

UNIVERSITÀ DEGLI STUDI DI CATANIA
SCUOLA SUPERIORE DI CATANIA

Roberta Sinatra

EVOLUTIONARY GAMES ON GRAPHS:
THE ULTIMATUM GAME

DIPLOMA DI LICENZA

RELATORI:

PROF. VITO LATORA

PROF. YAMIR MORENO VEGA

DOTT. JESUS GOMEZ GARDEÑES

ANNO ACCADEMICO 2007/2008

*Alle Donne del mondo
e a quelle della mia vita:
Giusy, Claudia e Emma*

Martin J. Osborne in his preface to *An introduction to Game Theory* (Oxford University Press, 2004) writes:

The English language lacks a third person singular pronoun widely interpreted to be sex neutral. In particular, many experiments have shown that “he” is not neutral, a finding consistent with the observation that whereas people may say “when an airplane pilot is working, he needs to concentrate”, they do not usually say “when a flight attendant is working, he needs to concentrate” or “when a secretary is working, he needs to concentrate”. To quote the *American Heritage Dictionary*, “Thus *he* is not really a gender-neutral pronoun; rather it refers to a male who is to be taken as the representative member of the group referred to by its antecedent. The traditional usage, then, is not simply a grammatical convention; it also suggests a particular pattern of thought”. Like many writers, I regard as unacceptable the bias implicit in the use “he” for individuals of unspecified sex. [...] A common solution has been to use “they”. In some contexts this usage sounds natural, but in others it does not; it can also create ambiguity when the pronoun follows references to more than one person. I choose a different solution: I use “she” exclusively. Obviously this usage, like that of “he” is not sex neutral, but it may help to counterbalance the widespread use of “he”, and seems unlikely to do any harm.

As I absolutely agree with Osborne, I follow his choice and use also “she” exclusively in this work.

It pays also a humble tribute to women, as too many times their role is forgotten and their importance underestimated.

Contents

Introduction	1
1 Complex networks in pills	3
1.1 Complex networks are all around us	3
1.2 The structure of complex networks	5
1.2.1 Definitions and notations	5
1.2.2 Node degree, degree distributions and correlations	8
1.2.3 Shortest path lengths, diameter and betweenness	9
1.2.4 Clustering	11
1.3 Topology of real networks	13
1.3.1 The small-world property	14
1.3.2 Scale-free degree distributions	15
1.4 Networks models	16
1.4.1 Random graphs	17
1.4.2 Small-world networks	19
1.4.3 Barabási-Albert model	21
1.4.4 Holme-Kim model	23
2 An introduction to game theory and evolutionary games	26
2.1 Overview	27

2.1.1	Classical game theory and its assumptions	27
2.1.2	Post-Nash game theory	28
2.1.3	Bounded rationality and evolutionary game theory	30
2.1.4	Beyond evolutionary game theory	31
2.2	Games, payoff, strategies	32
2.2.1	Notation	32
2.2.2	Matrix form	33
2.2.3	Strategies	33
2.2.4	Nash Equilibrium	34
2.3	Repeated Games	35
2.3.1	Subgame perfect Nash equilibrium	36
3	Ultimatum game on networks	37
3.1	Merging evolutionary game theory and complex networks	37
3.2	Ultimatum Game: differences between theory and experiments	38
3.3	Standard ultimatum game on networks: our model	40
3.4	Ultimatum Game on networks	42
3.4.1	Configuration $q_i = p_i$	43
3.4.2	Configuration $q_i = 1 - p_i$	48
	A semi-analytical approach	50
3.4.3	Scale-free networks with variable clustering coefficient	52
3.5	Eguíluz and Tessone's model	61
3.5.1	The model	61
3.5.2	Results	62
3.5.3	Discussion	63

Contents	iii
----------	-----

Conclusions	68
Acknowledgements	71
Bibliography	74

Introduction

*Let us try to teach generosity and altruism,
because we are born selfish.*

RICHARD DAWKINS

Since long time ago, great minds - Aristotle, Hobbes, Kant, Darwin, to cite some names - have been contemplating and speculating about a fundamental subject for life: the cooperative behaviour. The question initially was addressed mainly from a philosophical view point, but then interested diverse disciplines like political science, economics, sociology, anthropology, psychology and evolutionary biology. Nowadays also physics and mathematics deal with theories and models concerning cooperation. Definitely cooperation is something that can be encountered in many different environments: animal reign, gene-gene competition, economic equilibria, sociological systems are just some examples.

Nevertheless, cooperative behaviour can exist only as the counter-party to selfishness. We report the words of Richard Dawkins, in his discussion related to genes evolution but easily extendible also to other contexts, in order to better understand the role of this dichotomy [1]:

...A predominant quality to be expected in a successful gene is ruthless selfishness. This gene selfishness will usually give rise to selfishness in individual behaviour. However there are special circumstances in

which a gene can achieve its own selfish goals best by fostering a limited form of altruism at the level of individual animals. ‘Special’ and ‘limited’ are important words in the last sentence. Much as we might wish to believe otherwise, universal love and the welfare of the species as a whole are concepts that simply do not make evolutionary sense.

This is a key concept in the framework we will use to address the problem of the dualism cooperation-altruism. The theory we will deal with is the “evolutionary game theory” and we’ll study some particular application in the context of complex networks.

This work is structured as follows: the first chapter constitutes a brief introduction to the world of complex networks; the second chapter aims to give some key concepts of classical game theory and evolutionary game theory; the third and last chapter contains some original results we have found by implementing a particular evolutionary game, the Ultimatum Game, on different kinds of networks [2].

Chapter 1

Complex networks in pills

1.1 Complex networks are all around us

The last decade has witnessed the birth of a new movement of interest and research in the study of *complex networks*, i.e. networks whose structure is irregular, complex and dynamically evolving in time, with the main focus moving from the analysis of small networks to that of systems with thousands or millions of nodes, and with a renewed attention to the properties of networks of dynamical units. This flurry of activity, triggered by two seminal papers, that by Watts and Strogatz on small-world networks, appeared in Nature in 1998, and that by Barabási and Albert on scale-free networks appeared one year later in Science, has seen the physics' community among the principal actors, and has been certainly induced by the increased computing powers and by the possibility to study the properties of a plenty of large databases of real networks. These include transportation networks, phone call networks, the Internet and the World Wide Web, the actors' collaboration network in movie databases, scientific coauthorship and citation networks from the Science Citation Index, but also systems of interest in biology and medicine, as neural networks or genetic, metabolic and protein networks.

The massive and comparative analysis of networks from different fields has produced a series of unexpected and dramatic results. The first issue that has been faced is certainly structural. The research on complex networks begun with the effort of defining new concepts and measures to characterize the topology of real networks. The main result has been the identification of a series of unifying principles and statistical properties common to most of the real networks considered.

These empirical findings have initiated a revival of network modelling, since the models proposed in mathematical graph theory turned out to be very far from the real needs. Scientists had to do with the development of new models to mimic the growth of a network and to reproduce the structural properties observed in real topologies. The structure of a real network is the result of the continuous evolution of the forces that formed it, and certainly affects the function of the system. So that this stage of the research was motivated by the expectancy that understanding and modelling the structure of a complex network would lead to a better knowledge of its evolutionary mechanisms, and to a better cottoning on its dynamical and functional behavior.

At the same time, it outcropped for the first time the possibility of studying the dynamical behavior of large assemblies of dynamical systems interacting via complex topologies, as the ones observed empirically. This led to a series of evidences pointing to the crucial role played by the network topology in determining the emergence of collective dynamical behavior, such as synchronization, or in governing the main features of relevant processes that take place in complex networks, such as the spreading of epidemics, information and rumors.

1.2 The structure of complex networks

In this section we will introduce definitions and notations, and discuss the basic quantities used to describe the topology of a network. Then, we shall move to the analysis of the properties observed in real networks, and provide the reader with a brief review of the models motivated by the empirical observations.

1.2.1 Definitions and notations

Graph theory [4] is the natural framework for the exact mathematical treatment of complex networks and, formally, a complex network can be represented as a graph. A *undirected* (*directed*) graph $G = (\mathcal{N}, \mathcal{L})$ consists of two sets \mathcal{N} and \mathcal{L} , such that $\mathcal{N} \neq \emptyset$ and \mathcal{L} is a set of unordered (ordered) pairs of elements of \mathcal{N} . The elements of $\mathcal{N} \equiv \{n_1, n_2, \dots, n_N\}$ are the *nodes* (or *vertices*, or *points*) of the graph G , while the elements of $\mathcal{L} \equiv \{l_1, l_2, \dots, l_K\}$ are its *links* (or *edges*, or *lines*). The number of elements in \mathcal{N} and \mathcal{L} are denoted by N and K , respectively. In the following parts, we will indicate a graph as $G(N, K) = (\mathcal{N}, \mathcal{L})$, or simply $G(N, K)$ or $G_{N,K}$, whenever it is necessary to emphasize the number of nodes and links in the graph.

A node is usually referred to by its order i in the set \mathcal{N} . In a undirected graph, each of the links is defined by a couple of nodes i and j , and is denoted as (i, j) or l_{ij} . The link is said to be *incident* in nodes i and j , or to join the two nodes; the two nodes i and j are referred to as the *end-nodes* of link (i, j) . Two nodes joined by a link are referred to as *adjacent* or *neighboring*. In a directed graph, the order of the two nodes is important: l_{ij} stands for a link from i to j , and $l_{ij} \neq l_{ji}$. The usual way to picture a graph is by drawing a dot for each node and joining two dots by a line if the two corresponding nodes are connected by a link. How these dots and lines are drawn is irrelevant, and the only thing that

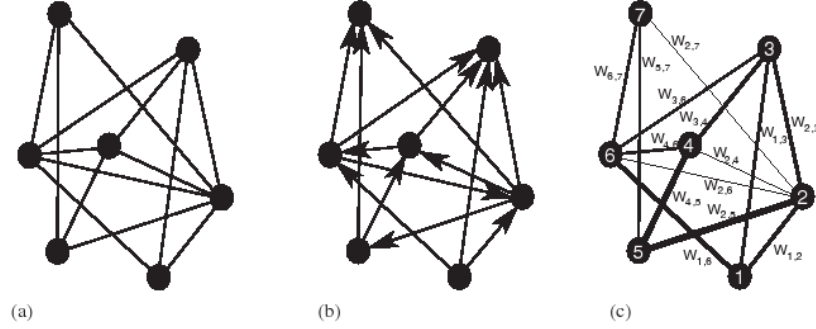


Figure 1.1: Graphical representation of an undirected (a), a directed (b), and a weighted undirected (c) graph with $N = 7$ nodes and $K = 14$ links. In the directed graph, adjacent nodes are connected by arrows, indicating the direction of each link. In the weighted graph, the values $w_{i,j}$ reported on each link indicate the weights of the links, and are graphically represented by the link thicknesses.

matters is which pairs of nodes form a link and which ones do not. Examples of a undirected graph and of a directed graph, both with $N = 7$ and $K = 14$, are shown in Fig. 1.1 (a) and (b), respectively. Note that the picture does not contain *loops*, i.e. links from a node to itself, nor *multiple edges*, i.e. couples of nodes connected by more than one link, since these elements are not allowed by the standard definition of graph given above. Graphs with either of these elements are called *multigraphs* [4]. In this work we deal only with undirected graphs, and not on multigraphs

For a graph G of size N , the number of edges K is at least 0 and at most $N(N-1)/2$ (when all the nodes are pairwise adjacent). G is said to be *sparse* if $K \ll N^2$ and *dense* if $K = \mathcal{O}(N^2)$. A graph $G_{N,K}$ is said a *complete* N -graph if $K = \binom{N}{2} = N(N-1)/2$, and is denoted by K_N .

A *subgraph* $G' = (\mathcal{N}', \mathcal{L}')$ of $G = (\mathcal{N}, \mathcal{L})$ is a graph such that $\mathcal{N}' \subseteq \mathcal{N}$ and $\mathcal{L}' \subseteq \mathcal{L}$. If G' contains all links of G that join two nodes in \mathcal{N}' , then G' is said to be the *subgraph induced* by \mathcal{N}' and is denoted as $G' = G[\mathcal{N}']$. A subgraph is said

to be *maximal* with respect to a given property if it cannot be extended without loosing that property. Of particular relevance for some of the definitions given in the following subsections is the subgraph of the neighbors of a given node i , denoted as G_i . G_i is defined as the subgraph induced by \mathcal{N}_i , the set of nodes adjacent to i , i.e. $G_i = G[\mathcal{N}_i]$.

A central concept in graph theory is that of reachability of two different nodes of a graph. In fact, two nodes that are not adjacent may nevertheless be reachable from one to the other. A *walk* from node i to node j is an alternating sequence of nodes and edges (a sequence of adjacent nodes) that begins with i and ends with j . The length of the walk is defined as the number of edges in the sequence. A *trail* is a walk in which no edge is repeated. A *path* is a walk in which no node is visited more than once. The walk of minimal length between two nodes is known as *shortest path* or *geodesic*. A *cycle* is a closed walk, of at least three nodes, in which no edge is repeated. A cycle of length k is usually said a k -cycle and denoted as C_k . C_3 is a triangle ($C_3 = K_3$), C_4 is called a quadrilater, C_5 a pentagon, and so on. A graph is said to be *connected* if, for every pair of distinct nodes i and j , there is a path from i to j , otherwise it is said *unconnected* or *disconnected*. A *component* of the graph is a maximally connected induced subgraph. A *giant component* is a component whose size is of the same order as N .

It is often useful to consider a matricial representation of a graph. A graph $G = (\mathcal{N}, \mathcal{L})$ can be completely described by giving the *adjacency* (or *connectivity*) matrix \mathcal{A} , a $N \times N$ square matrix whose entry a_{ij} ($i, j = 1, \dots, N$) is equal to 1 when the link l_{ij} exists, and zero otherwise. This is thus a symmetric matrix for undirected graphs. An alternative possibility is to use the *incidence* matrix \mathcal{B} , a $N \times K$ matrix whose entry b_{ik} is equal to 1 whenever the node i is incident with the link l_k , and zero otherwise.

1.2.2 Node degree, degree distributions and correlations

The *degree* (or *connectivity*) k_i of a node i is the number of edges incident with the node, and is defined in terms of the adjacency matrix \mathcal{A} as:

$$k_i = \sum_{j \in \mathcal{N}} a_{ij} . \quad (1.1)$$

If the graph is directed, the degree of the node has two components: the number of outgoing links $k_i^{out} = \sum_j a_{ij}$ (referred to as the out-degree of the node), and the number of ingoing links $k_i^{in} = \sum_j a_{ji}$ (referred to as the in-degree of the node). The total degree is then defined as $k_i = k_i^{out} + k_i^{in}$. A list of the node degrees of a graph is called the *degree sequence*.

The most basic topological characterization of a graph G can be obtained in terms of the *degree distribution* $P(k)$, defined as the probability that a node chosen uniformly at random has degree k or, equivalently, as the fraction of nodes in the graph having degree k . Alternatively, the degree distribution is denoted as P_k , or p_k , to indicate that the variable k assumes non negative integer values.

In the case of directed networks one needs to consider two distributions, $P(k^{in})$ and $P(k^{out})$. Information on how the degree is distributed among the nodes of a undirected network can be obtained either by a plot of $P(k)$, or by the calculation of the moments of the distribution. The n -moment of $P(k)$ is defined as:

$$\langle k^n \rangle = \sum_k k^n P(k). \quad (1.2)$$

The first moment $\langle k \rangle$ is the mean degree of G . The second moment measures the fluctuations of the connectivity distribution. The degree distribution completely determines the statistical properties of uncorrelated networks. However a large number of real networks are *correlated* in the sense that the probability that

a node of degree k is connected to another node of degree, say k' , depends on k . In these cases, it is necessary to introduce the *conditional* probability $P(k'|k)$, being defined as the probability that a link from a node of degree k points to a node of degree k' . $P(k'|k)$ satisfies the normalization $\sum_{k'} P(k'|k) = 1$, and the degree detailed balance condition $kP(k'|k)P(k) = k'P(k|k')P(k')$ [6]. For uncorrelated graphs, in which $P(k'|k)$ does not depend on k , the detailed balance condition and the normalization give $P(k'|k) = k'P(k')/\langle k \rangle$.

