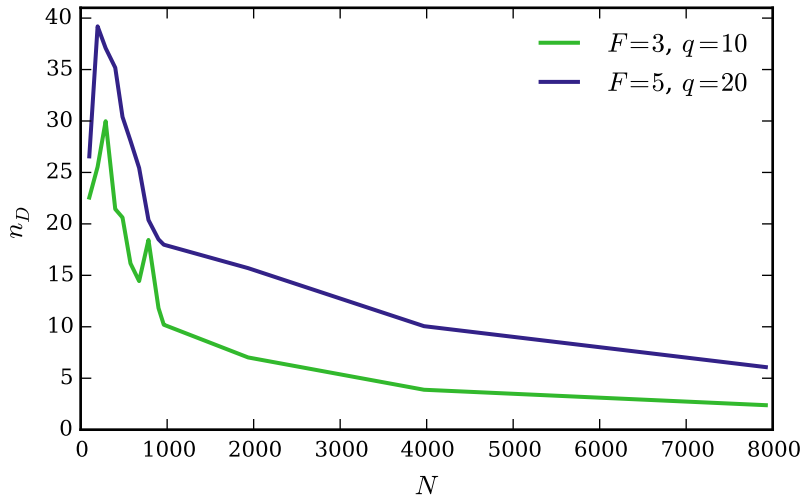


Figure 4.4: Absolute number of domains n_D for different system sizes $N = L^2$ in the Axelrod model. Results obtained for two parameter configurations and averaged over 50 realizations with periodic boundary conditions. We can see a finite size effect causing the initial growth of n_D for $N < 300$, which quickly vanishes for bigger networks.

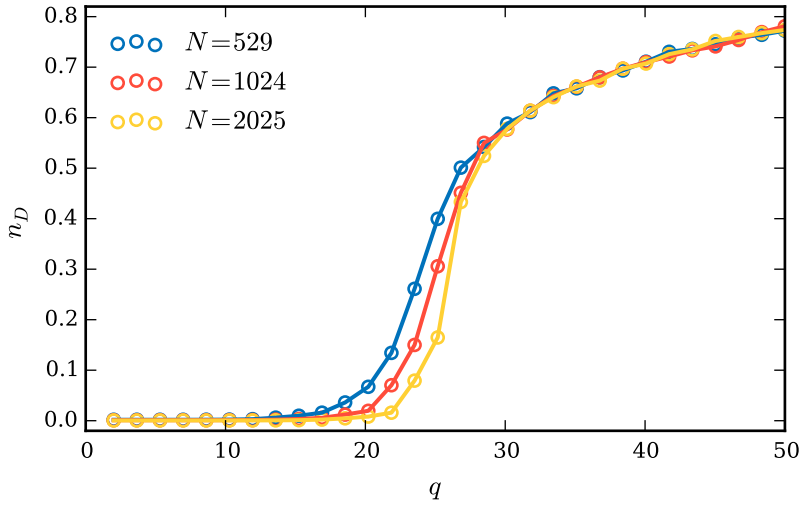


Figure 4.5: Number of domains n_D (normalized) vs parameter q for $F = 5$ and different network sizes. Results are averaged over 50 simulation runs with periodic boundary conditions. Two phases can be distinguished – for low values of q we have only one domain (ordered phase), for high values of q we observe many separate domains (disordered phase).

remote islands. The result is also problematic, because it is contradictory to empirical data, what was actually noticed already in the very first publication [234]. This issue shall be addressed in Section 4.3 together with a proposition of a simple solution. A short explanation of this phenomenon is that increasing the size of the network we are increasing the number of possible paths of interaction between two regions. Imagine a frozen configuration with small domains present at the border of the lattice, like in Figure 4.1d). If we increased the lattice, it would introduce a new direction of possible interaction with boundary domains. However, we can't use the same argument for periodic boundary condition, yet it displays qualitatively the same scaling behavior.

Later, a group of physicists thoroughly studied the Axelrod model and discovered a discontinuous order-disorder phase transition [240]. According to the previous argumentation, when increasing q we should observe decrease of homogeneity. However, it is not obvious that it would be an abrupt drop as it is visible in Figure 4.5 and Figure 4.6 rather than a slow decline. The phase transition in the number of domains n_D and the size of the largest domain D is clear. Both of them become steeper with bigger network, but remain in the same position, indicating a discontinuous transition in the thermodynamic limit. The transition point q_c is characterized by the highest fluctuation size and a power-law distribution of domain sizes, indicating long distance correlations. Reportedly, the behavior of the model is qualitatively the same for any $F > 2$ [240], whereas for $F = 2$ the transition becomes continuous. Additionally, for all $F > 2$ the power of the domain

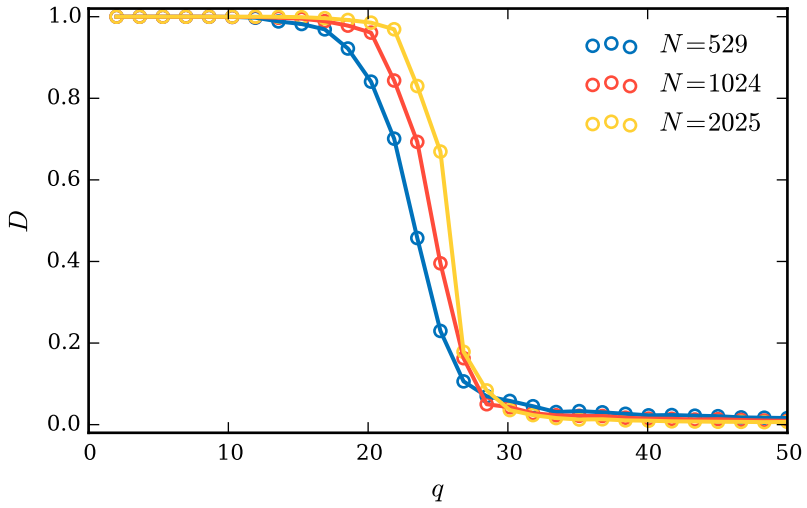


Figure 4.6: Size of the largest domain D (normalized) vs parameter q for $F = 5$ and different network sizes. Results are averaged over 50 simulation runs with periodic boundary conditions. Two phases can be distinguished – for low values of q the largest domain spans over the whole network (ordered phase), for high values of q it is much smaller than the system size (disordered phase).

size distribution is the same. These findings justify limiting further analysis to $F = 3$, as it is computationally faster. And so was done in many following extensions of the Axelrod model.

4.1.2 Selected extensions

The Axelrod model displayed interesting effects and most importantly explained how homophily can paradoxically lead to high diversity. The analysis, however, was restricted to square lattice. It is reasonable to ask how much does the model behavior depend on the chosen topology, especially since it was proved to be a major factor influencing results in most, if not all, network models.

Following the original publication, many physicists has developed the model and studied the dependencies further. It was discovered that the character of the phase transition in the largest domain is determined by the network's structure. After narrowing the network to one dimension it becomes a second-order phase transition with the order parameter changing continuously [241]. Originally, the transition was of the first order on a two-dimensional lattice. The standard model was later precisely described analytically, but only for a simple case of $F = q = 2$ [242].

As underlined many times in this thesis, real-world social networks are not square and regular, nor are they static. The second issue is addressed in

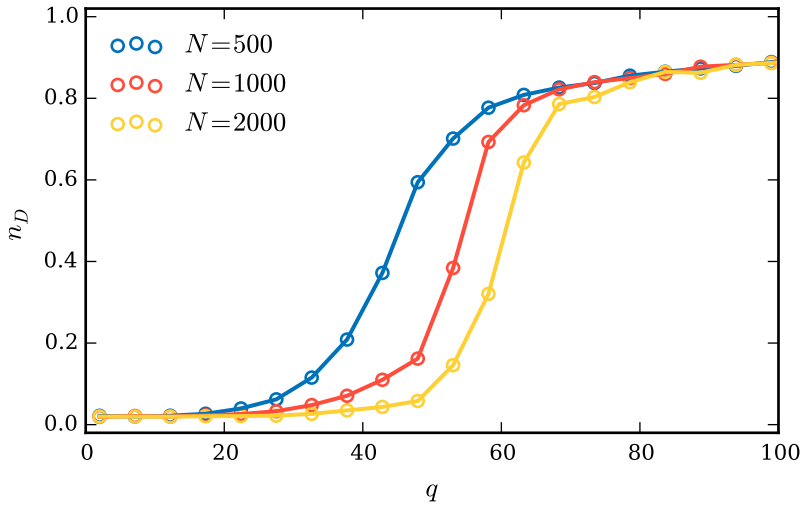


Figure 4.7: Number of domains n_D (normalized) vs parameter q in the ER graphs for $F = 5$. Results are averaged over 50 simulation runs. Similarly to the square lattice, two phases can be distinguished – for low values of q we have only one domain (ordered phase), for high values of q we observe many separate domains (disordered phase).

the next section, the first one has been resolved before. The model's behavior was investigated on ER random graphs, small-world networks, scale-free and clustered scale-free networks [243], [244]. Not surprisingly, the topology had a significant influence on the properties of the phase transition.

In random networks homogeneous ordered state persists for longer than in square lattices when increasing q. Therefore, in the small-world networks the transition point is being shifted to higher q when the ratio of random rewirings grows. In Figure 4.7 we can see the transition in n_D and in Figure 4.8 we can see the transition in D for a random ER graph compared with a square lattice[3]. As these are boundary cases of the WS model the general dependency can be deduced from it, but it was also explicitly simulated [243]. The higher the density of random long-distant connections, i.e. the less regular the network is, the further the transition is shifted to bigger values of q.

In scale-free networks generated by BA model scaling behavior suggests that the transition disappears for large networks and only the ordered configuration can be obtained in the thermodynamic limit. However, in structured scale-free networks, i.e. scale-free networks possessing high clustering, the order-disorder phase transition is restored. This is a particularly interesting result, because clustering is an important feature of social networks. Therefore, it gives another reason to distinguish its special contribution in shaping social interactions.

[3] After [243], the size of the largest domain D (usually normalized) became the typical order parameter for the Axelrod model, instead of the number of domains n_D.

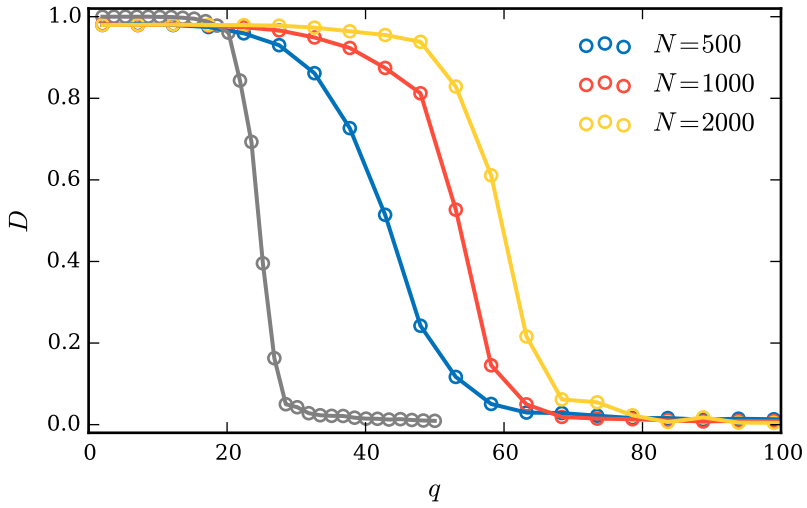


Figure 4.8: Size of the largest domain D (normalized) vs parameter q in the ER graphs for $F = 5$. Results are averaged over 50 simulation runs. The gray line shows the same results for a square lattice of $N = 1024$ vertices. The order-disorder phase transition in the ER network is shifted towards higher values of q, in comparison to the square lattice.

Axelrod suggested that his model could be extended to include *cultural drift*, perhaps in a form of random state changes. Social influence is not the only factor shaping culture. People can change their opinion, beliefs or habits also without an interaction with others. This fact can be simple accounted for by introducing noise into the model. It can be done in different ways, one of them is to perform a single feature perturbation with a given probability after every time step. This probability is called a noise rate. Of course, for any finite noise rate there can't be a frozen configuration. As a result, the heterogeneous multi-domain phase is no longer stable for fluctuations large enough. After a certain amount of feature perturbations two completely different domains can start interacting again to eventually merge. On the other hand, if the noise is too large it can on its own produce and sustain high diversity – simply by changing nodes states faster than they converge due to interactions. Therefore, one can expect a new order-disorder phase transition in the noise rate, and that is exactly what was discovered [245]–[247].

There have been other efforts to include noise in the model. One combined the noise rate with a confidence threshold [248]. Confidence threshold means that one common trait might not be enough to start interacting, and the required overlap can be chosen arbitrarily. Results, however, were qualitatively the same. Even an equilibrium version of the Axelrod model has been proposed [249], [250], where order-disorder phase transition occurs only at zero temperature.

As discussed in Section 3.3, a social interaction doesn't have to be always dyadic. It is not always one parson influencing the other, social pressure

can come from a group of peers or the whole environment, possibly in non-linear way. Non-dyadic interactions were studied producing more robust cultural diversity [251]. Among extensions addressing real-world phenomena, worth mentioning is also analysis of the influence of mass media with different strategies [252], [253], introducing tolerance thresholds [254], [255], adding acceptance or discussion in the interaction [256], applying conservatism or nonconformity [257], [258], taking into account also physical distance between agents [259], [260], changing the topology to a multilayer network [261].

Axelrod might have not accounted for every possible effect, but he did introduce two new ingredients into modeling of social interactions – discrete opinion and homophily. The second one being reflected in higher probability of interaction with individuals more alike. With these two mechanisms he showed how local convergence can produce global diversity, what had been absolutely unforeseen. It was later discovered that actually homophily on its own can produce heterogeneity in the macro scale [262], [263]. However, the original results still stand accurate and remarkable.

4.2 Coevolving Axelrod model

4.2.1 Random rewiring

The next natural extension of the Axelrod model is allowing the structure of the network to change over time. As stressed many times before, most of the real-world networks are not static and in particular social networks change their structure in the similar time scale as they change their state (the state of nodes). First attempt to incorporate this observations into the model gave birth to the coevolving Axelrod model [264], [265]. There were also further extensions introducing for instance physical location of agents [259], but first we shall discuss the basic definition.

In the coevolving Axelrod model we start with the same configuration as for the static version on random networks. Initially, it is an ER graph with randomly selected states of nodes σ_i. Then, in every time step:

1. select randomly an active node i from the network,
2. select randomly a node j from the neighbors of the active node i,
3. compare the state σ_i with σ_j and determine the overlap m, i.e. the number of traits such that $\sigma_{i,k} = \sigma_{j,k}$,
   (a) if all traits are equal ($m = F$), nothing happens,
   (b) if all traits are different ($m = 0$), the node i disconnects from the node j and creates a new connection to a randomly chosen node l,
   (c) in other cases, with probability equal m/F a positive interaction occurs, in which we randomly select one of not shared features f (from among $F - m$) and the active node i adopts its value from the node j, i. e. $\sigma_{i,f} \to \sigma'_{i,f} = \sigma_{j,f}$.

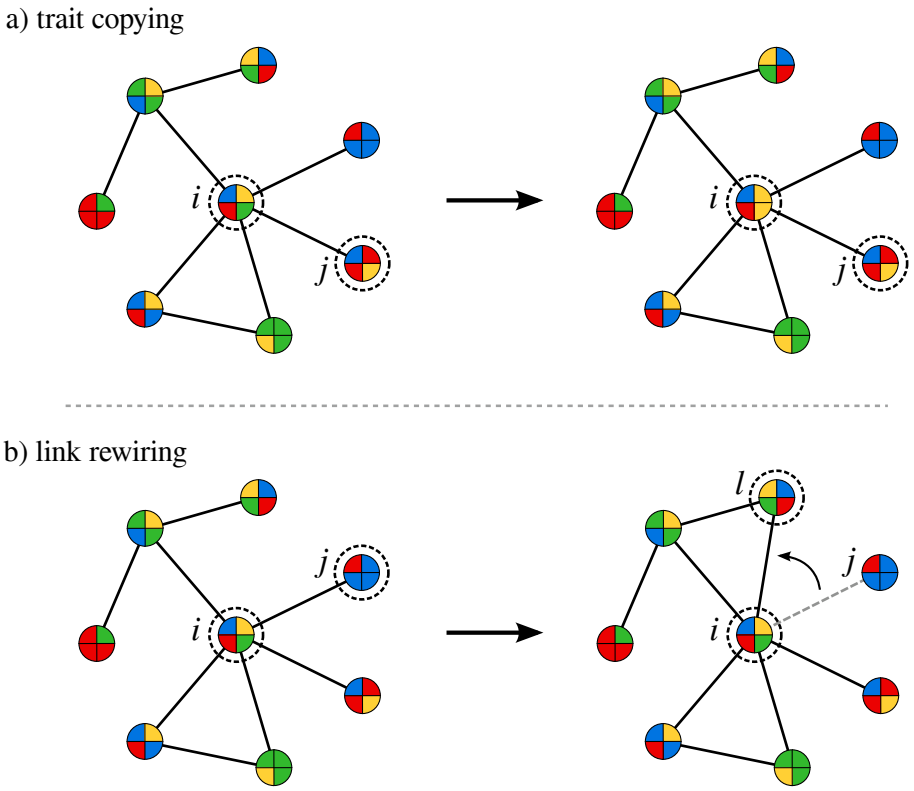


Figure 4.9: Schematic illustration of the coevolving Axelrod model algorithm. After choosing the active node i and one of its neighbors j their features are compared. Here, features are represented by four quarters of each node and possible values by four colors (red, green, blue, yellow). Therefore, $F = 4$ and $q = 4$ in this example. a) Nodes i and j share two traits $m = 2$ (both on the left), so the probability of interaction is 0.5. Interaction occurs and the active node copies the bottom right trait from the neighbor increasing similarity to $m = 3$. b) Nodes i and j do not share any trait $m = 0$ (there are the same colors, but on different positions) and so rewiring is performed. Node i cuts the link to the node j and connects to a randomly chosen node l.

A visual explanation of the algorithm is presented in Figure 4.9. Note, that multi- and auto-connections are not allowed in the model. Initial topology is not as important as in static models due to the rewiring procedure, which ultimately defines the network's structure.

In this version of the model network can become disconnected, small separated components or individual lonely nodes can appear. That's why one should take into account additional order parameter sensitive to these changes. Usually, it is either the number of components n_S or size of the largest component S (both normalized by the number of nodes). By definition, the number of domains can't be smaller than the number of components $n_D \geq n_S$ and the size of the largest domain can't be bigger than the size of

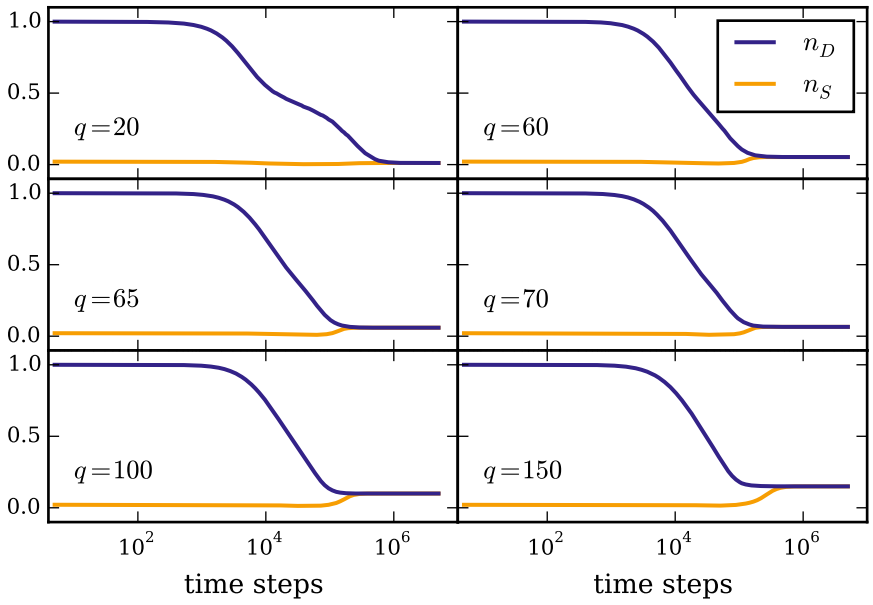


Figure 4.10: Average trajectories of the number of domains n_D and components n_S during simulations for different values of q, $F = 3$, $N = 500$ and $\langle k \rangle = 4$. Note that depending on the q value either n_D or n_S can settle at the final level first. Trajectories are averaged over 300 simulation runs.

the largest component $D \leq S$. Figure 4.10 presents average trajectories of n_D and n_S from the beginning until the end of the simulation, which is either a frozen configuration or a dynamical stationary state. It follows from the algorithm of the model that a frozen state can be obtained only if every link is connecting two nodes of identical set of traits. In other words, every component must contain exactly one domain. Otherwise, either trait copying or rewiring would be still possible, which means configuration is not frozen.

Increasing $q \to \infty$ the probability of sharing any trait between two neighbors drops to 0 and therefore we should observe infinite rewiring leading to a stationary state described in Section 2.2.5. In this limit state of the nodes can be neglected and the equilibrium configuration will be given by the formula 2.41. However, the most interesting part is the interplay between both aspects of the network.

In Figure 4.11 we can see a phase diagram of the coevolving Axelrod model in respect to the q parameter. Three phases are easily distinguishable. The first one is a consensus phase with one big component containing one domain – approximately all nodes have the same state. For intermediate values of q we observe disordered phase with network disintegration. The system divides into many small components, each being homogeneous inside, i.e having one domain. These two phases are characterized by a frozen final configuration, that's why we can see the same values for D and S in the plot. Increasing q further makes the network recombine into one connected component again, but with many different domains. In this dynamical phase the system reaches a stationary state with no changes in the state of nodes

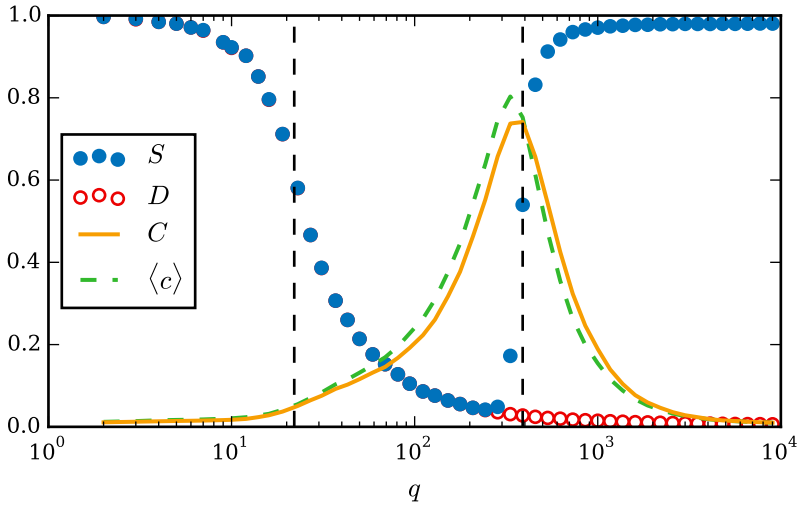


Figure 4.11: Phase diagram of the coevolving Axelrod model in respect to q. Transitions between phases are indicated by black dashed lines based on the maximal fluctuations. Note that in two first phases size of the largest component S is equal to the size of the largest domain D. Results are obtained for $N = 500$, $\langle k \rangle = 4$, and $F = 3$. Values are averaged over 400 realizations.

and constant rewiring. Both transition points are characterized by maximal fluctuations. Additionally at the first one, from the consensus to the disintegration phase, the distribution of component size becomes power-law [264]. At the second phase transition the clustering coefficient takes maximal value [266]. That new property of the transition between the disintegration and the recombination phase was discovered in this thesis. It is a universal indicator of the transition across all versions of the coevolving Axelrod model.

Analysing the phase diagram one can come to a conclusion that for small values of q (at least for the consensus phase) the dynamics of the system is dominated by the trait copying and from some point onwards it is mostly rewiring. A good way to investigate it is by looking at the thermalization times for the system as a whole τ, for domains τ_D and for components τ_S, as presented in Figure 4.12. From this plot is is clearly visible that there is a region where components settle first and a region with the opposite situation. Explanation of how it is possible to observe different thermalization times for n_S and n_D is fairly simple. If q is small (and/or F big) practically all connected pairs of nodes can share at least one trait. Therefore, there will be almost no rewiring (low τ_S), but a lot of trait copying to reach the consensus (high τ_D). For high values of q the opposite can happen – there is not much possibility for interaction, therefore all trait copying will finish quickly (low τ_D), but the links will keep rearranging to fit the domains (high τ_S). Interestingly, the point at which thermalization times for both become equal is between the transition points, at the beginning of the disintegration phase.

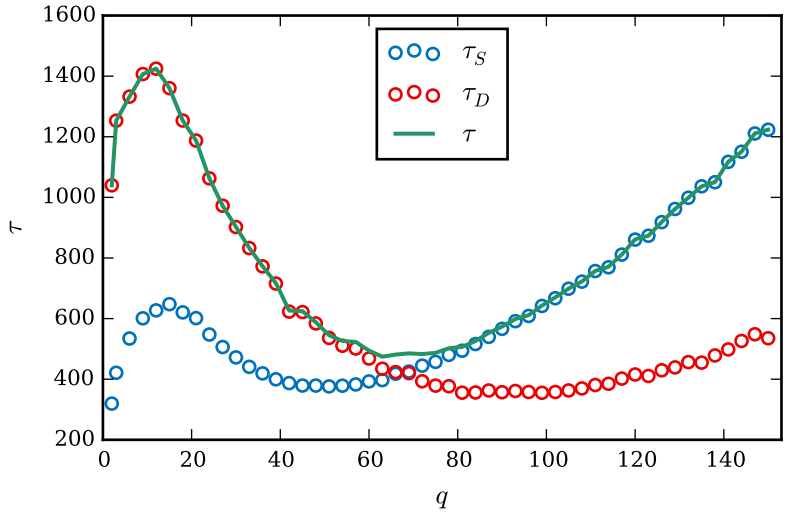


Figure 4.12: Thermalization time $\tau = \max\{\tau_S, \tau_D\}$ of the coevolving Axelrod model vs q for $N = 500$, $\langle k \rangle = 4$, $F = 3$. Values are averaged over 500 realization.

Coevolution in the voter model substantially changes behavior of the system. Topology of the network is no longer limited to the initially chosen configuration – now it is a result of the coupled dynamics of the states and links. In the simplest formulation of random and uniform rewiring it creates small communities in separate components (disordered phase) and causes reintegration of the network for higher q. Nonetheless, the resulting degree distribution is Poissonian, i.e. network has a structure of a random graph. Moreover, the system displays mostly small values of clustering coefficient. This two results are incoherent with empirical data about social networks. Solution of this problem proposed in this thesis is described in the following section.

4.2.2 Preferential attachment and triadic closure

Coevolving Axelrod model is certainly more realistic than its original definition on a static square lattice. It displays new effects and provides more general framework for social networks modeling. Topology not only plays a role by creating the playground for the interactions between individuals, but now it can also change itself being influenced by the state of nodes. The structures that are being created during the simulation is another aspect to analyse. What should be the main feature to look for? It has been showed many times that power-law degree distribution is a common characteristic of empirical networks [267], [268]. Additionally, in the context of social networks high clustering is important. Real-world social networks display values of the clustering coefficient (global and local) even at the level of 0.5 [26], [27], [63], [87], [88], [225], [226].

Both mentioned properties are purely topological, therefore they can be directly affected by the rewiring mechanism. Indirectly they are of course coupled also with the state dynamics. In the first coevolving definition of the model rewiring was random with uniform probability distribution for every node to be chosen as a new neighbor of the active vertex (excluding the active node itself and its existing neighbors). This led to a Poissonian degree distribution and mostly low values of the clustering coefficient. In order to improve this flaw four alternative possibilities are proposed in the thesis. This section contains original results obtained in mentioned alternative versions of the coevolving Axelrod model.

The most well known mechanism leading to scale-free degree distribution is preferential attachment [23], [109], [267], [268]. However, it does generate power laws in growing networks when a node can not lose connections once they are established. Situation in the coevolving networks is significantly different, because here the higher degree of a node means also a greater probability of loosing a connection due to choosing one of the neighbors as the active node and rewiring. On the other hand, in the case of pure rewiring, without taking into account states of nodes, linear preferential attachment also leads to a power-law degree distribution, as showed in Section 2.2.5. Importantly, the constant in this linear relation may decide between a power-law and exponential outcome. Nonetheless, neither growing network nor constant rewiring provides a precise description of a coevolving network, since the feed-back loop between rewiring and the states is the driving factor, which renders the analytical description non-trivial. Therefore, different types of preferential attachment are analysed in this work.