1.2.3 Shortest path lengths, diameter and betweenness

Shortest paths play an important role in the transport and communication within a network. For such a reason, shortest paths have also played an important role in the characterization of the internal structure of a graph [7]. It is useful to represent all the shortest path lengths of a graph G as a matrix \mathcal{D} in which the entry d_{ij} is the length of the geodesic from node i to node j . The maximum value of d_{ij} is called the *diameter* of the graph, and will be indicated in the following as $\text{Diam}(G)$. A measure of the typical separation between two nodes in the graph is given by the *average shortest path length*, also known as *characteristic path length*, defined as the mean of geodesic lengths over all couples of nodes [9]:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} d_{ij} . \quad (1.3)$$

A problem with this definition is that L diverges if there are disconnected components in the graph. One possibility to avoid the divergence is to limit the summation in formula (1.3) only to couples of nodes belonging to the largest connected component [9]. An alternative approach, that is useful in many cases, is to consider the harmonic mean [11] of geodesic lengths, and to define the so-

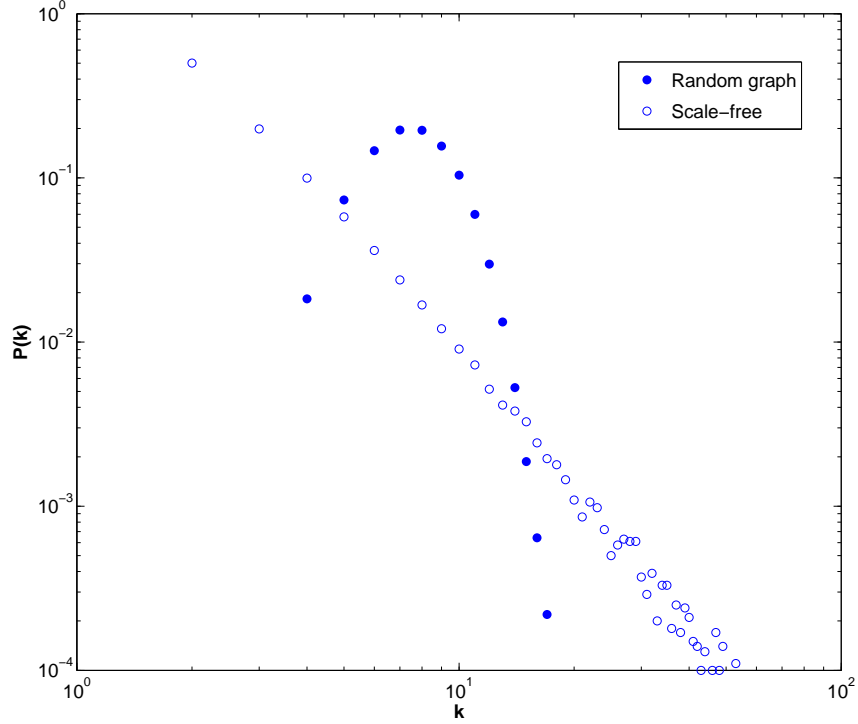


Figure 1.2: An important graph property is the degree distribution function $P(k)$, that describes the probability to find a node with k edges. A *random graph* is constructed by randomly linking N nodes with K edges, and has a Poissonian degree distribution ($P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$). That means that the majority of nodes have a degree close to the average degree $\langle k \rangle$. A *scale-free graph* is instead characterized by a power-law degree distribution ($P(k) = Ak^{-\gamma}$ usually with $2 < \gamma < 3$). A power-law distribution appears as a straight line in a double-logarithmic plot. In a scale-free graph, low degree nodes are the most frequent ones, but there are also a few highly connected nodes, usually called *hubs*, not present in a random graph.

called *efficiency* of G as [12, 13]:

$$E = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} \frac{1}{d_{ij}}. \quad (1.4)$$

Such a quantity is an indicator of the traffic capacity of a network, and avoids

the divergence of formula (1.3), since any couple of nodes belonging to disconnected components of the graph yields a contribution equal to zero to the summation in formula (1.4).

The communication of two non-adjacent nodes, say j and k , depends on the nodes belonging to the paths connecting j and k . Consequently, a measure of the relevance of a given node can be obtained by counting the number of geodesics going through it, and defining the so-called *node betweenness*. Together with the degree and the *closeness* of a node (defined as the inverse of the average distance from all other nodes), the betweenness is one of the standard *measures of node centrality*, originally introduced to quantify the importance of an individual in a social network [7, 10, 14]. More precisely, the *betweenness* b_i of a node i , sometimes referred to also as *load*, is defined as [7, 10]:

$$b_i = \sum_{j,k \in N, j \neq k} \frac{n_{jk}(i)}{n_{jk}}, \quad (1.5)$$

where n_{jk} is the number of shortest paths connecting i and j , while $n_{jk}(i)$ is the number of shortest paths connecting i and j and passing through i .

The concept of betweenness can be extended also to the edges. The *edge betweenness* is defined as the number of shortest paths between pairs of nodes that run through that edge [15].

1.2.4 Clustering

Clustering, also known as transitivity, is a typical property of acquaintance networks, where two individuals with a common friend are likely to know each other [7]. In terms of a generic graph G , transitivity means the presence of a high number of triangles. This can be quantified by defining the *transitivity* T of the graph as the relative number of transitive triples, i.e. the fraction of connected

triples of nodes (triads) which also form triangles [8]:

$$T = \frac{3 \times \# \text{ of triangles in } G}{\# \text{ of connected triples of vertices in } G} \quad (1.6)$$

The factor 3 in the numerator compensates for the fact that each complete triangle of three nodes contributes three connected triples, one centered on each of the three nodes, and ensures that $0 \leq T \leq 1$, with $T = 1$ for K_N .

An alternative possibility is to use the graph *clustering coefficient* C , a measure introduced by Watts and Strogatz [9], and defined as follows. A quantity c_i (the local clustering coefficient of node i) is first introduced, expressing how likely $a_{jm} = 1$ for two neighbors j and m of node i . Its value is obtained by counting the actual number of edges (denoted by e_i) in G_i (the subgraph of neighbors of i). Notice that G_i can be, in some cases, unconnected. The local clustering coefficient is defined as the ratio between e_i and $k_i(k_i - 1)/2$, the maximum possible number of edges in G_i [9]:

$$c_i = \frac{2e_i}{k_i(k_i - 1)} = \frac{\sum_{j,m} a_{ij}a_{jm}a_{mi}}{k_i(k_i - 1)} . \quad (1.7)$$

The clustering coefficient of the graph is then given by the average of c_i over all the nodes in G :

$$C = \langle c \rangle = \frac{1}{N} \sum_{i \in \mathcal{N}} c_i . \quad (1.8)$$

By definition, $0 \leq c_i \leq 1$, and $0 \leq C \leq 1$. It is also useful to consider $c(k)$, the clustering coefficient of a connectivity class k , which is defined as the average of c_i taken over all nodes with a given degree k . An alternative measure of the clustering properties of G is the *local efficiency*, defined as [12, 13]:

$$E_{loc} = \frac{1}{N} \sum_{i \in \mathcal{N}} E(G_i) , \quad (1.9)$$

where $E(G_i)$ is the efficiency of G_i , evaluated by formula (1.4).

1.3 Topology of real networks

Many systems in nature and in technology are made by a large number of highly interconnected dynamical units. Coupled biological and chemical systems, neural networks, social interacting species, the Internet or the World Wide Web, are only a few such examples. The first approach to capture the global properties of such systems is to model them as graphs whose nodes represent the dynamical units (for instance the neurons in the brain or the individuals in a social system) and the links stand for the interactions between the units. Of course, this is a very strong approximation, since it means translating the interaction between two dynamical units, which is usually depending on time, space and many more other details, into a simple binary number: the existence or not of a link between the two corresponding nodes. Nevertheless, in many cases of practical interest, such an approximation provides a simple but still very informative representation of the entire system.

During the last decade, the grown availability of large databases, the optimized rating of computing facilities, as well as the development of powerful and reliable data analysis tools, have constituted a better and better machinery to explore the topological properties of several networked systems from the real world. This has allowed to study the topology of the interactions in a large variety of systems as diverse as communication, social and biological systems. The main outcome of this activity has been to reveal that, despite the inherent differences, most of the real networks are characterized by the same topological properties, as for instance relatively small characteristic path lengths, high clustering coefficients, fat tailed shapes in the degree distributions, degree correlations, and the

presence of motifs and community structures. All these features make real networks radically different from regular lattices and random graphs, the standard models studied in mathematical graph theory. This has led to a large attention towards the understanding of the evolution mechanisms that have shaped the topology of a network, and to the design of new models retaining the most significant properties empirically observed.

1.3.1 The small-world property

The study of several dynamical processes over real networks has pointed out the existence of shortcuts, i.e. bridging links that connect different areas of the networks, thus speeding up the communication among otherwise distant nodes.

In regular hypercubic lattices in D dimensions, the mean number of vertices one has to pass by in order to reach an arbitrarily chosen node, grows with the lattice size as $N^{1/d}$. Conversely, in most of the real networks, despite of their often large size, there is a relatively short path between any two nodes. This feature is known as the *small-world* property and is mathematically characterized by an average shortest path length L , defined as in equation (1.3), that depends at most logarithmically on the network size N [9]. This property was first investigated, in the social context, by Milgram in the 1960s in a series of experiments to estimate the actual number of steps in a chain of acquaintances [7, 16, 17].

In its first experiment, Milgram asked randomly selected people in Nebraska to send letters to a distant target individual in Boston, identified only by his name, occupation and rough location. The letters could only be sent to someone whom the current holder knew by first name, and who was presumably closer to the final recipient. Milgram kept track of the paths followed by the letters and of the demographic characteristics of their handlers. Although the common guess was that it might take hundreds of these steps for the letters to reach their final

destination, Milgram's surprising result was that the number of links needed to reach the target person had an average value of just six.

The small-world property has been observed in a variety of other real networks, including biological and technological ones, and is an obvious mathematical property in some network models, as for instance in random graphs. At variance with random graphs, the small-world property in real networks is often associated with the presence of clustering, denoted by high values of the clustering coefficient, defined as in equation (1.8). For this reason, Watts and Strogatz, in their seminal paper, have proposed to define small-world networks as those networks having both a small value of L , like random graphs, and a high clustering coefficient C , like regular lattices [9]. In the efficiency-based formalism, such a definition corresponds to networks having a high value of global efficiency E_{glob} , defined as in equation (1.4), and a high value of local efficiency E_{loc} , defined as in equation (1.9), i.e. to networks extremely efficient in exchanging information both at a global and at a local scale [12, 13].

1.3.2 Scale-free degree distributions

The usual case until a few years ago was that of homogeneous networks. Homogeneity in the interaction structure means that almost all nodes are topologically equivalent, like in regular lattices or in random graphs. In these latter ones, for instance, each of the $N(N-1)/2$ possible links is present with equal probability, and thus the degree distribution is binomial or Poisson in the limit of large graph size (see sec. 1.4.1). It is not startling then that, when the scientists approached the study of real networks from the available databases, it was considered reasonable to find degree distributions localized around an average value, with a well-defined average of quadratic fluctuations. In contrast with all the expectancies, it was found that most of the real networks display power law shaped degree

distribution $P(k) \sim Ak^{-\gamma}$, with exponents varying in the range $2 < \gamma < 3$. The average degree $\langle k \rangle$ in such networks is therefore well defined and bounded, while the variance $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$ is dominated by the second moment of the distribution that diverges with the upper integration limit k_{max} as:

$$\langle k^2 \rangle = \int_{k_{min}}^{k_{max}} k^2 P(k) \sim k_{max}^{3-\gamma}. \quad (1.10)$$

Such networks have been named *scale-free* networks [19, 18], because power-laws have the property of having the same functional form at all scales. In fact, power-laws are the only functional form $f(x)$ that remains unchanged, apart from a multiplicative factor, under a rescaling of the independent variable x , being the only solution to the equation $f(\alpha x) = \beta f(x)$. Power-laws have a particular role in statistical physics because of their connections to phase transitions and fractals. In the following, when referring to scale-free networks, we will denote the class of graphs with power-laws in the degree distribution. Of course, this does not necessarily implies that such graphs are scale-free with respect to other measurable structural properties. These networks, having a highly inhomogeneous degree distribution, result in the simultaneous presence of a few nodes (the *hubs*) linked to many other nodes, and a large number of poorly connected elements.

1.4 Networks models

What has been said in the previous section clearly motivates the introduction of new concepts and models. In this section, we focus on some mathematical models of networks and some algorithms which have been used for the application described in the last chapter of this work.

1.4.1 Random graphs

The systematic study of random graphs was initiated by Erdős and Rényi in 1959 with the original purpose of studying, by means of probabilistic methods, the properties of graphs as a function of the increasing number of random connections. The term random graph refers to the disordered nature of the arrangement of links between different nodes. In their first article, Erdős and Rényi proposed a model to generate random graphs with N nodes and K links, that we will henceforth call *Erdős and Rényi (ER) random graphs* and denote as $G_{N,K}^{ER}$. Starting with N disconnected nodes, ER random graphs are generated by connecting couples of randomly selected nodes, prohibiting multiple connections, until the number of edges equals K [20]. We emphasize that a given graph is only one outcome of the many possible realizations, an element of the statistical ensemble of all possible combinations of connections. For the complete description of $G_{N,K}^{ER}$ one would need to describe the entire statistical ensemble of possible realizations, that is, in the matricial representation, the ensemble of adjacency matrices. An alternative model for ER random graphs consists in connecting each couple of nodes with a probability $0 < p < 1$. This procedure defines a different ensemble, denoted as $G_{N,p}^{ER}$ and containing graphs with different number of links (not exactly K links). The two models coincide in the limit of large N . Notice that the limit $N \rightarrow \infty$ is taken at fixed $\langle k \rangle$, which corresponds to fixing $2K/N$ in the first model and $p(N-1)$ in the second one. Although the first model seems to be more pertinent to applications, analytical calculations are easier and usually are performed in the second model.

ER random graphs are the best studied among graph models, although they do not reproduce most of the properties of real networks discussed in section 1.3. The structural properties of ER random graphs vary as a function of p showing,

in particular, a dramatic change at a critical probability $p_c = \frac{1}{N}$, corresponding to a critical average degree $\langle k \rangle_c = 1$. Erdős and Rényi proved that:

- if $p < p_c$, then almost surely, i.e. with probability tending to one as N tends to infinity, the graph has no component of size greater than $\mathcal{O}(\ln N)$, and no component has more than one cycle;
- if $p = p_c$, then almost surely the largest component has size $\mathcal{O}(N^{2/3})$;
- if $p > p_c$, the graph has a component of $\mathcal{O}(N)$ with a number $\mathcal{O}(N)$ of cycles, and no other component has more than $\mathcal{O}(\ln N)$ nodes and more than one cycle.

The transition at p_c has the typical features of a second order phase transition. In particular, if one considers as order parameter the size of the largest component, the transition falls in the same universality class as that of the mean field percolation transitions. Erdős and Rényi studied the distribution of the minimum and maximum degree in a random graph [20], while the full degree distribution was derived later by Bollobás [4]. The probability that a node i has $k = k_i$ edges is the binomial distribution $P(k_i = k) = C_{N-1}^k p^k (1-p)^{N-1-k}$, where p^k is the probability for the existence of k edges, $(1-p)^{N-1-k}$ is the probability for the absence of the remaining $N-1-k$ edges, and $C_{N-1}^k = \binom{N-1}{k}$ is the number of different ways of selecting the end points of the k edges. Since all the nodes in a random graph are statistically equivalent, each of them has the same distribution, and the probability that a node chosen uniformly at random has degree k has the same form as $P(k_i = k)$. For large N , and fixed $\langle k \rangle$, the degree distribution is well approximated by a Poisson distribution:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} . \quad (1.11)$$

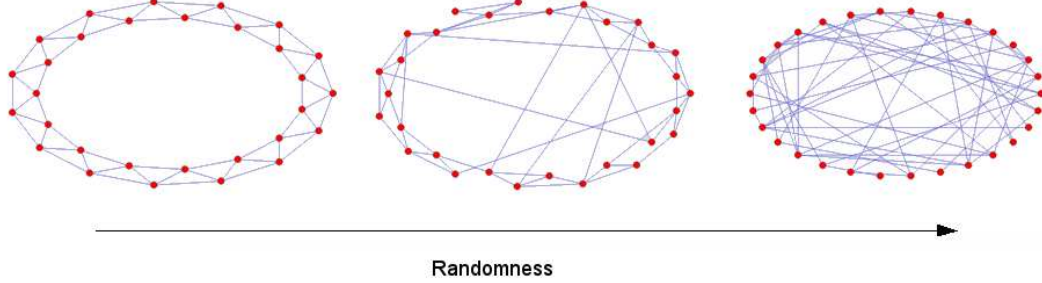


Figure 1.3: Small-world networks, as defined by Watts and Strogatz [9], have intermediate properties between regular lattices (such as the first graph in the figure) and random networks (such as the last graph in the figure). A regular lattice has high clustering but also a large average path length, while a random graph is characterized by a short path length together with a low clustering. A small-world network (in the middle in the figure) borrows a high clustering coefficient from the former and a short average path length from the latter.

For this reason, ER graphs are sometimes called *Poisson random graphs*. ER random graphs are, by definition, uncorrelated graphs, since the edges are connected to nodes regardless of their degree. Consequently, $P(k'|k)$ and $k_{nn}(k)$ are independent of k .

1.4.2 Small-world networks

The *Watts and Strogatz (WS)* model is a method to construct graphs, denoted as $G_{N,K}^{WS}$ having either the small-world property and a high clustering coefficient [9]. The model is based on a rewiring procedure of the edges implemented with a probability p . The starting point is a N nodes ring, in which each node is symmetrically connected to its $2m$ nearest neighbors for a total of $K = mN$ edges. Then, for every node, each link connected to a clockwise neighbor is rewired to a randomly chosen node with a probability p , and preserved with a probability $1 - p$. Notice that for $p = 0$ we have a regular lattice, while for $p = 1$ the model produces a random graph with the constraint that each node has a minimum

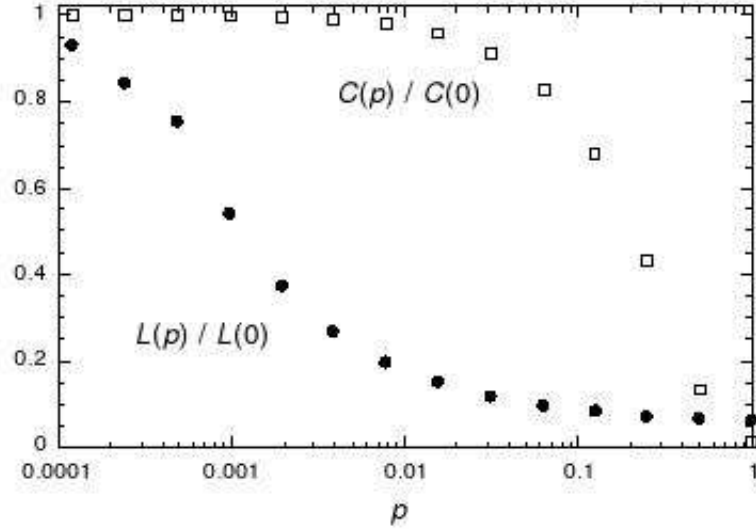


Figure 1.4: Characteristic path length $L(p)$ and clustering coefficient $C(p)$ for the family of randomly rewired graphs described in Fig. 1.3. Here L is defined as the number of edges in the shortest path between two vertices, averaged over all pairs of vertices. The clustering coefficient $C(p)$ is defined as described in (1.8). The data shown in the figure are averages over 20 random realizations of the rewiring process described in Fig.1.3, and have been normalized by the values $L(0)$, $C(0)$ for a regular lattice. All the graphs have $N = 1.000$ vertices and an average degree of $k = 10$ edges per vertex. Notice that a logarithmic horizontal scale has been used to resolve the rapid drop in $L(p)$, corresponding to the onset of the small-world phenomenon. During this drop, $C(p)$ remains almost constant at its value for the regular lattice, indicating that the transition to a small world is almost undetectable at the local level.

connectivity $k_{min} = m$. For intermediate values of p the procedure generates graphs with the small-world property and a non-trivial clustering coefficient.

The small-world property results from the immediate drop in $L(p)$ as soon as p is slightly larger than zero. This is because the rewiring of links creates long-range edges (shortcuts) that connects otherwise distant nodes. The effect of the rewiring procedure is highly nonlinear on L , and not only affects the nearest neighbors structure, but it also opens new shortest paths to the next-nearest neighbors and

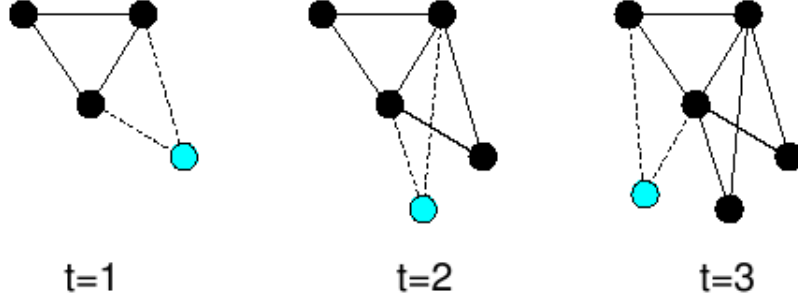


Figure 1.5: Illustration of the BA algorithm for $m_0 = 3$ and $m = 2$. At $t = 0$ we start with a complete graph of m_0 nodes. At every timestep a new node j is added, which is connected to $m = 2$ vertices, preferentially to the vertices with high connectivity, determined by the rule (1.12). Thus, at time t there are $m_0 + t$ vertices and $\binom{m_0}{2} + mt$ edges. At each time step, the new node n is in cyan, and the two new edges are drawn with dashed lines.

so on. Conversely, an edge redirected from a clustered neighborhood to another node has, at most, a linear effect on C . That is, the transition from a linear to a logarithmic behavior in $L(p)$ is faster than the one associated with the clustering coefficient $C(p)$. This leads to the appearance of a region of small (but nonzero) values of p , where one has both small path lengths and high clustering.

1.4.3 Barabási-Albert model

The *Barabási-Albert (BA)* model is a model of network growth inspired to the formation of the World Wide Web and is based on two basic ingredients: growth and preferential attachment [18]. The basic idea is that in the World Wide Web, sites with high degrees acquire new links at higher rates than low-degree nodes. More precisely, an undirected graph $G_{N,K}^{BA}$ with average degree $\langle k = 2m \rangle$ is constructed, starting with a complete graph with a small number $N(t = 0) = m_0$ of nodes and $K(t = 0) = \binom{m_0}{2}$ links. The graph growth obeys the two mechanisms illustrated in Fig. 1.5

- At each time step t ($t = 1, 2, 3, \dots$) a new node j is added. The new node has $m \leq m_0$ edges, that link j to m different nodes, already present in the system;
- When choosing the nodes to which the new node j connects, it is assumed that the probability $\Pi_{j \rightarrow i}$ that n will be connected to node i is linearly proportional to the degree k_i of node i , i.e.:

$$\Pi_{j \rightarrow i} = \frac{k_i}{\sum_l k_l} . \quad (1.12)$$

After t time steps, the algorithm results in a graph with:

$$N(t) = N(t=0) + t \quad \text{nodes} \quad (1.13)$$

$$K(t) = K(t=0) + mt \quad \text{edges} \quad (1.14)$$

For large times, this corresponds to a graph with an average degree $k = 2m$. The procedure is iterated until the desired final number of nodes N is reached.

The BA model has been solved in the mean-field approximation [18, 21] and, exactly, by means of rate equation [22] approach. In the limit $t \rightarrow \infty$, the model produces a degree distribution $P(k) \sim k^{-\gamma}$, with an exponent $\gamma = 3$. The case of a growing network with a constant attachment probability $\Pi_{j \rightarrow i} = \frac{1}{m_0+t-1}$ produces, instead, a degree distribution $P(k) = \frac{e}{m} \exp(-\frac{k}{m})$. This implies that the preferential attachment is an essential ingredient of the model.

The average distance in the BA model is smaller than in a ER-random graph with same N and K , and increases logarithmically with N [19]. The clustering coefficient vanishes with the system size as $C \sim N^{-0.75}$. This is a slower decay than that observed for random graphs, $C \sim N^{-1}$, but it is still different from the behavior in small-world models, where C is a constant independent of N .

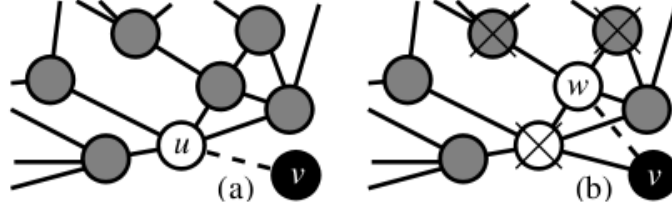


Figure 1.6: In the preferential attachment step (a) the new vertex v chooses a vertex u to attach to with a probability proportional to its degree. In the triad formation step (b) the new vertex v chooses a vertex w in the neighborhood of the one linked to in the previous preferential attachment step. \times symbolizes “not-allowed to attach to” (either since no triad would be formed, or that and edge already exists).

The BA model has attracted an exceptional amount of attention in the literature. In addition to analytic and numerical studies of the model itself, many authors have proposed modifications and generalizations to make the model a more realistic representation of real networks [19]. Various generalizations, such as models with nonlinear preferential attachment, with dynamic edge rewiring, fitness models and hierarchically and deterministically growing models, can be found in the literature. Such models yield a more flexible value of the exponent γ which is restricted to $\gamma = 3$ in the original BA construction. Furthermore, modifications to reinforce the clustering property, which the BA model lacks, have also been considered. In next section we will discuss one of the last kind of models, the *Holme-Kim* model, as we will be interested in studying how the clustering coefficient affect some games on graphs (see 3).

1.4.4 Holme-Kim model

The model proposed by Holme and Kim (HK model) [23] produces scale-free graphs with a tunable clustering coefficient. In order to achieve this, the model

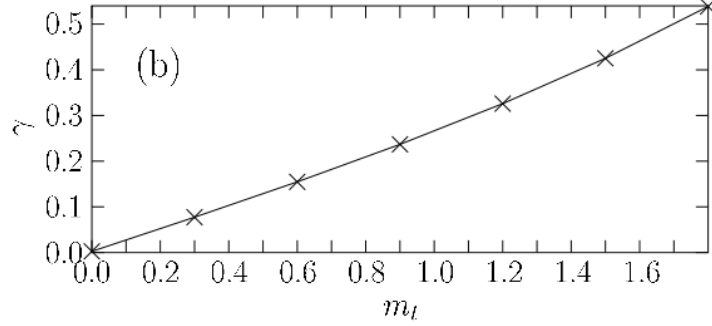


Figure 1.7: Clustering coefficient versus m_t , which is a quantity proportional to p and defined as $m_t = (m - 1)p$. It is evident the presence of a linear relation between these two quantities.

modifies slightly the BA model described in the previous section. The algorithm works as follows (see also Fig. 1.6):

1. the networks are constructed via a growing process that starts from an initial core of m_0 nodes, like in the BA model. At each time step, a new node i ($i = m_0 + 1, m_0 + 2, \dots, N$) is added to the network and is linked to m (with $m \leq m_0$) of the previously existent nodes;
2. the first link of the m to be drawn follows always a preferential attachment rule like (1.12) and is connected to a node j (preferential attachment step or PA step);
3. the remaining $m - 1$ links are attached in two different ways:
 - with probability p the new node i is connected to a randomly chosen neighbor of the node j , that is to say the first node to which i was attached (triad formation step or TF step);
 - with probability $1 - p$ the node i is attached doing another PA step, that is to say it is connected to one of the previously existent nodes following (1.12)

With such a procedure, once fixed the values of m and m_0 , one obtains scale-free networks with exponent $\gamma \sim 3$ and a tunable clustering coefficient depending on the value of p . In particular, for $p = 0$ the model coincides with the BA one (section 1.4.3 and [18]) where the clustering coefficient tends to zero as the network size N goes to infinity. For values of $p > 0$ the clustering coefficient grows monotonically with p (See Fig. 1.7).

Chapter 2

An introduction to game theory and evolutionary games

Game theory is the unifying paradigm behind many scientific disciplines. It is a set of analytical tools and solution concepts, which provide explanatory and predicting power in interactive decision situations, when the aims, goals and preferences of the participating players are potentially in conflict. It has successful applications in diverse fields such as evolutionary biology and psychology, computer science and operation research, political science and military strategy, cultural anthropology, ethics and moral philosophy, economics. The cohesive force of the theory stems from its mathematical structure which allows the practitioners to abstract away the common strategic essence of the actual biological, social or economic situation. Game theory creates a unified framework of abstract models and metaphors, together with a consistent methodology, in which these problems can be recast and analyzed.

The appearance of game theory as an accepted physics research field is a relatively late event. It required the mutual reinforcing of two important factors: the opening of physics, especially statistical physics, towards new interdisciplinary

research directions, and the sufficient maturity of game theory itself in the sense that it had started to tackle into complexity problems. Two new disciplines, socio- and econophysics were born, and the already existing field of biological physics got a new impetus with the clear mission to utilize the theoretical machinery of physics for making progress in questions whose investigation were traditionally connected to the social sciences or economics, and were formulated to a large extent using classical and evolutionary game theory.

In this chapter we present a brief introduction to some general concept of game theory and of evolutionary games. We will use some of these concepts in the following chapter, where evolutionary game theory is put in relation with complex networks in order to understand how particular topologies and dynamics affect the results obtained using mean-field evolutionary game theory.

2.1 Overview

2.1.1 Classical game theory and its assumptions

Classical (rational) game theory is based upon a number of assumptions about the structure of a game. Game theory assumes that agents (players) have well defined goals and preferences which can be described by a *utility function*. The utility is the measure of satisfaction the player derives from a certain outcome of the game, and the player's goal is to maximize her utility. Maximization (or minimization) principles abound in science. It is, however, important to underline that the maximization problem of game theory differs from the one of physics. In a physical theory the standard situation is to have a single function (say, a Hamiltonian or a thermodynamic potential) whose extremum condition characterizes the whole system. In game theory the number of functions to maximize is typically as much as the number of interacting agents. While physics tries to

optimize in a fixed landscape, the agents of game theory continuously restructure the landscape for each other in pursuit of their selfish individual optimum.

Another key assumption in the classical theory is that players are *perfectly rational* and this is *common knowledge*. “Perfect rationality” means that the players have well defined payoff functions, and they are fully aware of their own and the opponents’ strategy options and payoff values. They have no cognitive limitations in deducing the best possible way of playing whatever the complexity of the game is. In this sense computation is costless and instantaneous. “Common knowledge” implies that beyond the fact that all players are rational, they all know that all players are rational, and that all players know that all players are rational, etc.

Rationality, however, seems to be an ill-defined concept. There are extreme opinions arguing that the notion of perfect rationality is not more than pure tautology: rational behavior is the one which complies with directives of game theory, which in turn is based on the assumption of rationality. Any working definition of rationality is a negative definition, not telling us what rational agents do, but rather what they do not.

2.1.2 Post-Nash game theory

The post-Nash history of game theory is mostly the history of refinements. The Nash equilibrium concept seems to have enough predicting power in static games with complete information. The two mayor streams of extensions are toward dynamic games and games with incomplete information. Dynamic games are the ones where the timing of decision making plays a role. In these games the simple Nash equilibrium concept would allow outcomes which are based on non-credible threats or premises. In order to exclude these spurious equilibria the concept of a *subgame perfect Nash equilibrium* has been introduced (see 2.3.1).

It requires Nash-type optimality in all possible games. Incomplete information, on the other hand, means that the players' available strategy sets and associated payoffs (utility) are not common knowledge¹.

Despite the undoubted success of classical game theory, the paradigm has soon shown its limitations. In many specific cases further progress seemed to rely upon the relaxation of some of the key assumptions. A typical example where rational game theory seems to give an inadequate answer is the “backward induction paradox” related to repeated (iterated) social dilemmas like the Repeated Prisoner’s Dilemma. According to game theory the only subgame perfect Nash equilibrium in the finitely repeated game is the one determined by backward induction, i.e., when both players defect in all rounds. Nevertheless, cooperation is frequently observed in real-life psycho-economic experiments [45, 48]. This result either suggests that the abstract Prisoner’s Dilemma game is not the right model for the situation or that the players do not fulfill all the assumptions. Indeed, there are good reasons to believe that many realistic problems, in which the effect of an agent’s action depends on what other agents do, are more complex and perfect rationality of the players can’t be postulated. The standard deductive reasoning loses its appeal when agents have non-negligible cognitive limitations, there is a cost of gathering information about possible outcomes and payoffs or the agents do not have consistent preferences. A possible way out is *inductive reasoning*, i.e. a trial-and-error approach, in which agents continuously form hypotheses about their environment, build strategies accordingly, observe their performance in practice and verify or discard their assumptions based on empirical success rates. In

¹Incomplete information differs from the similar concept of *imperfect information*. The latter refers to the case when some of the history of the game is unknown to the players at the time of decision making. For example Chess is a game with perfect information because players know the whole previous history of the game, whereas the Prisoner dilemma is a game with imperfect information due to the simultaneity of the players’ decisions. Nevertheless, both are games with complete information

this approach the outcome, or in other words the solution, of a problem is determined by the evolving mental state (mental representation) of the constituting agents. Mind necessarily becomes an endogenous dynamic variable of the model. This kind of *bounded rationality* may explain that in many situations people respond instinctively, play according to heuristic rules and social norms rather than adopting the strategies indicated by rational game theory.

2.1.3 Bounded rationality and evolutionary game theory

Bounded rationality becomes a natural concept when the goal of the theory is to understand animal behaviour. Individuals in an animal population do not make conscious decisions about strategy, even though the incentive structure of the underlying formal game they “play” is identical to the ones discussed. They are genetically coded and maintained during the whole life-cycle, the strategy space is constrained, or strategy adoption or change is severely restricted by biologically predetermined learning rules or mutation rates. The success of a strategy applied is measured by biological fitness, which is usually related to reproductive success. *Evolutionary game theory* is an extension of the classical paradigm toward bounded rationality. There is however another aspect of the theory which has been left out in the classical approach, but gets special emphasis in the evolutionary version, namely dynamics. Dynamical issues were mostly neglected classically, because the assumption of perfect rationality made such questions irrelevant. Full deductive rationality allows the players to derive and construct the equilibrium solution instantaneously. In this spirit when dynamic methods were still applied they only served as a technical aid for deriving the equilibrium. Bounded rationality, on the other hand, is inseparable from dynamics. Contrary to perfect rationality, bounded rationality is always defined in a positive way, postulating what boundedly rational agents do. These behavioral rules are

dynamic rules, specifying howmuch of the game's earlier history is taken into consideration (memory), how long agents would think ahead (short sightdness), how they search for available strategies (search space), how they switch for more successful ones (adaptive learning), and what all these mean at the population level in terms of frequencies of strategies.

There is a static and a dynamic perspective of evolutionary game theory. The evolutionary stability of a Nash equilibrium is a static concept which does not require solving time-dependent dynamic equations. In simple terms evolutionary stability means that a rare mutant cannot successfully invade the population. The condition for evolutionary stability can be checked directly without incurring complex dynamic issues. The dynamic perspective, on the other hand, operates by explicitly postulating dynamical rules. These rules can be prescribed as deterministic rules at the population level for the rate of change strategy frequencies or as microscopic stochastic rules at the agent level (agent-based dynamics). Since bounded rationality may have different forms, there are many different dynamical rules one can consider. The most appropriate dynamics depends on the specificity of the actual biological or socio-economical situation under study. For example, in biological applications the *Replicator Dynamics* is the most natural choice; the rule is derived by assuming that payoffs are directly related to reproductive success. Socio-economic applications may require other adjustment or learning rules. Both the static and dunamic perspective of evolutionary game theory provide a basis for equilibrium selection when the classical form of the game has multiple Nash equilibria.