Second important structural aspect – high clustering – is addressed by introducing triadic closure, which by definition increases the clustering coefficient with every rewired link (or at least keeps it constant) and has been recognized as an important psychological and sociological mechanism [28], [29]. Having these two goals in mind four different models are proposed in this thesis. All the models are described by the same algorithm as the standard coevolving Axelrod model (Section 4.2.1), except having different rewiring procedures. When the active node cuts the connection to the neighbor with zero overlap in traits a new neighbor is chosen not with a uniform probability across the whole network, but according to a specific probability distribution $\Pi(k_i)$ of choosing a node i [266]:

- model A with preferential attachment: $\Pi(k_i) \sim k_i$,
- model B with preferential attachment: $\Pi(k_i) \sim k_i + 1$,
- model C with strong preferential attachment: $\Pi(k_i) \sim (k_i + 1)^2$,
- model D with triadic closure – uniform distribution among neighbors of neighbors.

As previously, multi- and auto-connections are prohibited. The model C has a very strong preference of high degree to compensate for the higher chances of losing connections by hubs. Note, that in models A and D once a

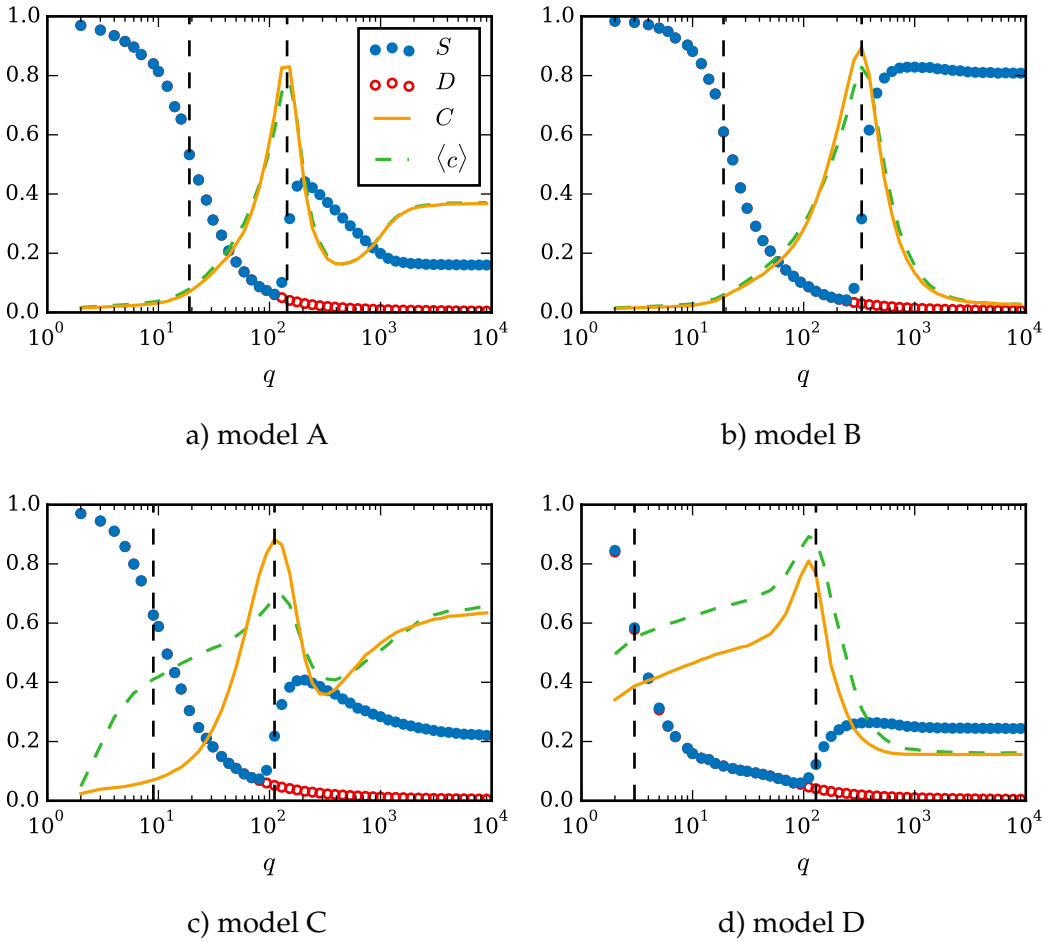


Figure 4.13: Phase diagram of the four versions of the coevolving Axelrod model in respect to q. Transitions between phases are indicated by black dashed lines based on the maximal fluctuations. Note that in two first phases size of the largest component S is equal to the size of the largest domain D. Results are obtained for $N = 500$, $\langle k \rangle = 4$, and $F = 3$. Values are averaged over 400 realizations.

Table 4.1: Degree distribution in different phases for four variants of the coevolving Axelrod model. The original model with random rewiring always results in a Poissonian degree distribution.

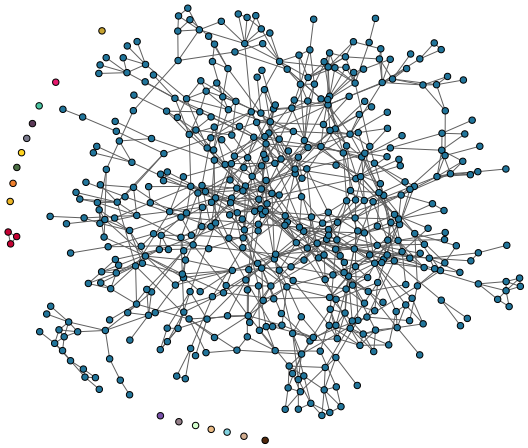
	phase I	phase II	phase III
model A	Poissonian/exponential	exponential	Gaussian
model B	Poissonian/exponential	exponential	exponential
model C	power-law	power-law	unclassified
model D	Poissonian	unclassified	Gaussian

node loses all connections it can not be reconnected with the rest of the network. Results obtained with alternative rewiring procedures are presented in Figure 4.13. General behavior is similar in all cases. We can distinguish three phases: the consensus phase with one big component and one domain (Figure 4.14a) for the model D), the disintegration phase with many small components containing one domain each (Figure 4.14b)), and the dynamical recombination phase where the network recombines into one relatively big component with the rest of the nodes scattered out (Figure 4.14c)). In case of many lonely nodes, what happens in all but the model B, the last phase is often called a shattered phase. It contains many small domains even in the main component, therefore rewiring never stops.

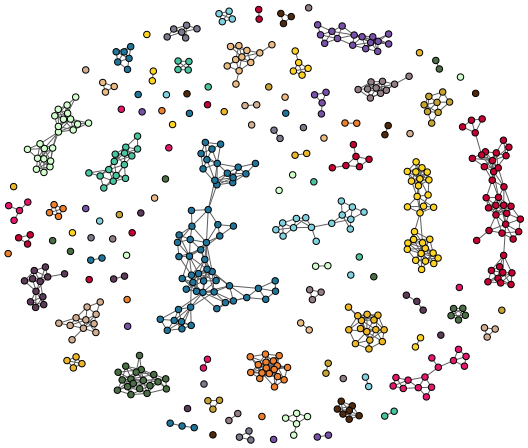
Other common feature across the models is the characteristics of transition points. Both are of course described by the maximal fluctuations. Additionally, at the first transition point q_D, detectable in the size of the largest domain and component, the distribution of component size is power-law, as showed in Figure 4.15. The second transition point q_S, observable only in the size of the largest component, is characterized by a peak of the clustering coefficient. This behavior could be expected in the model D, because at the transition point the thermalization time is usually the longest and with each time step there is a chance for rewiring and increasing the overall clustering. However, the rest of the models in general do not increase clustering during the rewiring procedure. The explanation of this phenomena can be found in the conversion from a frozen final configuration to a dynamical one at q_S. At this point there is just enough pairs of compatible nodes to fit all links between them. Therefore all small components consisting of compatible nodes will absorb as many links as possible, i.e. they will form complete graphs, which have clustering coefficient equal to 1.

What is different between the four models and the original coevolving Axelrod model is the position of transition points q_D and q_S, the level of recombination in the dynamical phase, i.e. the size of the largest component, and the value of the clustering coefficient out of the q_S point. As expected, the highest clustering is observed for the model D having $C \approx \langle c \rangle \approx 0.5$ in two first phases. Moreover, we can observe various degree distributions depending on the model and the phase (Figure 4.16). They are summarized in Table 4.1.

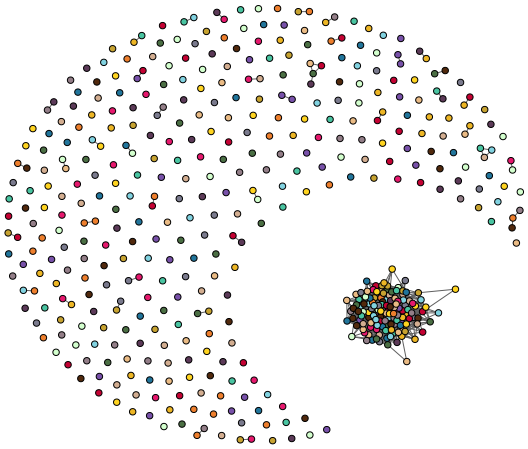
In the limit of $q \to \infty$ the coevolving Axelrod model can be described



a) $q = 2$



b) $q = 20$



c) $q = 400$

Figure 4.14: Visualisation of the network for each of three phases for the model D (with triadic closure), $N = 500$, $\langle k \rangle = 4$, and $F = 3$. Every color indicates a different set of traits of a node, therefore domains can be seen as areas of one color.

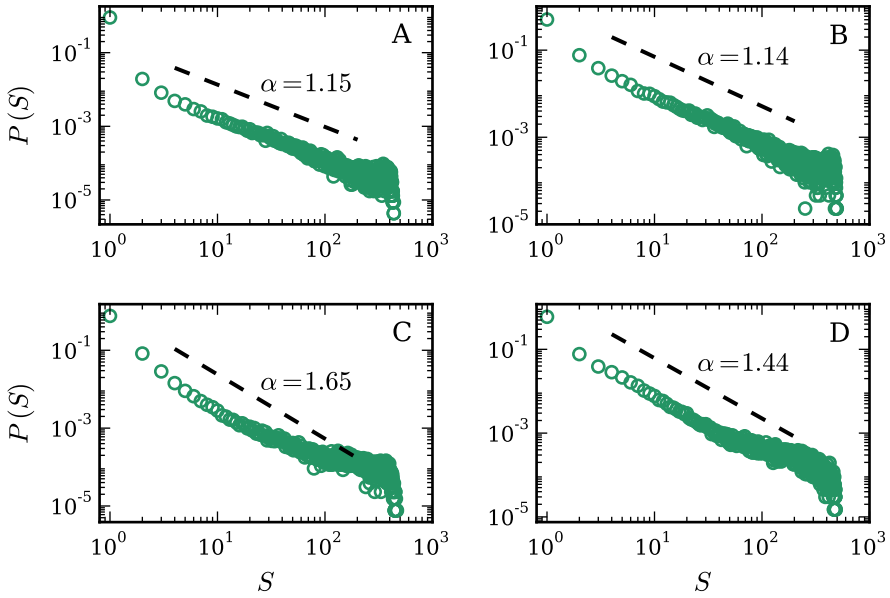


Figure 4.15: Size distribution of network components at the first transition point q_D for every model, $N = 500$, $\langle k \rangle = 4$, $F = 3$, collected over 400 realizations. Dashed lines represent the best fit of a power-law function with the power α given in the panels.

as a coevolving network with constant rewiring and no states of nodes. In such case the degree distribution is described by Equation 2.41. In particular, according to this relation the model B should have exponential degree distribution for $q \to \infty$ As we can see, this approximation works well also for finite values of q. On the other hand it does not work well in case of the model A, which does not display power-law distributions. It is worth mentioning that for $q = 5000$, for example in the model B, during the simulation only in 75% of time steps rewiring occurs. Therefore, states of nodes still influence the system at this level of q.

As presented in Figure 4.16, the goal of a power-law degree distribution is achieved in the model C with strong (quadratic) preferential attachment. In this particular choice of the q parameter values the exponent is equal $\alpha = 3.3$ in the first and $\alpha = 1.96$ in the second phase. Additionally, the bigger the network the longer is the straight slope in the log-log plot, indicating pure power law in the thermodynamic limit. Looking at all the models we can see a rich range of distributions.

It is clear that scale-free structures and high clustering can be obtained in the coevolving Axelrod model with proper rewiring. Another common feature of social networks is the small-world property, as described in Section 2.2.3. It can be observed in the scaling of the average path length l with the network size. This dependency is presented in Figure 4.17. Models A and B possess the standard small-world property – the average path length grows as a logarithm of the network size. The model C has a super small-world property – l initially decreases to stabilize at the level of approximately 3.

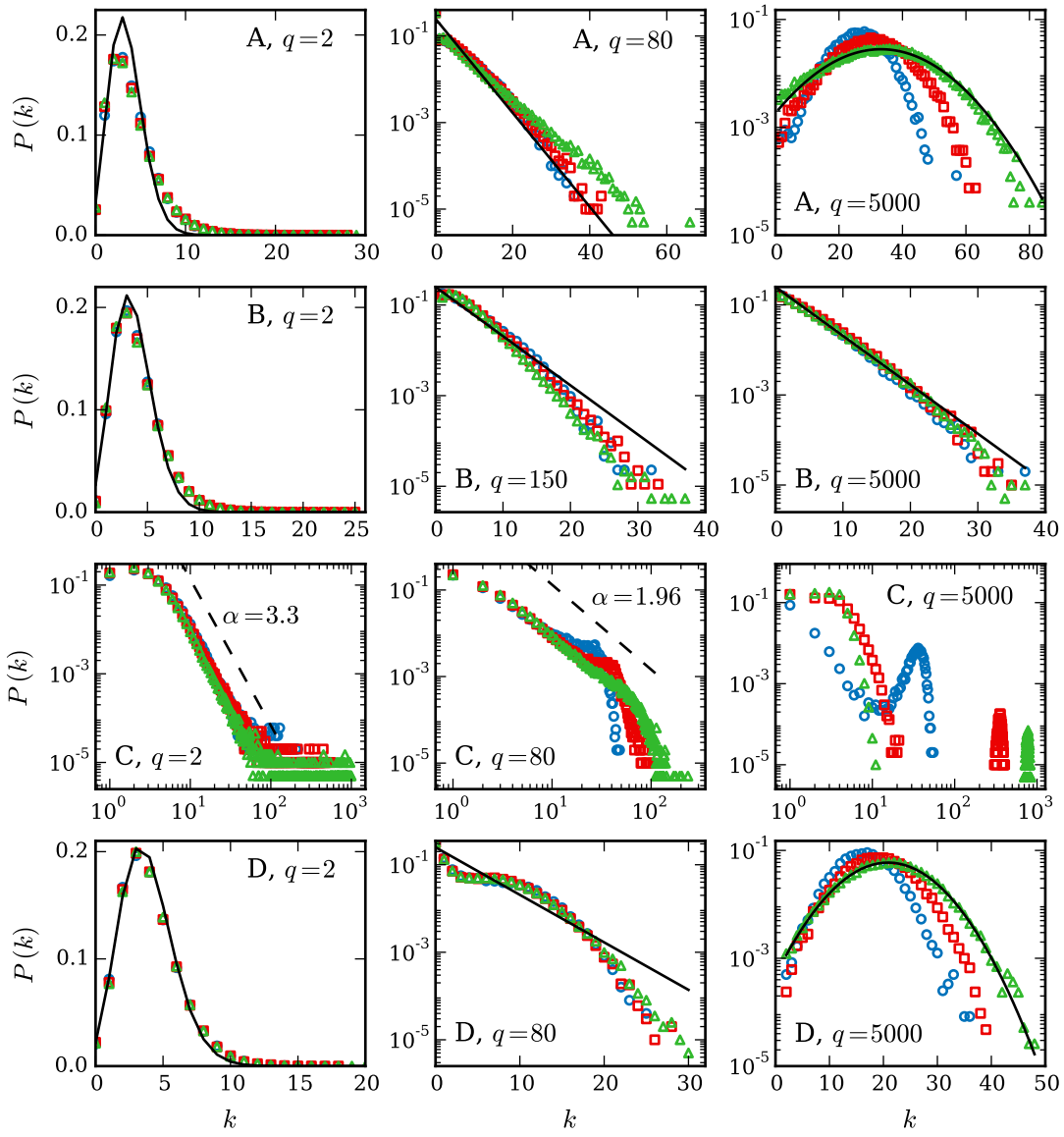


Figure 4.16: Degree distribution in the stationary state for $N = 500$ (blue circles), $N = 1000$ (red squares) and $N = 2000$ (green triangles), for every model and phase, $\langle k \rangle = 4$, $F = 3$, collected over 100 realizations. Dashed lines represent a power law, solid lines represent exponential, Gaussian or Poissonian distribution.

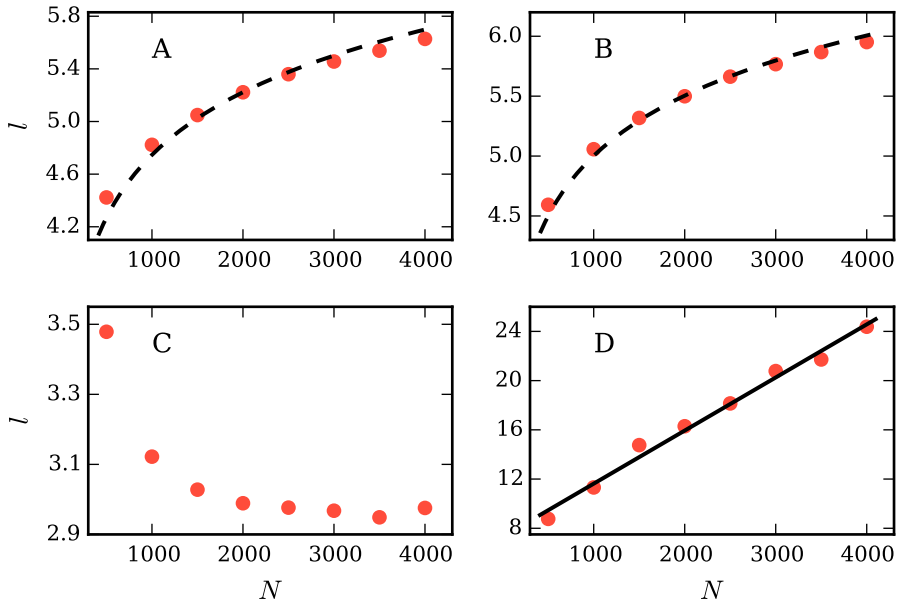


Figure 4.17: Average path length l vs. system size N for every model, $\langle k \rangle = 4$, $q = 3$, $F = 3$, averaged over 48 realizations. Dashed lines represent a logarithmic expansion, solid line represents a linear function.

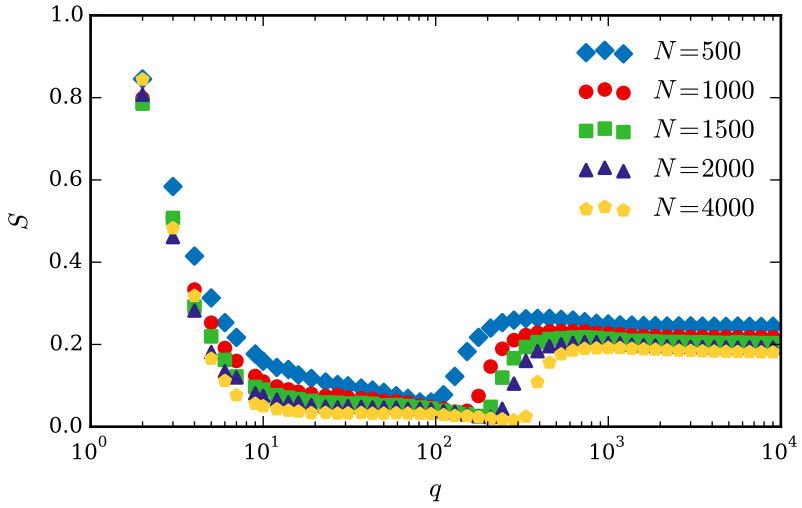


Figure 4.18: Average relative size of the largest network component S vs. parameter q for the model D, $\langle k \rangle = 4$, $F = 3$ and different network sizes, averaged over 400 realizations.

Such behavior is caused by the creation of big hubs with degree comparable to the number of nodes, but still smaller than $N-1$, because hubs of this size would decrease the average path length below 2. In the model D the average path length grows linearly with the number of nodes N, therefore it can not be described as small-world. Bigger distances between nodes are cased by the local rewiring procedure, which replaces long-distant connections with short links.

An additional difference between the model D and the rest of the models lays in the scaling behavior. In the models A, B and C, as in the standard coevolving Axelrod model, the first transition point q_D shifts towards bigger values of q when increasing the network size [264]. In the thermodynamic limit a continuous order-disorder phase transition is therefore expected. In the model D, however, the transition remains at the same value $q_D = 3$, regardless the system size, as visible in Figure 4.18. This indicates a discontinuous phase transition at this point in the limit of large N. The second transition point q_S shifts towards high values of q in every model.

These findings provide yet another proof of how important is the interplay between the state of nodes and the rewiring procedure in coevolving networks. Obtained results present a whole range of possible degree distributions, various average path length behaviors and different scaling of transition points. In a context of social networks, the discovery of the clustering coefficient peak at the second transition point is especially important, as well as higher general clustering in the model D. Majority of these phenomena is absent in static networks, what justifies generalizing network models to incorporate coevolution.

4.3 A solution of the problem with empirical data

Language, as one of the most important aspects of our culture, has been studied using numerous different approaches [269]. Significant part of the research in quantitative linguistics applies methods and ideas developed in physics. For instance, the classical gravity model explaining language change and spatial diffusion [270], [271]. Also more recent works borrow from physical sciences and use tools like agent based modeling [272], or scaling analysis [273]. Fokker-Planck equation, well known and studied in statistical physics, has been applied to simulate changes in language over time [274]. The term *statistical physics of language dynamics* is becoming popular as this branch of science is being developed [228], [275]. Research ranges from diffusion of modifications in spoken dialects [272], [276], to statistical and topological properties of written language [277], [278].

The Axelrod model has two basic interpretations. In the general one states of the nodes are assumed to describe culture in a broad meaning. It can be opinion, habits, attitude, religion, customs etc. Without going into details, as Robert Axelrod says „*culture is what social influence influences*" [234]. In this sense traits on every vertex can describe anything that can change during a social interaction.

A more particular interpretation, suggested already in the original paper [234], considers features of nodes as languages spoken by individuals. In such interpretation every feature should be understood as a relatively big part of a language, like a segment of vocabulary or general grammar rules. Then, different values of traits imply using different words to describe the same phenomena or different declination rules etc. It is worth noting that the Schulze model [279], which is very similar to the case of $q = 2$ in the Axelrod model, is considered a purely linguistic model describing language change.

The idea of social interactions driving the evolution and propagation of languages is well supported [280]–[283]. Multilingual society can be seen as a system of interacting agents [284]–[286], where the interaction leads to a modification of the language spoken by the parties [287], [288]. People interacting more often should reach the state of a linguistic compatibility. On the other hand, if they speak completely different languages they will not interact at all. Existence of these simple mechanisms makes the network science a suitable framework to analyse language change [129], [279], [289]. Many phenomena have been already explained using such approach [130], [210], [290], [291]. Therefore, it is a promising perspective to use the Axelrod model in description of the language diversity [292].

Despite the fact that the Axelrod model captures the essence of social interactions, its interpretation considering languages was abandoned after very first publication [234] due to a contradiction with the empirical data. Anthropological study of Solomon Islands in the late 70's [293], which provided the only empirical results comparable with the model at the time, showed that the number of languages existing on an island grows with the island's size. Behavior of the initial model defined on a static square lattice was exactly opposite – the number of domains, interpreted as separate languages, was decreasing with increasing size of the network (see Figure 4.3). Moreover, the first adaptive model [264], taking into account coevolution of the nodes states and the topology of the network, did not solve this issue – the number of domains was approximately constant for different sizes of the network.

First question that arises is whether the empirical findings for Solomon Islands still hold today and whether they are valid on a larger scale. Extrapolating results from the 70's we should expect a larger number of languages for bigger countries. In order to validate this dependency two independent databases are analyzed in the thesis [292]. The first one from 1996 with 6866 languages registered together with their 9130 dialects from 209 different countries [294], and the second one from 2013 (updated in 2015) consisting information about 2679 languages from 188 countries [295].

It should be noted that the difference between a language and a dialect is often subtle [269], [296]. In general, a dialect originates from a language and speakers of different dialects of the same language should be able to communicate without a big effort. There are measures invented to quantify the differences between dialects/languages, for example mutual intelligibility. Such criteria, however, are not foolproof and can classify different languages as dialects of one language [297], [298]. Sociolinguistic context also has to

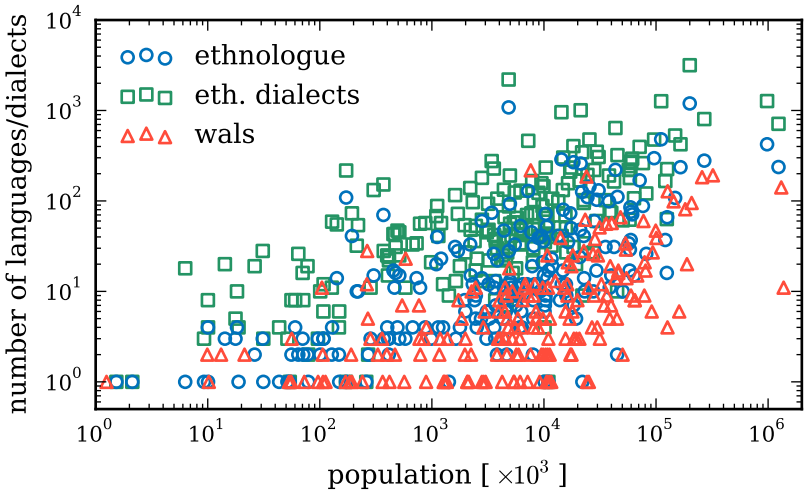


Figure 4.19: Empirical dependence of the number of languages on the size of a population. Red triangles represent languages from WALS [295] and population sizes for 2015 [299]. Blue circles represent languages from Ethnologue [294] and population sizes for 1996 [299], green squares represent all dialects from the latter source.

be considered in many cases. Nevertheless, this discussion shall be left for linguists. In this work, data on languages and dialects is used as given in the cited sources.

The gathered data is presented in Figure 4.19 as a dependency of the number of functioning languages (or dialects) vs. the population size in a given country. For each of the three datasets a general trend seems to be the same – the bigger the country the more languages are spoken there. However, the dependency is highly noisy. Social interactions is not the only factor driving language change. There are many others sometimes random factors, like language policy and legislation, colonization, border changes, demolition of the population during wars or epidemics, compulsory resettlement etc. Also television, radio and social media have a significant impact on the language diversity [300]–[302].

In order to see the average behavior it is better to aggregate the data. It is done in Figure 4.20 excluding, for the sake of clarity, four countries that have either the population size (China, India) or the number of languages (Indonesia, Papua New Guinea) grater by almost order of magnitude from the others. From this plot the average tendency can be clearly concluded – number of languages grows with the population size. The biggest dataset considering dialects displays a clear linear growth with a linear regression giving the coefficient of determination equal $R^2 = 0.94$. The language datasets can be also described by a linear function with a fair precision – in both cases the coefficient of determination is equal $R^2 = 0.85$.

These empirical findings confirm results obtained for Solomon Islands.