2.1.4 Beyond evolutionary game theory

The mission of evolutionary games theory was to remedy three key deficiencies of the classical theory: (1) bounded rationality, (2) the lack of dynamics and

(3) equilibrium selection in the case of multiple Nash equilibria. Although this mission was accomplished rather successfully, there was a series of weaknesses remaining. Evolutionary game theory in its early form considered population dynamics on the aggregate level. All the variables are averaged over the population (for example relative strategy abundances). Behavioral rules, on the other hand, control the system at the microscopic level (agent level). Agent decisions are frequently asynchronous, discrete and may contain stochastic elements. Moreover agents may have different individual preferences, payoffs, strategy options or be locally connected to well-defined other agents. All this requires for the theory some improvements and some refinements that can be achieved by coupling the evolutionary game theory with the theory of complex networks (see chapter 3).

2.2 Games, payoff, strategies

2.2.1 Notation

A game is an abstract formulation of an interactive decision situation with possibly conflicting interests. The *normal (strategic) form representation* of a game specifies the players of the game, their feasible actions (to be called pure strategies), and the payoffs received by them for each possible combination of actions (the action or strategy profile) that could be chosen by the players. Let $n = 1, \dots, N$ denote the players; $\mathcal{S}_n = \{e_{n_1}, e_{n_2}, \dots, e_{n_Q}\}$ the set of pure strategies available to player n , with $s_n \in \mathcal{S}_n$ an arbitrary element of this set; s_1, \dots, s_N a given strategy profile of all players; and $u_n(s_1, \dots, s_N)$ player n 's payoff function (utility function), i.e., the measure of her satisfaction if the strategy profile (s_1, \dots, s_N) gets realized. Such a game can be denoted as $G = \{\mathcal{S}_1, \dots, \mathcal{S}_N; u_1, \dots, u_N\}$.

2.2.2 Matrix form

When there are only two players $n = 1, 2$ and the set of available strategies is discrete, $\mathcal{S}_1 = \{e_1, e_2, \dots, e_Q\}$, $\mathcal{S}_2 = \{f_1, f_2, \dots, f_R\}$, the game can be written in a bi-matrix form $G = (A, B^T)$, which is a shorthand for the payoff table:

$$\begin{array}{c|ccc}
 & & \text{player 2} & \\
 & & f_1 & \dots & f_R \\
 \hline
 & e_1 & (A_{11}, B_{11}^T) & \dots & (A_{1R}, B_{1R}^T) \\
 \text{player 1} & \vdots & \vdots & \ddots & \vdots \\
 & e_Q & (A_{Q1}, B_{Q1}^T) & \dots & (A_{QR}, B_{QR}^T)
 \end{array} \tag{2.1}$$

Here the matrix $A_{ij} = u_1(e_i, f_j)$ denotes player 1's payoff for the strategy profile (e_i, f_j) . Two-players games can be *symmetric* (so-called *matrix games*) and *asymmetric* (*bi-matrix games*), with symmetry referring to the roles of players. For a symmetric game $R = Q$ and $B = A$. In this case the game is fully characterized by the single payoff matrix A and we can formally write $G = (A, A^T)$.

2.2.3 Strategies

The strategies that label the payoff matrices are *pure strategies*. In many games, however, players can also play *mixed strategies*, which are probability distributions over pure strategies. Playing a mixed strategy means that in each decision instance the player comes up with one of her feasible actions with a certain pre-assigned probability. Each mixed strategy corresponds to a point \mathbf{p} of the *mixed strategy simplex*:

$$\Delta_Q = \left\{ \mathbf{p} = (p_1, \dots, p_Q) \in \mathbb{R}^Q : p_q \geq 0, \sum_{q=1}^Q p_q = 1 \right\} \quad (2.2)$$

whose corners are the pure strategies.

A strategy s_i^* of player i is a *strictly dominant strategy* if for each strategy profile $s_{-i} = s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_N$ of the co-players happens:

$$u_i(s_i^*, s_{-i}) > u_i(s'_i, s_{-i}) \quad \forall s_{-i}, s'_i \neq s_i^* \quad (2.3)$$

The strategy s'_i is dominated if there exists $s''_i \in \mathcal{S}_i$ such that:

$$u_i(s''_i, s_{-i}) > u_i(s'_i, s_{-i}) \quad \forall s_{-i} \quad (2.4)$$

Dominance is important because, if utility payoff are correctly specified and players care only about their own utility, there is no good reason to violate strict dominance. One step of iterated dominance is a judgement by one player that the other player will not make a dumb mistake. This often tells a player what she herself should do.

2.2.4 Nash Equilibrium

The strategy profile $s^* = (s_i^*, s_{-i}^*)$ is a *Nash equilibrium* (NE) if each player's strategy is a best response to the other players' strategies. That is, no player has incentive to deviate, if no other player will deviate (if players find themselves in equilibrium, there is no reason to move away). Formally this can be expressed as:

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s'_i, s_{-i}^*) \quad \forall s'_i \quad (2.5)$$

When the inequality above is strict, s^* is called a *strict Nash equilibrium*. One

of the most fundamental results of classical game theory is *Nash's theorem*, which asserts that in normal-form games with a finite number of players and a finite number of pure strategies there exists at least one NE, possibly involving mixed strategies.

2.3 Repeated Games

Most of the inter-agent interactions that can be modelled by abstract games are not one-shot relationships but occur repeatedly on a regular basis. When a one-shot game is played between the same rational players iteratively, a single instance of this series cannot be singled out and treated separately. The whole series should be analyzed as one big “supergame”. What a player does early on can affect what others choose to do later on.

Assume that the same game G is played a number of times T . The set of feasible actions and payoffs in the game at time t ($t = 1, \dots, T$) are independent from t and from the former history of the game. This does not mean, however, that actions themselves should be chosen independently from time and history. When G is played at time t , all the game history so far is common knowledge. We will denote this repeated game as $G(T)$ and distinguish in *finitely repeated games* ($T < \infty$) and *infinitely repeated games* ($T = \infty$). For finitely repeated games the total payoff is simply the sum of the game G payoffs:

$$U = \sum_{t=1}^T u_t \quad (2.6)$$

In one-shot static games with complete information a strategy is simply an *action* a player can choose. For repeated games (and also for other kinds of dynamic games) the concept of strategy becomes more complex. In these games

a player's strategy is a complete plan of action, specifying a feasible action in any contingency in which the player may be called upon to act.

2.3.1 Subgame perfect Nash equilibrium

What is the prediction of rational game theory for the outcome of a repeated game? As always the outcome should correspond to a Nash equilibrium: no player can have a unilateral incentive to change its strategy (in the supergame sense), since this would induce immediate deviation from that profile. However, not all NEs are equally plausible outcomes in a dynamic game such as a repeated game. A stronger concept than Nash equilibrium is needed to exclude these spurious NEs. *Subgame perfection* [26] is a widely accepted criterion to solve this problem. Subgame perfect Nash equilibria are those that pass a credibility test.

A *subgame* of a repeated game is a subseries of the whole series that starts at period $t \geq 1$ and ends at the last period T . However, there are many subgames starting at t , one for each possible history of the game before t . Thus subgames are labeled by the starting period t and by the history of the game before t . Therefore when a subgame is reached in the game the players know the history of play. By definition a NE of the game is a subgame perfect if it is a NE in all subgames.

Chapter 3

Ultimatum game on networks

3.1 Merging evolutionary game theory and complex networks

A number of theoretical approaches have been developed to explain the arising of human altruism. Kin selection theory [27] accounts for situations in which it pays off to help relatives that shares some fraction of the genetic pool. In the absence of such kin relationships, repeated interactions have also been shown to lead to cooperation, as well as different kinds of reciprocity mechanisms [28, 29, 31, 32]. Recently, a series of behavioral experiments in which interactions are anonymous and one-shot have shown that humans can punish non-cooperators (altruistic punishment) and reward those individuals who cooperate (altruistic rewarding) [28, 33, 34, 35]. This so-called strong reciprocity can actually explain the observed cooperative behavior in terms of group and cultural selection. However, following standard evolutionary game theory, it is still far from being explained how cooperation may arise from selection at the individual level. Recent steps in this direction [37] have contributed to fill this gap, although a general theoretical

framework is still needed.

In the meanwhile, it is well-known that real-life interaction networks can possess a rather complex topology, which is far from the traditional mean-field case. Although the importance of this feature has been widely recognised already a long time ago, a systematic investigation of these questions is still in the forefront of research.

Very recently the research of evolutionary games has interfered with the extensive investigation of networks, because the actual social networks characterizing human interactions possess highly nontrivial topological properties. The first results clearly demonstrated that the topological features of these networks can influence significantly their behaviour. In many cases “games on graphs” differ qualitatively from their counterparts defined in a well-mixed (mean-field) population.

This is why in this chapter we will deal with evolutionary games coupled with complex networks. We will show this way how particular features, as altruism, can arise by merging a particular model of evolutionary game theory, the Ultimatum Game, with the investigation of complex networks. The high number of variables in the system makes most standard analytical techniques largely inapplicable. Therefore we will use extensive numerical simulations and analytical techniques going beyond the traditional mean-field level.

3.2 Ultimatum Game: differences between theory and experiments

The ultimatum game (UG) is a model extensively used to study altruistic behaviour. In the past years it has inspired dozens of theoretical and experimental investigations. The rules of the game are surprisingly simple. Two players have

to agree on how to split a sum of money, or in general a reward. One of the two players acts as proposer and makes an offer. The other player, then, plays the role of responder and she can either accept or reject the deal. If the responder accepts, the deal goes ahead. If she rejects, neither player gets anything. Obviously rational responders should accept even the smallest positive offer, since the alternative is getting nothing (subgame perfect Nash equilibrium solution). Proposers, therefore, should be able to claim almost the entire sum. In a large number of human studies, however, conducted with different incentives in different countries, the majority of proposers offer 40% to 50% of the total sum, and about half of all responders reject offers below 30% [40].

The irrational human emphasis on a fair division suggests that players have preferences which do not depend solely on their own payoff, and that responders are ready to punish proposers offering only a small share by rejecting the deal (which costs less to themselves than to the proposers).

Many solutions and explanations have been given to fill this gap between theory and experiments. Nowak et al. [30], for example, have shown how towards a mechanism of reputation (the proposer has some information on which deals the responder has accepted in the past) fairness can evolve. However this result is obtained in a well-mixed population, without taking in account the possibility of a microscopic structure or a particular pattern of interactions between agents. In fact it has been shown that social and biological networks, where the arising of cooperation is a relevant issue, display properties and structures (e.g. small-world property, scale-freeness, clustering, etc.) that haven't been considered in anyone of the models [41, 3].

In the following sections we study an UG model where:

1. players are placed on the nodes of a graph and play with their neighbors;
2. a selection rule in which the less fitted individuals are replaced together

with their neighbors is implemented.

We will show that particular properties, far from the ones obtained in the mean-field approximation, arise, even not introducing a mechanism of reputation, but as results merely due to the topological characteristics of the networks used.

3.3 Standard ultimatum game on networks: our model

In our model we consider N individuals associated to the nodes of a graph. At each time step, each individual plays a round robin with all of its neighbors. In each round, the individual plays the ultimatum game two times, with a different role, proposer and responder, on each. The reward to be divided in each game is equal to 1. An individual i ($i = 1, \dots, N$) is characterized by two parameters: $p_i, q_i \in [0, 1]$. When i acts as proposer it offers a division p_i of the reward, so that responder will earn p_i if it accepts the proposal. Instead, when agent i plays as responder, it will accept the offer if it is larger than the acceptance threshold, q_i . In general, p_i and q_i can be independent. However in the following we will consider two cases:

1. $p_i = q_i$;
2. $p_i = 1 - q_i$.

When two individuals (i, j) bargain, their payoffs, Π_i and Π_j , evolve according to the following rules:

- Player i offers the amount p_i to j . If $p_i \geq q_j$, the offer is accepted and the payoff of i and j are incremented by $\Delta\Pi_{ij}^O = (1 - p_i)$ and $\Delta\Pi_{ji}^R = p_i$

respectively. Conversely, if $p_i < q_j$, agreement is not possible and both players get nothing and their payoffs remain the same, $\Delta\Pi_{ij}^O = \Delta\Pi_{ji}^R = 0$.

- When player i is the responder, the same rules apply. Therefore, upon agreement, players i and j increase their payoffs by $\Delta\Pi_{ij}^R = p_j$ and $\Delta\Pi_{ji}^O = (1 - p_j)$ respectively.

The final payoffs of a node i after playing with all its neighbors is $\Pi_i = \sum_{l \in \Gamma} (\Delta\Pi_{il}^O + \Delta\Pi_{il}^R)$, where Γ_i denotes the set of i 's neighbors.

In addition to the rules for bargaining, we also implement a selection rule at the individual level in the same spirit of [37]. At each round, the player with lowest payoff in the whole population is removed and, moreover, all the agents in its neighborhood (no matter how wealthy they are) are also removed. The removal of the neighbors of the poorest agent implements a sort of social punishment (stigma) reinforced by law. This is quite different from the notion of punishment by individuals [43] and it is of no direct (differential) cost to individuals. All the removed agents are replaced in their nodes (so that they only inherit their contacts) by brand new randomly uniformly distributed p -strategists. With this evolutionary rule, and regarding evolutionary survival, it is not only important to earn as much as possible, but also players should take care of the neighbor's payoff. If an individual exploits its neighborhood so that she takes a large stake of the total reward, she would risk to be dropped out of the game as a result of one of its neighbors being that with the lowest payoff in the population of players. Consequently, what drives the evolution of the distribution of p values among the population is the balance between the conflicting interests of earning more (to avoid being the poorest) and earning less (to avoid being stigmatized). After the selection rule is implemented the payoffs of the agents are reset to zero.

3.4 Ultimatum Game on networks

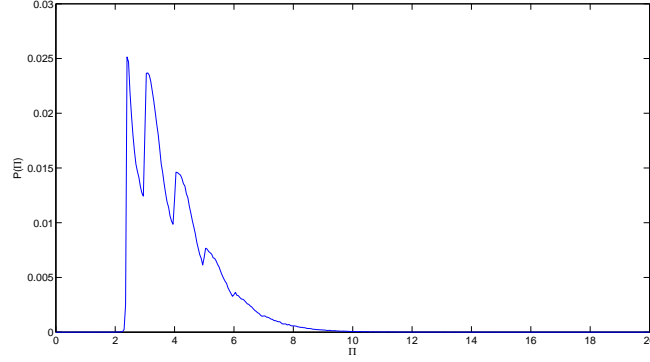
We have implemented both the cases $q_i = p_i$ and $q_i = 1 - p_i$ on random graphs and on scale-free graphs. The formers have been generated using the Watts and Strogatz algorithm (see 1.4.2) with the probability of rewiring p set equal to 1 and starting with a lattice where every node is connected to his first 4 neighbors; the latters have been constructed by using the Barabási-Albert model (see 1.4.3), setting $m = 2$. This way both the degree distribution for random and scale-free graphs have $k_{min} = 2$. All the graphs created are constituted by 10^4 nodes.

Every result we show is averaged over more than 100 different initial conditions and different graphs and for each of this we follow the evolution of the game for 10^7 time steps.

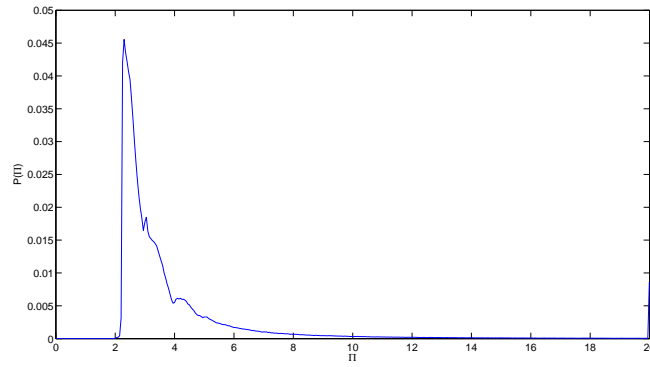
For every configuration of q and p and for every kind of graphs used, we report the plots of the following quantities:

- $P(\Pi)$, the distribution of the payoff at the stationary state (last time step computed);
- $\Pi(k)$, the average payoff for nodes with the same degree k ;
- $P(p)$, the distribution of the values p at different time-steps;
- $\frac{N_r}{N_k}$, the number of nodes removed over the number of nodes with same degree k .

Other particular quantities and plots will be shown in the detailed study of a particular configuration or of a particular topology. Particular emphasis has been given to the study of the configuration $p_i = 1 - q_i$ on scale-free graphs. The focus on this particular topology is due to the fundamental role scale-free graphs cover in the modelling of interactions in real-life (see chapter 1), while the



(a) Random graphs



(b) Scale-free graphs

Figure 3.1: Distribution of payoff Π after 10^7 time steps for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = p_i$.

configuration of p and q is taken in account as the most plausible to model the behaviour of players [30].

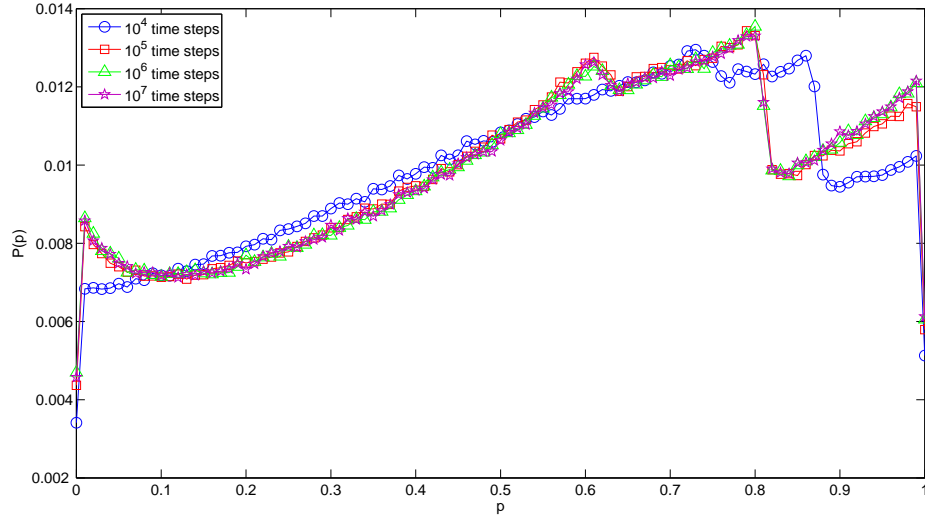
3.4.1 Configuration $q_i = p_i$

In this configuration of q and p , individuals do distinguish among roles, and moreover adjust the threshold for acceptance when responders with the proposal when proposers (i.e., $q_i = p_i$), so as to get half of the total stake on average. In fact, whatever is the relation between p_i and p_j , every player will increase her

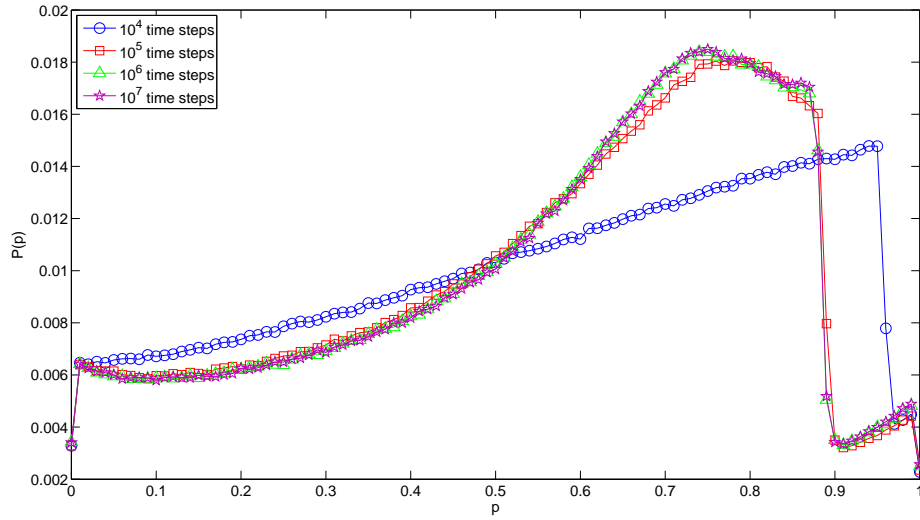
payoff in only one round, as proposer or as responder; in the other round in fact the inequality between p_i and p_j changes of sign and it is no longer satisfied.

We have first checked the distribution of payoff at the stationary step (in our simulations after 10^7 time steps), in order to see if there is a critical value Π_c creating avalanches phenomena like in [36] (see also 3.5) or in the Bak-Sneppen model [42]. Though the figures 3.1 show that for both ER and SF graphs it seems to exist a critical value around 2, the study of the avalanches ¹ show that there isn't a self-organized state (that is to say a power-law in distribution of the avalanche size) but rather a subcritical one.

¹An avalanche in this case is defined as follows: it starts when the lowest payoff gets larger than a preset critical value Π_c and stops when it drops below this value. The size, s , of an avalanche is the number of time steps it lasts.

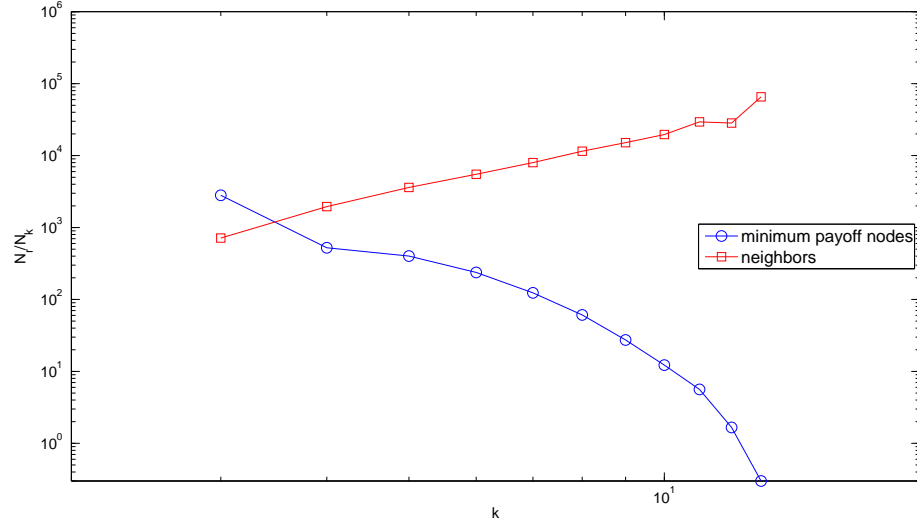


(a) Random graphs

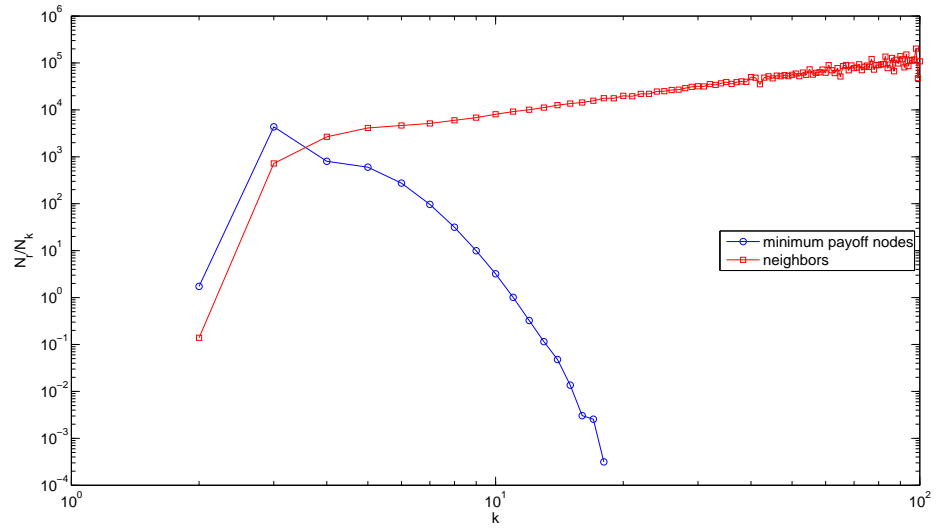


(b) Scale-free graphs

Figure 3.2: Distribution of p after 10^4 (blue circles), 10^5 (red squares), 10^6 (green triangles) and 10^7 (purple stars) time steps for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = p_i$.

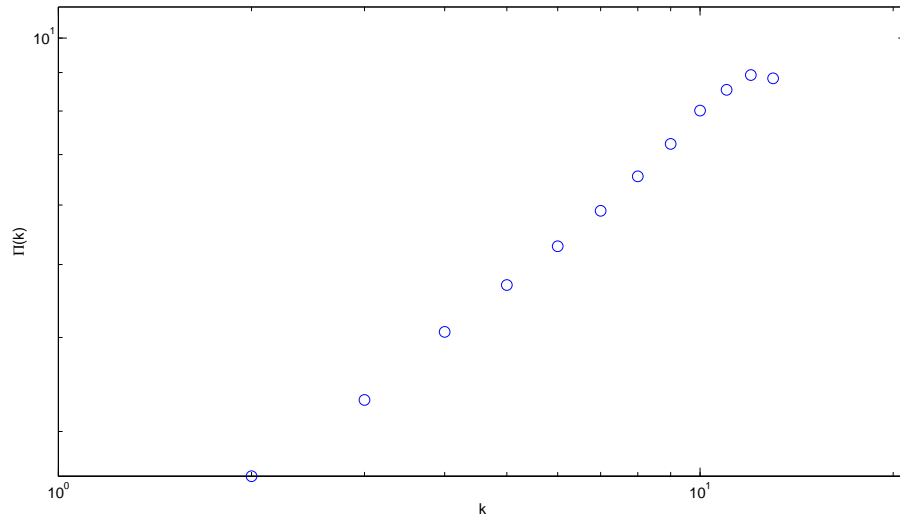


(a) Random graphs

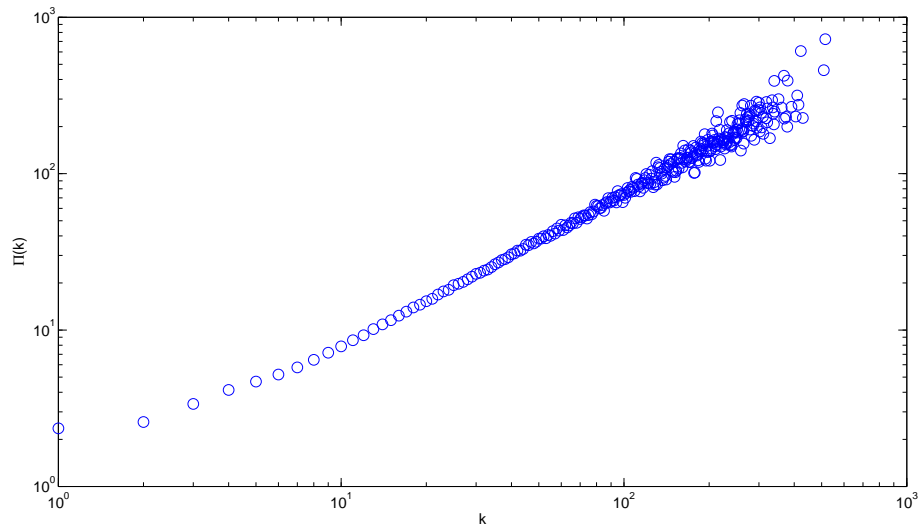


(b) Scale-free graphs

Figure 3.3: Fraction of nodes removed for every degree class in 10^7 time steps because minimum payoff nodes (blue circles) or their neighbors (red squares), for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = p_i$.



(a) Random graphs



(b) Scale-free graphs

Figure 3.4: Average payoff for every degree class after 10^7 time steps for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = p_i$.

3.4.2 Configuration $q_i = 1 - p_i$

We consider also the case in which the offer and the acceptance thresholds are such that $p_i + q_i = 1$. This amounts to consider that players do not establish correlations between role played and expectations, i.e. they do not differentiate between roles. They are *role-ignoring* agents that only accept, as responders, proposals above what they get when proposers, $q_i = 1 - p_i$ [30].

Note that both players will gain as proposers and responders only if the condition $p_i + p_j \geq 1$ is verified. In this case, their payoffs are incremented by $\Delta\Pi_{ij}^O + \Delta\Pi_{ij}^R = (1 - p_i) + p_j$ and $\Delta\Pi_{ji}^O + \Delta\Pi_{ji}^R = (1 - p_j) + p_i$. Therefore, the final payoffs of a node i after playing with all its neighbors is $\Pi_i = \sum_{l \in \Gamma} (\Delta\Pi_{il}^O + \Delta\Pi_{il}^R)$, where Γ_i denotes the set of i 's neighbors.

We have first monitored the payoff distribution and seen, as for the configuration $q_i = p_i$ that there is not a power-law distribution of avalanches. Instead, a particular phenomenon arising in this configuration is the presence of some “resonance peaks” in the payoff distribution for the random graphs. These peaks appear to be in correspondence to integer values of payoff. In order to understand this peculiarity, we have plotted how many nodes for every degree class own a certain integer payoff ($\Pi = 3, \Pi = 4, \Pi = 5, \dots$), as shown in fig. 3.5. The figure explains the presence of the peaks in the payoff distribution: the payoff is in relation with the degree, and in particular there is a high probability that a node of degree k owns a payoff $\Pi = k$, that is to say it is highly probable that there is a compensation effect for which every agent gains on average a number of “whole” rewards equal to her degree.

Then we have studied the evolution of the offers made by players by looking at the distribution $P(p)$ (i.e., how frequently a player proposes an amount p to its neighbors) over time. In Fig. 3.7 we report the distributions $P(p)$ after

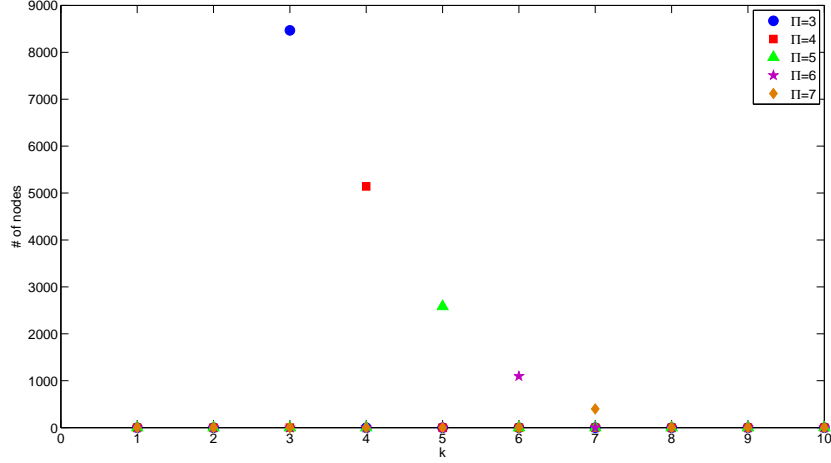


Figure 3.5: Number of nodes for every degree class holding payoff equal to three (blue circles), to four (red squares), to five (green triangles), to six (purple stars) and to seven (orange rhombus).

10^6 and 10^7 time steps for ER and SF topologies. From the figure it is evident that, for both ER and SF networks, the distribution $P(p)$ becomes stationary after a time no longer than 10^6 time steps. The stationary distributions $P(p)$ for ER and SF networks show the same average value for the offers $\langle p \rangle \simeq 0.5$. This points out, as observed in several experiments [40] the absence of rational behavior of players in both kind of network topologies. Though of equal average value, the distribution densities are strikingly different for both kind of networks. While $P(p)$ is almost flat, slowly decreasing at extremes ($p \leq 0.2$ and $p \geq 0.8$) for the ER network, it is bimodal for the SF network (peaked at $p \simeq 0.3$ and $p \simeq 1$). We see how the degree-heterogeneity of the network of interactions promotes a very different microscopic balance of conflicting aims, as reflected in the bimodal $P(p)$, respect to the mostly uniform density of strategies observed in near homogeneous networks (ER). Moreover, the degree of a node and its strategy are strongly correlated as shown in fig. 3.10.

In fig. 3.12, where we have plotted the probability that a node with degree

k is removed both as a consequence of being it the one with the lowest payoff in the population, or because it is a neighbor of the less fitted individual. It turns out that the altruistic behavior of the hubs increases their chances to remain in the game. Moreover, highly connected individuals are never exposed to direct extinction.

A semi-analytical approach

We now concentrate on the role that degree heterogeneity plays in the microscopic segregation, i.e. appearance of peaks in $P(p)$. Given that the relation $p_i + p_j \geq 1$ between the offers of two players i and j assures nonzero reward to both players, it is possible to show that, the way in which such a successful combination of the two offers is achieved, is strongly related to the degrees k_i and k_j .

Let us define the "interacting degree" of node i at time t , $k_i^{int}(t)$, as the number of neighbors of i with whom it interacts (*i.e.* those satisfying $p_j + p_i \geq 1$, the interacting neighborhood). If we consider the situation $k_i \gg 1$ (say a hub in a SF network), then under the assumption that p is distributed in the neighborhood of i following the same distribution as in the whole network, we obtain:

$$k_i^{int}(t) = k_i \int_{1-p_i}^1 P(p) dp = k_i (1 - F(1 - p_i)) , \quad (3.1)$$

where F is the (cumulative) distribution function of $P(p)$. Under the same assumptions, it follows that the payoff received by a hub is

$$\Pi_h = k_h [1 - F(1 - p_h)] \left[(1 - p_h) + \int_{1-p_h}^1 p dF \right] , \quad (3.2)$$

where the integral is the average of p in the "interacting" neighborhood of the

hub. Provided that this average is larger than θ_b/k_h , the limit when $p_h \rightarrow 1$ is

$$\lim_{p_h \rightarrow 1} \Pi_h > \theta_b . \quad (3.3)$$

If θ_b is an upper bound of $\min_i \Pi_i$, then a hub will not have the minimum payoff even if it offers the whole stake and accepts any offer. Therefore, *hubs can afford full generosity*. Moreover, they minimize the risk of being stigmatized by adopting high values of p . In other words, they not only can afford full generosity, but also better they do if they want their neighbors safe. In fact, one can give a simple estimate for the upper bound θ_b : For $k_{min} = 2$, the less connected nodes offering 0 and linked to two fully generous neighbors will obtain 4. That is, we can assume $\theta_b \leq 4$, in the argument above. In other words, if the average value of the hubs neighbors $p_{ave} > \theta_b/k_h$ (which at most is $4/k_h$), hubs can give away almost the whole stake. In particular, in the thermodynamic limit where k_h diverges, they can offer $p = 1$. Note that this argument also holds for other k classes, provided that the previous assumptions (before Eq. (3.2)) are satisfied.