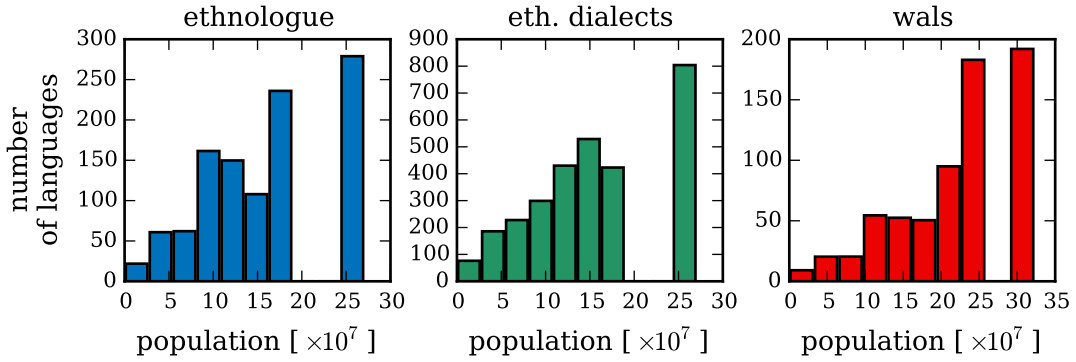


Figure 4.20: Dependence of the number of languages on the size of a population. Aggregated data from Figure 4.19, colors preserved. Height of a bar indicates the average number of languages in countries with a population lying within the width of the bar. China, India, Indonesia, and Papua New Guinea are excluded.

Unfortunately, it means they also confirm the problem of the Axelrod model. The scaling of the domains number, which is interpreted as the number of languages, with the system size is exactly opposite in the original model. In the coevolving Axelrod model behavior slightly depends on the phase. As described in Section 4.2.1, the model has three phases. The third one, however, does not make physical sense in the context of language diversity modeling, because traits of nodes do not play an important role there. In first two phases, which are the most plausible to describe language change, the number of domains remains approximately constant regardless the system size. Again, this behavior is contradictory to empirical analysis performed above.

However, the Axelrod model can be improved using modifications described in Section 4.2.2. In addition to four models (A, B, C, and D) introduced in that section a fifth one is proposed here. It is a combination of models C and D, therefore hereafter it is referred to as the model CD. As in the model D the rewiring is performed only within neighbors of neighbors, i.e. it is triadic closure. But in this set of nodes distant by two edges the probability distribution for being chosen as a new neighbor is not uniform. Instead it has the same form as in the model C: $\Pi(k_i) \sim (k_i + 1)^2$, but only if the node i is a neighbor of neighbor of the active node, otherwise $\Pi(k_i) = 0$.

Language scaling for all variants of the coevolving Axelrod model is presented in Figure 4.21 for $q = 5$. As we can see models A, C, D and CD have much better behavior, in the context of empirical findings, than the model with random rewiring. The model B displays similar results to the original coevolving model. Notably, models D and CD significantly improve scaling of the domains number. Therefore, this two models are analyzed further for different values of q, as showed in Figure 4.22.

We can observe a steady growth of the domains number n_D with the network size N. For every result of models D and CD in Figure 4.22 a fit of a

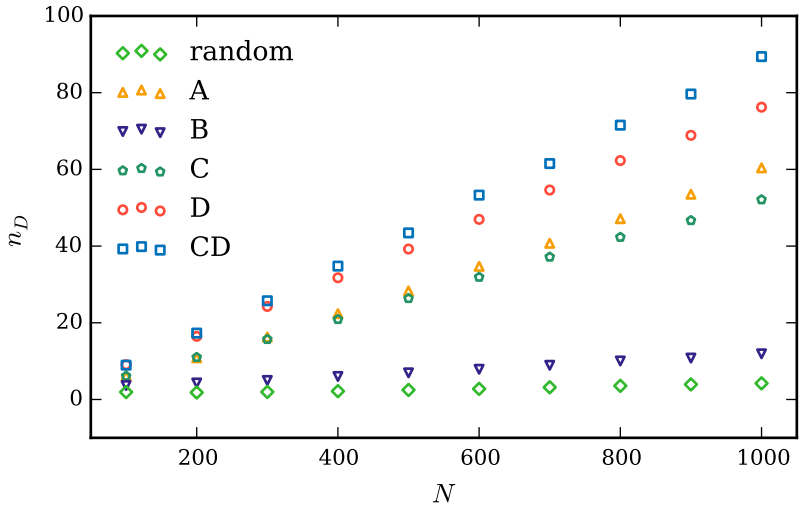


Figure 4.21: Number of domains n_D as a function of the network size N. For the original coevolving Axelrod model with random rewiring (green diamonds) and all the modifications, for $\langle k \rangle = 4$, $F = 3$, and $q = 5$, averaged over 400 realizations. Note that the original coevolving Axelrod model displays approximately constant number of domains.

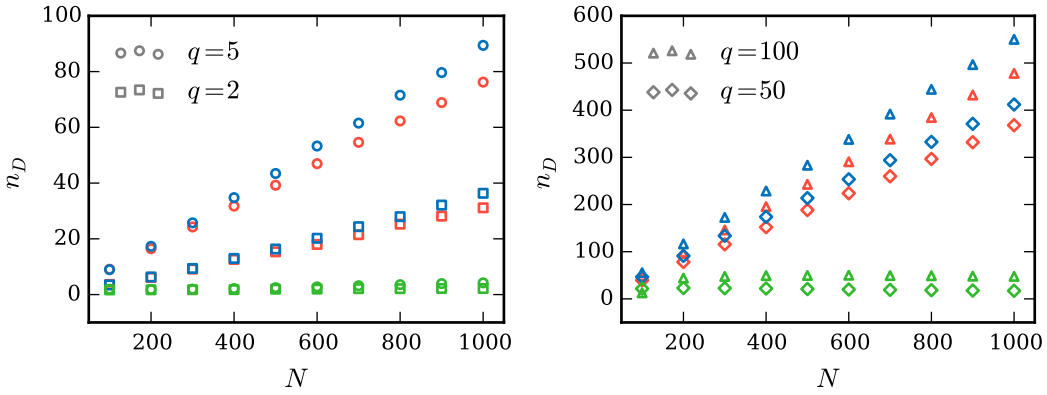


Figure 4.22: Number of domains n_D as a function of the network size N. For the original coevolving Axelrod model (green), the model D (red), and the model CD (blue), for $\langle k \rangle = 4$, $F = 3$, and different values of q indicated by the marker shape, averaged over 400 realizations. The original coevolving Axelrod model displays approximately constant number of domains, while the modifications have a clear linear growth, as in the empirical data.

linear function gives a value of the coefficient of determination $R^2 > 0.99$. Therefore, we can observe qualitatively the same behavior as for the empirical data on languages and dialects in Figure 4.20. Moreover, the slope of the function strongly depends on the q parameter, hence it can be adjusted to a particular empirical example giving quantitative agreement. In this way the contradiction between previous models [234], [264] and the Solomon Islands case can be erased.

The disagreement between the Axelrod model and empirical scaling of language diversity is eliminated for the first time in this thesis (based on [292]), to my best knowledge. It demonstrates the importance of the interplay between the topology of the network and states of the nodes, namely the coevolution, as well as particular type of rewiring.

It is worth noting that the studies on language diversity are very diverse themselves. For instance, properties of written texts are being extensively explored [303]–[305]. New words creation, language change or competition are being modeled on networks [272], [275], [286]. Network science, however, has its limitations in quantitative linguistics. It is hard, if possible, to take into account such essential aspects as geographical distribution and spatial diffusion in the process of language change [276], [306]–[308]. Nevertheless, results obtained using coevolving network models can be seen as one of the crucial parts of the complete description of language change or, more generally, of social interactions.

Chapter 5

Statistical mechanics of coevolving networks

So far models developed and described in this thesis belonged to the category of non-equilibrium models, or – using the division from [62] – to the category of algorithmic models. The voter model (see Chapter 3) and the Axelrod model (see Chapter 4) are defined by specifying the set of rules for each time step, i.e. by their *algorithm*. In both cases the system keeps evolving in time eventually reaching a stationary state, where basic parameters fluctuate around constant values, or a frozen configuration. The algorithmic approach is particularly useful when describing complex systems. It allows to incorporate the dynamical evolution rules of a system directly in the model, and complex systems are often naturally defined by such rules. It leads to a non-equilibrium behavior – the system is not characterized by a probability distribution over possible states realizing the most probable ones on average, but it rather collapses just into one (frozen configuration) or continues to change indefinitely, perhaps preserving certain quantities. For example, the voter model on static finite networks always reaches a gridlock (see Section 3.1.2). On the other hand, in the coevolving Axelrod model for high q values the network keeps changing maintaining the size of the largest component and domain (see Section 4.2.1).

Another way of looking at systems containing many elements comes directly from statistical mechanics and ideas of Ludwig Boltzmann [309]. We can assume a probability distribution over every microscopic configuration, as long as it satisfies constraints imposed by observation. This way we construct a statistical ensemble of the system considering all possible microscopic states at once. Of course, some of them might be extremely rare, other much more probable. In order to describe the system one must compute expected values of quantities of interest. An important step here is assumption that the ensemble does not change over time - the micro-states and the probability distribution put on them remain the same. For this reason such description is called *equilibrium* and models constructed in this manner analytic [62]. In this approach we are not interested in the evolution of the system, moreover we can not deduce any details of the evolution from the equilibrium approach. Instead, we obtain a description in terms of average values corresponding to macroscopic states of the system.

One may ask why to use the equilibrium approach in description of complex systems, if we don't obtain any knowledge about the evolution. Additionally, some of them can be naturally expressed in algorithmic manner. The main reason for this is limited availability of empirical realizations of such systems. There is only one World Wide Web, only one Internet, one financial network of bank loans, few examples of social online networks per country etc. Statistical mechanics approach still allows us to say something about the most probable realizations of these systems by means of constructing *the most ignorant* probability distribution (see Section 5.1). The second advantage is a natural analytical formulation, which gives a possibility of precise description. In algorithmic models the analytical description many times must be approximated *by definition*, since the model is described by the algorithm. Equilibrium approach may prove difficult to be fully solved, however is solvable in principle. It can also provide insight into quantities not directly observed thanks to the constructed probability distribution.

Nevertheless, the statistical mechanics approach can be more complicated when applied to complex systems than the algorithmic one. In certainly is for coevolving networks. For that reason this framework is more rarely chosen by researchers and is presented as the last one in this thesis.

5.1 Maximal entropy approach

The main problem in constructing a statistical ensemble is the proper choice of the probability distribution over the set of possible configurations. First of all, what does it mean that a distribution is *proper*? It is usually stated that such probability distribution has to (i) satisfy imposed constraints and (ii) include the least possible amount of assumptions that are not included in the constraints. The first point is rather straightforward. If we observe a particular average value $\langle f(x)\rangle = \sum_i f(x_i)p_i$ of some function $f(x_i)$ of the ensemble elements x_i, we want the distribution $\{p_i\}$ to recover the exact same value. Otherwise it would be simply in contradiction to the empirical evidence. The second point, however, is much more vague. It says that the distribution should be as little biased or as ignorant as possible. But how does this help to chose between two distributions both fulfilling the constraints?

This question dates back to the times of Jacob Bernoulli and and his *principle of insufficient reason* and later Pierre-Simon de Laplace's *indifference principle*. These principles state that having no prior knowledge about a system, except that is has n possible realizations, we should regard every realization as equally likely, for there is no reason to believe that one is more probable than any other one. Thus, in a situation with no information we should favor a uniform probability distribution. However reasonable and intuitive this principle sounds, the choice of the uniform distribution can be as arbitrary as any other one. If there is truly no prior knowledge about the system, no average value is known, why to put any distribution in front of the others?

The solution is provided by the maximal entropy principle (MaxEnt). One can define a functional $S(\{p_i\}) \equiv S(p_1, p_2, ..., p_n)$, called entropy, taking as an argument probability distribution $\{p_i\}$ and returning a number evaluating

amount of assumption the distribution contains. Not willing to make unjustified assumptions, we should chose the distribution of the maximal entropy. It would seem that this merely creates yet another issue of subjective criteria, since we have to select a particular form of such functional. However, it can be shown that the most reasonable form is the well-known Boltzmann-Gibbs-Shannon (BGS) entropy [310]:

$$S(\{p_i\}) = -K \sum_{i=1}^{n} p_i \ln p_i, \tag{5.1}$$

where K is a positive constant setting units, which can be omitted without loss of generality.

The BGS entropy can be derived based on a set of axioms treating about continuity, symmetry (or zero probability), maximum value and additivity [311]. It can be proved that the only functional dependence that satisfies these axioms is given by Equation 5.1 [310], [312]. Interestingly, the maximum value axiom states that entropy should have the maximal value for a uniform distribution, therefore it makes exactly the same assumption as the principle of insufficient reason. The maximal entropy approach, however, has two advantages. It can be derived in a strictly combinatorial way, without any reference to uncertainty or uninformativeness, just looking for the most probable distribution describing observed evidences [313]. Additionally, the entropy concentration theorem provides a quantitative justification for choosing the probability distribution maximizing BGS entropy. More precisely, a certain fraction r of possible distributions $\{p_i\}$ will result in such value of entropy $S(\{p_i\})$ that:

$$S_{max} - \Delta S \leq S(\{p_i\}) \leq S_{max}, \tag{5.2}$$

where S_{max} is the maximal possible value of the BGS entropy. The concentration theorem shows that for any fixed fraction r the margin decreases with the system size as $\Delta S \sim 1/N$ [314]. Take $r = 0.99$, for a large system it's simply unreasonable to believe that the true probability distribution is far from the one maximizing the entropy. When N becomes very large, like Avogadro's number, deviations from the MaxEnt principle are so improbable that they are just never observed. In this sense, the second law of thermodynamics can be seen as purely combinatorial result, what paradoxically makes it even stronger[1] [312], [316].

It is important to make a distinction between the probability of a given realization and the probability of observing a certain (probability) distribution. It's often stated that all realizations of a system are equally possible, but we observe only the regular or disordered ones because there is much more

[1] Perhaps that's why Arthur S. Eddington wrote: „*The law that entropy always increases holds, I think, the supreme position among the laws of Nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations – then so much the worse for Maxwell's equations. If it is found to be contradicted by observation – well, these experimentalists do bungle things sometimes. But if your theory is found to be against the Second Law of Thermodynamics I can give you no hope; there is nothing for it to collapse in deepest humiliation*" [315].

of them (e.g. homogeneous distribution of gas particles vs. high concentration in one place). This is basically assuming *a priori* the Laplace indifference principle and uniform distribution. On the other hand, the maximal entropy principle, supported by the concentration theorem, states that the outcome very close to the uniform distribution (in absence of constraints) is so overwhelmingly more probable that it's unlikely to observe any other. In a way, the MaxEnt principle is en extension of the principle of insufficient reason, justifying choice of the uniform distribution and showing how to incorporate constraints in this reasoning.

Having agreed on this logical framework of statistical inference, the next step is finding the probability distribution maximizing the formula 5.1. This can be done using Lagrange multipliers. For the full derivation see Appendix A. Assuming a constraint fixing the value $\langle H(x) \rangle$ we will obtain a probability distribution:

$$p(x_i) = \frac{1}{Z} \, \mathrm{e}^{-\beta H(x_i)}, \tag{5.3}$$

where $Z = \sum_{i=1}^{n} \mathrm{e}^{-\beta H(x_i)}$ is a partition function playing a role of normalization factor. Function $H(x_i)$ is called Hamiltonian and represents energy of the system in physics or likeliness of particular realizations x_i in general. Having the probability distribution gives a complete stochastic description of the system and allows for prediction of quantities not observed directly. It represents the best estimates that could have been made on the basis of the available information. An example of application of this approach is the Exponential Random Graphs Model (see Section 2.2.2).

The distribution obtained by the MaxEnt principle is exponential in Hamiltonian. However, it doesn't always have to be of this form depending on the imposed constraints [317]. The universality of this approach can be observed in the range of its application. MaxEnt has been used not only in originally associated fields, like statistical mechanics, information theory, or statistical inference [62], [314], but also biology, economics and social sciences [85], [318].

Despite the generality and unquestionable versatility of the Boltzmann-Gibbs-Shannon entropy researchers have tried to extend it in order to obtain different probability distribution, for instance a power law. MaxEnt principle is usually maintained in such approaches, however, with different form of the entropy. Popular examples are Rényi entropy [319] and Tsallis entropy [320], both recovering BGS entropy as a special case. The latter one can be even further generalized for a class of statistics, called superstatistics [321], [322]. Nevertheless, the only entropy satisfying the additivity axiom is BGS. Proposing any other functional form than in Equation 5.1 we must give up on this condition. In fact, once we agree on violating additivity axiom we can derive a whole class of entropies with different asymptotic behaviors [323], [324]. Still, only BGS entropy has such strong justification. Any other form of entropy will poses a bias not enforced by the data and should be used with caution [325], [326].

5.2 The Ising model

One of the most famous examples of the maximum entropy approach is the Ising model [14], [15]. It's a theoretical model inspired by the ferromagnetic phenomena. At the same time it's very useful in analysis of phase transitions and critical phenomena. Additionally, it surprisingly well represents some physical systems in a qualitative manner, and with certain extensions even quantitative. The Ising model is a binary-state equilibrium model. It consists of a set of spins $\{s_i\}$ placed on a certain network – initially it was a one-dimensional square lattice (a chain of spins). Each of the spins can take one of two values $s_i \in \{-1, +1\}$. Two adjacent spins s_i and s_j interact with each other with a strength J_{ij} promoting parallel alignment ($J_{ij} > 0$, ferromagnetic interactions), or anti-parallel alignment ($J_{ij} < 0$, anti-ferromagnetic interactions). Additionally, there might be an external field h_i interacting with every spin in not necessarily homogeneous manner. The model is then defined by its Hamiltonian, describing the energy of the system, or the preference towards particular configurations:

$$H(\{s_i\}) = -\sum_{\langle i,j \rangle} J_{ij}\, s_i s_j - \sum_i h_i s_i, \tag{5.4}$$

where $\langle i, j \rangle$ stands for summation over nearest neighbors.

This is somewhat inverse approach than the one described in the previous section. We are not given constraints upon which we construct probability distribution, by finding the proper Hamiltonian including those constraints. Here we first postulate the Hamiltonian arguing that it's form resembles the phenomena we are willing to describe, that it's particular terms are corresponding to the actual system's behavior. Then, we take the Boltzmann statistics as the optimal one, based on the maximal entropy principle:

$$p(\{s_i\}) = \frac{1}{Z}\, \mathrm{e}^{-\beta H(\{s_i\})}, \tag{5.5}$$

where $Z = \sum_{\{s_i\}} e^{-\beta H(\{s_i\})}$ is a partition function and β is a Lagrange multiplier.

This approach can be understood as setting a "flexible" constraint on the total energy $E = H(\{s_i\})$ given by the Hamiltonian. However, instead of finding a proper value of the Lagrange multiplier β corresponding to the assumed energy, the energy is adjusted to a fixed value of β. In this way, by varying the value of the Lagrange multiplier we can cover a whole range of energies. Therefore, we can study the system under different circumstances. In this approach the Lagrange multiplier has a physical interpretation. As it accounts for the amount of fluctuations in the system, it can be associated with the inverse temperature $\beta = 1/T$.

The most famous interpretation of the Ising model, being as well the original one [327], is a ferromagnetic material. The lattice represents a crystal structure and states of the nodes represent a quantum intrinsic angular momentum carried by electrons, i.e. the spin. That's actually why we call states

s_i in the Ising model spins. J_{ij} stands then for the spin-spin interactions and h_i for an external magnetic field. Nonetheless, the definition of the model is universal and it was used in description of other systems, like spin glass, alloys, lattice gas, neural networks or even bacterial vortexes [328]–[330]. It is worth to note that the voter model (Chapter 3) can be regarded as a non-equilibrium algorithmic version of the Ising model.

The Hamiltonian 5.4 takes the simplest form under the assumption of isotropic and homogeneous space:

$$H(\{s_i\}) = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_i s_i, \tag{5.6}$$

where additionally the external field is often excluded $h = 0$ and $J = 1$. There is a fistful of questions to ask about a system described by such Hamiltonian, perhaps the main one being: how does a typical configuration look like? Of course, the configuration of spins $\{s_i\}$ may depend on the parameters choice, therefore we can also wonder how does it depend on the temperature T. Originally the model was analyzed only on a one-dimensional static square lattice (a chain of spins) and only one phase was found [15], [331]. For any temperature $T > 0$ the system does not display any order and spins do not prefer any direction. Therefore, there is no phase transition in finite temperatures.

Although initially discarded due to uninteresting results, the Ising model came back to favors after discovery of an order-disorder phase transition [327]. First a proof of existence of an ordered phase was given by Rudolf Peierls [332] and later a full analytical solution was provided by Lars Onsager [333]. It was shown that for a square lattice (as for any higher dimension) a critical temperature T_c exists, below which the system orders with most of the spins pointing the same direction and magnetization $|m| = |\sum_i s_i/N| \approx 1$ [334]. For two dimensions and isotropic homogeneous space the critical temperature reads $T_c = 2J/\ln(1+\sqrt{2})$. In Figure 5.1 we can observe a time evolution of the system for the ferromagnetic phase in a square lattice of a linear size $L = 200$ ($L^2 = N$).

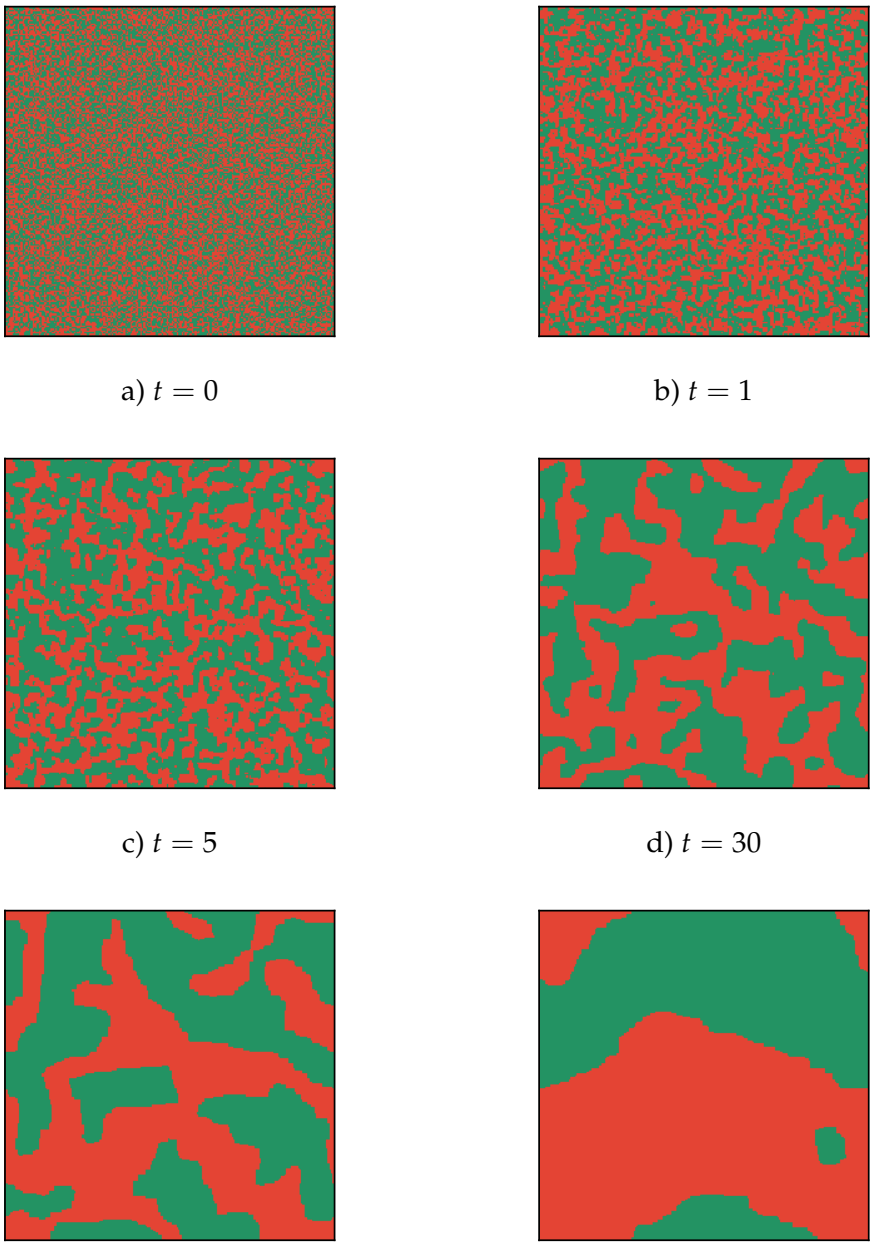
Remarkably, the definition of the model is fully symmetrical (for the Hamiltonian 5.6), yet the outcome is not. This phenomenon is referred to as *spontaneous symmetry breaking*. Both directions $+1$ and -1 are equally likely, but the system realises one of them.

A convenient quantity to use is a logarithm of the partition function called *free energy*[2] $F = -T \ln Z$ and the free energy per site $\mathcal{F} = \lim_{N\to\infty} F/N$. Then, the zero-field magnetisation is given by $m = \frac{\partial \mathcal{F}}{\partial h}|_{h=0}$ [309], what for a static square lattice results in [333]:

$$m = \left(1 - \sinh^{-4}(2\beta J)\right)^{1/8}. \tag{5.7}$$

Other quantities of interest can be also computed from the free energy, for

[2]The Boltzmann constant k_B is assumed to be unity in these calculations.



a) $t = 0$ b) $t = 1$

c) $t = 5$ d) $t = 30$

e) $t = 100$ f) $t = 1000$

Figure 5.1: Snapshots of the lattice during evolution of the Ising model (Hamiltonian 5.6) from a random configuration at $T < T_c$. Simulation was ran for $L = \sqrt{N} = 200$, $J = 1$, $h = 0$, and $T = 0.4$. Time is given in a number of MC time steps. States $+1$ and -1 are indicated by different colors. Note the emergence of macroscopic domains.

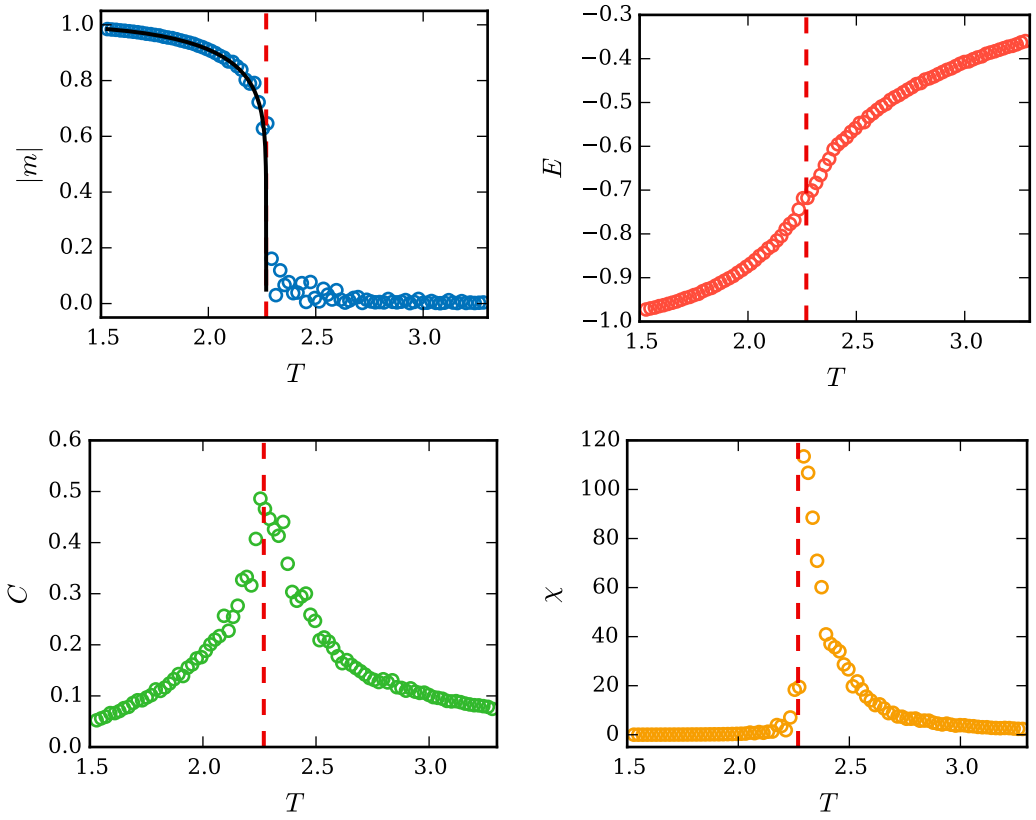


Figure 5.2: Absolute magnetization $|m|$, energy E, specific heat C, and magnetic susceptibility χ as a function of the temperature $T = 1/\beta$, for $L = 25$ ($N = 625$), $J = 1$, $h = 0$, averaged over 10^4 MC time steps. Analytical solution (Equation 5.7) for the magnetization is indicated by a solid black line. Critical temperature T_c is marked with a dashed red line in every panel.