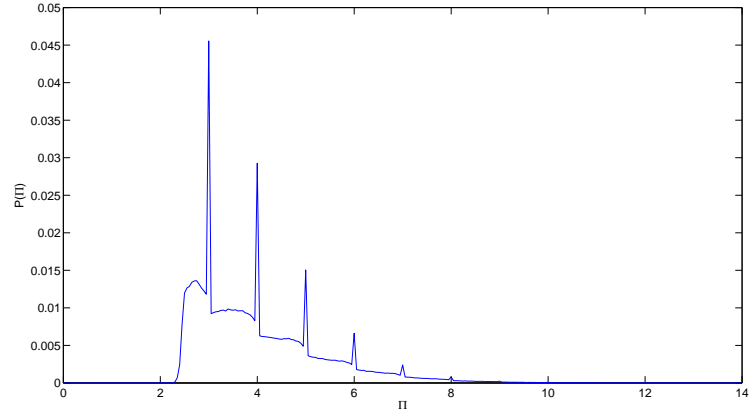
As for lowly connected nodes, the analysis is less precise because we must consider samples of a few individuals from a given distribution, which is a highly fluctuating situation. Avoiding being the less fitted individuals is now much harder and can be better reached by low values of p , if, correspondingly, the neighbors are generous. These nodes cannot afford being generous and, moreover, they need neighbors generosity to survive. The higher the values of $P(p)$ near $p = 1$, the lower the expected values of the offers of lowly connected nodes. On the other hand, the absence of the low- p peak, along with a decaying profile of $P(p)$ in the generous range in the ER case with respect to the SF case can be simply understood from the previous arguments. The increasing density profile at high p values is associated to the existence of large degrees in the network, and

thus should not be expected in homogeneous networks. Moreover, insofar as the peak at low values of p needs a high density profile in the generous range of high p values, it should neither be expected. In other words, in homogeneous networks one should expect the more uniform stationary $P(p)$ observed in fig. 3.7(a).

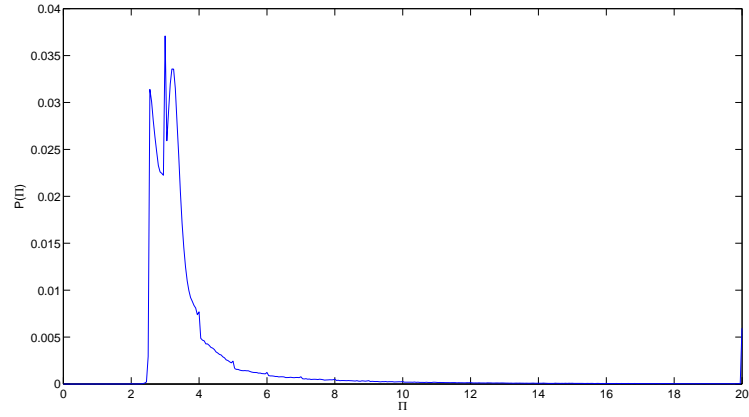
Figure 3.10 confirms that this is indeed the mechanism at work. In this figure we plot, for SF networks, the average value of the offers made by players of degree k , $\langle p \rangle_k$, as a function of the degree. The results reveal that highly connected nodes offer a large portion of the reward to its neighbors. Moreover, not only the highest connected players display altruistic behavior, but also those players with $k \geq k^* \approx 6$ show an average offer larger than 0.8. On its turn, lowly connected nodes ($k < k^*$) are more egoist and play harder offering a small piece of the whole stake. However, this behavior does not guarantee the highest payoffs to low degree nodes. In the inset of Fig. 3.10 we have reported the average payoffs of agents as a function of their degrees, $\langle \Pi \rangle_k$. It turns out that the larger the connectivity of the players, the larger the Π_i 's are on average. Note that the overabundance of low degree agents manifests itself by the peak at $p \simeq 0.3$ in Fig. 3.7. Thus, the segregation between low p strategists (egoists) and highly generous ($p \geq 0.8$) strategists takes place on the *degree scale* dimension. Therefore, the average generosity $\langle p \rangle \approx 0.5$ is not a characteristic value due to the bimodal shape of $D(p)$. The quickly saturating profile of $\langle p \rangle_k$ produces the observed segregation of the peaks.

3.4.3 Scale-free networks with variable clustering coefficient

The BA model produces a graph with a scale-free degree distribution but with clustering coefficient equal to zero. In order to see whether such a topological property can affect the evolution of the game or not, and in case how the game

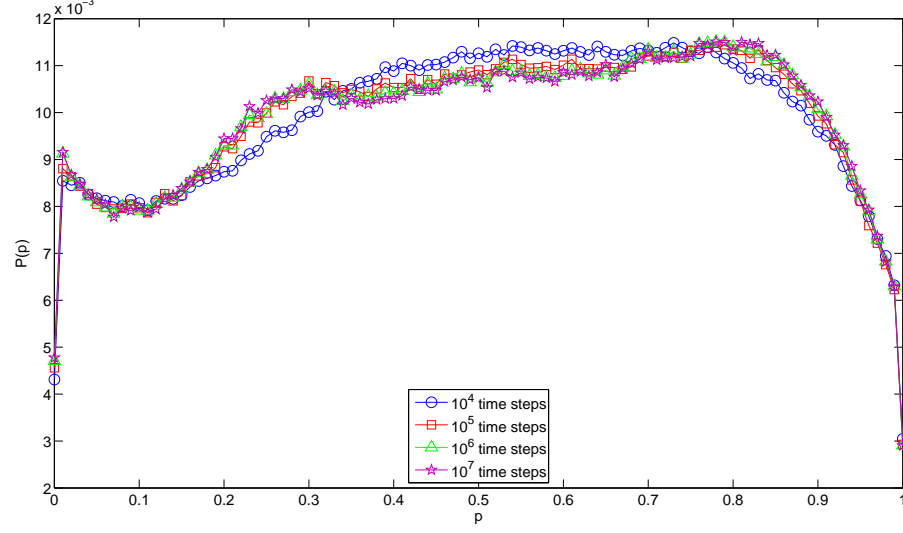


(a) Random graphs

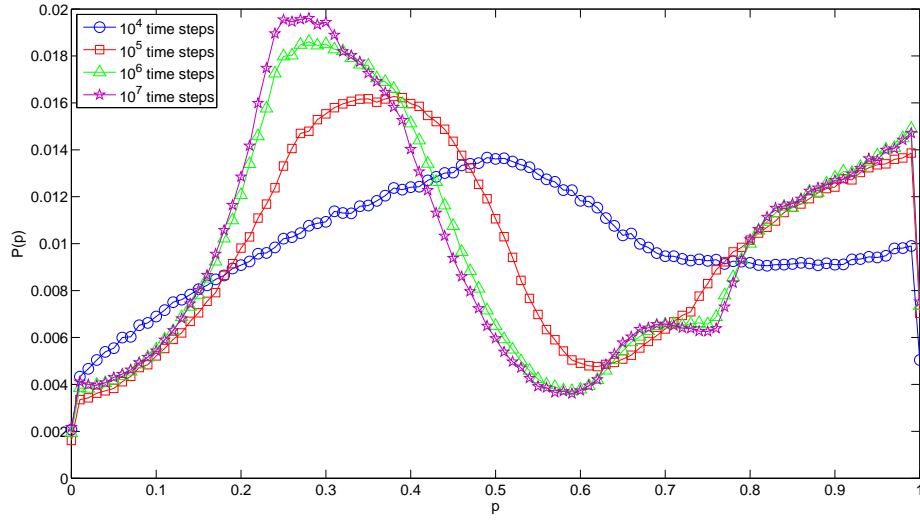


(b) Scale-free graphs

Figure 3.6: Distribution of payoff Π after 10^7 time steps for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = 1 - p_i$.

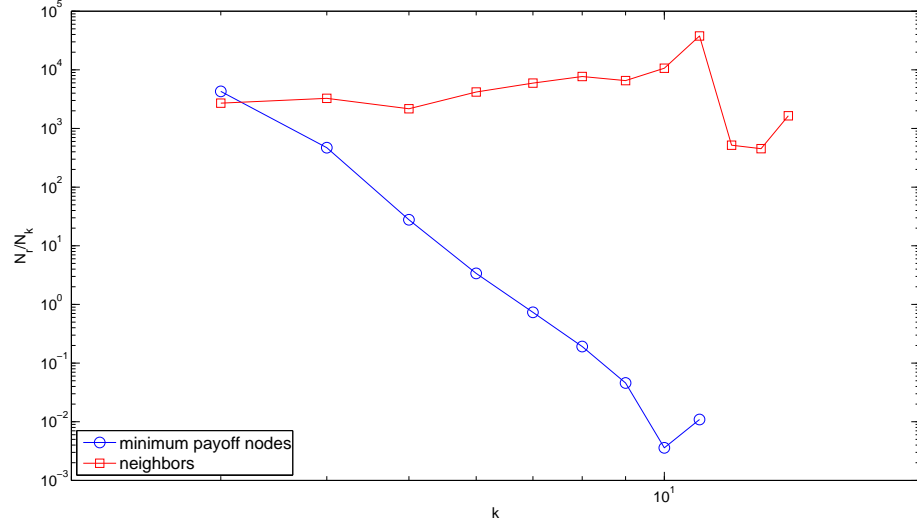


(a) Random graphs

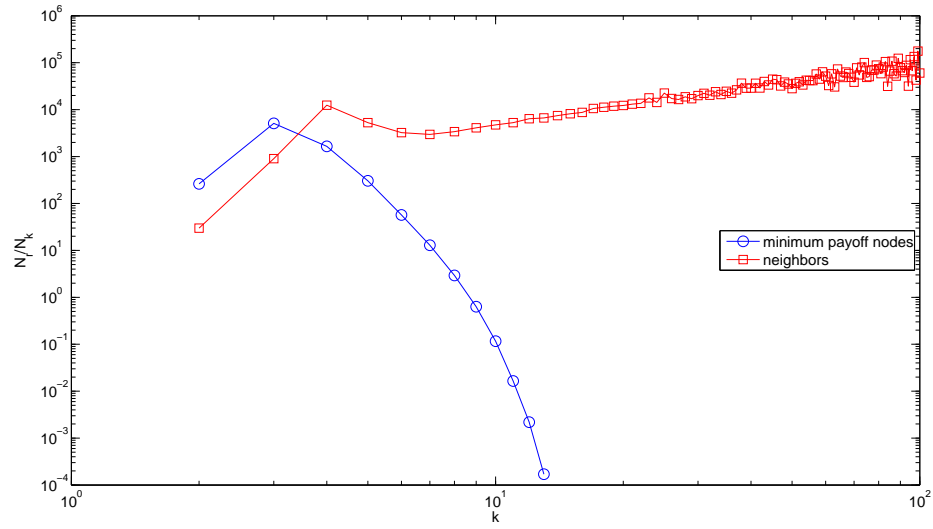


(b) Scale-free graphs

Figure 3.7: Distribution of p after 10^4 (blue circles), 10^5 (red squares), 10^6 (green triangles) and 10^7 (purple stars) time-steps time steps for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = 1 - p_i$.

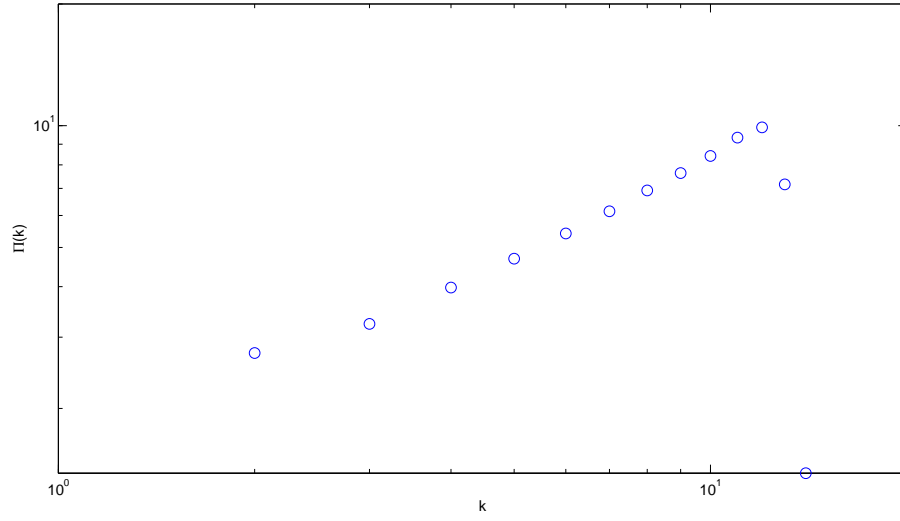


(a) Random graphs

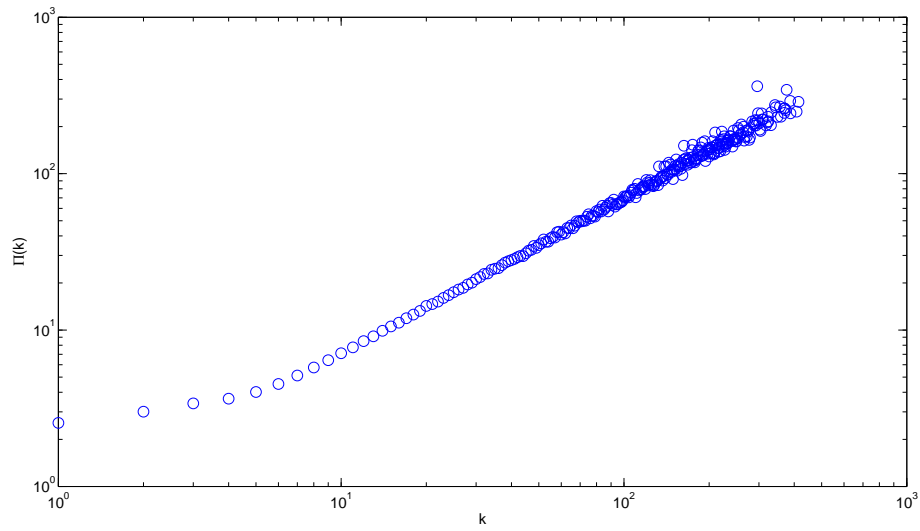


(b) Scale-free graphs

Figure 3.8: Fraction of nodes removed for every degree class in 10^7 time steps because minimum payoff nodes (blue circles) or their neighbors (red squares), for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = 1 - p_i$.



(a) Random graphs



(b) Scale-free graphs

Figure 3.9: Average payoff for every degree class after 10^7 time steps for (a) random graphs and (b) scale-free graphs when using the configuration $q_i = 1 - p_i$.

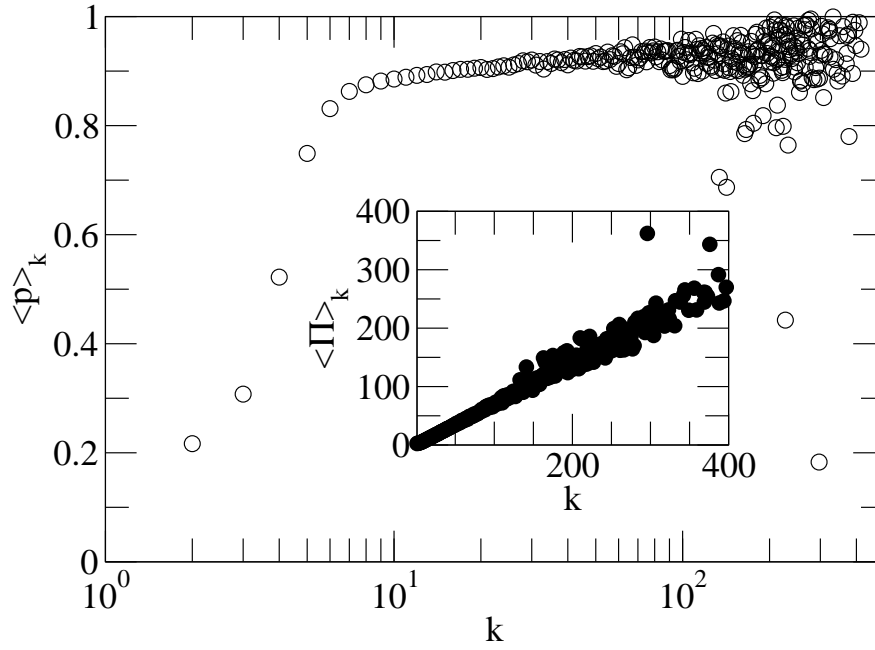


Figure 3.10: The main figure shows the average offers of proposers, $\langle p \rangle_k$, as a function of k in scale-free networks after 10^7 time steps when using the configuration $q_i = 1 - p_i$. The curve indicates that highly connected individuals offer very high portions of the stake, a signal of altruistic behavior.

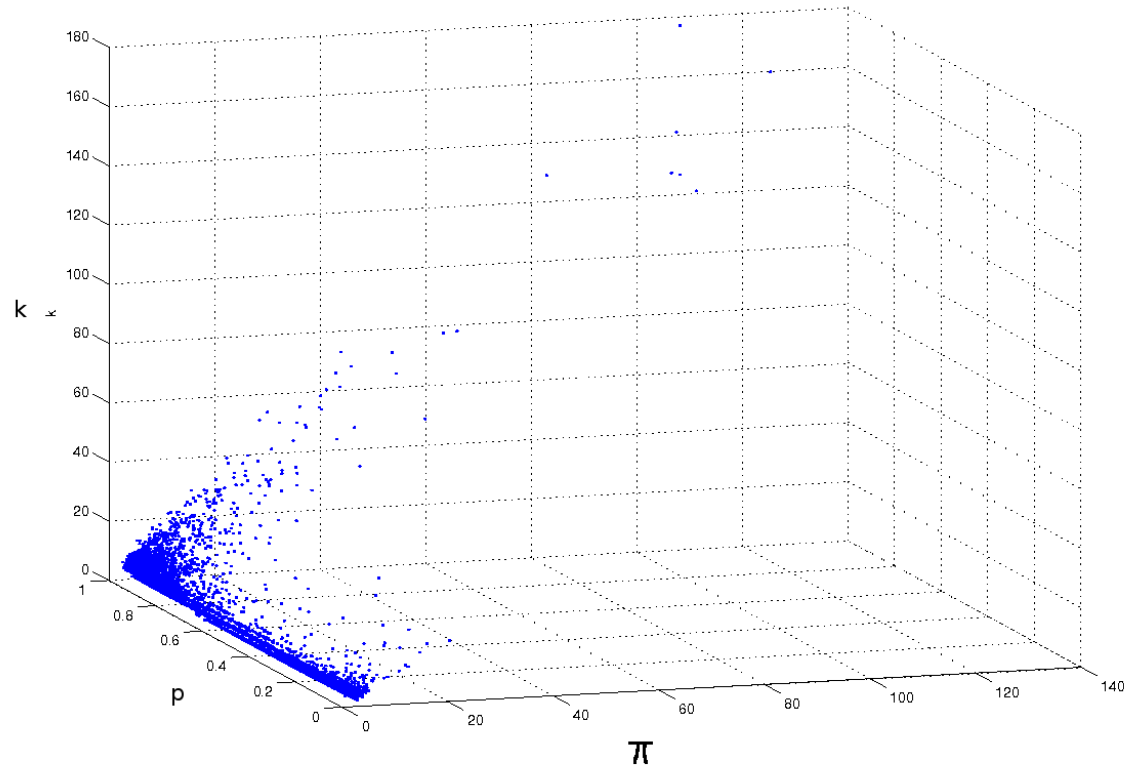


Figure 3.11: The tridimensional version of fig. 3.10, seen as a scatter plot. The figure shows the offers p of proposers as a function of k in scale-free networks after 10^7 time steps when using the configuration $q_i = 1 - p_i$. The points curve indicate that highly connected individuals offer very high portions of the stake, a signal of altruistic behavior. Despite this, they continue in having higher payoffs than other nodes.

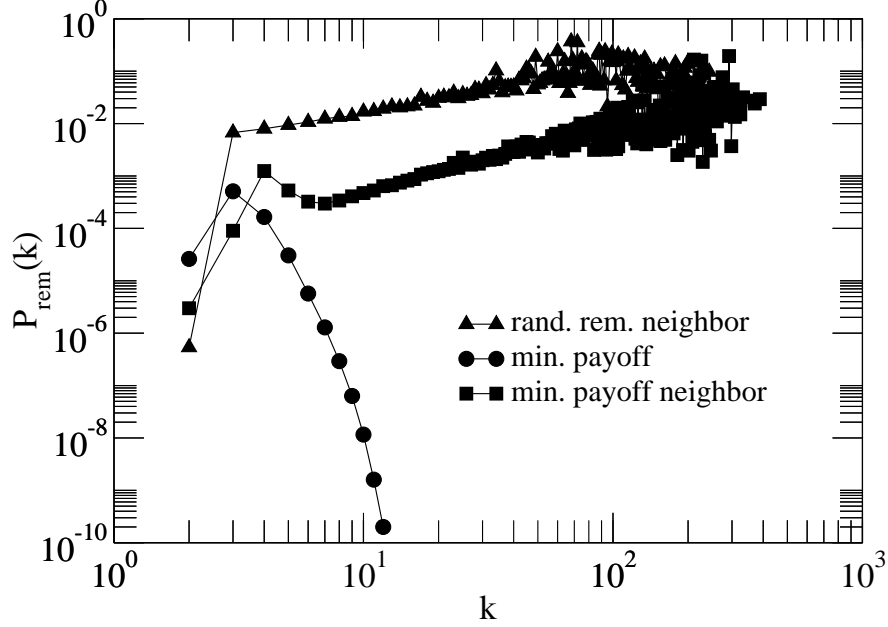


Figure 3.12: Effects of the individual selection mechanism. We have plotted the probability of removing players from the game as a result of being the less fitted individual (circles) or by being a neighbor of the latter (squares) as a function of their degree k . We have also represented the same probability but considered that bargainers are removed from the network by a random deletion process (triangles). The results show that highly connected nodes increases their chances of remaining in the game by roughly and order of magnitude. The underlying network of contacts is in all cases a SF network with the configuration $q_i = 1 - p_i$.

changes, we have implemented the UG on graphs created with the Holme-Kim model (see 1.4.4).

After averaging over an ensemble of graphs with different initial conditions, we have looked at the same quantities as done for graphs created with the BA model. It came out that the game is robust in respect to these topological characteristics, showing that there are not significant changes (in p distribution, payoff distribution, removed nodes, etc) compared to scale-free graphs without clustering.

The only result we report here is a scatter plot of the values of p vs the degree k for different values of the clustering coefficient, proportional to the values of mt

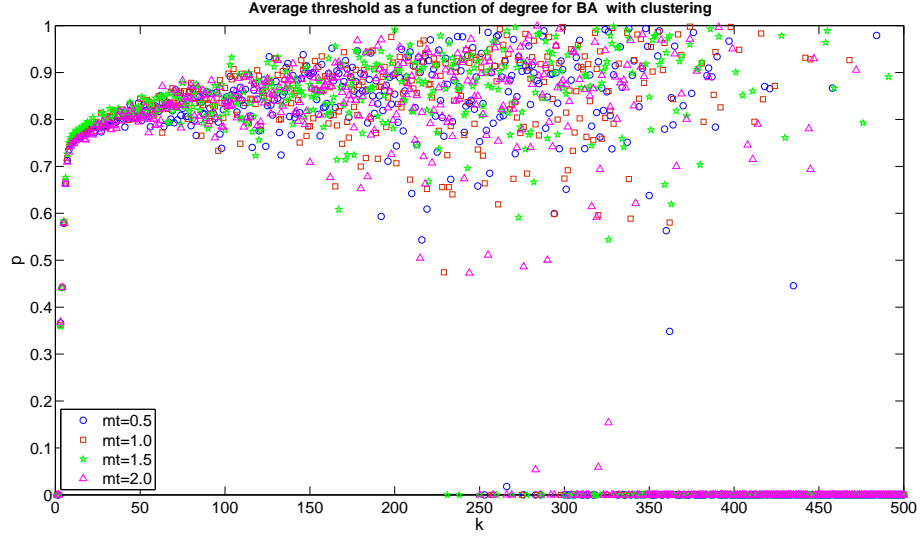


Figure 3.13: Scatter plot of the values of p vs the degree k for scale-free graphs with different values of clustering coefficient mt (configuration used: $q_i = 1 - p_i$).

listed in the legend (fig. 3.13). It is clear that there aren't significant differences varying mt .

3.5 Eguíluz and Tessone's model

A paradigmatic model similar to the Ultimatum Game, but non exactly the same one, has been studied by Eguíluz and Tessone [36]. We have implemented the same model and reproduced their results which are deeply different from the ones obtained analyzing the standard Ultimatum Game. We report this study in order to show how sensitive the system is to the set of variables chosen.

3.5.1 The model

A set of N agents arranged in the nodes of a network. An amount of 1 unit is set to be shared in each interactions between two agents. Each agent i is characterized by a threshold $p_i \in [0, 1]$: as responder it indicates the minimum amount she will accept; as proposer, it also defines the amount of money she will offer.

The model runs as follows: at each time step all agents play with all their neighbors synchronously. Thus for each interaction link between two neighbors agents (i, j) :

1. If the offer p_i is above the threshold of agent j , that is to say $p_i > p_j$, then the offer p_i is accepted: agent j increases his payoff Π_j by p_i while agent i 's payoff Π_i increases by $1 - p_i$. Therefore $\Delta\Pi_{ij}^O = (1 - p_i)$ and $\Delta\Pi_{ji}^R = p_i$.
2. Otherwise if the offer p_j is above the threshold of agent i , $p_j > p_i$, then the offer p_j is accepted: agent i 's payoff increases by p_j while agent j 's payoff increases by $1 - p_j$. In formulas: $\Delta\Pi_{ji}^O = (1 - p_j)$ and $\Delta\Pi_{ij}^R = p_j$.

Note that we this rules in each interaction every agent increases his payoff, whatever is the relation between his threshold and his neighbors' ones. On this relation only depends the amount she earns. In this case there is a complete symmetry between the two agents in the sense that there is no distinction between the roles

of proposer and responder: in each interaction the highest offer, that is to say the highest value of p , is chosen.

The payoff obtained by the agent i after interacting with all his neighbors $l \in \Gamma_i$ is given by $\Pi_i = \sum_{l \in \Gamma} (\Delta \Pi_{il}^O + \Delta \Pi_{il}^R)$. After each round, a selection rule is applied to the system: the agent with the lowest payoff in the population and its immediate neighbors, determined by the topology of the network, are replaced by new agents with randomly chosen new thresholds. The payoffs of all agents are then reset to zero and the system is let evolve again.

3.5.2 Results

All the results presented in this section are obtained using networks holding a scale-free degree distribution. In order to do this, the Barabási-Albert algorithm has been used (see 1.4.3 and [18]), setting $m = 2$.

In fig. 3.14, the stationary distribution of payoffs in the population is plotted. There is a well defined critical payoff, below which the agents are removed. This critical value is $\Pi_c = 1.75$. The results suggest that the system self-organizes in a critical state where the distribution of avalanches is also a power law. The critical state would emerge despite the non uniform distribution of threshold and payoffs (see fig. 3.14 and fig. 3.17). An avalanche is typically defined as follows: it starts when the lowest payoff gets larger than a preset value Π_c and stops when it drops below this value. The size, s , of an avalanche is thenumber of time steps it lasts. In fig. 3.15 we show the distribution of avalanche sizes when we use a payoff $\Pi_c = 1.75$ as the indication of an avalanche. The probability distribution displays a power-law decay:

$$P(s) \sim s^{-\alpha} \quad (3.4)$$

with an exponent $\alpha = 1.67$.

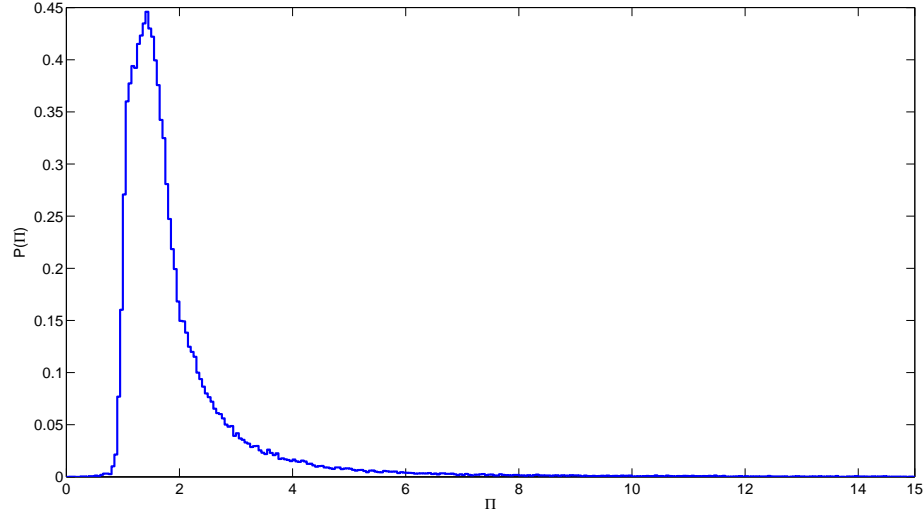


Figure 3.14: Asymptotic distribution of the payoff after 10^6 time-steps averaged over 10^2 different scale-free graphs.

In order to characterize further the dynamics, we have measured also the first return time distribution, $P_f(t)$, the time elapsed between two mutations affecting the same agent. The results are plotted in fig. 3.18 and it comes out that the tail of the distribution is well fitted also by a power-law:

$$P_f(t) \sim t^{-\alpha} \quad (3.5)$$

3.5.3 Discussion

The model proposed by Eguíluz and Tessone tries to explain altruistic behavior by introducing local interactions which lead the system to self-organization in a critical state. The model anyway is not a proper Ultimatum Game.

The amount of altruistic behavior is reflected in the distribution of p in the population. In scale-free networks the distribution of thresholds displays a maxi-

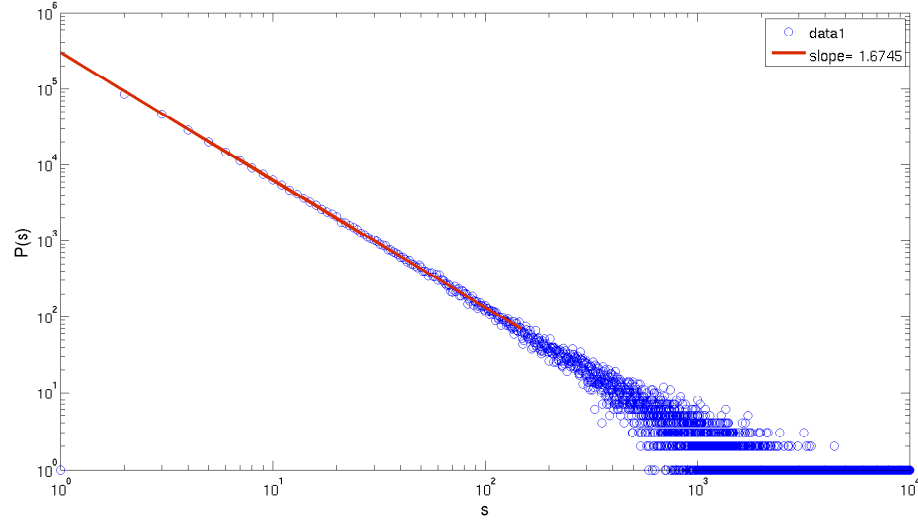


Figure 3.15: Avalanche size $P(s)$ distribution of the payoff after 10^6 time-steps averaged over 10^2 different scale-free graphs. The solid line is a power-law fit with exponent $\alpha = 1.67$.

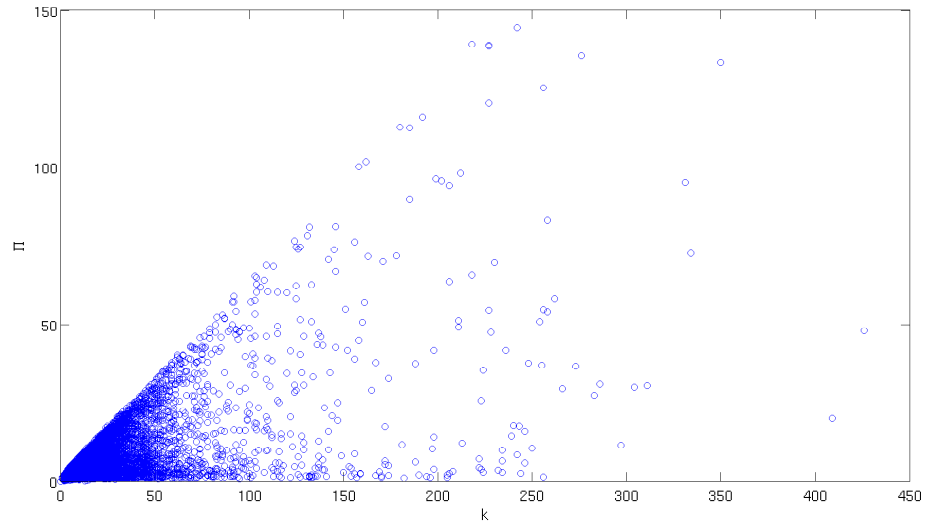


Figure 3.16: Scatter plot payoff vs degree. Each point stands for an agent, and its position on the plan indicates its degree (x -axis) and its payoff at the 10^6 th time step.

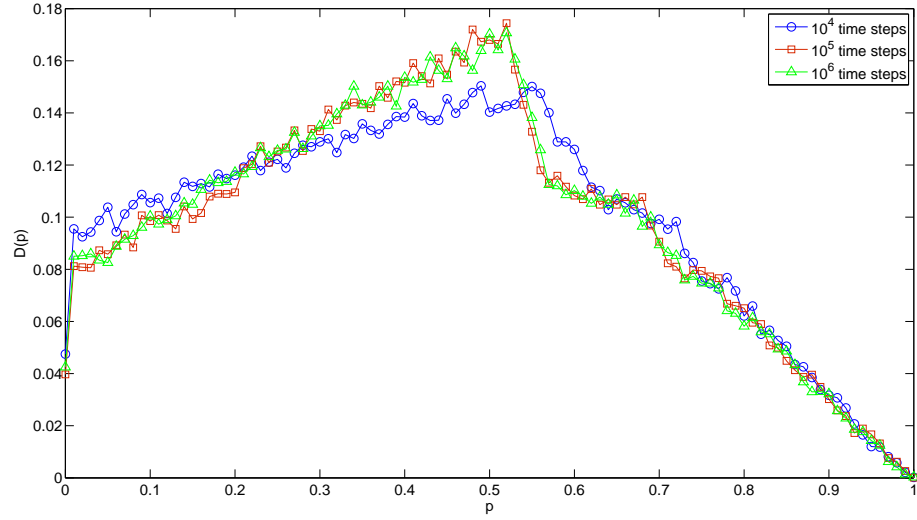


Figure 3.17: Distribution of the values of p after 10^4 (blue circles), 10^5 (red squares) and 10^6 (green triangles) time steps.