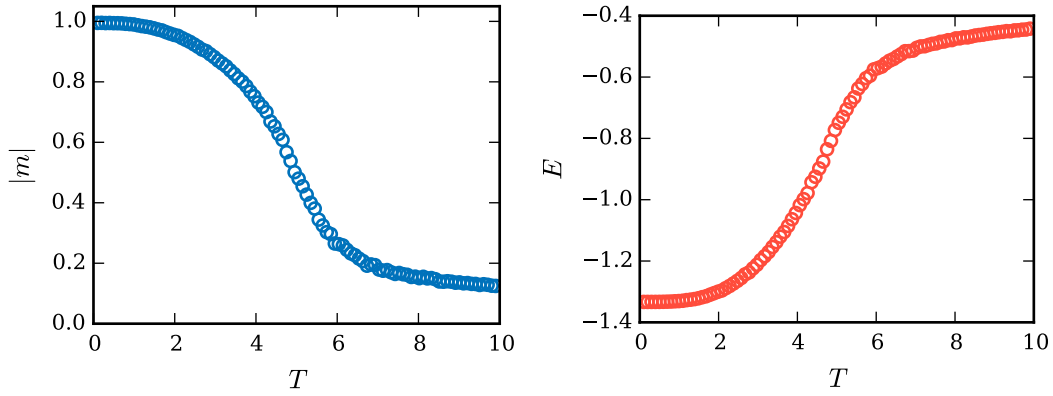


Figure 5.3: Absolute magnetization $|m|$ and energy E as a function of the temperature $T = 1/\beta$, for $N = 100$, $M = 300$, $J = 1$, $h = 0$, averaged over 10^5 MC time steps. Results for the model from [340] with not static network defined by the Hamiltonian 5.8

instance the specific heat given by $C = -T\frac{\partial^2 \mathcal{F}}{\partial T^2}$ and the magnetic susceptibility $\chi = \frac{\partial^2 \mathcal{F}}{\partial h^2}|_{h=0}$. All of them are plotted vs temperature in Figure 5.2. The data in the figure was obtained from a Monte Carlo simulations of the Ising model using Metropolis algorithm [335]. As can be noted, the system diplays special behavior at the critical temperature. Described quantities exhibit power-law scaling close to the point T_c. If we define reduced temperature as $\tau = (T - T_c)/T_c$, we can define critical exponents from relations[3]: $m \sim |\tau|^\beta$ for $\tau < 0$, $C \sim |\tau|^\alpha$, and $\chi \sim |\tau|^\gamma$ (note that the exponents may differ depending on the sign of τ, i.e. the side from which the critical point is approached). Additionally, the system possesses long-term correlations – the correlation length also scales as a power law. However, a complete description of critical phenomena goes beyond the scope of this thesis. More on this topic can be found in [336].

Since the Ising model uses the general framework of maximal entropy principle and has a lot of spare room for parameters adjustment, it has been extended into different more complex typologies. In many social, economical, financial, biological, ecological and technical systems a square lattice is rare, if ever observed. However, the equilibrium approach with a binary state and nearest neighbor interactions can be applied more widely. For that reason the Ising model was studied on complex networks, like small-world and scale-free [337]–[339]. In the small-world networks generated from square lattice of two and three dimensions the transition depends on the rewiring probability p as a power law, approaching mean-field description in the thermodynamic limit. In Barabási-Albert networks, as expected, hubs have a crucial role being able to force the system into a particular ferromagnetic state.

Finally and most importantly in the context of this thesis, the Ising model

[3] Unfortunately, it's the usual convention to use the same latter for the critical exponent describing the scaling behavior of magnetization as for the inverse temperature.

was studied on coevolving networks [340] (it is worth mentioning that the Potts model was considered with a similar approach [341]). The Hamiltonian of such model depends on the adjacency matrix as well as on the spin configuration:

$$H(\{a_{ij}\},\{s_i\}) = -\sum_{i<j} a_{ij}s_i s_j, \tag{5.8}$$

where $\{a_{ij}\}$ represents all possible configurations of the network with a fixed number of nodes N and edges M. Links can be rewired, adding a new possibility of interaction between spins. However, the dependence on the topology is weak, since no particular topological structure is preferred. The transition in this model is presented in Figure 5.3. The ferromagnetic phase displays dependence on the average degree in such model. But the topology of the system was not in the main focus of the article, since it produces only ER graphs. Additionally, this approach should be perhaps called *semi-coevolution*, because the feedback loop between the topology of the network and spins is weak, although both can change. For instance, in low temperatures the possibility of link rewiring can not change the outcome of the model – all nodes are in the same state, therefore it doesn't matter which node is connected to which one. A generalization of this model, incorporating topological traits, was developed in this thesis and is described in Section 5.4.

5.3 Statistical mechanics of network's topology

Maximal entropy approach was applied also in networks. The first example was already discussed in Section 2.2.2. Exponential Random Graphs Model (ERGM) is a direct application of the MaxEnt principle – first a constraint (or constraints) is chosen, then a probability distribution over an ensemble of graphs is constructed. Since the Boltzmann distribution in respect to the energy is a general result, to find the particular form of the distribution one must calculate values of Lagrange multipliers satisfying the constraint. It is therefore a single-point distribution.

In the Ising model the approach is reversed in some sense. The Hamiltonian is of course constructed to depend on quantities of interest, but there is no fixed constraint. For example, magnetization could be fixed and then, as in the ERGM, temperature adjusted to satisfy the choice of the magnetization. But this is not the usual method of analyzing the Ising model. Instead, the Lagrange multiplier, which means temperature in this case, is varied to see what outcomes it will produce. This direction of analysis is not only simpler computationally, it also covers a whole range of parameters.

A concept similar to the one of the Ising model, and reverse to that from the ERGM, has been applied in description of networks without any states of the nodes. As only the topological traits are taken into account it can be seen as statistical mechanics of network's topology. Usually, this kind of models consider a set of networks, which can be represented by all possible configurations $\{a_{ij}\}$ of the adjacency matrix, with a fixed number of nodes N and

edges M. In general, the partition function is given by [342]:

$$Z = \prod_{i<j} \sum_{a_{ij}=0}^{1} \delta(M - \mathrm{Tr}A^2/2) e^{-\beta H(\{a_{ij}\})}, \tag{5.9}$$

where A is an adjacency matrix, therefore $\mathrm{Tr}A^2/2$ gives the actual number of edges in the network.

The main point of this methodology is choosing a particular form of the Hamiltonian reflecting properties of graphs or microscopic interactions, as those between spins in the Ising model. Several possibilities have been studied in the literature. On one hand, Hamiltonians dependent on global features of networks were proposed [343], for example:

$$H(\{a_{ij}\}) = f(S), \tag{5.10}$$

where $f(S)$ is a function of the largest network component to be determined. Also an extension to all the components of the network has been investigated: $H(\{a_{ij}\}) = \sum_i f(S_i)$, where S_i is the size of the i-th component. In this case a fragmentation transition was reported between an ordered phase with one network component and disordered phase with many small components.

On the other hand, Hamiltonians dependent on local topological features, i.e. features of individual nodes, were studied. The main topological trait of a node is it's degree, therefore the postulated Hamiltonian took the following form [343]–[345]:

$$H(\{a_{ij}\}) = \sum_{i=1}^{N} f(k_i). \tag{5.11}$$

Among analyzed functions were $f(k_i) = -k_i^2$ and $f(k_i) = -k_i \ln k_i$. In this way a preference of high (or low) degree can be incorporated in the model. For Hamiltonians of the form 5.11 scale-free networks were reported, as well as star configurations and random graphs.

Finally, correlations were taken into account [342]. Since many networks are driven by interaction between neighbors the proposed general Hamiltonian was:

$$H(\{a_{ij}\}) = \sum_{i<j} a_{ij} g(k_i, k_j). \tag{5.12}$$

An example of the function is $g(k_i, k_j) = \delta_{k_i,1}\delta_{k_j,1}$ forcing the network to disintegrate into individual edges, or $g(k_i, k_j) = \delta_{k_i,k_j}$ enhancing assortativity mixing. Long distance correlations were considered as well by analyzing a Hamiltonian dependent on path lengths:

$$H(\{a_{ij}\}) = \sum_{i<j} d_{ij}, \tag{5.13}$$

where d_{ij} is the shortest path length between nodes i and j. These models, as expected, produced much higher correlations between degrees of the nodes.

Other aspects were also studied. Worth mentioning is an attempt to obtain a systematic theory of clustering in networks [80]. However, equilibrium description doesn't have to be always appropriate. Many real-world networked systems are far from equilibrium or are literally defined by the dynamical evolution (like in the models described in Chapters 3 and 4). Nevertheless, studying equilibrium models, especially originating from the maximal entropy principle, is useful for at least two reasons. Firstly, this approach still can be fruitful in many cases. Even if not fully explaining the process which led to a particular network configuration, it can well characterize the configuration and give clues to what other possibilities one could expect. Secondly, it is an interesting theoretical framework worth exploring.

5.4 Coevolution of spins and topology

We have already seen applications of the maximum entropy principle in spin models (Section 5.2), as well as network models (Section 5.3). A natural extension of previous studies is to apply the same methodology in a model with both: distinguishable states of the nodes and a variable topology, evolving together. That step is taken in this thesis, to my best knowledge, for the first time [346]. The following section contains original results of a model developed in the thesis.

As mentioned before, a semi-coevolving model of spins placed on a network without a fixed structure has been already studied [340]. It is disputable whether a network in an equilibrium model can be called coevolving. To talk about coevolution there should be evolution (in time) in the first place, while equilibrium models don't say anything about the system's evolution, solely about the equilibrium state. Of course, some kind of evolution is often assumed for the purpose of performing numerical simulations, for instance the Metropolis-Hastings algorithm (or the Glauber algorithm) in the Ising model [347], [348]. However, the rules of such evolution are not a part of the model itself and are tailored in order to satisfy the detailed balance principle. Therefore, all that matters is the equilibrium state reached after the thermalization, not how it was reached, since the choice of the algorithm may be arbitrary. Nevertheless, the most important feature of coevolving networks is the feedback between the topology and states of the nodes, and it can be obtained by using a proper Hamiltonian of a general form:

$$H(\{a_{ij}\},\{s_i\}) = \sum_{i<j} a_{ij} f(k_i, k_j, s_i, s_j) + \sum_i g(k_i, s_i). \tag{5.14}$$

With the presence of spins and topological traits of the nodes at the same time the feedback can be achieved in a sense of the equilibrium state of the network's topology influencing the equilibrium state of the spins, and vice versa. In this spirit, the Hamiltonian 5.14 represents equilibrium models of coevolving networks or *statistical mechanics of coevolving networks*.

Having a Hamiltonian we define the partition function as:

$$Z = \sum_{\{a_{ij}\},\{s_i\}} e^{-\beta H(\{a_{ij}\},\{s_i\})}, \tag{5.15}$$

where $\{a_{ij}\}$ represents all possible configurations of the network with a fixed number of nodes N and edges M. From the partition function one can obtain quantities of interest. For instance, the energy $E \equiv \langle H \rangle$ which is the average value of the Hamiltonian is given by:

$$E = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}, \tag{5.16}$$

what is straightforward to prove by performing the differentiation. The relation between average values and derivatives of the partition function is derived in Appendix A.

A particular form of the Hamiltonian studied here can be written as [346]:

$$H(\{a_{ij}\},\{s_i\}) = -\sum_{i<j} a_{ij}\left(\frac{k_i k_j}{\alpha}\right)^{\phi} s_i s_j - \sum_i k_i^{\gamma} - h\sum_i s_i, \tag{5.17}$$

where ϕ and γ are the main model parameters, α is a normalization of the interaction term and h is the external field interacting with spins. This simple Hamiltonian allows to continuously switch from complex topological interactions to the classical Ising model by tuning the parameters ϕ and γ. The weight of interaction ϕ, in the Ising model framework, can be treated as a non-homogeneous weight J_{ij} assigned to an edge (i,j). A multiplication of degrees is one of the simplest interaction expressions involving topological traits. Moreover, $J_{ij} = \left(k_i k_j/\alpha\right)^{\phi}$ is consistent with real-world weighted networks characteristics, as the weight can depend exactly as a power law on the degrees product [42]. The normalization term α is added to make the interaction and the external topological field effects comparable. Additionally, it changes scaling behavior as we will see. The simplest choice is obviously $\alpha = 1$, however to be able to compare effects for $\phi = 1$ and $\gamma = 2$ the best choice is $\alpha = \langle k \rangle = 2M/N$, which sets both sums in the same order of magnitude. In some cases it is more convenient to describe the network's density by the connectivity $c = M/N = \langle k \rangle/2$, instead of the average degree or the number of edges. The second sum in the Hamiltonian can be seen as an external field γ interacting with topological traits of the nodes, i.e. the degree. It introduces a preference for high or low degree in the system.

Finding exact solutions of the model in a form of average values for equilibrium state is a difficult task. Therefore, in addition to the analytical approach it is worth exploring it by means of numerical simulations. It is, however, not obvious how to perform such simulations. In the Ising model the most common approach is a Monte Carlo simulation with the Metropolis algorithm [347], [348], which satisfies the detailed balance condition. But

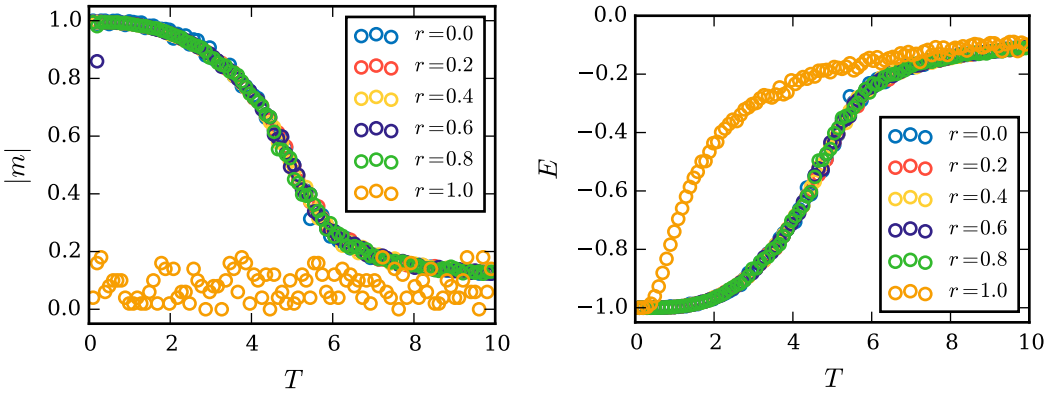


Figure 5.4: Absolute magnetization $|m|$ and energy E as a function of temperature T for the Hamiltonian 5.18, for $N = 100$, $M = 300$, averaged over 10^5 MC time steps. Results for different probabilities r of edge rewiring are represented by different colors.

in the standard Ising model only spins can change. Here we have an additional degree of freedom associated with the possibility of realizing any topology of the network. Hence, the algorithm of a simulation must include edge rewiring to enable covering the whole space of possible configurations. The question is how to perform the rewiring in practice.

In every time step there has to be a possibility for both – a spin update and an edge update. The update of a spin means changing the spin of a randomly chosen node to the opposite one. The update of an edge means rewiring a randomly chosen link, what can be done by randomly changing either one or two attachment points of it, i.e. only one side or the whole link can be rewired. After any update the energy, given by the value of the Hamiltonian, is compered between new and old configurations. If the energy change is $\Delta E \leq 0$ the new configuration is accepted. Otherwise, for $\Delta E > 0$ the new configuration is accepted only with a probability $e^{-\beta \Delta E}$, where $\beta = 1/T$ is the inverse temperature. In practice, it doesn't matter which type of rewiring is performed, if both cover the whole space of possible configurations [64], i.e. we can obtain any possible topology. However, rewiring the whole edge leads to a shorter thermalization time, therefore it is used in this work for computational efficiency reasons.

The next issue is choosing a proper ratio of edge update to spin update frequency. Intuitively, it should not matter. If both are performed at any finite ratio, any configuration of spins and topology can be obtained, perhaps after longer time. Nevertheless, it is an important aspect and in order to have certainty that the algorithm will reflect the equilibrium state given by the Hamiltonian it should be investigated. We shall compare results generated for $\gamma = 0$, $\phi = 0$, and $h = 0$, what gives a Hamiltonian consistent with the

one from Equation 5.8 [340]:

$$H(\{a_{ij}\},\{s_i\}) = -\sum_{i<j} a_{ij}s_is_j - N, \tag{5.18}$$

where N is an additional factor coming from the second sum in 5.17. However, the equilibrium state is invariant under additive constants in the Hamiltonian, because they do not change the energy difference ΔE between configurations. Similarly, $\gamma = 1$ would result in the second sum in 5.17 giving a constant factor equal $2M$ instead of N, therefore this parameter has two neutral values. The model described by the Hamiltonian 5.18 is simulated with a probability r of edge rewiring in every time step and $1 - r$ of spin update. If $r = 0$ the network is static and we obtain the Ising model on a random graph (every simulation is initialized with a random network and random spins). On the other hand, for $r = 1$ there is only edge rewiring, hence the magnetization can not change. The results are shown in Figure 5.4. As we can see, the outcome is the same for all ratios except $r = 1$. Additionally, for $r = 0$ the topology can not evolve, therefore this value also has to be excluded in analysis of the complex Hamiltonian 5.17. However, the thermalization time differs significantly between different values of r, raising abruptly close to the boundary values. For this reason in the final algorithm edge rewiring is performed as frequently as spin update – both once in every time step. In equilibrium models it is especially important to specify the algorithm used for simulations to enable repeating and verification of obtained results (in algorithmic models the algorithm is the definition of the model). Therefore, the algorithm used to simulate models described in this section is provided in Appendix C written in a pseudo-code.

Figure 5.4 leads to yet another conclusion. The magnetization and energy are perfectly the same for $r = 0$ as for $0 < r < 1$. In other words, the freedom to realize any topology (with constant N and M) does not change the outcome of the Ising model on a static random network, if topological traits are not present in the Hamiltonian. It doesn't even change the topology of the network, as in both cases it is a random graph. This is why such approach was referred to as *semi-coevolution* previously – topology and spins evolve on the same network, but without a strong feedback loop between them. The picture, however, changes dramatically after including the topological effects of the Hamiltonian 5.17. Not only the structure of the network can differ, but also the magnetization, since it is coupled with the topology through the interaction term.

5.4.1 Topological field and spins

In order to better understand the behavior of the system the topological effects corresponding to parameters γ and ϕ are studied separately. In this section we will focus on the topological external field interacting with the degree of the nodes. After setting the interaction parameter to the neutral

value $\phi = 0$, and removing the magnetic field $h = 0$ for the sake of simplicity, the resulting Hamiltonian is:

$$H(\{a_{ij}\}, \{s_i\}) = -\sum_{i<j} a_{ij} s_i s_j - \sum_i k_i^\gamma. \tag{5.19}$$

In Figure 5.5 we can see the behavior of typical order parameters[4] at different temperatures for $\gamma = 2$, corresponding to a preference of high degree. We can clearly observe a transition in every one of them, however the relative change in the number of components n_S and the size of the largest component S is marginal. The density of active links ρ is anti-correlated with the absolute magnetization. Therefore we will focus on three parameters which can comprehensively describe the system: the absolute magnetization $|m|$, the energy E, and the largest degree k_{max}.

Magnetization exhibits a continuous transition from an ordered ferromagnetic phase ($|m| = 1$) to a disordered paramagnetic phase ($|m| = 0$), except a small jump at $T \approx 15$. Energy displays exceptional behavior – it has four levels of approximately constant values and at certain temperatures jumps between them. First of these discontinuous changes takes place at the same temperature as the small jump of the magnetization. The largest degree also discontinuously changes from a maximal possible value ($k_{max} = 1$) for low temperatures to a small value corresponding to a random graph. The jump of k_{max} overlaps with the last jump of the energy.

The discontinuous jumps visible in the order parameters can be explained by assuming that star configurations are formed – hubs with a degree comparable to the size of the whole network emerge. The number of edges $M = 300$ being three times bigger than the number of nodes $N = 100$ allows for at most 3 stars to emerge. Consequently, 4 different levels of energy in Figure 5.5 correspond to, going from low to high temperatures, 3 stars, 2 stars, 1 star, and no hubs. This description is consistent with observed values of the largest degree – it drops from maximal possible value at the last jump of energy. Three possible star configurations are presented in Figure 5.6. The snapshots of the network state show how the links from destroyed stars are distributed among the rest of the nodes. Note also, that the magnetization is influenced by the disappearance of of the first star when decreasing T. It exhibits a small drop, what is in accordance with results of the Ising model on scale-free networks, where it was proved that hubs can drive the ferromagnetic state of the system [338].

Finally, the hypothesis of star formation is confirmed by the behavior of the system for a changing connectivity. In Figure 5.7 we have three main order parameters plotted for the same network size $N = 100$ and different numbers of edges $M = 100$, 300, and 600. For $M = 100$ only one jump in the energy is observed, for $M = 300$ three jumps, and for $M = 600$ six jumps, i.e. the number of jumps in the energy corresponds to the maximal number of stars possible to obtain in the network. Additionally, the drop

[4]Note that, as everywhere in this thesis, order parameters except the energy are normalized to fit the range $[0, 1]$. The energy is given in arbitrary units.

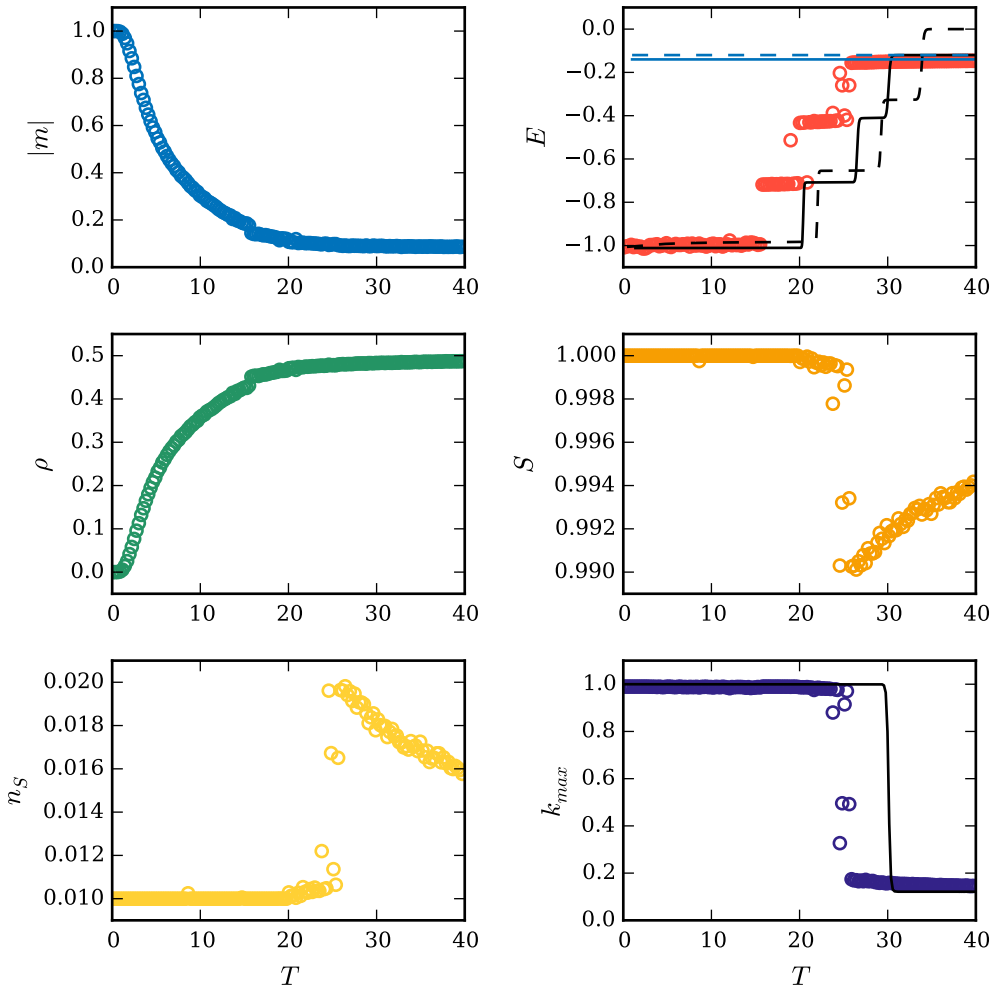
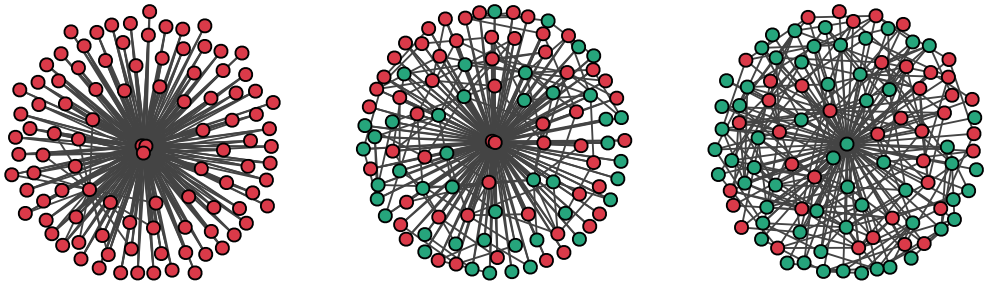


Figure 5.5: Absolute magnetization $|m|$, energy E, interface density ρ, the largest component S, number of components n_S, and the largest degree k_{max} as a function of temperature T for the Hamiltonian 5.19, for $\gamma = 2$, $N = 100$, $M = 300$, averaged over 10^5 MC time steps. Analytical approximations for the energy are given by Equations 5.20 (blue dashed line), 5.21 (blue solid line), 5.26 (black dashed line), 5.29 (black solid line). Approximation of the largest degree (solid black line) comes from Equation 5.31.



a) $T = 0.5$ b) $T = 18$ c) $T = 20$

Figure 5.6: Network configurations for different temperatures with $\gamma = 2$, $N = 100$, $M = 300$. Red color indicates spin -1, green $+1$. Note that the number of hubs in the middle changes from a) three, through b) two, to c) one.

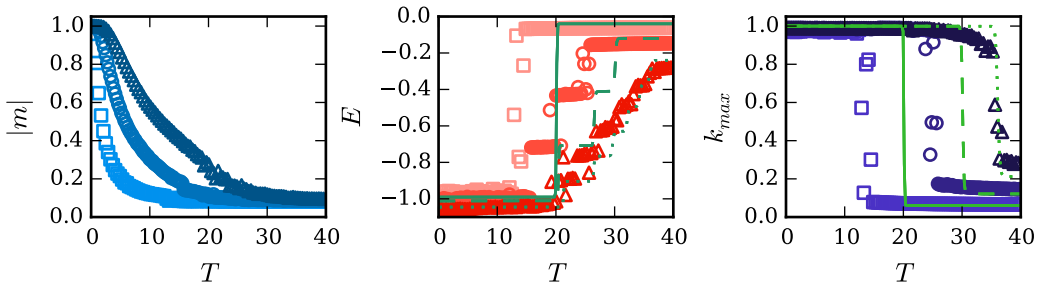


Figure 5.7: Absolute magnetization $|m|$, energy E, and the largest degree k_{max} as a function of temperature T, for $\gamma = 2$, $N = 100$, averaged over 10^5 MC time steps. Number of edges M is indicated by marker and color: $M = 100$ squares (brightest color), $M = 300$ circles, $M = 600$ triangles (darkest color). Analytical approximation for the energy is given by Equation 5.29 and for the largest degree by Equation 5.31. Solid line represents $M = 100$, dashed $M = 300$, dotted $M = 600$.

in the largest degree always overlaps with the last jump of the energy. In general, connectivity has a significant influence on the model's outcome.

Having the empirical knowledge, we can construct analytical approximation. Starting from the simplest task, we shall approximate the energy level for high temperatures. The magnetization is zero for $T \to \infty$, hence the first sum in the Hamiltonian 5.19 does not provide any contribution to the energy. There is the same number of connections between different spins as between identical ones (it can be also seen looking at ρ in Figure 5.5). Therefore, one must find the contribution from the second sum only. In the most naive mean-field approach we can assume that every node has the same degree equal to the average degree in the network $\langle k \rangle = 2M/N$. In such case we immediately find the formula for the energy:

$$E = N\langle k \rangle^{\gamma}. \tag{5.20}$$

This result is represented by the blue dashed line in Figure 5.5. It slightly overestimates the energy level in $T \to \infty$. To improve it we should take into account the real degree distribution $P(k)$, which for random graphs is binomial or Poissonian in the limit of big network. The number of nodes with a degree k is equal $NP(k)$, and after summing over all possible degrees we obtain:

$$E = -\sum_{k=0}^{N-1} NP(k)k^{\gamma} = -\sum_{k=0}^{N-1} N\frac{\langle k \rangle^{k} e^{-\langle k \rangle}}{k!}k^{\gamma}, \tag{5.21}$$

which is represented by the blue solid line in Figure 5.5. It slightly better predicts the energy level for high temperatures, however the main challenge is to explain the behavior of the system for low temperatures.

The next step is to compute the partition function Z (or approximate it), because then we would obtain the probability distribution of the ensemble and we could calculate average value of any quantity of interest. For instance, energy can be calculated from the relation 5.16. Analyzing the Hamiltonian 5.19 we can conclude that the biggest contribution to it's value comes from the hubs placed in the center of stars and having degree equal $k_m = N - 1$. Neglecting other contributions we can approximate the Hamiltonian by:

$$H \approx -n_h k_m^{\gamma}, \tag{5.22}$$

where n_h is the number of hubs with the maximum possible degree k_m. Note that $n_h \leq \lfloor c \rfloor = \lfloor M/N \rfloor$. In order to find the partition function, we must calculate all possible configurations corresponding to a given energy value. However, the energy level is fully determined by the number of stars n_h in this approach, so we can find all possible configurations for every given value of n_h. The spin configuration is easy to get – assuming it doesn't depend on the number of hubs we just have 2^N different possible configurations. The nodes to be placed in the center of stars can be chosen in $\binom{N}{n_h}$ different ways. Every hub has a determined degree k_m, being connected to every other node. Therefore, the last element differentiating between configurations in the distribution of remaining edges between the rest of the nodes.

There is $N - n_h$ of them. To calculate the remaining links (those not connected to hubs) it's sufficient to notice that the first hub takes k_m edges, the second one $k_m - 1$, since one edge is shared with the first hub, and so on. Finally, n_h hubs will take $L(n_h)$ edges:

$$\begin{aligned} L(n_h) &= k_m + k_m - 1 + ... + k_m - (n_h - 1) = n_h \frac{k_m + k_m - (n_h - 1)}{2} \\ &= n_h \frac{2k_m + 1 - n_h}{2} = n_h \frac{2N - 1 - n_h}{2}. \end{aligned} \tag{5.23}$$

Therefore, there is $M - L(n_h)$ free edges to be placed among $N - n_h$ vertices. The number of pairs among those vertices, which is the number of possible placements for the edges, is equal $(N - n_h)(N - n_h - 1)/2$. Consequently, there are $R(n_h)$ possible ways of placing remaining edges:

$$R(n_h) = \binom{(N - n_h)(N - n_h - 1)/2}{M - L(n_h)}. \tag{5.24}$$

With the exception of $n_h = \lfloor c \rfloor$, this approximation overestimates the number of configurations, due to including a possibility of existence of another hub except those related to n_h stars [45]. However, there is much fewer such heterogeneous configurations than the random disordered ones, and therefore the overestimation is negligible. Collecting all above factors together, the Hamiltonian 5.22 leads to a following partition function:

$$Z = 2^N \sum_{n_h=0}^{\lfloor c \rfloor} \binom{N}{n_h} R(n_h) e^{\beta n_h k_m^\gamma}, \tag{5.25}$$

and using Equation 5.16 we immediately obtain a formula for the average energy of the system:

$$\begin{aligned} E &= -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{2^N}{Z} \sum_{n_h=0}^{\lfloor c \rfloor} \binom{N}{n_h} A(n_h) n_h k_m^\gamma e^{\beta n_h k_m^\gamma} \\ &= -\frac{2^N}{Z} (N-1)^\gamma \sum_{n_h=0}^{\lfloor c \rfloor} \binom{N}{n_h} A(n_h) n_h e^{\beta n_h (N-1)^\gamma}. \end{aligned} \tag{5.26}$$

The energy given by 5.26 is presented in Figure 5.5 with a black dashed line. It captures stairs-like behavior of the energy, but predicts lower than observed absolute values of the energy for each level. Discrepancy increases with temperature and for $T \gtrsim 35$, when there are no hubs, the absolute level of E is significantly underestimated. This is because in that regime there is no contribution into the Hamiltonian 5.22, as it doesn't take into account nodes with a degree smaller than $N - 1$.

To improve the analytical approximation we should include the contribution from nodes with a small degree as well. A simple way to achieve it is by assuming that all nodes except the hubs have the same degree equal k_a and taking into account their interaction with the topological field. In this way

we obtain a new Hamiltonian:

$$H \approx -n_h k_m^\gamma - (N - n_h)k_a^\gamma, \tag{5.27}$$

where k_a depends on the number of hubs and $k_a \equiv k_a(n_h) = n_h + \frac{2(M-L(n_h))}{N-n_h}$. This dependence comes from the fact that every node is connected to every star and the free edges contribute to the degree of two nodes each. Possible configurations are the same as in the previous approach, hence we only need to incorporate the new Hamiltonian 5.27. Doing so we obtain a partition function:

$$Z = 2^N \sum_{n_h=0}^{\lfloor c \rfloor} \binom{N}{n_h} R(n_h) e^{\beta[n_h k_m^\gamma + (N-n_h)k_a^\gamma]}, \tag{5.28}$$

and using Equation 5.16 once again we arrive at a formula for the energy:

$$\begin{aligned}
E &= -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{2^N}{Z}\sum_{n_h=0}^{\lfloor c \rfloor}\binom{N}{n_h} R(n_h)\left[n_h k_m^\gamma + (N-n_h)k_a^\gamma\right] e^{\beta[n_h k_m^\gamma + (N-n_h)k_a^\gamma]} \\
&= -\frac{2^N}{Z}\sum_{n_h=0}^{\lfloor c \rfloor}\binom{N}{n_h} R(n_h)\left[n_h (N-1)^\gamma + (N-n_h)k_a^\gamma\right] e^{\beta[n_h (N-1)^\gamma + (N-n_h)k_a^\gamma]} \\
&= -\frac{2^N}{Z}\left[(N-1)^\gamma \sum_{n_h=0}^{\lfloor c \rfloor}\binom{N}{n_h} R(n_h) n_h e^{\beta[n_h (N-1)^\gamma + (N-n_h)k_a^\gamma]}\right. \\
&\quad \left. + \sum_{n_h=0}^{\lfloor c \rfloor}\binom{N}{n_h} R(n_h)(N-n_h)k_a^\gamma e^{\beta[n_h (N-1)^\gamma + (N-n_h)k_a^\gamma]}\right].
\end{aligned} \tag{5.29}$$

This result is plotted with a black solid line in Figure 5.5. It now not only captures the jumps of the energy, but also properly predicts energy levels for all temperatures, with a small shift in T in relation to the simulation results. From Figure 5.7 we can see that the approximation works for different values of the average degree and shift becomes smaller with growing $\langle k \rangle$.

It is possible to approximate behavior of the largest degree in the same manner. When $n_h > 0$ there is at least one star, thus the maximal degree is equal $N - 1$. When $n_h = 0$ there is no star. The simulation suggests that the structure of network in this regime is random. Therefore, we can approximate the degree distribution to be Poissonian. Dependence of the maximal degree $\bar{k}(n_h)$ in the network on the number of stars can be summarized by:

$$\bar{k}(n_h) = \begin{cases} N-1 & \text{for } n_h > 0 \\ F^{-1}_{Poiss,\lambda}(1-\frac{1}{N}) & \text{for } n_h = 0 \end{cases}, \tag{5.30}$$

where $F^{-1}_{Poiss,\lambda}$ is an inverse of the Poisson cumulative distribution function with a parameter $\lambda = \langle k \rangle$. Together with the partition function 5.28 it yields

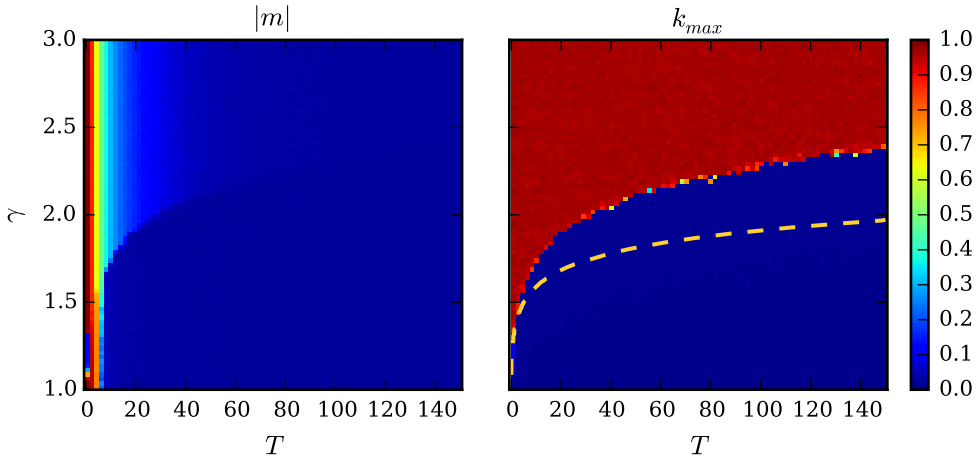


Figure 5.8: Absolute magnetization $|m|$ and the largest degree k_{max} as a function of temperature T and parameter γ, for $N = 1000$ and $M = 3000$, averaged over $5 \cdot 10^5$ MC time steps. Analytical approximation for the phase transition comes from Equation 5.31 (orange dashed line). Phase transitions can be observed in both order parameters.

a formula for the average maximum degree:

$$\begin{aligned}\langle k_{\max}\rangle &= \frac{2^N}{Z}\sum_{n_h=0}^{\lfloor c\rfloor}\binom{N}{n_h}R(n_h)\bar{k}(n_h)e^{\beta[n_h k_m^\gamma+(N-n_h)k_a^\gamma]}\\ &= \frac{2^N}{Z}\Bigg\{R(0)F^{-1}_{Poiss,\lambda}\left(1-\frac{1}{N}\right)e^{\beta N(2M/N)^\gamma}\\ &\quad+(N-1)\sum_{n_h=1}^{\lfloor c\rfloor}\binom{N}{n_h}R(n_h)e^{\beta[n_h(N-1)^\gamma+(N-n_h)k_a^\gamma]}\Bigg\}.\end{aligned} \tag{5.31}$$

This approximation was presented in Figures 5.5 and 5.7. It properly captures the discontinuous behavior of the largest degree. The approximation becomes more accurate with increasing density of the network.

Finally, we can ask a question how does the system behave under different values of γ. This aspect is studied numerically and results are presented in Figure 5.8. In this two-dimensional phase diagram we can distinguish four phases. They are described by: (i) full magnetization with star configurations, (ii) zero magnetization with star configurations, (iii) full magnetization with a random network, and (iv) zero magnetization with a random network. Note, that there are transitions of the magnetization without any influence from the topology, but also transitions driven by the change of the topology. It underlines the importance of the coevolution, i.e. the feedback loop between the topology and state of the nodes, which can not be obtained without presence of both in the Hamiltonian.

An important aspect of the model is it's scaling behavior with varying network sizes, since we want to know what to expect in large real-world

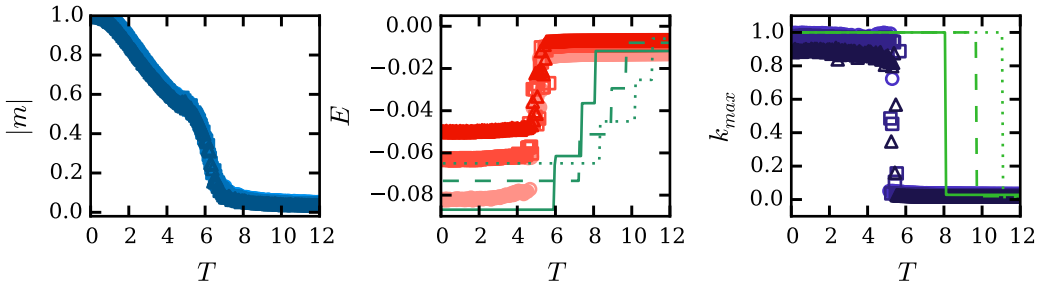


Figure 5.9: Absolute magnetization $|m|$, energy E, and the largest degree k_{max} as a function of temperature T, for $\gamma = 1.6$, $c = M/N = 3$, averaged over 10^6 MC time steps. Size of the networks N is indicated by marker and color: $N = 500$ circles (brightest color), $N = 750$ squares, $N = 1000$ triangles (darkest color). Analytical approximation for the energy is given by Equation 5.29 and for the largest degree by Equation 5.31. Solid line represents $N = 500$, dashed $N = 750$, dotted $N = 1000$.

systems. The scaling is presented in Figure 5.9. Interestingly, order parameters remain at the same positions indicating a discontinuous phase transition in k_{max} and a continuous phase transition in m in the thermodynamic limit. Unfortunately, the analytical approximation does not capture this behavior properly. It can be caused by either the approximation simply being too rough, or a hysteresis in the system. It is possible that starting simulations from a random configuration we end up on one side of the hysteresis and the analytical approximation describes the other side. It is a phenomena worth exploring in future extensions of the work.

5.4.2 Topological correlations and spins

In this section we will focus on the topological interactions between nodes fused with the interaction between the spins. After setting the topological field parameter to the neutral value $\gamma = 0$ (or $\gamma = 1$), and removing the magnetic field $h = 0$, the resulting Hamiltonian is:

$$H(\{a_{ij}\}, \{s_i\}) = -\sum_{i<j} a_{ij} \left(\frac{k_i k_j}{\alpha}\right)^{\phi} s_i s_j. \tag{5.32}$$

To be precise, for neutral values of γ in the Hamiltonian 5.17 we obtain additional constant. However, a constant value added to a Hamiltonian does not change the outcome of the model in any way (except shifting the energy values by the said constant), therefore it is omitted in 5.32 for the sake of simplicity.

In Figure 5.10 we can see the behavior of order parameters at different temperatures for the Hamiltonian 5.32 without normalization, i.e. $\alpha = 1$.

A phase transition can be observed in every one of them. Note that in absence of normalization the temperature range necessary to observe the transitions is of order of magnitude greater than for the topological field (Hamiltonian 5.19). Here, the magnetization has two transitions – from a stable value of approximately 0.3 for low temperatures it abruptly increases to 0.7, and then a continuous transition towards $|m| = 0$ follows. The abrupt change can be also observed in the energy, as well as n_S and S. A very interesting behavior is visible in the plot of the largest degree k_{max}. From a stable value around 0.25 it changes into small values corresponding to a random graph, however at the transition point it raises and takes the maximal value.

Configuration of the network at low temperatures can be deduced form Figure 5.10. From the small size of the largest component S and high number of components n_S we should expect a shattered network. The largest degree $k_{max} \approx 0.25$ indicates a relatively big main component and many very small ones. Magnetization is low and the density of active links is $\rho \approx 0$. It implies that the main component must be homogeneous containing spins of the same orientation. These predictions are confirmed when looking at snapshots of the network in the equilibrium state in Figure 5.11. There is one clique in the network possessing all the links. Its structure is close to a complete graph. With increasing temperature the clique grows obtaining more nodes and becoming more sparse to finally become a random graph at $T \to \infty$. The dense main component displays very strong homogeneity. Surprisingly, at the transition point when the network becomes fully connected the degree distribution becomes highly heterogeneous, as the value of the largest degree raises significantly. Behavior of the model is qualitatively the same for normalization $\alpha = \langle k \rangle$, although scaling with the number of edges is different.

Value of the average degree $\langle k \rangle$ has a very strong impact on the system's behavior. In Figure 5.12 we can see three main order parameters for different values of the connectivity. For the same number of edges as nodes magnetization doesn't raise above approximately 0.2. This is caused by a big number of lonely separated nodes, since the main component attracting all the links is very small. Separated nodes can not interact, therefore their spins can not align. For $\langle k \rangle = 12$, however, magnetization can reach almost the full ferromagnetic state. Transition in the largest degree is qualitatively similar across different connectivities, but the maximal obtained value grows with the number of edges. In other words, bigger hubs are generated for more dense networks. Notably, the transition point is being shifted towards higher and higher temperatures when increasing the average degree. To avoid this, the normalization in the Hamiltonian 5.32 is set to $\alpha = \langle k \rangle$. For a constant value of α it can be extracted from the sum in the Hamiltonian and effectively rescale the temperature in the equilibrium distribution.

Previous observations provide a starting point in developing an analytical description of the system. As previously, the method will be based on approximating the Hamiltonian 5.32, taking into account only the most important effects. For low temperatures we observe a shattered network with one component having all the links and many lonely nodes. These nodes,

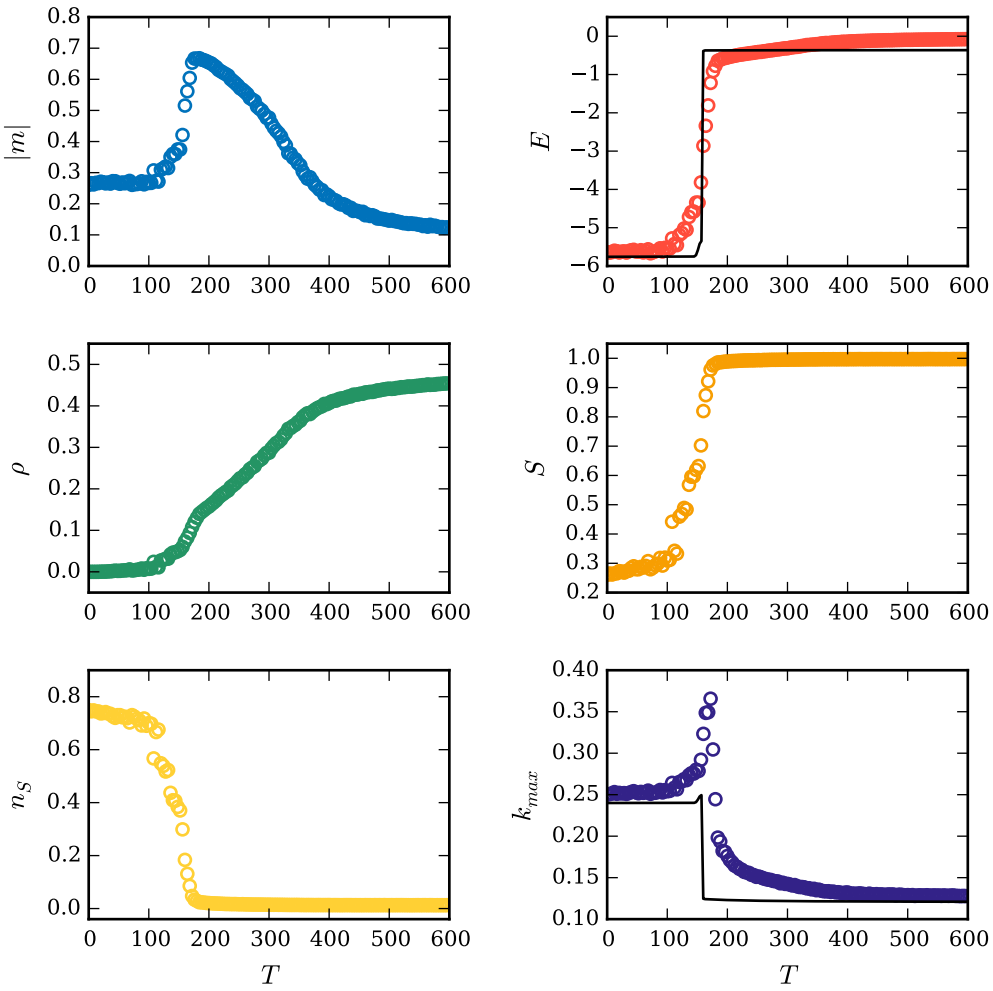
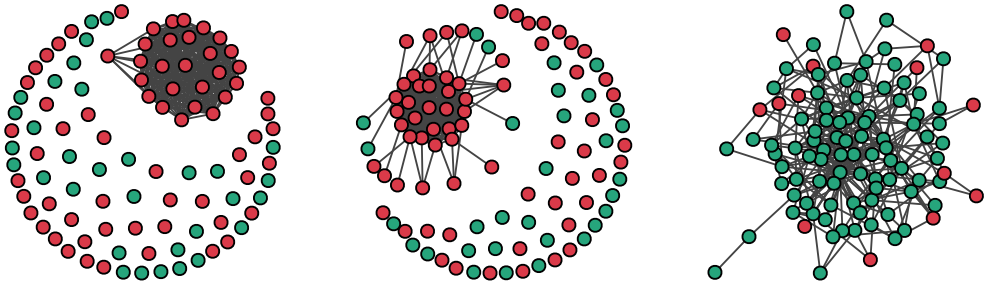


Figure 5.10: Absolute magnetization $|m|$, energy E, interface density ρ, the largest component S, number of components n_S, and the largest degree k_{max} as a function of temperature T for the Hamiltonian 5.32, for $\phi = 1$, $\alpha = 1$, $N = 100$, $M = 300$, averaged over 10^5 MC time steps. Analytical approximation for the energy is given by Equation 5.37 and for the largest degree comes from Equation 5.39 (solid black lines).



a) $T = 1$ b) $T = 120$ c) $T = 175$

Figure 5.11: Network configurations for different temperatures with $\phi = 1$, $\alpha = 1$, $N = 100$, $M = 300$. Red color indicates spin -1, green $+1$. Note that for low temperatures a very dense clique is formed, close to a complete subgraph.

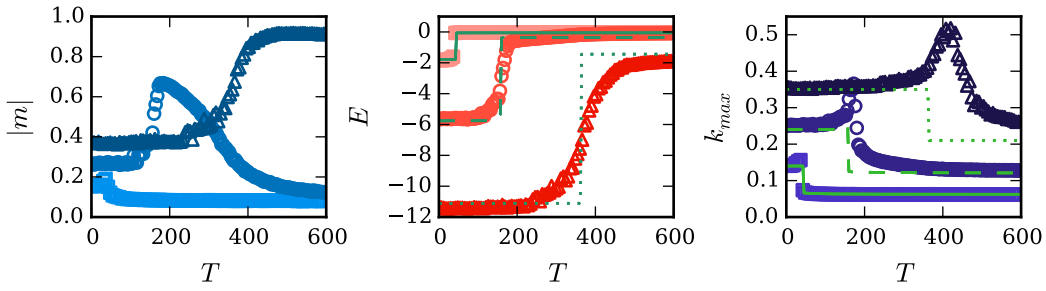


Figure 5.12: Absolute magnetization $|m|$, energy E, and the largest degree k_{max} as a function of temperature T, for $\phi = 1$, $\alpha = 1$, $N = 100$, averaged over 10^5 MC time steps. Number of edges M is indicated by marker and color: $M = 100$ squares (brightest color), $M = 300$ circles, $M = 600$ triangles (darkest color). Analytical approximation for the energy is given by Equation 5.37 and for the largest degree by Equation 5.39. Solid line represents $M = 100$, dashed $M = 300$, dotted $M = 600$.

having degree equal 0, do not provide any contribution to the energy. Therefore, we need to estimate the contribution from the component. In order to distinguish this component from the topological meaning of the word *component* we shall call it *the active component*. The active component is a set of nodes that can obtain links, in contrast to the shattered inactive ones. Assuming it has a structure of a random graph it can be approximated in a mean-field spirit, treating all nodes as if they had the same degree k_a. Then, the Hamiltonian 5.32 reduces to:

$$H \approx -M\frac{k_a^{2\phi}}{\alpha^{\phi}}. \tag{5.33}$$

The active component contains all edges, so there is M pairs of interacting nodes in it.

Let n_a denote the number of nodes the active component contains. Obviously, this number is restricted to $n_a \in \{\lceil n_{\min}\rceil, \lceil n_{\min}\rceil + 1, \ldots, N\}$, where $n_{\min} = (1+\sqrt{1+8M})/2$. The minimal value $n_{\min}$ comes from the fact that all connections must fit in the active component and the limit of compressing them is a complete graph with $n_a(n_a-1)/2$ connections. The average degree in the active component is related to the number of nodes as $k_a = 2M/n_a$, what allows to rewrite the Hamiltonian 5.33 in a different form:

$$H \approx -2^{2\phi}M^{2\phi+1}\alpha^{-\phi}n_a^{-2\phi}, \tag{5.34}$$

what for $\alpha = \langle k\rangle$ would result in $H \approx -2^{\phi}M^{\phi+1}N^{\phi}n_a^{-2\phi}$. The energy of the system is determined by the number of nodes in the active component n_a. To calculate the partition function we also need the number of topological configurations that can be realized for a given size of the active component. Links within the active component can be arranged in $R(n_a)$ ways:

$$R(n_a) = \binom{\frac{n_a(n_a-1)}{2}}{M}. \tag{5.35}$$

Additionally, we need to take into account all possible assignments of nodes into the active component, what can be done in $\binom{N}{n_a}$ ways. We also assume that the active component is homogeneous inside having only two possible configuration of spins – all spins positive or all negative. The rest of the spins can be arranged in 2^{N-n_a} ways. Putting all these parts together we obtain the partition function:

$$Z = \sum_{n_a=\lceil n_{\min}\rceil}^{N} 2^{N-n_a+1}\binom{N}{n_a}R(n_a)e^{\beta 2^{2\phi}M^{2\phi+1}\alpha^{-\phi}n_a^{-2\phi}}. \tag{5.36}$$

From Equation 5.16 we calculate the energy:

$$E = -\frac{1}{Z} \sum_{n_a=\lceil n_{\min} \rceil}^{N} 2^{N-n_a+1+2\phi} M^{2\phi+1} \alpha^{-\phi} n_a^{-2\phi} \binom{N}{n_a} \binom{\frac{n_a(n_a-1)}{2}}{M} e^{\beta 2^{2\phi} M^{2\phi+1} \alpha^{-\phi} n_a^{-2\phi}}. \tag{5.37}$$

The maximum degree k_{max} in the network is the maximum degree in the active component. We assumed the active component to be a random graph, so we can approximate its maximum degree in the same way as in Section 5.4.1 for high temperatures, using the inverse of the Poisson cumulative distribution function $F^{-1}_{Poiss,\lambda}$, where the parameter $\lambda = k_a$. The maximal degree of the active component is obviously limited by $n_a - 1$, hence its full formula is:

$$\bar{k}(n_a) = \min \left\{ n_a - 1,\ F^{-1}_{Poiss,\lambda} \left(1 - \frac{1}{n_a} \right) \right\}. \tag{5.38}$$

Using this equation and the partition function 5.36 we obtain the average maximal degree in the model:

$$\langle k_{\max} \rangle = \frac{1}{Z} \sum_{n_a=\lceil n_{\min} \rceil}^{N} \bar{k}(n_a)\, 2^{N-n_a+1} \binom{N}{n_a} \binom{\frac{n_a(n_a-1)}{2}}{M} e^{\beta 2^{2\phi} M^{2\phi+1} \alpha^{-\phi} n_a^{-2\phi}}. \tag{5.39}$$

Analytical approximation for the energy (Equation 5.37) and the largest degree (Equation 5.39) are compared with the simulation results in Figure 5.10. As we can see, it predicates the transition point very well. In the approximation of k_{max} we can even observe a small peak at the transition, as in the numerical results. However, its size is much smaller. Values of the maximal degree and energy for high and low temperatures are also well described by the analytical approximation. Only at the transition the discrepancy becomes significant. From Figure 5.12 we see that the approximation works better for sparse networks.

The next step in the analysis of the model defined by the Hamiltonian 5.32 is studying its behavior for different values of the parameter ϕ. Numerically it is done in Figure 5.13, where a two-dimensional phase diagram is presented. We can distinguish four phases: (i) with a small and dense active component and relatively small magnetization, (ii) with a random network structure and high magnetization, (iii) with a random network structure and zero magnetization, and (iv) with full magnetization and $k_{max} \approx 1$. The last phase was previously undetected, because it does not exist for $\phi = 1$. From the value of the largest degree in the network we can deduce that in the phase (iv) we obtain star configurations as in the case of the external topological field γ. Interestingly, the critical value ϕ_c above which the star configuration is preferred over the clique does not depend on the temperature.