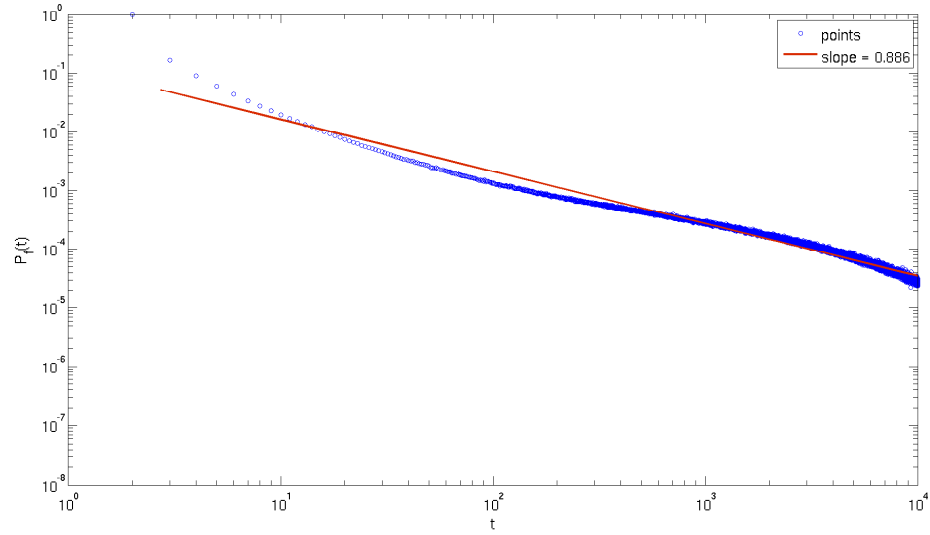


Figure 3.18: First return time distribution. $P_f(t)$ is the distribution of time intervals occurring between two consecutive removals of the same node in the graph.

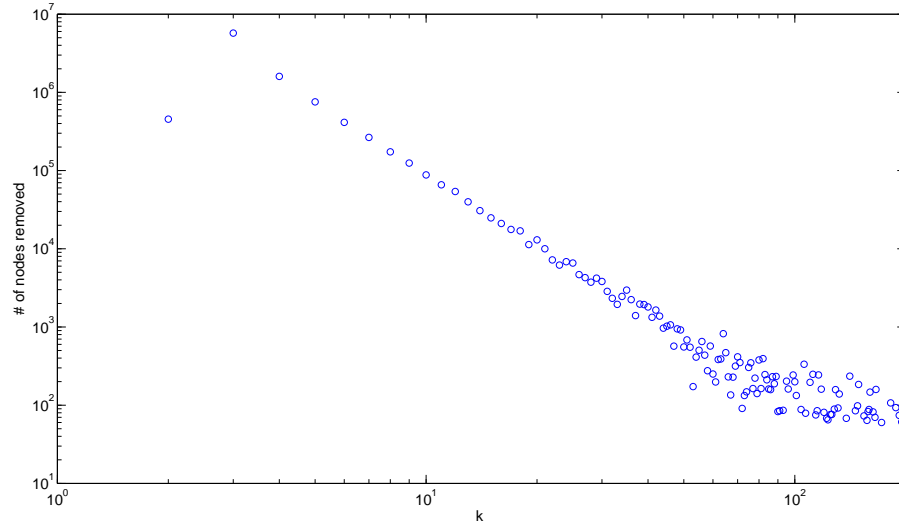


Figure 3.19: Degree distribution of nodes with minimum payoff.

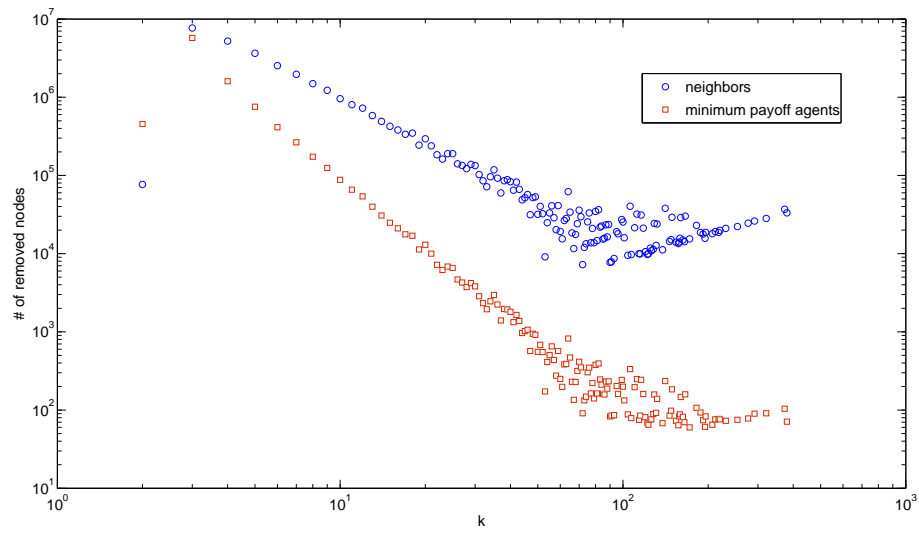


Figure 3.20: Degree distribution of nodes removed because of the minimum payoff (blue circles) or because neighbors of minimum payoff nodes (red squares).

mum around a value of 50%, decaying for lower and larger p . This case captures only some experimental findings.

From a dynamical view point, the distribution of avalanches displays a power-law scaling with an exponent that depends on network topology. This feature is a typical signature of *self-organized criticality*.

Conclusions

*A conclusion is the place
where you got tired of thinking.*

MARTIN HENRY FISCHER

We have studied an ultimatum game without empathy in ER and SF networks. The ultimatum game is a model extensively used to explain cooperative behavior. Implementing it on graphs, we have seen that a particular equilibrium arises in SF networks, usually more interesting than other topologies because of their role in modelling real-life interactions. In this kind of networks when playing an UG with an evolution rule based on altruistic punishment, it comes out a sort of “reciprocal altruism”. The existence of different connectivity classes in SF networks is at the root of this behavior. In other words, the fate of each node is determined by its interaction with low and/or highly connected nodes. We stress that altruism is an emergent property of the system as the rules are stochastic and no memory mechanism is explicitly assumed. Admittedly, altruism arises in a self-organized manner with selection acting locally: highly connected agents optimize their chances to survive by increasing their generosity, without risking at all being the poorest in town, due to bounds in Eq. (3.3). In support of the validity of this argument we have also calculated the probability that a node with degree k is removed (figure not shown), distinguishing the two possible causes for removal at the selection stage: (i) as a consequence of being the node with the

lowest payoff in the population, *(ii)* as a consequence of being a neighbor of the less fitted individual. It turns out that highly connected nodes are never exposed to direct extinction and their probability of being removed as a consequence of being neighbors of the less fittest individual is one order of magnitude lower than that due to random removal.

The mere possibility of the observed bimodal distribution density lies in the degree heterogeneity of the network. This strategic segregation into low k -low p and high k -high p agents is a genuine topological effect. Our results also provide new insights into the relation between the behavior of wealthy individuals and altruism. Very recently [43], experiments in which in each round of a repeated game players have three available strategies have been conducted. The participants were able to choose between cooperation, defection or costly punishment. The authors found that those individuals with the highest payoffs are not likely to get involved in punishment acts. In other words, winners don't punish. In the context of the model discussed here, our results suggest a different mechanism by which wealthy individuals can afford full generosity. The abundance of highly altruistic individuals does not arise due to reputation [30], nor costly individuals' punishment [43], but from a purely scale free effect combined with a social reinforcement of altruism. In other words, if you are well connected, you are assured a minimal payoff above the lower bound and then you can increase your likelihood of survival by helping those that can drop you out of the game. Additionally, the net result of this highly polarized strategic distribution is that extreme egoists ($p \leq 0.2$) and mild-generous ($0.5 \leq p \leq 0.8$) agents are looser strategists and are outcompeted by the rest of the population.

In any kind of network, anyway, our results show that typical offers are more generous than those expected from a rational player. And, as the most important result, we observe that altruism naturally emerges in SF networks as a

consequence of their heterogeneous degree distribution together with social enforcement: highly connected players survive better if they give away a large part of the reward for the benefit of their lowly connected neighbors, which now have a chance of increasing their total rewards.

Acknowledgements

La gratitudine è la più squisita forma di cortesia.

FRANÇOIS DE LA ROCHEFOUCAULD

Though the entire work has been written in English, I can't write in a foreign language when dealing with something so personal like acknowledgments and feelings of gratitude...

Per cui passo all'italiano, la mia lingua. Appunto perché non si può usare una lingua diversa dalla propria lingua madre, se si vuole esprimere qualcosa di tanto personale e che viene dal cuore, come la riconoscenza e la gratitudine; diversamente mi sembrerebbe un artefatto.

Questo lavoro è nato e cresciuto (ma non morto! :-P) a Zaragoza, presso l'Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), sotto l'attenta e paziente supervisione del Prof. Yamir Moreno. Devo a lui se questo lavoro ha potuto prendere forma e ha potuto raggiungere un livello tale da poter essere sottomesso alla comunità scientifica internazionale. Grande riconoscenza va naturalmente anche agli altri due supervisor di questo lavoro, cioè al Prof. Vito Latora e al Dott. Jesús Gómez Gardeñes, i quali mi hanno continuamente consigliato e impartito insegnamenti su come procedere. In particolare al primo, al mio "maestro", va il mio grazie più sentito per avermi insegnato tutto ciò che ho

appreso riguardo la ricerca e per avermi trasmesso la passione per questo mondo.

Un particolare riconoscimento va anche al Prof. Mario Floría e alle utili conversazioni avute con lui, grazie alle quali questo lavoro è ulteriormente cresciuto e maturato.

Un sentito ringraziamento va anche a tutto lo staff e agli amici del BIFI, non solo per aver agevolato il mio lavoro, ma anche per aver facilitato e allietato la mia permanenza in Spagna.

Questo lavoro conclude anche formalmente il mio iter di allieva della Scuola Superiore di Catania, all'interno della quale ho trascorso i miei anni universitari. La Scuola ha significato tanto per la mia formazione e per la mia crescita accademica; voglio pertanto ringraziare tutti coloro che ogni giorno lavorano per mandarla avanti e perché funzioni al meglio: grazie a tutti i membri della segreteria, degli uffici, della reception, al personale di pulizia, allo staff della mensa... E chiedo perdono se ho dimenticato qualcuno.

Tuttavia ciò che di più bello la scuola mi ha dato non riguarda il percorso accademico né la formazione in senso stretto, ma l'esperienza della vita di residenza. Durante questi anni ho condiviso la mia vita di ogni giorno, le conquiste, le delusioni, le gioie, i malumori, i successi, con tutti gli altri ragazzi e ragazze che, come me, affrontavano il percorso universitario alla SSC. Con loro sono cresciuta, passando da diciottenne appena diplomata, a laureata alle prime prese con il mondo del lavoro. Ognuno di loro mi ha lasciato qualcosa e non sarei quella che sono senza questa esperienza. A tutti loro sono affezionata e riconoscente, non per meriti specifici, ma per aver costituito per cinque anni la mia seconda "famiglia". Tra di loro, un grazie speciale per tutte le risate e l'allegria regalatemi, sia nei momenti lieti che in quelli difficili, va a Federico, Alessandra, Alessia, Filippo B., Daniele M., Salvo S., Violetta, Salvo A., Antonio, Elio, Ivan, Mimmo, Luca,

Agata, Ivano, Mario T., Beniamino, Paolo, Rachele, Ciccio P., Pantaleo, Mario S., Tony, Emanuele, Marco L., Paolo G.: grazie a tutti voi!

Il contributo maggiore, tuttavia, perché questo iter fosse portato a compimento è di coloro che hanno sempre creduto in me e che per me hanno sempre tifato, perché senza "supporters" è impossibile andare avanti... Quindi grazie Mamma, Claudia e Emma, Donne della mia vita. I miei successi sono anche vostri e a voi va sempre il mio pensiero e la mia gratitudine.

Bibliography

- [1] R. Dawkins, *The Selfish Gene*, 3rd edition, Oxford University Press, Oxford (2006).
- [2] R. Sinatra, Y. Moreno, J. Gómez Gardéñes, V. Latora and M. Florà, arXiv:0807.0750v1 (2008).
- [3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, D. U., *Phys. Rep.* **424**, 175 (2006).
- [4] B. Bollobàs, *Modern Graph Theory*, Graduate Texts in Mathematics, Springer, New York (1998).
- [5] D. B. West, *Introduction to Graph Theory*, Prentice-Hall, Englewood Cliffs, NJ (1955).
- [6] M. Boguñá and R. Pastor-Satorras, *Phys. Rev. E* **66**, 047104 (2002).
- [7] K. Faust Wasserman, *Social Networks Analysis*, Cambridge University Press, Cambridge (1994).
- [8] M.E.J. Newman, *SIAM Review* **45**, 167 (2003).
- [9] D.J. Watts and S.H. Strogatz, *Nature* **393**, 440 (1998).

-
- [10] J. Scott, *Social Network Analysis: A Handbook*, (Sage Publications, London, 2nd edition, 2000).
 - [11] M. Marchiori and V. Latora, *Physica A* **285**, 539 (2000).
 - [12] V. Latora and M. Marchiori, *Phys. Rev. Lett.* **87**, 198701 (2001).
 - [13] V. Latora and M. Marchiori, *Eur. Phys. J.* **B32**, 249 (2003).
 - [14] V. Latora and M. Marchiori, Preprint cond-mat/0402050.
 - [15] M. E. J. Newman and M. Girvan, *Phys. Rev. E* **69**, 026113 (2004).
 - [16] S. Milgram, *Psychol. Today* **1**, 60 (1967).
 - [17] J. Travers and S. Milgram, *Sociometry* **32**, 425 (1969).
 - [18] A.L. Barabási and R. Albert, *Science*, **286**, 509 (1999).
 - [19] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
 - [20] P. Erdős and A. Rényi, *Publ. Math. Debrecen* **6**, 290 (1959).
 - [21] A.-L. Barabási, R. Albert and H. Jeong, *Physica A* **272**, 173 (1999).
 - [22] P. L. Krapivsky, S. Redner and F. Leyvraz, *Phys. Rev. Lett.* **85**, 4629 (2000).
 - [23] P. Holme and B.J. Kim, *Phys. Rev. E* **65**, 026107 (2002).
 - [24] S. J. Gould, *The Structure of Evolutionary Theory*, Harvard University Press, Cambridge, 2002.
 - [25] G. Szabó and G. Fáth, *Phys. Rep.* **446**, 97-216 (2007).
 - [26] R. Selten, *Spieltheoretische behandlung eines oligopolmodells mit nachfrage-tragheit*, *Z. Gesamte Staatswiss* **121**, 301-324.

-
- [27] W. D. Hamilton, *J. Theor. Biol.* **7**, 1 (1964).
 - [28] G. Vogel, *Science* **303**, 1128 (2004).
 - [29] M. A. Nowak and K. Sigmund, *Nature* **393**, 573 (1998).
 - [30] M. A. Nowak, K. M. Page, and K. Sigmund, *Science* **289**, 1773 (2000).
 - [31] M. Milinski, D. Semmann, and H. J. Krambeck, *Nature* **415**, 424 (2002).
 - [32] M. A. Nowak *Evolutionary Dynamics. Exploring the Equations of Life*. (Harvard University Press, Harvard, MA) (2006).
 - [33] E. Fehr and S. Gächter, *Nature* **415**, 137 (2002).
 - [34] H. Gintis, S. Bowles, R. Boyd, and E. Fehr, *Evol. Hum. Behav.* **24**, 153 (2003).
 - [35] E. Fehr and U. Fischbacher, *Nature* **425**, 785 (2003).
 - [36] V. M. Eguiluz and C. J. Tessone, *arXiv:0705.1976 [nlin.AO]* (2007).
 - [37] A. Sánchez and J. A. Cuesta *J. Theor. Biol.* **235**, 233 (2005).
 - [38] W. Güth, R. Schmittberger, and B. Schwarze, *J. Econ. Behavior and Organization* **3**, 367 (1982).
 - [39] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis and R. McElreath, *Am. Econ. Rev.* **91**, 73 (2001).
 - [40] C. F. Camerer, *Behavioral Game Theory*. (Russell Sage Foundation, Princeton University Press, Princeton, NJ) (2003).
 - [41] M. E. J. Newman *SIAM Rev.* **45**, 167 (2003).
 - [42] P. Bak and K. Sneppen, *Phys. Rev. Lett.* **71** 4083 (1993).

-
- [43] A. Dreber, D. G. Rand, D. Fudenberg, and M. A. Nowak, *Nature* **452**, 348 (2008).
 - [44] J. Gómez-Gardeñes and Y. Moreno, *Phys. Rev. E* **73**, 056124 (2006).
 - [45] J. Gómez-Gardeñes, M. Campillo, L. M. Floría, and Y. Moreno, *Phys. Rev. Lett.* **98**, 108103 (2007).
 - [46] F. C. Santos, and J. M Pacheco, *Phys. Rev. Lett.* **95**, 098104 (2005).
 - [47] F. C. Santos, J. M. Pacheco, and T. Lenaerts, *Proc. Nat. Acad. Sci. USA* **103**, 3490 (2006).
 - [48] J. Gómez-Gardeñes, J. Poncela, L. M. Floría, and Y. Moreno, *J. Theor. Biol.* **253**, 296 (2008).