In order to find an analytical formula for ϕ_c we shall approximate the energy of the system below and above it. The energy of a configuration corresponds to the probability of realizing it, therefore at the critical line both energies should be equal. For $\phi > \phi_c$ the energy is dominated by the hubs of a degree equal $N - 1$. As previously, let us denote the number of hubs by n_h.

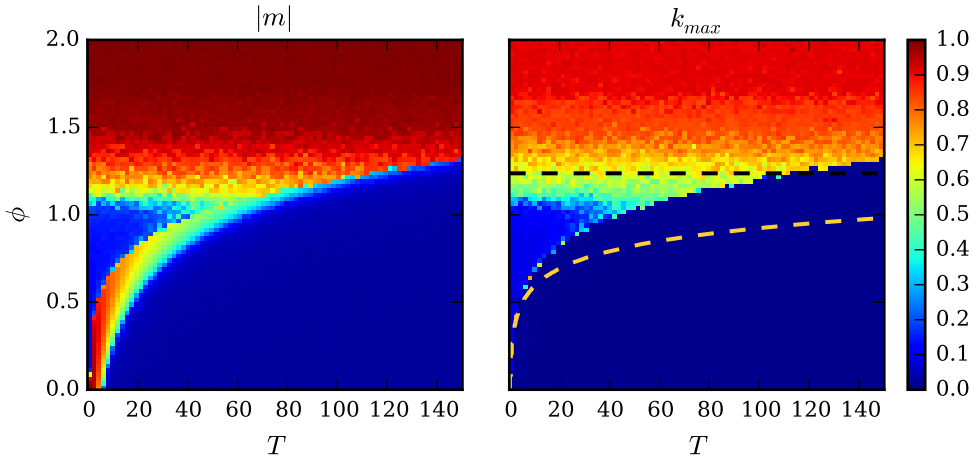


Figure 5.13: Absolute magnetization $|m|$ and the largest degree k_{max} as a function of temperature T and parameter ϕ, for $\alpha = \langle k \rangle$, $N = 1000$ and $M = 3000$, averaged over $5 \cdot 10^5$ MC time steps. Analytical approximation for the phase transition comes from Equation 5.39 (orange dashed line) and the critical value ϕ_c (Equation 5.42) is indicated by a black dashed line. Phase transitions can be observed in both order parameters.

In the first approximation we can take into account interactions between the hubs (there is $n_h(n_h-1)/2$ such interactions) and interactions between the hubs and the rest of the nodes (there is $n_h(N-n_h)$ of them). We can also assume that the only links that non-star nodes have are those connecting them to the hubs. In other words, the degree of all nodes except hubs is equal n_h. This assumption is precise for number of links being multiplication of the number of nodes. Since the magnetization is equal 1 for $\phi > \phi_c$, every interaction increases absolute energy. Putting all of these considerations together we obtain the energy above the critical point:

$$\begin{aligned} E_{\phi>\phi_c} &= -\sum_{i=1}^{\frac{n_h(n_h-1)}{2}} \alpha^{-\phi}(N-1)^{2\phi} - \sum_{j=1}^{n_h(N-n_h)} \alpha^{-\phi}(N-1)^{\phi} n_h^{\phi} \\ &= -\frac{n_h(n_h-1)}{2}\alpha^{-\phi}(N-1)^{2\phi} - n_h(N-n_h)\alpha^{-\phi}(N-1)^{\phi} n_h^{\phi} \\ &= -n_h(N-1)^{\phi}\alpha^{-\phi}\left[\frac{n_h-1}{2}(N-1)^{\phi} + n_h^{\phi}(N-n_h)\right], \end{aligned} \tag{5.40}$$

where i iterates over the links between hubs and j over the links between the hubs and the rest of the nodes. When $T \to 0$ we have $n_h = \lfloor M/N \rfloor = \lfloor c \rfloor$. Using a continuous approximation we can assume that $n_h \approx c$.

For $\phi < \phi_c$ and low temperatures we have the clique configuration described before. The clique contains n_a nodes. We can assume that all nodes in the cluster are connected and of the same spin, therefore every one of them has degree $n_a - 1$. Disconnected nodes do not provide any contribution to the

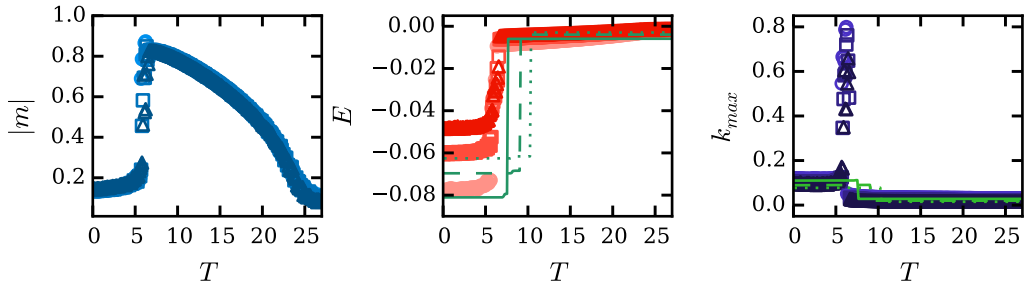


Figure 5.14: Absolute magnetization $|m|$, energy E, and the largest degree k_{max} as a function of temperature T, for $\phi = 0.6$, $\alpha = \langle k \rangle$, $c = M/N = 3$, averaged over 10^6 MC time steps. Size of the networks N is indicated by marker and color: $N = 500$ circles (brightest color), $N = 750$ squares, $N = 1000$ triangles (darkest color). Analytical approximation for the energy is given by Equation 5.37 and for the largest degree by Equation 5.39. Solid line represents $N = 500$, dashed $N = 750$, dotted $N = 1000$.

energy, because they have degree equal zero. Consequently, we obtain:

$$E_{\phi<\phi_c} = -\sum_{i=1}^{\frac{n_a(n_a-1)}{2}} \alpha^{-\phi}(n_a-1)^{2\phi} = -\frac{n_a(n_a-1)}{2}\alpha^{-\phi}(n_a-1)^{2\phi}. \tag{5.41}$$

where i iterates over all connections in the active component. For $T \to 0$ we have $n_a = \lceil n_{\min} \rceil = \lceil (1+\sqrt{1+8M})/2 \rceil$. Again, we can use a continuous approximation and assume $n_a \approx n_{\min} = (1+\sqrt{1+8M})/2$, or equivalently $M \approx n_a(n_a-1)/2$. At the critical line between the clique phase and the star phase probabilities of realizing them should be equal, hence energies as well. Assuming that $E_{\phi<\phi_c} = E_{\phi>\phi_c}$ we obtain:

$$\begin{aligned}
M\alpha^{-\phi_c}(n_{min}-1)^{2\phi_c} &= c(N-1)^{\phi_c}\alpha^{-\phi_c}\left[\frac{c-1}{2}(N-1)^{\phi_c} + c^{\phi_c}(N-c)\right], \\
\frac{M}{c}\left(\frac{(n_{min}-1)^2}{N-1}\right)^{\phi_c} &= \frac{c-1}{2}(N-1)^{\phi_c} + c^{\phi_c}(N-c), \\
\phi_c \ln\frac{(n_{min}-1)^2}{N-1} + \ln\frac{M}{c} &= \ln\left(\frac{c-1}{2}(N-1)^{\phi_c} + c^{\phi_c}(N-c)\right).
\end{aligned} \tag{5.42}$$

From this equation we can compute the value of ϕ_c for a given combination of parameters. As we can see in Figure 5.13 the approximation works very well.

Last element necessary for a comprehensive study of the model is scaling analysis. It is presented in Figure 5.14. Both transitions in the magnetization $|m|$, the discontinuous raise and continuous decrease, are present at the same point for different network sizes. The transition in the largest degree k_{max}

also remains at the same position with the maximal value obtained always at the transition point. Therefore, in the thermodynamic limit we should expect those transitions to exist. Analytical approximation scales differently – transitions are being shifted towards higher temperatures with growing network. However, the discrepancy is smaller than in the case of the topological field γ.

As we could see, statistical mechanics approach to coevolving networks may lead to intriguing results and generate complex structures. On the abstract level, we can observe an interplay between the topology of the network and state of the nodes resulting in different transitions. The crucial observation is that those transitions can influence each other, i.e. a transition in magnetization can be triggered by a topological transition. It is yet another argument for using coevolving network models in order to grasp the nuanced phenomena coming from the feedback loop. On the practical level, the most important result is emergence of structures similar to the empirical ones and to those generated in algorithmic models. In statistical mechanics of coevolving networks we observed big hubs being created, as it happens in many real-world networks. We obtained a shattered network phase very similar to the shattered phase from the nonlinear coevolving voter model (Section 3.3) or the coevolving Axelrod model (Section 4.2). There was emergence of consensus (high magnetization) in the system. These results confirm previously argued importance of equilibrium models and their analysis.

Chapter 6

Summary

6.1 Review of obtained results

Many complex systems can be naturally described using network science due to their inherent discrete nature. This approach proved successful regardless of the system's origin – it can be physical, biological, socio-technical, economical, financial, ecological... In most of the empirical systems coevolution of the network's topology and states of the nodes is a crucial aspect. Therefore, coevolving network models are of a broad interest. The aim of this thesis was to explore them.

Complex systems are also large. They are made of many separate elements interacting which each other. From the microscopic dynamics the macroscopic patterns emerge. This connection between the micro and macro scale is long known in statistical physics. Physicists have developed mathematical and numerical methods for studying it. Hence, it is a natural step to merge both approaches from network science and statistical physics in description of complex systems. This is the origin of *statistical physics of coevolving networks*. The goal was to provide a relatively broad coverage of coevolving network models with emphasis on development of three particular ones.

Chapter 1 of the thesis is a conceptual introduction to the further content. It discusses connections between physics and network science, provides empirical examples of networks, and describes the foundations of the theory on the abstract level. At the end of the chapter an outline of the work is provided.

Chapter 2 contains a proper introduction to network science. In Section 2.1 the basics of graph theory are presented – from mathematical tools used in the description of networks to different measures and types of graphs. In Section 2.2 the fundamental network models are described, together with a reconstruction of the most important analytical and computational results. These basics of network science are put in a context of the state-of-the-art research and assisted by a range of references. Finally, in Section 2.3 the mechanism of coevolution is described and an overview of coevolving network models is given.

In Chapter 3 the voter model is introduced. Different interpretations and the original definition of the model are discussed. In Section 3.1 the voter model on static networks is covered. Its early version on a complete graph

is characterized. The behavior in this example can be well explained using Fokker-Planck equation, as presented in the section. The picture is completed in Appendix B, where the stationary solution of the Fokker-Planck equation is derived. The extension to complex networks is also depicted. In Section 3.2 the coevolving voter model with its basic properties is described. Section 3.3 provides further extensions taking into account a possible non-linearity of interactions and the triadic closure rewiring. It contains original results obtained in this thesis. The last Section 3.4 extends the previously proposed nonlinear coevolving voter model by including noisy state updates. This is one of the most general versions of the voter model taking into account three important aspects of real-world networks – coevolution, non-linearity, and noise. Together with previously analyzed triadic closure, Chapter 3 provides a comprehensive study of the voter model and its extensions. The main results obtained by myself in this context are:

- discovery of a new phase in the nonlinear CVM with triadic closure – a shattered phased not observed before,
- characterization of phases and convergence times in the nonlinear CVM with triadic closure,
- achievement of high clustering values in the nonlinear CVM,
- development of a new and general version of the voter model – the nonlinear CVM with noise,
- discovery of two different types of the consensus phase in the nonlinear CVM with noise,
- discovery of internal differences in the coexistence phase of the nonlinear CVM with noise,
- discovery of topological community structures driven by the state of the nodes,
- derivation of the analytical approximation of the nonlinear CVM with noise,
- derivation of the analytical approximation of the behavior of the dynamical fragmentation phase in the nonlinear CVM with noise.

Chapter 4 treats about another important algorithmic model – the Axelrod model. The historical development of the model is briefly covered in Section 4.1, as well as its motivation and interpretation. Basic simulation results are reproduced for the original version of the model and selected extensions. Section 4.2 introduces the coevolving Axelrod model, first with random rewiring, then with preferential attachment and triadic closure. This section contains new results obtained by me. A possibility of generating complex structures and real-world networks features in the Axelrod model is discussed. In Section 4.3 a long standing issue of the Axelrod model is reviewed and solved. Firstly, an analysis of empirical data is performed. Then,

modifications of the model incorporating empirical results are proposed. The original results obtained by myself in this chapter are:

- development of the coevolving Axelrod model incorporating real-world dynamics, i.e. preferential attachment and triadic closure,
- achievement of high clustering values in the coevolving Axelrod model,
- achievement of a power-law degree distribution in the coevolving Axelrod model,
- achievement of the small-world effect in the coevolving Axelrod model,
- characterization of phases in the coevolving Axelrod model and discovery of a new characteristic of the second transition point,
- collection and analysis of empirical data for validation of the model,
- modification of the model solving a long standing contradiction with the empirical data.

In Chapter 5 the equilibrium description of coevolving networks is introduced. Section 5.1 gives a general description of the maximal entropy principle and its applications. Appendix A supplements this description providing a step-by-step entropy maximization method via Lagrange multipliers. In Section 5.2 one of the most famous applications of said approach is covered, i.e. the Ising model. Basic results are reproduced and selected extensions are discussed. Section 5.3 covers the same methodology applied to the network's topology, instead of state of the nodes. Finally, in Section 5.4 the original model developed in this thesis is introduced. The model provides equilibrium description of coevolving networks with a Hamiltonian depending on both: state of the nodes and their topological traits. It's the first model of this kind to my best knowledge. The model is explored by means of extensive numerical simulations. Appendix C provides a detailed description of the algorithm used for simulations. Additionally, precise analytical approximations of observed phenomena are derived. The main original achievements obtained in this chapter are:

- application of the statistical mechanics approach in description of coevolving networks,
- development of an equilibrium model of coevolving networks incorporating state of the nodes and topology of the network in the Hamiltonian,
- achievement of complex structures, often observed in algorithmic models, in the equilibrium model,
- characterization of the phase diagrams of the model,
- derivation of the analytical description of the system's behavior.

For every model analyzed in this thesis I reproduce the most important historical results, usually together with several extensions. Motivation, applications, interpretation, and meaning of every model are broadly discussed. All descriptions are supported by an extensive list of references to the essential articles and books in the field. The thesis contains also a comprehensive introduction to the field for non-specialists.

6.2 Conclusions

Based on the review of obtained results, the main goal of the thesis has been achieved. Effects of coevolution in network models were thoroughly studied in algorithmic and analytical approach. The influence of coevolution on two important non-equilibrium models was examined and a new equilibrium model of coevolving networks was proposed. By including effects inspired by real-world networks in the models, the thesis brings the network analysis closer to the empirical observation. It provides a binding between the macroscopic dynamics and the macroscopic effects observed in networks.

When creating a model defined by microscopic rules it is important to relate these rules to the empirical observation, but also to generate a macroscopic outcome consistent with the observation as well. In a sense, both scales should be associated with the real-world phenomena. This gives a strong argument for the relevance of a new phenomena observed in the model, suggesting it can be realized as well empirically. The models studied in the thesis were developed in that spirit. The microscopic effects like the triadic closure or preferential attachment have a strong justification in numerous fields where network science is applied. Additionally, the macroscopic patterns like high clustering, power-law degree distribution, or the small-world effect are obtained.

Coevolving network models are still a young branch of science, gaining more interest recently. The assumption justifying coevolution is simple – in real-world networks it is usually impossible to separate time scales of the evolution of states the nodes and of the topology. These two aspects of every network *coevolve* together creating a feed back loop, which may and mostly does influence the outcome of the model in a complex way. In order to fully describe complex systems, inclusion of coevolution in network models is inevitable. For that reason it is crucial to determine the essential characteristics of coevolving network models. This thesis adds a new brick in the foundations of a general theory of coevolving networks.

Appendix A

Entropy maximization via Lagrange multipliers

The task is to maximize the Boltzmann-Gibbs-Shannon entropy:

$$S(\{p_i\}) = -\sum_{i=1}^{n} p_i \ln p_i \tag{A.1}$$

by finding the optimal probability distribution $p(x_i) = p_i$ of a random variable $x \in \{x_1, x_2, ..., x_n\}$ with n possible realizations. We will restrict the analysis to the discrete case, as it is relevant to the content of the thesis. Obviously, the distribution must be normalized:

$$\sum_{i=1}^{n} p_i = 1. \tag{A.2}$$

Additionally, we want to impose m constraints of the form:

$$\sum_{i=1}^{n} f_j(x_i) p_i = \langle f_j(x) \rangle, \tag{A.3}$$

where the right-hand side is a fixed value and $j = 1, 2, ..., m$. These constraints can be understood as the available information about the system. Together with the normalization it gives $m + 1$ constraints.

A standard method of solving an optimization problem of this form is by using Lagrange multipliers. Let us define a function $\mathcal{L}$ as:

$$\mathcal{L}(\{p_i\}, \mu, \lambda_1, ..., \lambda_m) = S(\{p_i\}) - \mu h(\{p_i\}) - \sum_{j=1}^{m} \lambda_j g_j(\{p_i\}), \tag{A.4}$$

where μ and λ_j are *Lagrange multipliers*, $h(\{p_i\})$ is a function representing normalization and $g_j(\{p_i\})$ are functions representing the rest of constraints:

$$\begin{aligned} h(\{p_i\}) &= \sum_{i=1}^{n} p_i - 1, \\ g_j(\{p_i\}) &= \sum_{i=1}^{n} f_j(x_i) p_i - \langle f_j(x) \rangle. \end{aligned} \tag{A.5}$$

Therefore all constraints can be represented by a condition:

$$h(\{p_i\}) = g_j(\{p_i\}) = 0. \tag{A.6}$$

Finally, it can be shown that a set of n equations:

$$\frac{\partial \mathcal{L}}{\partial p_i} = 0 \tag{A.7}$$

provides a necessary condition for the maximum of $S(\{p_i\})$ under constraints A.2 and A.3 [349]. Using A.1, A.4 and A.5 the equations A.7 give us:

$$-\ln(p_i) - 1 - \mu - \sum_{j=1}^{m} \lambda_j f_j(x_i) = 0, \tag{A.8}$$

which yields a solution:

$$p_i = \mathrm{e}^{-1-\mu-\sum_{j=1}^{m} \lambda_j f_j(x_i)}. \tag{A.9}$$

The constants μ and λ_j are recovered from the constraints. From A.2 we have:

$$1 = \sum_{i=1}^{n} \mathrm{e}^{-1-\mu-\sum_{j=1}^{m} \lambda_j f_j(x_i)} = \mathrm{e}^{-1-\mu} \sum_{i=1}^{n} \mathrm{e}^{-\sum_{j=1}^{m} \lambda_j f_j(x_i)}, \tag{A.10}$$

therefore:

$$\mathrm{e}^{1+\mu} = \sum_{i=1}^{n} \mathrm{e}^{-\sum_{j=1}^{m} \lambda_j f_j(x_i)} \equiv Z, \tag{A.11}$$

where Z is called *partition function* [312]. We can now rewrite the equation for the probability distribution as:

$$p_i = \frac{1}{Z} \mathrm{e}^{-\sum_{j=1}^{m} \lambda_j f_j(x_i)}. \tag{A.12}$$

From the remaining constraints A.3 we obtain m equations for m constants λ_j:

$$\frac{\sum_{i=1}^{n} f_j(x_i) \mathrm{e}^{-\sum_{j=1}^{m} \lambda_j f_j(x_i)}}{\sum_{i=1}^{n} \mathrm{e}^{-\sum_{j=1}^{m} \lambda_j f_j(x_i)}} = \langle f_j(x) \rangle. \tag{A.13}$$

Note, that it can be rewritten as:

$$\langle f_j(x) \rangle = \frac{\partial \ln Z}{\partial \lambda_j}, \tag{A.14}$$

which is often a particularly useful form. Finally, one can easily find that the maximal value of entropy reads:

$$S_{max} = 1 + \mu + \sum_{j=1}^{m} \lambda_j \langle f_j(x) \rangle. \tag{A.15}$$

If we impose a constraint on the average value of a function $\langle H(x) \rangle$ and denote the related Lagrange multiplier as β we obtain:

$$p_i = \frac{1}{Z} \, \mathrm{e}^{-\beta H(x_i)}. \tag{A.16}$$

In this way we recover Boltzmann's statistics of a canonical ensemble, where the function $H(x)$ represents energy of the system and is called Hamiltonian, while the constant β is related to the inverse temperature.

Appendix B

Stationary solution of the Fokker-Planck equation

The Fokker-Planck equation, also known as Smoluchowski equation in applications to Brownian motion, is a special type of master equation for the evolution of probability distribution density $P(x,t)$ of a random variable x in time t [177]. It's general form in one dimension is:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\Big[\mu(x)P(x,t)\Big] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big[D(x)P(x,t)\Big], \tag{B.1}$$

where $\mu(x)$ and $D(x)$ are real differentiable functions with a restriction of $D(x) > 0$. Function $\mu(x)$ is usually referred to as drift coefficient and $D(m)$ as diffusion coefficient. The equation can be written in a from of continuity equation for probability density:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial j(x,t)}{\partial x}, \tag{B.2}$$

where $j(x,t)$ is the probability current (or flux). It is therefore obvious that the probability current must be given by:

$$j(x,t) = \mu(x)P(x,t) - \frac{1}{2}\frac{\partial}{\partial x}\Big[D(x)P(x,t)\Big]. \tag{B.3}$$

Condition for the stationary solution of the Fokker-Planck equation is disappearance of the probability current $j(x,t) = 0$, what leads to the equation for the stationary probability distribution $P_{st}(x)$:

$$\mu(x)P_{st}(x) = \frac{1}{2}\frac{\partial}{\partial x}\Big[D(x)P_{st}(x)\Big]. \tag{B.4}$$

Dividing both sides by $D(x)P_{st}(x)$ we get:

$$\frac{\mu(x)}{D(x)} = \frac{1}{2D(x)P_{st}(x)}\,\frac{\partial}{\partial x}\Big[D(x)P_{st}(x)\Big] = \frac{1}{2}\frac{\partial}{\partial x}\ln\Big[D(x)P_{st}(x)\Big], \tag{B.5}$$

and further by integrating both sides in respect to x we obtain:

$$\ln\Big[D(x)P_{st}(x)\Big] = 2\int_{x_0}^{x}\frac{\mu(x')}{D(x')}\mathrm{d}x' + \text{const.}, \tag{B.6}$$

what after applying exponential function to both sides and simple rearrangement takes the well known form of the stationary solution:

$$P_{st}(x) = \frac{\text{const.}}{D(x)} \exp\left[2\int_{x_0}^{x} \frac{\mu(x')}{D(x')} \mathrm{d}x'\right]. \tag{B.7}$$

After substituting particular forms of $\mu(x)$ and $D(x)$ for a given problem and calculating the integral we obtain the analytical formula for the stationary solution $P_{st}(x)$ of the Fokker-Planck equation. The last requirement for the existence of the solution is that $P_{st}(x)$ is integrable so that it can be normalized to represent a probability distribution.

Appendix C

Main algorithm of the coevolving spin system

The main part of the algorithm for simulations of the coevolving spin system from Section 5.4 written in pseudo-code. It presents the initialization of the network and procedures performed in every time step.

```
// generate random graph
graph = random_Erdos_Renyi_graph(n, m)
for i in [0, ..., n-1]:
    // assign random spin to every node
    graph[i][spin] = random_chice(-1, +1)

for step in time_steps:
    v_index = rand_int(n)  // chose one of n nodes
    e_index = rand_int(m)  // chose one of m edges

    // spin switching
    neighbors = graph.get_neighbors(v_index)
    // compute energy difference
    delta = graph.energy_change_spin(v_index, neighbors)
    if delta <= 0 or rand() < exp(- delta / T):
        // flip spin according to Metropolis rule
        graph[i][spin] = graph[i][spin] * -1

    // edge rewiring
    old_from, old_to = graph.get_edge_ends(e_index)
    new_from, new_to = graph.draw_new_edge(exclude_existing=true)
    // compute energy difference
    delta = graph.energy_change_edge(old_from, old_to, new_from, new_to)
    if delta <= 0 or rand() < exp(- delta / T):
        // rewire edge according to Metropolis rule
        graph.delete_edge([v_from, v_to])
        graph.add_edge([new_from, new_to])
```

Bibliography

[1] M. Newman, "The physics of networks", *Physics today*, vol. 61, no. 11, pp. 33–38, 2008.

[2] M. Newman *et al.*, *Networks: An introduction*. Oxford University Press, 2016, pp. 1–12, 36–37.

[3] A.-L. Barabási *et al.*, *Network science*. Cambridge university press, 2016, pp. 43–45.

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics", *Physics reports*, vol. 424, no. 4, pp. 175–308, 2006.

[5] A. Fronczak and P. Fronczak, *Świat sieci złożonych: Od fizyki do internetu*. Wydawnictwo Naukowe PWN, 2009.

[6] J.-B. Michel, Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, J. P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M. A. Nowak, and E. L. Aiden, "Quantitative analysis of culture using millions of digitized books", *Science*, vol. 331, no. 6014, pp. 176–182, 2011, ISSN: 0036-8075. DOI: 10.1126/science.1199644.

[7] N. Biggs, E. K. Lloyd, and R. J. Wilson, *Graph theory, 1736-1936*. Oxford University Press, 1986, pp. 3–8.

[8] G. Alexanderson, "About the cover: Euler and königsberg's bridges: A historical view", *Bulletin of the american mathematical society*, vol. 43, no. 4, pp. 567–573, 2006.

[9] S. Gießmann, *Henry charles beck, material culture and the london tube map of 1933*, 2013 (accessed January 30, 2019). [Online]. Available: http://amodern.net/article/henry-c-beck-material-culture-and-the-london-tube-map-of-1933/#rf8-2028.

[10] S. Kirchhoff, "Ueber den durchgang eines elektrischen stromes durch eine ebene, insbesondere durch eine kreisförmige", *Annalen der Physik*, vol. 140, no. 4, pp. 497–514, 1845.

[11] A. T. Balaban, "Applications of graph theory in chemistry", *Journal of Chemical Information and Computer Sciences*, vol. 25, pp. 334–343, 1985.

[12] G. Caldarelli and M. Catanzaro, *Networks: A very short introduction*. Oxford University Press, 2012, vol. 335.

[13] R. Thomas, "An update on the four-color theorem", *Notices of the AMS*, vol. 45, no. 7, pp. 848–859, 1998.

[14] W. Lenz, "Beitrag zum verständnis der magnetischen erscheinungen in festen körpern", *Z. Phys.*, vol. 21, pp. 613–615, 1920.

[15] E. Ising, "Beitrag zur theorie des ferromagnetismus", *Zeitschrift für Physik*, vol. 31, no. 1, pp. 253–258, 1925.

[16] R. J. Baxter, *Exactly solved models in statistical mechanics*. Elsevier, 2016.

[17] J. L. Moreno, "Who shall survive?: A new approach to the problem of human interrelations.", 1934.

[18] P. Erdős and A Rényi, "On random graphs", *Publicationes Mathematicae*, vol. 6, no. 290-297, p. 5, 1959.

[19] P. Erdős, A Rényi, *et al.*, "On the evolution of random graphs", *Publ. Math. Inst. Hungar. Acad. Sci*, vol. 5, pp. 17–61, 1960.

[20] R. Diestel, *Graph theory*. Graduate Texts in Math, 1997.

[21] J Wilson Robin, "Wprowadzenie do teorii grafów", *Warszawa. Wydawnictwo Naukowe PWN*, 2012.

[22] J. Travers and S. Milgram, "An experimental study of the small world problem", *Sociometry*, pp. 425–443, 1969.

[23] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks", *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[24] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: Simple building blocks of complex networks", *Science*, vol. 298, no. 5594, pp. 824–827, 2002.

[25] S. Itzkovitz and U. Alon, "Subgraphs and network motifs in geometric networks", *Physical Review E*, vol. 71, no. 2, p. 026 117, 2005.

[26] G. Kossinets and D. J. Watts, "Empirical analysis of an evolving social network", *Science*, vol. 311, no. 5757, pp. 88–90, 2006.

[27] N. S. Contractor, S. Wasserman, and K. Faust, "Testing multitheoretical, multilevel hypotheses about organizational networks: An analytic framework and empirical example", *Academy of Management Review*, vol. 31, no. 3, pp. 681–703, 2006.

[28] G. Simmel, *Soziologie: Untersuchungen über die formen der vergesellschaftung*. BoD–Books on Demand, 2015.

[29] F. Heider, "Attitudes and cognitive organization", *The Journal of psychology*, vol. 21, no. 1, pp. 107–112, 1946.

[30] M. Bianchini, M. Gori, and F. Scarselli, "Inside pagerank", *ACM Transactions on Internet Technology (TOIT)*, vol. 5, no. 1, pp. 92–128, 2005.

[31] A. N. Langville and C. D. Meyer, "Deeper inside pagerank", *Internet Mathematics*, vol. 1, no. 3, pp. 335–380, 2004.

[32] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli, "Debtrank: Too central to fail? financial networks, the fed and systemic risk", *Scientific reports*, vol. 2, p. 541, 2012.

[33] S. Thurner and S. Poledna, "Debtrank-transparency: Controlling systemic risk in financial networks", *Scientific reports*, vol. 3, p. 1888, 2013.

[34] S. P. Borgatti, "Centrality and network flow", *Social networks*, vol. 27, no. 1, pp. 55–71, 2005.

[35] W.-K. Chen, *Applied graph theory*. North-Holland Publishing Company, 1971.

[36] D. B. West *et al.*, *Introduction to graph theory*. Prentice hall Upper Saddle River, NJ, 1996, vol. 2.

[37] F. Harary, *Graph theory and theoretical physics*. Academic Press, London and New York, 1967.

[38] R. J. Wilson, *Applications of graph theory*. Academic Press, 1979.

[39] E. A. Leicht and M. E. Newman, "Community structure in directed networks", *Physical review letters*, vol. 100, no. 11, p. 118 703, 2008.

[40] M. Wilinski, P. Mazzarisi, D. Tantari, and F. Lillo, "Detectability of macroscopic structures in directed asymmetric stochastic block model", *Physical Review E*, vol. 99, no. 4, p. 042 310, 2019.

[41] M. E. Newman, "Analysis of weighted networks", *Physical review E*, vol. 70, no. 5, p. 056 131, 2004.

[42] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, "The architecture of complex weighted networks", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 11, pp. 3747–3752, 2004.

[43] R. N. Mantegna, "Hierarchical structure in financial markets", *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 11, no. 1, pp. 193–197, 1999.

[44] M. Wilinski, Y. Ikeda, and H. Aoyama, "Complex correlation approach for high frequency financial data", *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2018, no. 2, p. 023 405, 2018.

[45] M. Wiliński, "Przemiany fazowe w empirycznych, korelacyjnych sieciach złożonych", 2019.

[46] T. Zhou, J. Ren, M. Medo, and Y.-C. Zhang, "Bipartite network projection and personal recommendation", *Physical review E*, vol. 76, no. 4, p. 046 115, 2007.

[47] P. G. Lind, M. C. Gonzalez, and H. J. Herrmann, "Cycles and clustering in bipartite networks", *Physical review E*, vol. 72, no. 5, p. 056 127, 2005.

[48] K. Kuratowski, "Sur le problème des courbes gauches en topologie", *Fundamenta Mathematicae*, vol. 15, no. 1, 271–283, 1930, ISSN: 0016-2736.

[49] M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna, "A tool for filtering information in complex systems", *Proceedings of the National Academy of Sciences*, vol. 102, no. 30, pp. 10 421–10 426, 2005, ISSN: 0027-8424. DOI: 10.1073/pnas.0500298102.

[50] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin, "The structure and dynamics of multilayer networks", *Physics Reports*, vol. 544, no. 1, pp. 1–122, 2014.

[51] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, "Mathematical formulation of multilayer networks", *Physical Review X*, vol. 3, no. 4, p. 041 022, 2013.

[52] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer networks", *Journal of complex networks*, vol. 2, no. 3, pp. 203–271, 2014.

[53] A. Aleta and Y. Moreno, "Multilayer networks in a nutshell", *Annual Review of Condensed Matter Physics*, vol. 10, pp. 45–62, 2019.

[54] M. De Domenico, V. Nicosia, A. Arenas, and V. Latora, "Structural reducibility of multilayer networks", *Nature communications*, vol. 6, p. 6864, 2015.

[55] M. Diakonova, V. Nicosia, V. Latora, and M. San Miguel, "Irreducibility of multilayer network dynamics: The case of the voter model", *New Journal of Physics*, vol. 18, no. 2, p. 023 010, 2016.

[56] Y. Murase, J. Török, H.-H. Jo, K. Kaski, and J. Kertész, "Multilayer weighted social network model", *Physical Review E*, vol. 90, no. 5, p. 052 810, 2014.

[57] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, "The physics of spreading processes in multilayer networks", *Nature Physics*, vol. 12, no. 10, p. 901, 2016.

[58] P. Holme, "Modern temporal network theory: A colloquium", *The European Physical Journal B*, vol. 88, no. 9, p. 234, 2015.

[59] A. Li, S. P. Cornelius, Y.-Y. Liu, L. Wang, and A.-L. Barabási, "The fundamental advantages of temporal networks", *Science*, vol. 358, no. 6366, pp. 1042–1046, 2017.

[60] O. E. Williams, F. Lillo, and V. Latora, "Effects of memory on spreading processes in non-markovian temporal networks", *New Journal of Physics*, vol. 21, no. 4, p. 043 028, 2019.

[61] L. d. F. Costa, O. N. Oliveira Jr, G. Travieso, F. A. Rodrigues, P. R. Villas Boas, L. Antiqueira, M. P. Viana, and L. E. Correa Rocha, "Analyzing and modeling real-world phenomena with complex networks: A survey of applications", *Advances in Physics*, vol. 60, no. 3, pp. 329–412, 2011.

[62] S. Thurner, R. Hanel, and P. Klimek, *Introduction to the theory of complex systems*. Oxford University Press, 2018.

[63] S. N. Dorogovtsev and J. F. Mendes, "Evolution of networks", *Advances in physics*, vol. 51, no. 4, pp. 1079–1187, 2002.

[64] S. N. Dorogovtsev and J. F. Mendes, *Evolution of networks: From biological nets to the internet and www*. OUP Oxford, 2013.

[65] R. Solomonoff and A. Rapoport, "Connectivity of random nets", *The bulletin of mathematical biophysics*, vol. 13, no. 2, pp. 107–117, 1951.

[66] P. Erdős and A. Rényi, "On the strength of connectedness of a random graph", *Acta Mathematica Hungarica*, vol. 12, no. 1-2, pp. 261–267, 1961.

[67] A. Fronczak, P. Fronczak, and J. A. Hołyst, "Average path length in random networks", *Physical Review E*, vol. 70, no. 5, p. 056 110, 2004.

[68] D. Stauffer and A. Aharony, *Introduction to percolation theory*. Taylor & Francis, 2018.

[69] D. S. Callaway, M. E. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: Percolation on random graphs", *Physical review letters*, vol. 85, no. 25, p. 5468, 2000.

[70] I. Derényi, G. Palla, and T. Vicsek, "Clique percolation in random networks", *Physical review letters*, vol. 94, no. 16, p. 160 202, 2005.

[71] D. Achlioptas, R. M. D'Souza, and J. Spencer, "Explosive percolation in random networks", *Science*, vol. 323, no. 5920, pp. 1453–1455, 2009.

[72] R. A. da Costa, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Explosive percolation transition is actually continuous", *Physical review letters*, vol. 105, no. 25, p. 255 701, 2010.

[73] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks", *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.

[74] J. Park and M. E. Newman, "Statistical mechanics of networks", *Physical Review E*, vol. 70, no. 6, p. 066 117, 2004.

[75] S. Chatterjee, P. Diaconis, *et al.*, "Estimating and understanding exponential random graph models", *The Annals of Statistics*, vol. 41, no. 5, pp. 2428–2461, 2013.

[76] J. Karmeshu, *Entropy measures, maximum entropy principle and emerging applications*. Springer Science & Business Media, 2003, vol. 119.

[77] A. Ben-Naim, *A farewell to entropy: Statistical thermodynamics based on information: S*. World Scientific, 2008.

[78] J. N. Kapur, *Maximum-entropy models in science and engineering*. John Wiley & Sons, 1989.

[79] J. Park and M. E. Newman, "Solution of the two-star model of a network", *Physical Review E*, vol. 70, no. 6, p. 066 146, 2004.

[80] Z. Burda, J. Jurkiewicz, and A. Krzywicki, "Network transitivity and matrix models", *Physical Review E*, vol. 69, no. 2, p. 026 106, 2004.

[81] J. Park and M. Newman, "Solution for the properties of a clustered network", *Physical Review E*, vol. 72, no. 2, p. 026 136, 2005.

[82] A. Fronczak, P. Fronczak, and J. A. Hołyst, "Fluctuation-dissipation relations in complex networks", *Physical Review E*, vol. 73, no. 1, p. 016 108, 2006.

[83] P. Fronczak, A. Fronczak, and J. A. Hołyst, "Phase transitions in social networks", *The European Physical Journal B*, vol. 59, no. 1, pp. 133–139, 2007.

[84] D. Garlaschelli and M. I. Loffredo, "Generalized bose-fermi statistics and structural correlations in weighted networks", *Physical review letters*, vol. 102, no. 3, p. 038 701, 2009.

[85] G. Robins, P. Pattison, Y. Kalish, and D. Lusher, "An introduction to exponential random graph (p*) models for social networks", *Social networks*, vol. 29, no. 2, pp. 173–191, 2007.

[86] G. Robins, T. Snijders, P. Wang, M. Handcock, and P. Pattison, "Recent developments in exponential random graph (p*) models for social networks", *Social networks*, vol. 29, no. 2, pp. 192–215, 2007.

[87] M. E. Newman and J. Park, "Why social networks are different from other types of networks", *Physical Review E*, vol. 68, no. 3, p. 036 122, 2003.

[88] D. V. Foster, J. G. Foster, P. Grassberger, and M. Paczuski, "Clustering drives assortativity and community structure in ensembles of networks", *Physical Review E*, vol. 84, no. 6, p. 066 117, 2011.

[89] P. Colomer-de Simón, M Serrano, M. G. Beiró, J. I. Alvarez-Hamelin, and M. Boguná, "Deciphering the global organization of clustering in real complex networks", *Scientific Reports*, vol. 3, no. 2517, 2013.

[90] S. Milgram, "The small world problem", *Psychology today*, vol. 2, no. 1, pp. 60–67, 1967.

[91] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, "Classes of small-world networks", *Proceedings of the national academy of sciences*, vol. 97, no. 21, pp. 11 149–11 152, 2000.

[92] A. Wagner and D. A. Fell, "The small world inside large metabolic networks", *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 268, no. 1478, pp. 1803–1810, 2001.

[93] V. Latora and M. Marchiori, "Is the boston subway a small-world network?", *Physica A: Statistical Mechanics and its Applications*, vol. 314, no. 1-4, pp. 109–113, 2002.

[94] D. J. Watts and S. H. Strogatz, "Collective dynamics of ?small-world?networks", *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[95] V. Latora and M. Marchiori, "Efficient behavior of small-world networks", *Physical review letters*, vol. 87, no. 19, p. 198 701, 2001.

[96] J. M. Kleinberg, "Navigation in a small world", *Nature*, vol. 406, no. 6798, p. 845, 2000.

[97] J. M. Kleinberg *et al.*, "Small-world phenomena and the dynamics of information", in *Advances in neural information processing systems*, 2002, pp. 431–438.

[98] M. Newman, A.-L. Barabasi, and D. J. Watts, *The structure and dynamics of networks*. Princeton university press, 2006.

[99] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: Diameter of the world-wide web", *Nature*, vol. 401, no. 6749, pp. 130–131, 1999.

[100] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, "The large-scale organization of metabolic networks", *Nature*, vol. 407, no. 6804, p. 651, 2000.

[101] A.-L. Barabâsi, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek, "Evolution of the social network of scientific collaborations", *Physica A: Statistical mechanics and its applications*, vol. 311, no. 3-4, pp. 590–614, 2002.

[102] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology", in *ACM SIGCOMM computer communication review*, ACM, vol. 29, 1999, pp. 251–262.

[103] H Jeong, Z Néda, and A. L Barabási, "Measuring preferential attachment in evolving networks", *Europhysics Letters (EPL)*, vol. 61, no. 4, pp. 567–572, 2003. DOI: 10.1209/epl/i2003-00166-9.

[104] G. U. Yule, "Ii.—a mathematical theory of evolution, based on the conclusions of dr. jc willis, fr s", *Philosophical transactions of the Royal Society of London. Series B, containing papers of a biological character*, vol. 213, no. 402-410, pp. 21–87, 1925.

[105] H. A. Simon, "On a class of skew distribution functions", *Biometrika*, vol. 42, no. 3/4, pp. 425–440, 1955.

[106] D. J.D. S. Price, "Networks of scientific papers", *Science*, pp. 510–515, 1965.

[107] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks", *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[108] A.-L. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks", *Physica A: Statistical Mechanics and its Applications*, vol. 272, no. 1-2, pp. 173–187, 1999.

[109] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Structure of growing networks with preferential linking", *Physical review letters*, vol. 85, no. 21, p. 4633, 2000.

[110] P. L. Krapivsky, S. Redner, and F. Leyvraz, "Connectivity of growing random networks", *Physical review letters*, vol. 85, no. 21, p. 4629, 2000.

[111] R. Cohen and S. Havlin, "Scale-free networks are ultrasmall", *Physical review letters*, vol. 90, no. 5, p. 058 701, 2003.

[112] B. Bollobás and O. Riordan, "The diameter of a scale-free random graph", *Combinatorica*, vol. 24, no. 1, pp. 5–34, 2004.

[113] K. Klemm and V. M. Eguiluz, "Growing scale-free networks with small-world behavior", *Physical Review E*, vol. 65, no. 5, p. 057 102, 2002.

[114] B. Bollobás and O. M. Riordan, "Mathematical results on scale-free random graphs", *Handbook of graphs and networks: From the genome to the internet*, pp. 1–34, 2003.

[115] M. P. Stumpf and M. A. Porter, "Critical truths about power laws", *Science*, vol. 335, no. 6069, pp. 665–666, 2012.

[116] A. D. Broido and A. Clauset, "Scale-free networks are rare", *Nature communications*, vol. 10, no. 1, p. 1017, 2019.

[117] I. Voitalov, P. van der Hoorn, R. van der Hofstad, and D. Krioukov, "Scale-free networks well done", *ArXiv preprint arXiv:1811.02071*, 2018.

[118] P. Holme, "Rare and everywhere: Perspectives on scale-free networks", *Nature communications*, vol. 10, no. 1, p. 1016, 2019.

[119] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Principles of statistical mechanics of uncorrelated random networks", *Nuclear Physics B*, vol. 666, no. 3, pp. 396–416, 2003.

[120] A. Jędrzejewski and K. Sznajd-Weron, "Statistical physics of opinion formation: Is it a spoof?", *Comptes Rendus Physique*, 2019.

[121] M. Mäs, A. Flache, and D. Helbing, "Individualization as driving force of clustering phenomena in humans", *PLoS computational biology*, vol. 6, no. 10, e1000959, 2010.

[122] M. W. Macy, J. A. Kitts, A. Flache, and S. Benard, "Polarization in dynamic networks: A hopfield model of emergent structure", *Dynamic social network modeling and analysis*, pp. 162–173, 2003.

[123] Y. Moreno, M. Nekovee, and A. F. Pacheco, "Dynamics of rumor spreading in complex networks", *Physical Review E*, vol. 69, no. 6, p. 066 130, 2004.

[124] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili, "Theory of rumour spreading in complex social networks", *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 1, pp. 457–470, 2007.

[125] P Smaga, M Wiliński, P Ochnicki, P Arendarski, and T Gubiec, "Can banks default overnight? modelling endogenous contagion on the o/n interbank market", *Quantitative Finance*, vol. 18, no. 11, pp. 1815–1829, 2018.

[126] D. Di Gangi, D Sardo, V. Macchiati, T. P. Minh, F. Pinotti, A. Ramadiah, M. Wilinski, and G. Cimini, "Network sensitivity of systemic risk", *ArXiv preprint arXiv:1805.04325*, 2018.

[127] D. Stauffer, X. Castelló, V. M. Eguiluz, and M. San Miguel, "Microscopic abrams–strogatz model of language competition", *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 2, pp. 835–842, 2007.

[128] X. Castelló, V. M. Eguíluz, and M. San Miguel, "Ordering dynamics with two non-excluding options: Bilingualism in language competition", *New Journal of Physics*, vol. 8, no. 12, p. 308, 2006.

[129] X. Castelló, V. Eguíluz, M. Miguel, L. Loureiro-Porto, R. Toivonen, J Saramäki, and K Kaski, "Modelling language competition: Bilingualism and complex social networks", in *The Evolution of Language: Proceedings of the 7th International Conference. Singapore: World Scientific Publishing Co*, Citeseer, 2008, pp. 59–66.

[130] M. Patriarca, X. Castelló, J. Uriarte, V. M. Eguíluz, and M. San Miguel, "Modeling two-language competition dynamics", *Advances in Complex Systems*, vol. 15, no. 03n04, p. 1 250 048, 2012.

[131] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks", *Physical review letters*, vol. 86, no. 14, p. 3200, 2001.

[132] R. M. May and A. L. Lloyd, "Infection dynamics on scale-free networks", *Physical Review E*, vol. 64, no. 6, p. 066 112, 2001.

[133] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics and endemic states in complex networks", *Phys. Rev. E*, vol. 63, p. 066 117, 6 2001. DOI: 10.1103/PhysRevE.63.066117.

[134] P. L. Krapivsky and S. Redner, "Organization of growing random networks", *Physical Review E*, vol. 63, no. 6, p. 066 123, 2001.

[135] T. Gross and B. Blasius, "Adaptive coevolutionary networks: A review", *Journal of the Royal Society Interface*, vol. 5, no. 20, pp. 259–271, 2008.

[136] K. Christensen, R. Donangelo, B. Koiller, and K. Sneppen, "Evolution of random networks", *Physical Review Letters*, vol. 81, no. 11, p. 2380, 1998.

[137] S. Bornholdt and T. Rohlf, "Topological evolution of dynamical networks: Global criticality from local dynamics", *Physical Review Letters*, vol. 84, no. 26, p. 6114, 2000.

[138] S. Bornholdt and K. Sneppen, "Robustness as an evolutionary principle", *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 267, no. 1459, pp. 2281–2286, 2000.

[139] V. M. Eguíluz, M. G. Zimmermann, C. J. Cela-Conde, and M. S. Miguel, "Cooperation and the emergence of role differentiation in the dynamics of social networks", *American journal of sociology*, vol. 110, no. 4, pp. 977–1008, 2005.

[140] S. Gil and D. H. Zanette, "Coevolution of agents and networks: Opinion spreading and community disconnection", *Physics Letters A*, vol. 356, no. 2, pp. 89–94, 2006.

[141] I. Glauche, W. Krause, R. Sollacher, and M. Greiner, "Distributive routing and congestion control in wireless multihop ad hoc communication networks", *Physica A: Statistical Mechanics and its Applications*, vol. 341, pp. 677–701, 2004.

[142] P. Gong and C. van Leeuwen, "Evolution to a small-world network with chaotic units", *EPL (Europhysics Letters)*, vol. 67, no. 2, p. 328, 2004.

[143] P. Holme and G. Ghoshal, "Dynamics of networking agents competing for high centrality and low degree", *Physical review letters*, vol. 96, no. 9, p. 098 701, 2006.

[144] P. Holme and M. E. Newman, "Nonequilibrium phase transition in the coevolution of networks and opinions", *Physical Review E*, vol. 74, no. 5, p. 056 108, 2006.

[145] S. Jain and S. Krishna, "A model for the emergence of cooperation, interdependence, and structure in evolving networks", *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 543–547, 2001.

[146] M. Liu and K. E. Bassler, "Emergent criticality from coevolution in random boolean networks", *Physical Review E*, vol. 74, no. 4, p. 041 910, 2006.

[147] B. Luque, F. J. Ballesteros, and E. M. Muro, "Self-organized critical random boolean networks", *Physical Review E*, vol. 63, no. 5, p. 051 913, 2001.

[148] J. M. Pacheco, A. Traulsen, and M. A. Nowak, "Coevolution of strategy and structure in complex networks with dynamical linking", *Physical review letters*, vol. 97, no. 25, p. 258 103, 2006.

[149] M. Rosvall and K. Sneppen, "Modeling self-organization of communication and topology in social networks", *Physical Review E*, vol. 74, no. 1, p. 016 108, 2006.

[150] H. Sayama, "Generative network automata: A generalized framework for modeling complex dynamical systems with autonomously varying topologies", in *2007 IEEE Symposium on Artificial Life*, IEEE, 2007, pp. 214–221.

[151] A. Scire, I. Tuval, and V. M. Eguíluz, "Dynamic modeling of the electric transportation network", *EPL (Europhysics Letters)*, vol. 71, no. 2, p. 318, 2005.

[152] A. M. Seufert and F. Schweitzer, "Aggregate dynamics in an evolutionary network model", *International Journal of Modern Physics C*, vol. 18, no. 10, pp. 1659–1674, 2007.

[153] B. Skyrms and R. Pemantle, "A dynamic model of social network formation", in *Adaptive networks*, Springer, 2009, pp. 231–251.

[154] D. van den Berg and C. van Leeuwen, "Adaptive rewiring in chaotic networks renders small-world connectivity with consistent clusters", *EPL (Europhysics Letters)*, vol. 65, no. 4, p. 459, 2004.

[155] D. H. Zanette and S. Gil, "Opinion spreading and agent segregation on evolving networks", *Physica D: Nonlinear Phenomena*, vol. 224, no. 1-2, pp. 156–165, 2006.

[156] M. J. Krawczyk and K. Kułakowski, "Adaptive rewiring of links in a network: Jammed states", *Physica A: Statistical Mechanics and its Applications*, p. 121 716, 2019.

[157] C. Zhou and J. Kurths, "Dynamical weights and enhanced synchronization in adaptive complex networks", *Physical review letters*, vol. 96, no. 16, p. 164 102, 2006.

[158] E. T. Tchuisseu, D. Gomila, and P. Colet, "Reduction of power grid fluctuations by communication between smart devices", *International Journal of Electrical Power & Energy Systems*, vol. 108, pp. 145–152, 2019.

[159] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: An explanation of the 1/f noise", *Physical review letters*, vol. 59, no. 4, p. 381, 1987.

[160] P. Bak, C. Tang, K. Wiesenfeld, *et al.*, "Self-organized criticality", *Physical review A*, vol. 38, no. 1, p. 364, 1988.

[161] P. Fronczak, A. Fronczak, and J. A. Hołyst, "Self-organized criticality and coevolution of network structure and dynamics", *Physical Review E*, vol. 73, no. 4, p. 046 117, 2006.

[162] T. Gross, C. J. D. D'Lima, and B. Blasius, "Epidemic dynamics on an adaptive network", *Physical review letters*, vol. 96, no. 20, p. 208 701, 2006.

[163] M. G. Zimmermann, V. M. Eguíluz, and M. San Miguel, "Coevolution of dynamical states and interactions in dynamic networks", *Physical Review E*, vol. 69, no. 6, p. 065 102, 2004.

[164] M. G. Zimmermann and V. M. Eguíluz, "Cooperation, social networks, and the emergence of leadership in a prisoner's dilemma with adaptive local interactions", *Physical Review E*, vol. 72, no. 5, p. 056 118, 2005.

[165] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics", *Reviews of modern physics*, vol. 81, no. 2, p. 591, 2009.

[166] P. Clifford and A. Sudbury, "A model for spatial conflict", *Biometrika*, vol. 60, no. 3, pp. 581–588, 1973.

[167] R. A. Holley and T. M. Liggett, "Ergodic theorems for weakly interacting infinite systems and the voter model", *The annals of probability*, pp. 643–663, 1975.

[168] A. Kirman, "Ants, rationality, and recruitment", *The Quarterly Journal of Economics*, vol. 108, no. 1, pp. 137–156, 1993.

[169] S. Alfarano, T. Lux, and F. Wagner, "Estimation of agent-based models: The case of an asymmetric herding model", *Computational Economics*, vol. 26, no. 1, pp. 19–49, 2005.

[170] K. Suchecki, V. M. Eguíluz, and M. San Miguel, "Voter model dynamics in complex networks: Role of dimensionality, disorder, and degree distribution", *Physical Review E*, vol. 72, no. 3, p. 036 132, 2005.

[171] P. Przybyła, K. Sznajd-Weron, and R. Weron, "Diffusion of innovation within an agent-based model: Spinsons, independence and advertising", *Advances in Complex Systems*, vol. 17, no. 01, p. 1 450 004, 2014.

[172] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, "Is the voter model a model for voters?", *Physical review letters*, vol. 112, no. 15, p. 158 701, 2014.

[173] P. Krapivsky, "Kinetics of a monomer-monomer model of heterogeneous catalysis", *Journal of Physics A: Mathematical and General*, vol. 25, no. 22, p. 5831, 1992.

[174] L. Frachebourg and P. L. Krapivsky, "Exact results for kinetics of catalytic reactions", *Physical Review E*, vol. 53, no. 4, R3009, 1996.

[175] D. D. Chinellato, I. R. Epstein, D. Braha, Y. Bar-Yam, and M. A. de Aguiar, "Dynamical response of networks under external perturbations: Exact results", *Journal of Statistical Physics*, vol. 159, no. 2, pp. 221–230, 2015.

[176] S. Alfarano, T. Lux, and F. Wagner, "Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach", *Journal of Economic Dynamics and Control*, vol. 32, no. 1, pp. 101–136, 2008.

[177] N. G. Van Kampen, *Stochastic processes in physics and chemistry*. Elsevier, 1992, vol. 1.

[178] A. Carro, R. Toral, and M. San Miguel, "Markets, herding and response to external information", *PloS one*, vol. 10, no. 7, e0133287, 2015.

[179] A. Kirman, "Epidemics of opinion and speculative bubbles in financial markets", *Money and financial markets chap*, vol. 17, 1991.

[180] A. Kononovicius and V. Gontis, "Agent based reasoning for the nonlinear stochastic models of long-range memory", *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 4, pp. 1309–1314, 2012.

[181] J. Klamut and T. Gubiec, "Directed continuous-time random walk with memory", *The European Physical Journal B*, vol. 92, no. 4, p. 69, 2019.

[182] J. Klamut, R. Kutner, T. Gubiec, and Z. R. Struzik, "The new face of multifractality: Multi-branchness and the phase transitions in time series of inter-event times", *ArXiv preprint arXiv:1809.02674*, 2018.

[183] J. Klamut and T. Gubiec, "'continuous' time random walk in continuous time random walk. the crucial role of inter-event times in volatility clustering", *ArXiv preprint arXiv:1909.04986*, 2019.

[184] N. Khalil, M. San Miguel, and R. Toral, "Zealots in the mean-field noisy voter model", *Physical Review E*, vol. 97, no. 1, p. 012 310, 2018.

[185] P. Nyczka, K. Sznajd-Weron, and J. Cisło, "Phase transitions in the q-voter model with two types of stochastic driving", *Physical Review E*, vol. 86, no. 1, p. 011 105, 2012.

[186] S. Alfarano and M. Milaković, "Network structure and n-dependence in agent-based herding models", *Journal of Economic Dynamics and Control*, vol. 33, no. 1, pp. 78–92, 2009.

[187] K. Suchecki, V. M. Eguiluz, and M. San Miguel, "Conservation laws for the voter model in complex networks", *EPL (Europhysics Letters)*, vol. 69, no. 2, p. 228, 2004.

[188] C. Castellano, V. Loreto, A. Barrat, F. Cecconi, and D. Parisi, "Comparison of voter and glauber ordering dynamics on networks", *Physical review E*, vol. 71, no. 6, p. 066 107, 2005.

[189] F. Vazquez, V. M. Eguíluz, and M. San Miguel, "Generic absorbing transition in coevolution dynamics", *Physical review letters*, vol. 100, no. 10, p. 108 702, 2008.

[190] I. Dornic, H. Chaté, J. Chave, and H. Hinrichsen, "Critical coarsening without surface tension: The universality class of the voter model", *Physical Review Letters*, vol. 87, no. 4, p. 045 701, 2001.

[191] O. Al Hammal, H. Chaté, I. Dornic, and M. A. Munoz, "Langevin description of critical phenomena with two symmetric absorbing states", *Physical review letters*, vol. 94, no. 23, p. 230 601, 2005.

[192] C. Castellano, D. Vilone, and A. Vespignani, "Incomplete ordering of the voter model on small-world networks", *EPL (Europhysics Letters)*, vol. 63, no. 1, p. 153, 2003.

[193] D. Vilone and C. Castellano, "Solution of voter model dynamics on annealed small-world networks", *Physical Review E*, vol. 69, no. 1, p. 016 109, 2004.

[194] B. L. Granovsky and N. Madras, "The noisy voter model", *Stochastic Processes and their applications*, vol. 55, no. 1, pp. 23–43, 1995.

[195] A. Peralta, A. Carro, M. San Miguel, and R. Toral, "Stochastic pair approximation treatment of the noisy voter model", *New Journal of Physics*, vol. 20, no. 10, p. 103 045, 2018.

[196] O. Artime, A. F. Peralta, R. Toral, J. J. Ramasco, and M. San Miguel, "Aging-induced continuous phase transition", *Physical Review E*, vol. 98, no. 3, p. 032 104, 2018.

[197] V. Sood and S. Redner, "Voter model on heterogeneous graphs", *Physical review letters*, vol. 94, no. 17, p. 178 701, 2005.

[198] V. Sood, T. Antal, and S. Redner, "Voter models on heterogeneous networks", *Physical Review E*, vol. 77, no. 4, p. 041 121, 2008.

[199] C. Castellano, M. A. Muñoz, and R. Pastor-Satorras, "Nonlinear q-voter model", *Physical Review E*, vol. 80, no. 4, p. 041 129, 2009.

[200] A. Abramiuk, J. Pawłowski, and K. Sznajd-Weron, "Is independence necessary for a discontinuous phase transition within the q-voter model?", *Entropy*, vol. 21, no. 5, p. 521, 2019.

[201] W. Radosz, A. Mielnik-Pyszczorski, M. Brzezińska, and K. Sznajd-Weron, "Q-voter model with nonconformity in freely forming groups: Does the size distribution matter?", *Physical Review E*, vol. 95, no. 6, p. 062 302, 2017.

[202] A. Jedrzejewski, "Pair approximation for the q-voter model with independence on complex networks", *Physical Review E*, vol. 95, no. 1, p. 012 307, 2017.

[203] P. Nyczka and K. Sznajd-Weron, "Anticonformity or independence?—insights from statistical physics", *Journal of Statistical Physics*, vol. 151, no. 1-2, pp. 174–202, 2013.

[204] A. Jędrzejewski and K. Sznajd-Weron, "Nonlinear *q*-voter model from the quenched perspective", *ArXiv preprint arXiv:2001.01059*, 2020.

[205] J. Toruniewska, K. Kułakowski, K. Suchecki, and J. A. Hołyst, "Coupling of link-and node-ordering in the coevolving voter model", *Physical Review E*, vol. 96, no. 4, p. 042 306, 2017.

[206] G. Demirel, F. Vazquez, G. Böhme, and T. Gross, "Moment-closure approximations for discrete adaptive networks", *Physica D: Nonlinear Phenomena*, vol. 267, pp. 68–80, 2014.

[207] R. Durrett, J. P. Gleeson, A. L. Lloyd, P. J. Mucha, F. Shi, D. Sivakoff, J. E. Socolar, and C. Varghese, "Graph fission in an evolving voter model", *Proceedings of the National Academy of Sciences*, vol. 109, no. 10, pp. 3682–3687, 2012.

[208] M. Diakonova, M. San Miguel, and V. M. Eguíluz, "Absorbing and shattered fragmentation transitions in multilayer coevolution", *Physical Review E*, vol. 89, no. 6, p. 062 818, 2014.

[209] A. Nowak, J. Szamrej, and B. Latané, "From private attitude to public opinion: A dynamic theory of social impact.", *Psychological Review*, vol. 97, no. 3, p. 362, 1990.

[210] D. M. Abrams and S. H. Strogatz, "Linguistics: Modelling the dynamics of language death", *Nature*, vol. 424, no. 6951, pp. 900–900, 2003.

[211] F. Vazquez, X. Castelló, and M. San Miguel, "Agent based models of language competition: Macroscopic descriptions and order–disorder transitions", *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2010, no. 04, P04007, 2010.

[212] D. Nettle, "Using social impact theory to simulate language change", *Lingua*, vol. 108, no. 2, pp. 95–117, 1999.

[213] B. Min and M. San Miguel, "Fragmentation transitions in a coevolving nonlinear voter model", *Scientific Reports*, vol. 7, no. 1, p. 12 864, 2017.

[214] D. Centola, V. M. Eguíluz, and M. W. Macy, "Cascade dynamics of complex propagation", *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 1, pp. 449–456, 2007.

[215] A. Czaplicka, R. Toral, and M. San Miguel, "Competition of simple and complex adoption on interdependent networks", *Physical Review E*, vol. 94, no. 6, p. 062 301, 2016.

[216] J. P. Gleeson, "Binary-state dynamics on complex networks: Pair approximation and beyond", *Physical Review X*, vol. 3, no. 2, p. 021 004, 2013.

[217] E. Pugliese and C. Castellano, "Heterogeneous pair approximation for voter models on networks", *EPL (Europhysics Letters)*, vol. 88, no. 5, p. 58 004, 2009.

[218] A. Jędrzejewski, J. Toruniewska, K. Suchecki, O. Zaikin, and J. A. Hołyst, "Spontaneous symmetry breaking of active phase in coevolving nonlinear voter model", *ArXiv preprint arXiv:2004.08927*, 2020.

[219] D Lee, K.-I Goh, B Kahng, and D Kim, "Complete trails of co-authorship network evolution", *Phys. Rev. E*, p. 026 112, 2010.

[220] P. Klimek and S. Thurner, "Triadic closure dynamics drives scaling laws in social multiplex networks", *New Journal of Physics*, vol. 15, no. 6, p. 063 008, 2013.

[221] G. Bianconi, R. K. Darst, J. Iacovacci, and S. Fortunato, "Triadic closure as a basic generating mechanism of communities in complex networks", *Physical Review E*, vol. 90, no. 4, p. 042 806, 2014.

[222] P. Holme and B. J. Kim, "Growing scale-free networks with tunable clustering", *Physical review E*, vol. 65, no. 2, p. 026 107, 2002.

[223] J. Davidsen, H. Ebel, and S. Bornholdt, "Emergence of a small world from local interactions: Modeling acquaintance networks", *Physical Review Letters*, vol. 88, no. 12, p. 128 701, 2002.

[224] P. Klimek, M. Diakonova, V. M. Eguíluz, M. San Miguel, and S. Thurner, "Dynamical origins of the community structure of an online multilayer society", *New Journal of Physics*, vol. 18, no. 8, p. 083 045, 2016.

[225] G. Palla, A.-L. Barabási, and T. Vicsek, "Quantifying social group evolution", *Nature*, vol. 446, no. 7136, pp. 664–667, 2007.

[226] M. Szell and S. Thurner, "Measuring social dynamics in a massive multiplayer online game", *Social networks*, vol. 32, no. 4, pp. 313–329, 2010.

[227] T. Raducha, B. Min, and M. San Miguel, "Coevolving nonlinear voter model with triadic closure", *EPL (Europhysics Letters)*, vol. 124, no. 3, p. 30 001,

[228] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki, "Statistical physics of human cooperation", *Physics Reports*, 2017.

[229] K. Sznajd-Weron, M Tabiszewski, and A. M. Timpanaro, "Phase transition in the sznajd model with independence", *EPL (Europhysics Letters)*, vol. 96, no. 4, p. 48 002, 2011.

[230] T. Raducha and M. S. Miguel, "Emergence of complex structures from nonlinear interactions and noise in coevolving networks", *ArXiv preprint arXiv:2004.06515*, 2020.

[231] M. Diakonova, V. M. Eguíluz, and M. San Miguel, "Noise in coevolving networks", *Physical Review E*, vol. 92, no. 3, p. 032 803, 2015.

[232] A. F. Peralta, A. Carro, M San Miguel, and R Toral, "Analytical and numerical study of the non-linear noisy voter model on complex networks", *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 7, p. 075 516, 2018.

[233] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks", *Physical review E*, vol. 70, no. 6, p. 066 111, 2004.

[234] R. Axelrod, "The dissemination of culture a model with local convergence and global polarization", *Journal of conflict resolution*, vol. 41, no. 2, pp. 203–226, 1997.

[235] P. Ball, *Critical mass: How one thing leads to another*. Arrow Books, 2005.

[236] D. Byrne, "Attitudes and attraction", in *Advances in experimental social psychology*, vol. 4, Elsevier, 1969, pp. 35–89.

[237] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks", *Annual review of sociology*, vol. 27, no. 1, pp. 415–444, 2001.

[238] R. B. Potts, "Some generalized order-disorder transformations", in *Mathematical proceedings of the cambridge philosophical society*, Cambridge University Press, vol. 48, 1952, pp. 106–109.

[239] F.-Y. Wu, "The potts model", *Reviews of modern physics*, vol. 54, no. 1, p. 235, 1982.

[240] C. Castellano, M. Marsili, and A. Vespignani, "Nonequilibrium phase transition in a model for social influence", *Physical Review Letters*, vol. 85, no. 16, p. 3536, 2000.

[241] K. Klemm, V. M. Eguılluz, R. Toral, and M. San Miguel, "Role of dimensionality in axelrod's model for the dissemination of culture", *Physica A: Statistical Mechanics and its Applications*, vol. 327, no. 1, pp. 1–5, 2003.

[242] N. Lanchier *et al.*, "The axelrod model for the dissemination of culture revisited", *The Annals of Applied Probability*, vol. 22, no. 2, pp. 860–880, 2012.

[243] K. Klemm, V. M. Eguíluz, R. Toral, and M. San Miguel, "Nonequilibrium transitions in complex networks: A model of social interaction", *Physical Review E*, vol. 67, no. 2, p. 026 120, 2003.

[244] M. Laguna, G. Abramson, and D. H. Zanette, "Vector opinion dynamics in a model for social influence", *Physica A: Statistical Mechanics and its Applications*, vol. 329, no. 3, pp. 459–472, 2003.

[245] K. Klemm, V. M. Eguíluz, R. Toral, and M. San Miguel, "Global culture: A noise-induced transition in finite systems", *Physical Review E*, vol. 67, no. 4, p. 045 101, 2003.

[246] K. Klemm, V. M. Eguíluz, R. Toral, and M. San Miguel, "Globalization, polarization and cultural drift", *Journal of Economic Dynamics and Control*, vol. 29, no. 1-2, pp. 321–334, 2005.

[247] Y. Kim, M. Cho, and S.-H. Yook, "Effects of the underlying topology on perturbation spreading in the axelrod model for cultural dissemination", *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 21, pp. 3989–3995, 2011.

[248] L. De Sanctis and T. Galla, "Effects of noise and confidence thresholds in nominal and metric axelrod dynamics of social influence", *Physical Review E*, vol. 79, no. 4, p. 046 108, 2009.

[249] Y. Gandica, E Medina, and I Bonalde, "A thermodynamic counterpart of the axelrod model of social influence: The one-dimensional case", *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 24, pp. 6561–6570, 2013.

[250] J. Genzor, V. Bužek, and A. Gendiar, "Thermodynamic model of social influence on two-dimensional square lattice: Case for two features", *Physica A: Statistical Mechanics and its Applications*, vol. 420, pp. 200–211, 2015.

[251] A. Flache and M. W. Macy, "Local convergence and global diversity: From interpersonal to social influence", *Journal of Conflict Resolution*, vol. 55, no. 6, pp. 970–995, 2011.

[252] A. H. Rodríguez and Y Moreno, "Effects of mass media action on the axelrod model with social influence", *Physical Review E*, vol. 82, no. 1, p. 016 111, 2010.

[253] S. Pinto, P. Balenzuela, and C. O. Dorso, "Setting the agenda: Different strategies of a mass media in a model of cultural dissemination", *Physica A: Statistical Mechanics and its Applications*, vol. 458, pp. 378–390, 2016.

[254] C. Gracia-Lázaro, F. Quijandría, L. Hernández, L. M. Floría, and Y. Moreno, "Coevolutionary network approach to cultural dynamics controlled by intolerance", *Physical Review E*, vol. 84, no. 6, p. 067 101, 2011.

[255] C Gracia-Lázaro, L. Floria, and Y Moreno, "Selective advantage of tolerant cultural traits in the axelrod-schelling model", *Physical Review E*, vol. 83, no. 5, p. 056 103, 2011.

[256] B. Dybiec, N. Mitarai, and K. Sneppen, "Axelrod model: Accepting or discussing", *The European Physical Journal B*, vol. 85, no. 10, pp. 1–5, 2012.

[257] B. Dybiec, "Axelrod model with extended conservativeness", *International Journal of Modern Physics C*, vol. 23, no. 12, p. 1 250 086, 2012.

[258] A Parravano, H Rivera-Ramirez, and M. Cosenza, "Intracultural diversity in a model of social dynamics", *Physica A: Statistical Mechanics and its Applications*, vol. 379, no. 1, pp. 241–249, 2007.

[259] J. Pfau, M. Kirley, and Y. Kashima, "The co-evolution of cultures, social network communities, and agent locations in an extension of axelrod's model of cultural dissemination", *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 2, pp. 381–391, 2013.

[260] A. Stivala, G. Robins, Y. Kashima, and M. Kirley, "Ultrametric distribution of culture vectors in an extended axelrod model of cultural dissemination", *Scientific reports*, vol. 4, 2014.

[261] F. Battiston, V. Nicosia, V. Latora, and M. San Miguel, "Layered social influence promotes multiculturality in the axelrod model", *Scientific reports*, vol. 7, no. 1, p. 1809, 2017.

[262] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents", *Advances in Complex Systems*, vol. 3, no. 01n04, pp. 87–98, 2000.

[263] R. Hegselmann, U. Krause, *et al.*, "Opinion dynamics and bounded confidence models, analysis, and simulation", *Journal of artificial societies and social simulation*, vol. 5, no. 3, 2002.

[264] F. Vazquez, J. C. González-Avella, V. M. Eguíluz, and M. San Miguel, "Time-scale competition leading to fragmentation and recombination transitions in the coevolution of network and states", *Physical Review E*, vol. 76, no. 4, p. 046 120, 2007.

[265] D. Centola, J. C. Gonzalez-Avella, V. M. Eguiluz, and M. San Miguel, "Homophily, cultural drift, and the co-evolution of cultural groups", *Journal of Conflict Resolution*, vol. 51, no. 6, pp. 905–929, 2007.

[266] T. Raducha and T. Gubiec, "Coevolving complex networks in the model of social interactions", *Physica A: Statistical Mechanics and its Applications*, vol. 471, pp. 427–435, 2017.

[267] M. E. Newman, "Clustering and preferential attachment in growing networks", *Physical Review E*, vol. 64, no. 2, p. 025 102, 2001.

[268] H. Jeong, Z. Néda, and A.-L. Barabási, "Measuring preferential attachment in evolving networks", *EPL (Europhysics Letters)*, vol. 61, no. 4, p. 567, 2003.

[269] M. Wieling and J. Nerbonne, "Advances in dialectometry", 2015.

[270] P. Trudgill, "Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography", *Language in society*, vol. 3, no. 2, pp. 215–246, 1974.

[271] G. Bailey, T. Wikle, J. Tillery, and L. Sand, "Some patterns of linguistic diffusion", *Language variation and change*, vol. 5, no. 3, pp. 359–390, 1993.

[272] J. N. Stanford and L. A. Kenny, "Revisiting transmission and diffusion: An agent-based model of vowel chain shifts across large communities", *Language Variation and Change*, vol. 25, no. 2, pp. 119–153, 2013.

[273] A. M. Petersen, J. N. Tenenbaum, S. Havlin, H. E. Stanley, and M. Perc, "Languages cool as they expand: Allometric scaling and the decreasing need for new words", *Scientific reports*, vol. 2, p. 943, 2012.

[274] G. J. Baxter, R. A. Blythe, W. Croft, and A. J. McKane, "Utterance selection model of language change", *Physical Review E*, vol. 73, no. 4, p. 046 118, 2006.

[275] V. Loreto, A. Baronchelli, A. Mukherjee, A. Puglisi, and F. Tria, "Statistical physics of language dynamics", *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 04, P04006, 2011.

[276] J. Burridge, "Unifying models of dialect spread and extinction using surface tension dynamics", *Royal Society Open Science*, vol. 5, no. 1, p. 171 446, 2018.

[277] D. R. Amancio, "Authorship recognition via fluctuation analysis of network topology and word intermittency", *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2015, no. 3, P03005, 2015.

[278] A. Kulig, S. Drożdż, J. Kwapień, and P. Oświęcimka, "Modeling the average shortest-path length in growth of word-adjacency networks", *Physical Review E*, vol. 91, no. 3, p. 032 810, 2015.

[279] C. Schulze, D. Stauffer, and S. Wichmann, "Birth, survival and death of languages by monte carlo simulation", *Communications in Computational Physics*, vol. 3, no. 2, pp. 271–294, 2008.

[280] L. Bloomfield, *Language (new york: Holt, rinehart and winston)*, 1933.

[281] C. Beckner, R. Blythe, J. Bybee, M. H. Christiansen, W. Croft, N. C. Ellis, J. Holland, J. Ke, D. Larsen-Freeman, and T. Schoenemann, "Language is a complex adaptive system: Position paper", *Language learning*, vol. 59, no. s1, pp. 1–26, 2009.

[282] S. S. Mufwene, "Competition and selection in language evolution", *Selection*, vol. 3, no. 1, pp. 45–56, 2002.

[283] M. Tomasello, *Origins of human communication*. MIT press, 2010.

[284] P. Eckert, *Language variation as social practice: The linguistic construction of identity in belten high*. Wiley-Blackwell, 2000.

[285] G. J. Baxter, R. A. Blythe, W. Croft, and A. J. McKane, "Modeling language change: An evaluation of trudgill's theory of the emergence of new zealand english", *Language Variation and Change*, vol. 21, no. 02, pp. 257–296, 2009.

[286] A. Carro, R. Toral, and M. San Miguel, "Coupled dynamics of node and link states in complex networks: A model for language competition", *New Journal of Physics*, vol. 18, no. 11, p. 113 056, 2016.

[287] E. Lieberman, J.-B. Michel, J. Jackson, T. Tang, and M. A. Nowak, "Quantifying the evolutionary dynamics of language", *Nature*, vol. 449, no. 7163, pp. 713–716, 2007.

[288] J. L. Bybee, "From usage to grammar: The mind's response to repetition", *Language*, vol. 82, no. 4, pp. 711–733, 2006.

[289] D. J. Hruschka, M. H. Christiansen, R. A. Blythe, W. Croft, P. Heggarty, S. S. Mufwene, J. B. Pierrehumbert, and S. Poplack, "Building social cognitive models of language change", *Trends in cognitive sciences*, vol. 13, no. 11, pp. 464–469, 2009.

[290] V. Loreto and L. Steels, "Social dynamics: Emergence of language", *Nature Physics*, vol. 3, no. 11, pp. 758–760, 2007.

[291] W. J. Sutherland, "Parallel extinction risk and global distribution of languages and species", *Nature*, vol. 423, no. 6937, pp. 276–279, 2003.

[292] T. Raducha and T. Gubiec, "Predicting language diversity with complex networks", *PloS one*, vol. 13, no. 4, e0196593, 2018.

[293] J. Terrell, "Human biogeography in the solomon islands", *Fieldiana. Anthropology*, vol. 68, no. 1, pp. 1–47, 1977.

[294] B. F. Grimes *et al.*, *Ethnologue language name index*. Summer Institute of linguistics, 1996. [Online]. Available: http://www.ethnologue.com/13/names/.

[295] M. S. Dryer and M. Haspelmath, Eds., *Wals online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. [Online]. Available: http://wals.info/.

[296] J. Grieve, D. Speelman, and D. Geeraerts, "A statistical method for the identification and aggregation of regional linguistic variation", *Language Variation and Change*, vol. 23, no. 2, pp. 193–221, 2011.

[297] C. Gooskens, "The contribution of linguistic factors to the intelligibility of closely related languages", *Journal of Multilingual and multicultural development*, vol. 28, no. 6, pp. 445–467, 2007.

[298] M. J. Munro, T. M. Derwing, and S. L. Morton, "The mutual intelligibility of l2 speech", *Studies in second language acquisition*, vol. 28, no. 1, pp. 111–131, 2006.

[299] U. N. D. of Economic and S. Affairs, *World population prospects: The 2015 revision*. United Nations, 2015. [Online]. Available: https://esa.un.org/unpd/wpp/.

[300] J. Eisenstein, B. O'Connor, N. A. Smith, and E. P. Xing, "Diffusion of lexical change in social media", *PloS one*, vol. 9, no. 11, e113114, 2014.

[301] U. Schmitz, "Semiotic economy, growth of mass media discourse, and change of written language through multimodal techniques.", *Mediatization and Sociolinguistic Change*, vol. 36, p. 279, 2014.

[302] D. Sayers, "The mediated innovation model: A framework for researching media influence in language change", *Journal of Sociolinguistics*, vol. 18, no. 2, pp. 185–212, 2014.

[303] S. Zhou, G. Hu, Z. Zhang, and J. Guan, "An empirical study of chinese language networks", *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 12, pp. 3039–3047, 2008.

[304] D. R. Amancio, O. N. Oliveira Jr, and L. d. F. Costa, "Structure–semantics interplay in complex networks and its effects on the predictability of similarity in texts", *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 18, pp. 4406–4419, 2012.

[305] D. R. Amancio, "A complex network approach to stylometry", *PLoS One*, vol. 10, no. 8, e0136076, 2015.

[306] J. Burridge, "Spatial evolution of human dialects", *Physical Review X*, vol. 7, no. 3, p. 031 008, 2017.

[307] J. Fort and J. Pérez-Losada, "Can a linguistic serial founder effect originating in africa explain the worldwide phonemic cline?", *Journal of The Royal Society Interface*, vol. 13, no. 117, p. 20 160 185, 2016.

[308] J. Nerbonne, "Measuring the diffusion of linguistic change", *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1559, pp. 3821–3828, 2010.

[309] K. Huang, "Statistical mechanics", *Stme*, 1987.

[310] C. E. Shannon, "A mathematical theory of communication", *Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[311] A. I. Khinchin, "Mathematical foundations of information theory", *Dover Publications, New York*, 1957.

[312] E. T. Jaynes, "Information theory and statistical mechanics", *Physical review*, vol. 106, no. 4, p. 620, 1957.

[313] E. T. Jaynes *et al.*, *Probability theory: The logic of science*. Cambridge university press, 2003.

[314] E. T. Jaynes, "On the rationale of maximum-entropy methods", *Proceedings of the IEEE*, vol. 70, no. 9, pp. 939–952, 1982.

[315] A. Eddington, *New pathways in science: Messenger lectures (1934)*. Cambridge University Press, 2012.

[316] E. T. Jaynes, "Information theory and statistical mechanics. ii", *Physical review*, vol. 108, no. 2, p. 171, 1957.

[317] J. Peterson, P. D. Dixit, and K. A. Dill, "A maximum entropy framework for nonexponential distributions", *Proceedings of the National Academy of Sciences*, vol. 110, no. 51, pp. 20 380–20 385, 2013.

[318] A. De Martino and D. De Martino, "An introduction to the maximum entropy approach and its application to inference problems in biology", *Heliyon*, vol. 4, no. 4, e00596, 2018.

[319] A. Rényi *et al.*, "On measures of entropy and information", in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, The Regents of the University of California, 1961.

[320] C. Tsallis, "Possible generalization of boltzmann-gibbs statistics", *Journal of statistical physics*, vol. 52, no. 1-2, pp. 479–487, 1988.

[321] C. Beck and E. G. Cohen, "Superstatistics", *Physica A: Statistical mechanics and its applications*, vol. 322, pp. 267–275, 2003.

[322] R. Hanel, S. Thurner, and M. Gell-Mann, "Generalized entropies and the transformation group of superstatistics", *Proceedings of the National Academy of Sciences*, vol. 108, no. 16, pp. 6390–6394, 2011.

[323] R. Hanel and S. Thurner, "Generalized boltzmann factors and the maximum entropy principle: Entropies for complex systems", *Physica A: Statistical Mechanics and its Applications*, vol. 380, pp. 109–114, 2007.

[324] R. Hanel, S. Thurner, *et al.*, "A comprehensive classification of complex statistical systems and an axiomatic derivation of their entropy and distribution functions", *EPL (Europhysics Letters)*, vol. 93, no. 2, p. 20 006, 2011.

[325] J. Shore and R. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy", *IEEE Transactions on information theory*, vol. 26, no. 1, pp. 26–37, 1980.

[326] T. Oikonomou and G. B. Bagci, "Misusing the entropy maximization in the jungle of generalized entropies", *Physics Letters A*, vol. 381, no. 4, pp. 207–211, 2017.

[327] S. G. Brush, "History of the lenz-ising model", *Reviews of modern physics*, vol. 39, no. 4, p. 883, 1967.

[328] A. Coolen, R. Penney, and D Sherrington, "Coupled dynamics of fast spins and slow interactions: An alternative perspective on replicas", *Physical Review B*, vol. 48, no. 21, p. 16 116, 1993.

[329] A. Allahverdyan and K. Petrosyan, "Statistical networks emerging from link-node interactions", *EPL (Europhysics Letters)*, vol. 75, no. 6, p. 908, 2006.

[330] H. Wioland, F. G. Woodhouse, J. Dunkel, and R. E. Goldstein, "Ferromagnetic and antiferromagnetic order in bacterial vortex lattices", *Nature physics*, vol. 12, no. 4, pp. 341–345, 2016.

[331] D. Stauffer and H. E. Stanley, "From newton to mandelbrot", *Translated by S. Miyajima and H. Nishihara (Asakura Book Publishers, Tokyo, 1993) p*, vol. 197, pp. 164–169, 1996.

[332] R. Peierls, "On ising's model of ferromagnetism", in *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, vol. 32, 1936, pp. 477–481.

[333] L. Onsager, "Crystal statistics. i. a two-dimensional model with an order-disorder transition", *Physical Review*, vol. 65, no. 3-4, p. 117, 1944.

[334] B. A. Cipra, "An introduction to the ising model", *The American Mathematical Monthly*, vol. 94, no. 10, pp. 937–959, 1987.

[335] M Newman and G Barkema, *Monte carlo methods in statistical physics chapter 1-4*. Oxford University Press: New York, USA, 1999.

[336] H. E. Stanley, *Introduction to phase transitions and critical phenomena*. International Series of Monographs on Physics, OUP USA, 1971.

[337] C. P. Herrero, "Ising model in small-world networks", *Physical Review E*, vol. 65, no. 6, p. 066 110, 2002.

[338] A. Aleksiejuk, J. A. Hołyst, and D. Stauffer, "Ferromagnetic phase transition in barabási–albert networks", *Physica A: Statistical Mechanics and its Applications*, vol. 310, no. 1, pp. 260–266, 2002.

[339] B. Tadić, K. Malarz, and K. Kułakowski, "Magnetization reversal in spin patterns with complex geometry", *Physical review letters*, vol. 94, no. 13, p. 137 204, 2005.

[340] C. Biely, R. Hanel, and S. Thurner, "Socio-economical dynamics as a solvable spin system on co-evolving networks", *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 67, no. 3, pp. 285–289, 2009.

[341] J. Toruniewska, K. Suchecki, and J. A. Hołyst, "Unstable network fragmentation in co-evolution of potts spins and system topology", *Physica A: Statistical Mechanics and its Applications*, vol. 460, pp. 1–15, 2016.

[342] J. Berg and M. Lässig, "Correlated random networks", *Physical review letters*, vol. 89, no. 22, p. 228 701, 2002.

[343] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Statistical mechanics of topological phase transitions in networks", *Physical Review E*, vol. 69, no. 4, p. 046 117, 2004.

[344] Z. Burda, J. D. Correia, and A. Krzywicki, "Statistical ensemble of scale-free random graphs", *Physical Review E*, vol. 64, no. 4, p. 046 118, 2001.

[345] Z Burda and A. Krzywicki, "Uncorrelated random networks", *Physical Review E*, vol. 67, no. 4, p. 046 118, 2003.

[346] T. Raducha, M. Wilinski, T. Gubiec, and H. E. Stanley, "Statistical mechanics of a coevolving spin system", *Physical Review E*, vol. 98, no. 3, p. 030 301, 2018.

[347] 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*, vol. 21, no. 6, pp. 1087–1092, 1953.

[348] I. Beichl and F. Sullivan, "The metropolis algorithm", *Computing in Science & Engineering*, vol. 2, no. 1, pp. 65–69, 2000.

[349] A. E. Bryson, *Applied optimal control: Optimization, estimation and control*. CRC Press, 1975.